Automation of concept induction based CNN neuron interpretation using high performance computing

by

Samatha Ereshi Akkamahadevi

B.E., University B D T College of Engineering, India, 2016

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Computer Science
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Approved by:

Major Professor
Dr. Pascal Hitzler

Copyright

Abstract

Interpreting hidden neuron activations in convolutional neural networks (CNNs) is a critical objective in the field of Explainable Artificial Intelligence (XAI), aiming to transform "black box" models into transparent systems. Previously, our research group addressed this challenge by employing concept induction and semantic reasoning using a concept hierarchy derived from the Wikipedia knowledge graph. However, that process was manual and required several days to complete, limiting its practicality and scalability. In this study, we introduce a fully automated pipeline that streamlines model training, data preparation, concept induction, image retrieval, classification, and statistical validation, reducing execution time from days to approximately 1 hour and 30 minutes while ensuring consistent, reproducible results. The automation enables flexible adjustments, such as incorporating a broader range of training images and classes and examining additional concept induction results across various neuron layers with minimal code changes. The automation efficiently processes large datasets by utilizing parallel processing capabilities on high-performance computing resources, ensuring that the pipeline operates quickly and effectively. The results confirm that our automated approach effectively bridges the gap between deep learning models and human-understandable concepts, enhancing model interpretability without compromising performance. This work advances the field of Explainable AI (XAI) by offering a practical solution that balances interpretability with computational efficiency, contributing to the broader goal of making AI systems more transparent and trustworthy.

Table of Contents

List of Figures

Nomenclature

Acronyms and Abbreviations

1. AI – Artificial Intelligence
2. XAI - Explainable AI
3. CNN - Convolutional Neural Networks
4. ECII - Efficient Concept Induction from Instances
5. HPC – High-Performance Computing
6. SLURM – Simple Linux Utility for Resource Management
7. CPU – Central Processing Unit
8. GPU – Graphics Processing Unit
9. RAM – Random Access Memory
10. I/O – Input/Output
11. MaxRSS – Maximum Resident Set Size

Neural Network Terms

1. Neuron Activation – The process by which a neural network node (neuron) becomes activated in response to input data.
2. Target Activation – Percentage of activations corresponding to the expected class (target).
3. Non-target Activation – Percentage of activations corresponding to the unexpected or non-target class.
4. Label Hypotheses – Predicted labels for neuron activations based on concept induction.
5. Coverage Score – A metric indicating the extent to which a neuron covers or explains the input data.

Statistical Terms

1. p-value – The probability of obtaining a result at least as extreme as the one observed, under the assumption that the null hypothesis is true.
2. z-score – A measure of how many standard deviations an element is from the mean.
3. Mean Activation – The average activation value for a group of neurons.

4. Median Activation – The middle value in the sorted list of neuron activations.
5. Mann-Whitney U Test – A statistical test used to determine whether there is a significant difference between two independent groups.

Computational Terms

1. Batch Processing – A method of running multiple tasks simultaneously on a computer system.
2. Parallelization – Distributing tasks across multiple CPUs or nodes to reduce overall computation time.
3. Job Array – A SLURM feature that allows the execution of multiple independent tasks simultaneously.
4. SLURM – A job scheduling system used to manage computing resources on high-performance clusters.
5. MaxRSS – Maximum Resident Set Size; a measure of the maximum memory usage by a job.
6. Beocat – The high-performance computing cluster used for executing neural network training, processing, and statistical analysis tasks.

Acknowledgements

I would like to express my deepest gratitude to my major advisor, Dr. Pascal Hitzler, for his unwavering support, insightful guidance, and continuous encouragement throughout the course of this research. I am especially grateful to him for providing me with the opportunity to work as a Graduate Research Assistant, where I was able to focus on this automation project. His expertise and vision were invaluable in shaping the direction of this work.

I extend my heartfelt thanks to my committee members, Dr. Mitchell Larry Neilsen and Dr. Torben Amtoft, for their constructive feedback, advice, and support. Their expertise and thoughtful insights significantly contributed to the refinement and success of this research.

I would also like to thank Abhilekha Dalal for her thoughtful feedback, valuable reviews, and constant encouragement during the course of this research. Her insights helped me improve and refine various aspects of my work.

I would also like to thank Dr. Scott DeLoach, Head of the Department, for his leadership and support throughout my Master's Program at Kansas State University. I am deeply appreciative of my academic advisor, Sheryl Cornell, for her ongoing guidance and for providing invaluable support throughout my academic journey. A special thanks to Dr. Bill Hsu for his helpful clarifications, which helped me in overcoming specific challenges I encountered during this project. I am particularly grateful to Dr. Dennis Lang for the opportunity to work as a Graduate Teaching Assistant for the Programming Fundamentals course, which not only broadened my teaching experience but also enhanced my coding skills significantly.

Finally, I acknowledge the entire faculty and staff of the Department of Computer Science at Kansas State University for providing a supportive academic environment that has greatly contributed to my personal and professional growth.

Dedication

I dedicate this thesis to my husband, Dr. Rakshith Kumar, whose boundless encouragement and wisdom have been a guiding light throughout my academic journey. His unwavering belief in my abilities, paired with his thoughtful guidance, has given me the strength to overcome every challenge I faced. Rakshith's emotional support has been my anchor, and his intellectual insights have continually inspired me to push beyond my limits. His love, patience, and belief in my potential made this achievement possible, and for that, I am deeply grateful.

I also dedicate this work to my beloved parents, Ereshi R and Akkamahadevi D M, whose unwavering love and belief in me have been the foundation of all my successes. Their constant support, sacrifices, and encouragement have carried me through every challenge, and their guidance has shaped the person I am today. A heartfelt thank you to my wonderful sisters, Srinidhi A E and Sinchana A E, whose endless love and joyful presence have always filled my life with happiness and motivation. I am also thankful to my dearest friends, Cathrine Joy, Priyanka Bijjargi, and Harish Kammar, who stood by my side through every step of my life.

I am deeply grateful to my in-laws G R Kumar and Pushpa Kumar for their constant support and encouragement throughout this journey.

A special mention to Dr. M. M. Chengappa and Muthu Chengappa, who have shown me kindness and guidance akin to family. Their constant support and encouragement have always been a source of strength to me.

I would also like to extend my gratitude to my computer science friends, Mana Ranabhat, Kavya Kompella, Anika Nawar, Niketa Penumajji, Annet Kolenchery, Raviteja Vempati, Bharaneeshwar Balasubramaniyam, and Pavan Kumar Reddy Lakkireddy, whose camaraderie and shared experiences have made this journey all the more rewarding.

Chapter 1 - Introduction and Related Work

Artificial Intelligence (AI) has become an integral part of our daily lives, influencing decisions in various subtle yet impactful ways. From personalized product recommendations on platforms like Amazon and Netflix to friend suggestions on Facebook and targeted advertisements on Google, AI systems operate seamlessly behind the scenes, shaping our digital experiences. While we have become accustomed to these automated decisions, the need for transparency becomes crucial when AI is applied to life-changing decisions, such as disease diagnosis. In these contexts, it is essential to understand not only the outcome but also the reasoning behind AI-driven conclusions. This highlights the importance of Explainable AI (XAI), which aims to make AI systems more interpretable and trustworthy by offering insights into their decision-making processes[1].

Explainable Artificial Intelligence (XAI), first coined in 2004 by Van Lent et al. [2] refers to a collection of techniques and processes aimed at making the decision-making mechanisms of machine learning models transparent and understandable to humans [3], [4]. It provides insights into how AI models arrive at their predictions or decisions, enabling users to comprehend and trust the outcomes. XAI is crucial for applications where accountability, fairness, and interpretability are important, as it bridges the gap between the "black-box" nature of AI models and human understanding, fostering confidence in AI-driven systems [3], [5].

There is a growing interest in Explainable AI (XAI) within the industrial sector, with several leading companies developing solutions to make AI more transparent. H2O.ai offers its driverless AI product, while Microsoft has integrated XAI into its next-generation Azure ML Workbench. Kyndi provides an XAI platform tailored for government, financial services, and healthcare, and FICO applies XAI to its credit risk models [1].

Deep learning is a branch of artificial intelligence (AI) that focuses on training large-scale neural networks to learn and make predictions from data autonomously, without being explicitly programmed for specific tasks [6]. It has become a driving force behind innovations in various domains, from healthcare and autonomous vehicles to voice recognition and personalized recommendations. The ability to learn from vast datasets and make decisions without human intervention has driven technological breakthroughs across industries [7]. These models consist of multiple layers of interconnected neurons, processing data hierarchically, capturing increasingly complex features and patterns as they progress through each layer [8], [9].

A fundamental component of these neural networks is the neuron, which serves as the basic computational unit. Neurons are organized into layers, with each layer receiving and processing inputs before passing outputs to the next layer. Between these, hidden layers process the data while passing information from one neuron to the next. A neuron in any layer receives input signals, processes them using a mathematical function, and generates an output. Each output has a specified threshold or weight, and if the output exceeds this threshold, the neuron is activated, passing data to the next neuron in the network; otherwise, no data is transmitted. This transmission through layers enables neural networks to learn complex patterns and make accurate predictions from the input data [10].

Convolutional Neural Networks (CNNs) represent a specialized form of deep learning optimized for tasks involving spatially structured data, such as image recognition and computer vision. It uses convolutional layers to extract local features from smaller regions of the input. Each convolutional layer applies a set of filters (or kernels) to these regions, effectively capturing important spatial features like edges, textures, and shapes. As the network deepens, these layers learn more abstract patterns, allowing the model to recognize increasingly complex

structures in the data [11]. LeNet-5 was the first CNN architecture developed using backpropagation algorithm [12], and later, many CNN architectures were introduced to improve performance and handle more complex tasks. Some notable ones include AlexNet, ZFNet, VGGNet, GoogleNet, and ResNet. ResNet is remarkable for its greater depth and its ability to handle complex data representations [11], [13].

While CNNs have revolutionized fields like image recognition, their lack of interpretability has raised concerns, especially in applications requiring transparency. This has led to growing interest in Explainable AI (XAI). One approach to enhancing the interpretability of these AI models is semantic reasoning, which enables AI systems to infer new information based on existing data and predefined inference rules or ontologies. Unlike traditional machine learning approaches that rely solely on data-driven patterns, semantic reasoning introduces an additional layer of intelligence by utilizing structured knowledge bases, such as ontologies and knowledge graphs.

Knowledge bases provide a formalized structure for understanding relationships between concepts, allowing AI systems to make inferences that go beyond the data explicitly provided. By integrating this structured knowledge into the reasoning process, semantic reasoning allows AI models to generate contextually relevant and interpretable conclusions. For example, a system trained to recognize objects in images might not only detect a "car" but also infer that it is a "vehicle" or identify its relationship to other objects in the image, such as "roads" or "traffic signs." This capacity to reason about relationships and abstract concepts enriches the raw data with deeper, human-understandable insights, making the AI model’s decisions more transparent and meaningful. This approach, often termed Semantic AI [14].

Knowledge graphs and ontologies play a pivotal role in enhancing the interpretability of AI models, especially in the context of semantic reasoning and concept induction. A knowledge graph organizes information in a structured manner, representing relationships between entities in a way that mimics human understanding. Ontologies, on the other hand, define a set of concepts and categories within a domain and the relationships between them. Here, in our research, Wikipedia concept hierarchy serves as a vast knowledge base from which human-understandable concepts can be derived to explain neuron activations. By integrating such structured knowledge, AI models can move beyond mere statistical patterns and offer more meaningful, context-aware interpretations of their behavior[15]. This combination of structured background knowledge with deep learning enhances the transparency and explainability of complex AI systems, making it possible to associate abstract neural activations with understandable concepts.

Li et al. (2019) utilized semantic reasoning by employing Graph Convolutional Networks (GCNs) to model how different image regions are connected. Their approach captured key objects and their relationships between them and refining these connections through a broader reasoning process which helped to create a comprehensive visual representation that effectively aligned with text descriptions [16]. Dalal et al. (2023) further advanced the application of semantic reasoning by using concept induction based on description logics. They employed Wikipedia Concept Hierarchy to hypothesize and verify neuron activations in Convolutional Neural Networks (CNNs). This approach enabled the generation of meaningful explanations for hidden neuron activations, making these deep learning models a little less of a "black box" by associating neuron responses with human-understandable concepts [17], [18], [19]. Building on this foundation, our research automates this entire interpretability process to improve efficiency

and ensure scalability across large datasets. By leveraging Beocat's high-performance computing infrastructure, we eliminate the need for manual intervention, enabling large-scale analysis of CNN neurons and creating a fully automated pipeline for neuron interpretation.

Concept induction, a key element in semantic reasoning, is a method used to identify and generate meaningful concepts or categories based on patterns observed in data. It plays a critical role in making AI systems interpretable by systematically associating data points with understandable labels. It typically involves using a predefined set of rules and a large-scale knowledge base, such as ontologies or hierarchies, to systematically associate data points with human-understandable labels. The core idea behind concept induction is to enable an AI system to infer new concepts or refine existing ones based on observed relationships, ultimately making the system's behavior interpretable.

Sarker et al. (2020) utilized concept induction as a technique to generate complex class expressions from a description logic knowledge base. They applied this approach to distinguish between sets of positive and negative examples by creating expressions that correctly represent positive cases while excluding negative ones. They introduced an algorithm named ECII to enhance the efficiency of concept induction. It was designed to improve runtime efficiency by combining atomic concepts to create more complex expressions, without relying on traditional refinement operators. They tested ECII on various datasets, including the ADE20K dataset, and demonstrated that it significantly reduced processing time by two to three orders of magnitude while maintaining comparable levels of accuracy [20], [21].

Automation offers numerous advantages across various industries, improving efficiency, accuracy, and scalability while reducing human error and operational costs. By streamlining repetitive tasks, automation allows human workers to focus on more strategic and creative

endeavors, enhancing overall productivity. One of the primary benefits is the ability to handle large volumes of data or processes quickly and consistently, making automation indispensable in industries that demand precision and speed. For example, in manufacturing, automation has revolutionized production lines with robotics and AI-driven systems, significantly boosting output while ensuring consistent product quality. In software development, automation tools like continuous integration and deployment (CI/CD) pipelines have accelerated the release of reliable software updates, reduced time-to-market and minimizing manual intervention. Similarly, in the healthcare sector, automation in diagnostic systems and administrative tasks has improved patient care by speeding up diagnosis and reducing paperwork, enabling healthcare professionals to focus on critical aspects of patient treatment.

Manually interpreting and analyzing the activations of neurons in a CNN is a laborious and time-intensive task, often involving significant human intervention to extract and label relevant patterns. This manual process is not only inefficient but also introduces the risk of inconsistencies, as different individuals may interpret the same activations differently. The growing complexity of neural networks, with their increasing depth and scale, makes manual interpretation virtually infeasible in many practical applications. This paper presents a comprehensive overview of our automated approach, detailing its technical components, performance evaluation, and broader implications for advancing Explainable AI (XAI). Automation is, therefore, a crucial advancement in AI workflows. By automating the interpretation process and analysis, we can significantly reduce execution time, enhance reproducibility, and scale the analysis to handle large datasets. This is where the concept induction approach becomes invaluable, as it provides a systematic way of labeling neuron activations based on predefined rules and a structured knowledge base. Automating this process

ensures that the results are consistent and scalable, eliminating the bottleneck of human interpretation.

Moreover, high-performance computing (HPC) plays a vital role in enabling large-scale AI systems to function effectively. HPC utilizes powerful systems to process vast amounts of data and perform complex calculations at high speeds. In fields such as scientific research, engineering, and data analysis, HPC enables faster simulations, accelerates problem-solving, and allows for efficient handling of large datasets, leading to more accurate and timely results. Beocat, a Beowulf high-performance computing cluster at Kansas State University, consists of 362 compute nodes with over 10,980 cores and 5.57 terabytes of memory. It uses the SLURM job scheduler and is designed to support research requiring large-scale computational resources. Beocat serves a wide range of applications, from data-intensive simulations to machine learning tasks. Its infrastructure allows researchers to efficiently manage and execute complex workflows, offering a robust platform for computationally demanding projects [22].

AI workflows, particularly those involving deep learning models like CNNs, are computationally intensive, requiring significant processing power to handle vast amounts of data and perform complex calculations. High-performance computing (HPC) resources are indispensable in such scenarios, allowing researchers to train models, conduct experiments, and analyze results at scales that would be impossible with standard computing systems. Beocat, the HPC cluster at Kansas State University, plays a crucial role in enabling our research by providing the computational infrastructure necessary for automating tasks like model training, concept induction, image retrieval, and statistical analysis. The use of SLURM job scheduling further enhances the efficiency of resource allocation, ensuring that the pipeline runs smoothly and rapidly, even when processing large datasets[23]. HPC allows for the parallelization of tasks,

significantly reducing execution time and enabling the exploration of more complex models and datasets, which is vital for ensuring scalability and efficiency in AI research.

The automation pipeline developed in this research has significant real-world applications, particularly in fields where transparency in AI decisions is critical. In healthcare, for example, automated interpretation of CNN activations could aid in medical image analysis, helping clinicians better understand how AI models arrive at diagnoses from radiographic images or MRI scans. By providing interpretable explanations for AI-driven decisions, the system could improve trust and acceptance of AI in life-critical domains. Similarly, in autonomous systems, understanding the decisions made by AI—such as why an autonomous vehicle chooses one route over another—can prevent costly errors and ensure safety. The ability to automate and interpret neuron activations also has applications in other fields like finance, where XAI can help explain decisions in credit scoring models, fraud detection, and risk assessment. The integration of automation and explainability opens up new possibilities for deploying AI in high-stakes environments where trust and transparency are paramount.

Chapter 2 - System Architecture

2.1 Overview of the Pipeline

The pipeline consists of four main stages:

1. **Stage 1 - Model Training and Data Configuration:** handles the training and preparation of a CNN model using the ADE20k dataset. The output is a trained ResNet50V2 model and neuron activation data for further analysis.
2. **Stage 2 - Parallelized Concept Induction and Label Hypothesis Generation:** focuses on concept induction using a pre-defined knowledge source (OWL file). This stage generates hypothesized labels for neuron activations using semantic reasoning tools.
3. **Stage 3 - Parallelized Image Retrieval and Classification:** automates the retrieval and classification of images based on the neuron-specific concept files generated in Stage 2. These images are used to evaluate the accuracy of neuron labels.
4. **Stage 4 - Statistical Analysis and Verification of Neuron Activations:** involves statistical validation, where the neuron activations are analyzed using metrics such as target and non-target activations, and statistical tests (e.g., Mann-Whitney U test) are applied to verify the quality of the neuron labels.

2.2 System Architecture and Data Flow

The system architecture of the proposed automated pipeline is designed to streamline the process of interpreting CNN neuron activations. This section outlines the architecture and provides a detailed flow (Figure 2.1) of how data is transformed and passed through each stage.

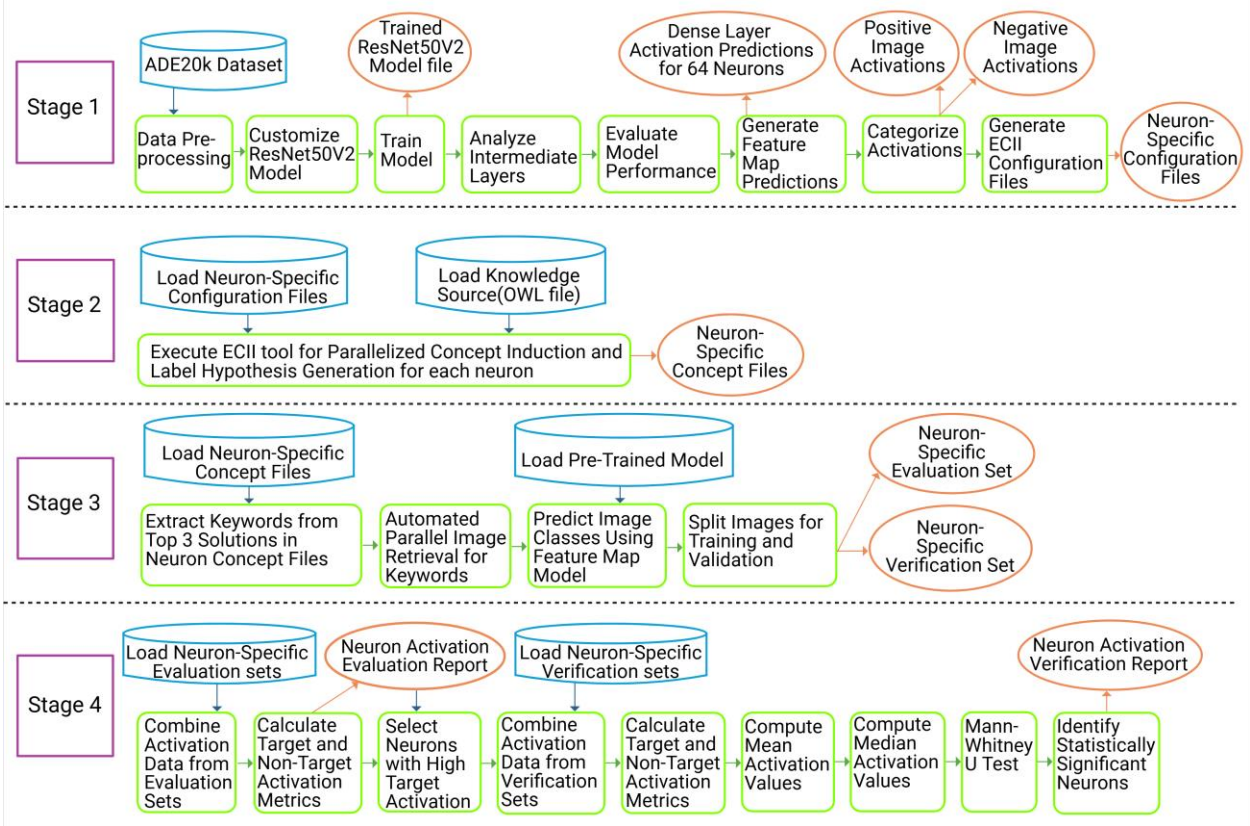


Figure 2.1: Automated Four-Stage Pipeline for Neuron Activation Analysis and Concept Induction (created with BioRender.com).

2.2.1 Stage 1: Model Training and Data Configuration

In Stage 1, the process begins with the ADE20k dataset as the primary input for training the CNN model. The system first loads the training configuration, which defines essential parameters such as learning rate, batch size, and number of epochs. Next, the data preprocessing step ensures the dataset is normalized, resized, and split into training and validation sets, preparing it for model training. The ResNet50V2 model is then customized by adding fully connected layers, adapting the pre-trained model to the specific task. The model is subsequently trained on the ADE20k dataset, producing a trained ResNet50V2 model file as output. Once trained, the system proceeds to analyze the intermediate layers, extracting neuron activations to

gain insights into the feature representations learned by the model. Following this, the model's performance is evaluated using validation data, with accuracy and loss metrics computed. The pipeline then generates feature map predictions by activating 64 neurons in the dense layer, resulting in dense layer activation predictions. These activations are categorized into positive and negative image activations based on predefined thresholds. Finally, ECII configuration files are generated for each neuron, creating neuron-specific configuration files that will be used in subsequent stages for concept induction and hypothesis generation.

2.2.2 Stage 2: Parallelized Concept Induction and Label Hypothesis Generation

In Stage 2, the process begins by loading the neuron-specific configuration files, which were generated in Stage 1 and serve as input for the concept induction process. These files contain information about the neuron activations and the categories they belong to. Next, the system loads the knowledge source, specifically a combined OWL file, which contains a structured semantic hierarchy (such as the Wikipedia knowledge graph) that is used for reasoning and label hypothesis generation. The core of Stage 2 involves executing the ECII tool (Extended Concept Induction and Interpretation), which processes each neuron in parallel. The tool performs concept induction and label hypothesis generation by leveraging the knowledge graph to find meaningful semantic concepts that match the neuron activations. The output of this stage is a set of neuron-specific concept files, each containing the hypothesized labels and coverage scores for every neuron, ready for further validation in the next stages.

2.2.3 Stage 3: Parallelized Image Retrieval and Classification

In Stage 3, the process begins by loading the neuron-specific concept files, which contain the hypothesized labels for each neuron generated in Stage 2. Simultaneously, the pre-trained ResNet50V2 model, trained in Stage 1, is loaded to perform image classification tasks. For each neuron, the system extracts keywords from the top 3 solutions in the concept files, which represent the most likely labels for the neuron activations. These keywords are then used in an automated parallel image retrieval process, where relevant images are retrieved from external sources based on the extracted keywords. Once the images are retrieved, the feature map model from Stage 1 is employed to predict the image classes, validating whether the retrieved images activate the neurons as expected. The classified images are then split into training and validation sets, creating neuron-specific evaluation sets and neuron-specific verification sets, which will be used in the final stage for statistical analysis and verification of neuron labels.

2.2.4 Stage 4: Statistical Analysis and Verification of Neuron Activations

In Stage 4, the process begins by loading the neuron-specific evaluation sets, which were generated in Stage 3, and contain the activation data for each neuron. The system then combines the activation data from these evaluation sets, allowing for the calculation of key metrics. Target activation metrics are calculated for neurons that strongly respond to the target images, followed by the calculation of non-target activation metrics for neurons that respond to non-relevant images. The system then filters neurons with high target activation, producing a neuron activation evaluation report that highlights the neurons with significant activations for further analysis.

Next, the system proceeds by loading the neuron-specific verification sets, also generated in Stage 3, and selects neurons with high target activation from the evaluation phase for verification. The system combines the activation data from the verification sets and recalculates target activation metrics and non-target activation metrics to further validate the neurons' activations. To provide more detailed insights, the system computes the mean and median activation values for both target and non-target activations. A Mann-Whitney U test is then performed to statistically compare the activations, identifying neurons that show a significant difference between target and non-target activations. The final output is a neuron activation verification report, which identifies the statistically significant neurons, confirming the validity of the neuron labels generated throughout the pipeline.

Chapter 3 - Implementation

3.1 Stage 1: Model Training and Data Configuration

3.1.1 Technical Requirements

The model training and data configuration process was executed on the Beocat high-performance computing cluster, leveraging the SLURM workload manager for job scheduling and resource allocation. The following specifications were used for the job:

1. **Job time limit:** 2 hours
2. **Memory:** 5 GB
3. **Compute resources:** 4 nodes, each with 6 processing cores, totaling 24 cores.
4. **Parallel execution:** Managed through SLURM to distribute tasks efficiently across multiple nodes, enhancing processing speed and resource utilization.

The workflow was automated using a Bash script, that initialized the environment, loaded the necessary modules, and triggered the Python scripts responsible for model training and data analysis. The Bash script ensured that each task was completed without manual oversight, making the process highly scalable and repeatable. Key steps in the Bash script included:

1. **Environment setup:** Python 3.11.5 was loaded using the module system, and a virtual environment was created to manage dependencies, ensuring that all required libraries (e.g., TensorFlow, Pandas, scikit-learn, SciPy) were installed in an isolated environment.
2. **Repository access:** The necessary Python project repositories were securely cloned from GitHub using an authentication token. This automation ensured data integrity and allowed for version control, enabling seamless updates to the codebase without manual intervention.

3. **Execution of Python scripts:** The Bash script automated the execution of Python scripts responsible for model configuration, training, and evaluation. By triggering the Python training script (main.py), the automation pipeline handled the entire process, from data preparation to output generation, ensuring that each stage of the workflow was completed consistently and reproducibly.

The Python script executed within this automated environment handled:

1. **Data preprocessing:** Automated loading, augmentation, and splitting of the dataset.
2. **Model customization:** Automated configuration of the ResNet50V2 architecture.
3. **Model training:** Automated training of the model, including the application of early stopping to prevent overfitting.
4. **Performance evaluation:** Automated generation of predictions and model evaluation metrics.

The outputs, including the trained model and activation results, were saved in a predefined directory for subsequent stages of the analysis. This automated workflow ensured seamless and reproducible execution across multiple runs eliminating human errors.

The source code and scripts for implementing the pipeline is available in GitHub repository https://github.com/Samatha1994/ExAI_Automation_Project. This repository contains all configuration files, code modules, and instructions for replicating the pipeline stages. A demonstration video illustrating the pipeline's functionality and execution process is also available at https://www.youtube.com/watch?v=a_tHVwexlEE.

3.1.2 Inputs and Data Setup

We employed the ADE20K dataset, consisting of over 27,000 images across 365 scene categories. For this study, we selected 10 scene categories based on the number of available images in each class and relevance for hidden neuron analysis: “bathroom,” “bedroom,” “building facade,” “conference room,” “dining room,” “highway,” “kitchen,” “living room,” “skyscraper,” and “street.” The dataset was organized into separate training and validation folders for each scene category to ensure effective data management throughout the pipeline (Figures 3.1 to 3.6).

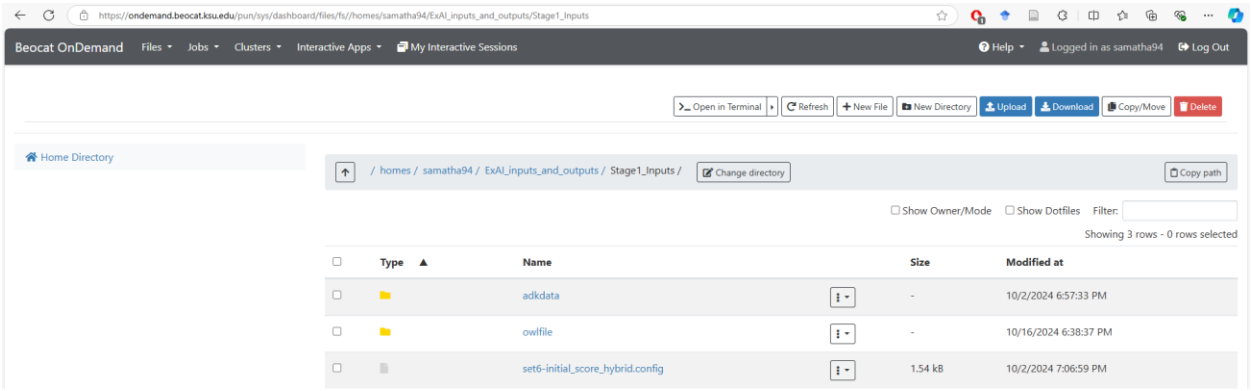


Figure 3.1: Beocat OnDemand interface displaying the initial input files required for training the ResNet50V2 model, including the ADE20K dataset, owlfile and configuration files used for processing image datasets.

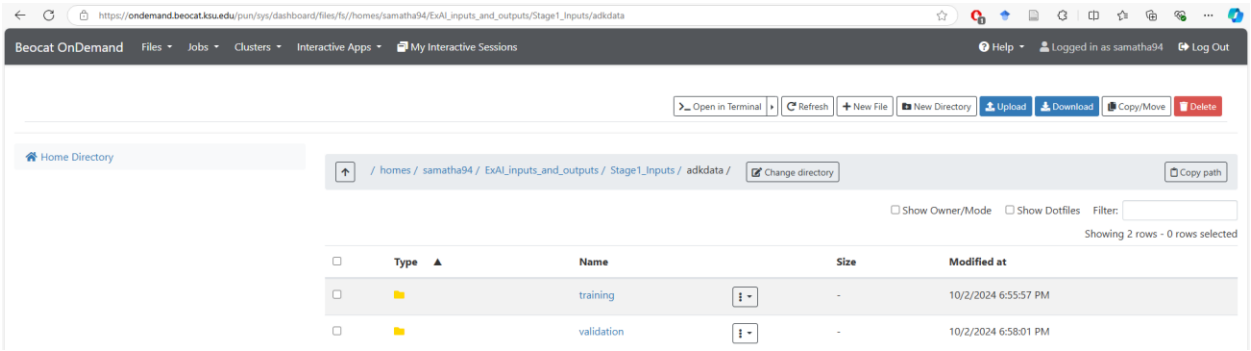


Figure 3.2: Directory structure showing the training and validation data split under the Stage1_Inputs/addkdata folder, which contains the image categories for analysis, emphasizing organized data handling in the pipeline.

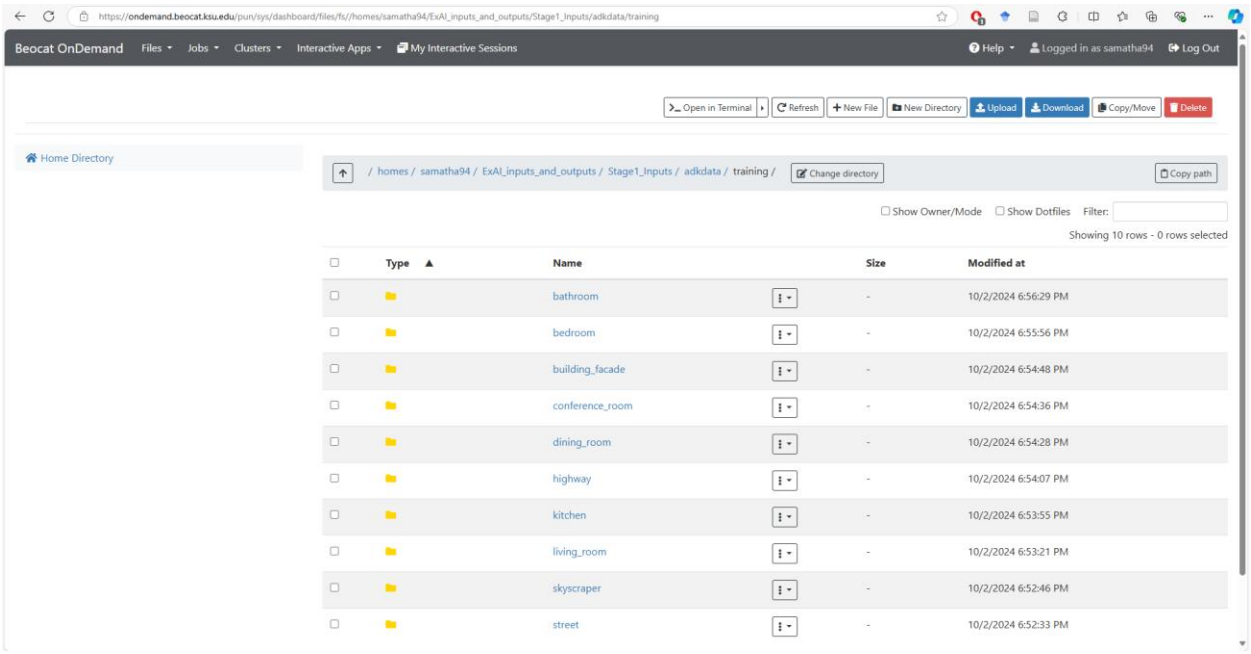


Figure 3.3: View of the training folder containing subfolders for each of the selected scene categories from the ADE20K dataset. This highlights the classification of training images for each environment type used in the initial model training stage.

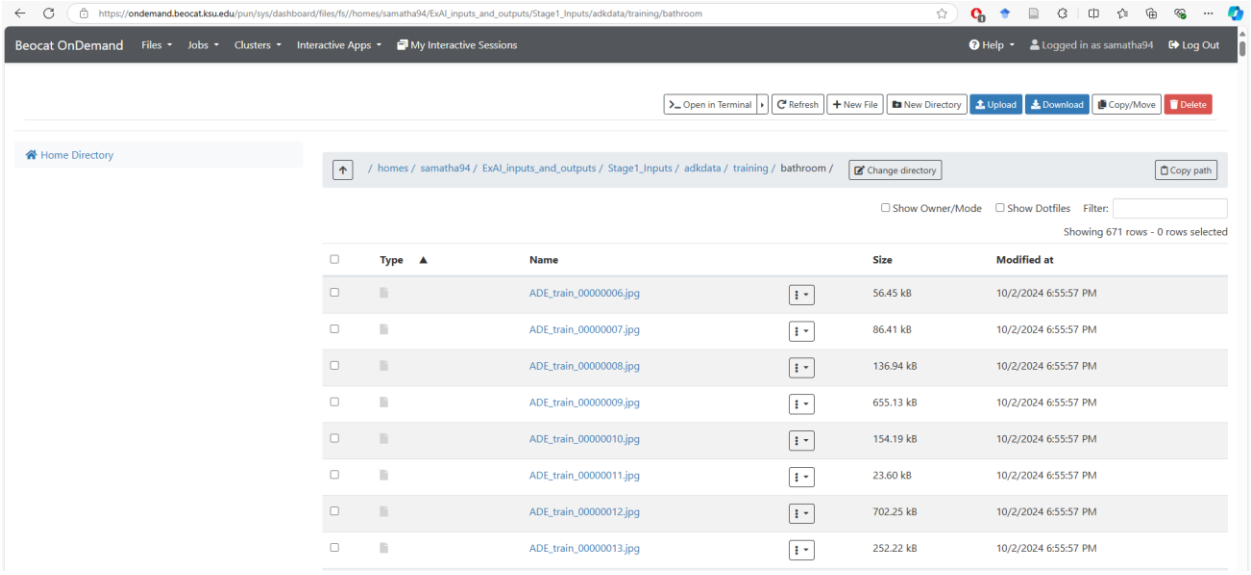


Figure 3.4: The “bathroom” subfolder under the training folder, displaying individual image files from the ADE20K dataset.

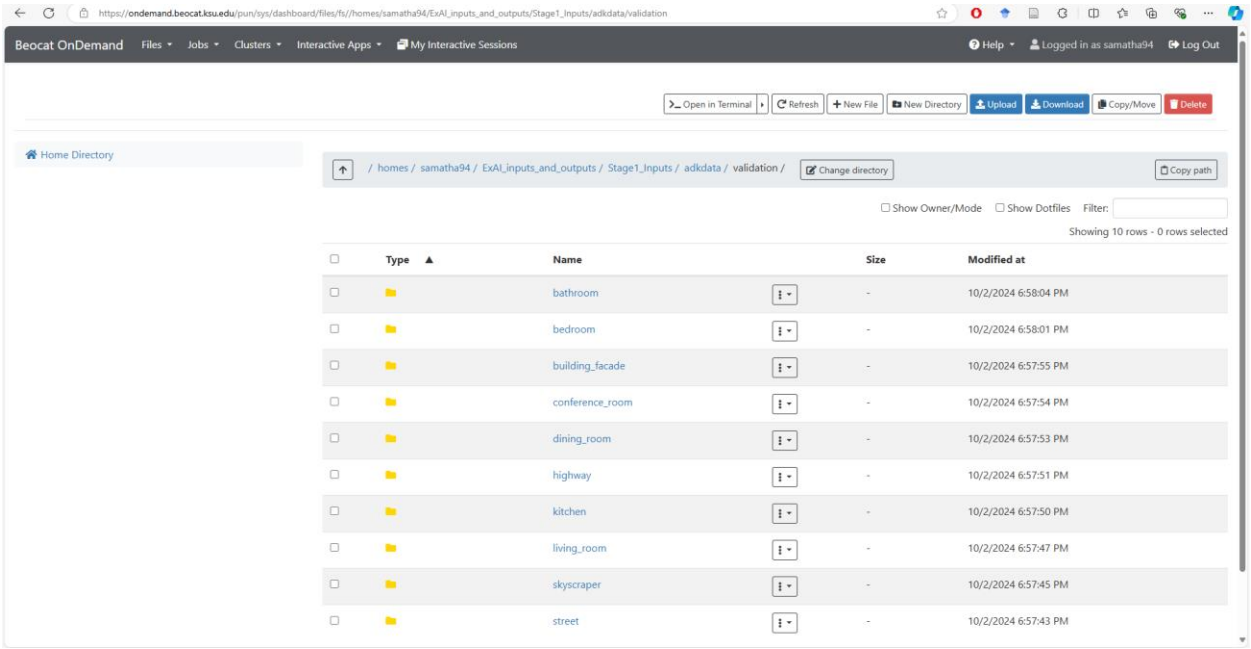


Figure 3.5: The validation folder, similar to the training folder, showing the subfolders containing images for each selected scene category. This structure was used to validate model performance and avoid overfitting during training.

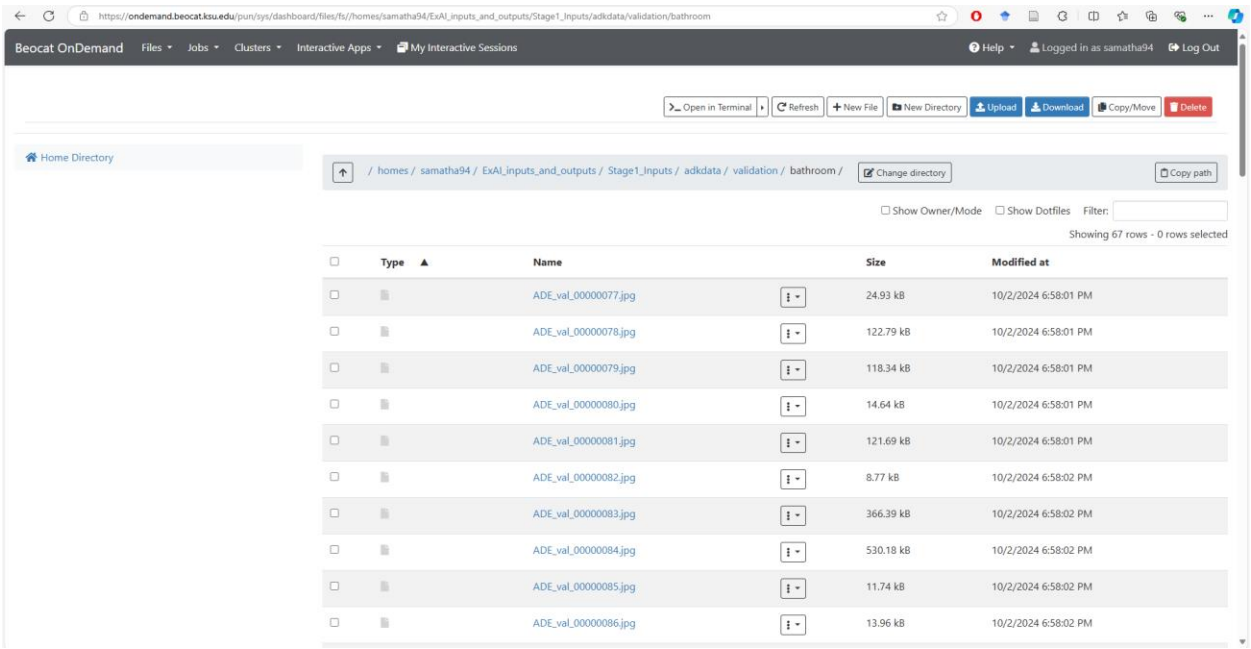


Figure 3.6: The “bathroom” subfolder under the validation folder, displaying individual image files from the ADE20K dataset.

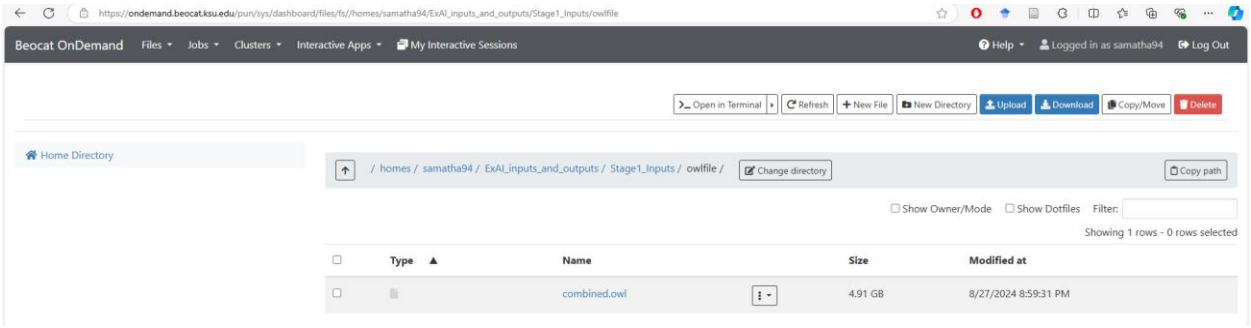


Figure 3.7: Combined.owl file shown in Beocat, representing the OWL (Web Ontology Language) file loaded during concept induction. This file contains structured semantic information used to infer neuron-label relationships through semantic reasoning.

The dataset was split into training and validation sets with a total of 7,557 images. Each image was resized to 224x224 pixels. Training was conducted over 30 epochs with a batch size of 32, using early stopping to prevent overfitting. The pixel-level annotations provided in the ADE20K dataset were not utilized during training, but rather for post-training analysis to generate label hypotheses for neuron activations.

The training parameters, including dataset paths, batch size, image size, and number of epochs, were managed through a configuration file (config.json), ensuring consistency and reproducibility. Additionally, an OWL file containing structured semantic information was incorporated to infer neuron-label relationships through semantic reasoning (Figure 3.7).

3.1.3 Process

The entire process of model training, data configuration, and analysis was fully automated through a streamlined pipeline. This automation eliminated the need for manual intervention at each stage, ensuring consistency, reproducibility, and efficiency across multiple runs. The pipeline was designed to manage everything from data preprocessing to the generation

of outputs (Figure 3.8), making the workflow scalable and reliable. Below are the key steps involved in this automated process:

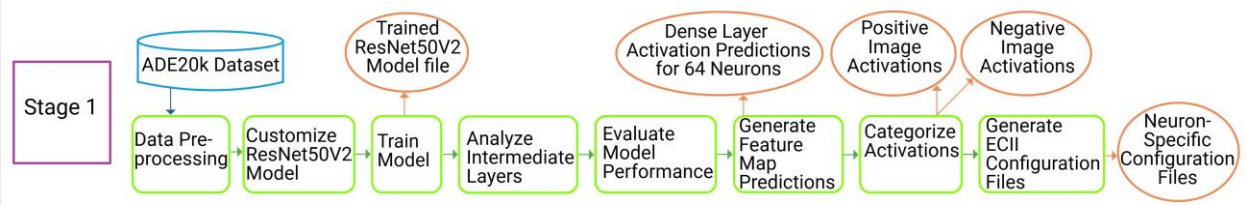


Figure 3.8: Stage 1 workflow for model training and data configuration using the ADE20k dataset and ResNet50V2 model.

1. **Data Preprocessing:** The dataset was preprocessed using TensorFlow's ImageDataGenerator, which normalized pixel values by rescaling them to [0, 1] and applied real-time augmentation to the training set. Augmentation techniques included random shear, zoom, and horizontal flips to improve model generalization and prevent overfitting. The data was split into training (80%) and validation (20%) sets. Both sets were resized to 224x224 pixels, with a batch size of 32. The validation set was left unaltered, while augmentation was applied only to the training set during runtime.
2. **Customize ResNet50V2 Model:** For the model architecture, we fine-tuned a pre-trained ResNet50V2 model. The base model was initialized with ImageNet weights and modified to exclude the top classification layers. This allowed us to retain the pre-trained feature extraction layers, which were kept frozen to prevent retraining. The custom head of the model consisted of an Average Pooling layer, followed by a Flatten layer, a 64-unit Dense layer with ReLU activation, and a Dropout layer (0.5) to reduce overfitting. The final output layer was a Dense layer with softmax activation, where the number of units matched the number of scene categories (10), allowing for multi-class classification. By freezing the base layers and adding a lightweight custom head, we ensured that the model

could leverage pre-trained features while efficiently learning scene-specific classifications.

3. **Train Model:** The customized ResNet50V2 model was trained using the Adam optimizer, with a learning rate of 0.001 and decay applied over 30 epochs. The model was compiled with categorical cross-entropy as the loss function and accuracy as the evaluation metric. Training was performed using the prepared training and validation datasets, with an early stopping mechanism implemented to avoid overfitting. The early stopping callback monitored the validation loss and restored the model weights from the epoch with the best validation performance after 3 consecutive epochs of no improvement. The model was trained for up to 30 epochs, though the use of early stopping typically reduced the total training time. Upon completion, the trained ResNet50V2 model was saved as a .h5 file, ensuring the model could be reused for further analysis and evaluations.
4. **Analyze Intermediate Layers:** To gain insights into the behavior of the model during feature extraction, we analyzed the outputs of the intermediate layers of the trained ResNet50V2 model. The saved model (model_resnet50V2_10classes_retest2023June.h5) was loaded, and outputs from specific layers were extracted for further examination. In this study, we focused on the outputs of the last three layers of the model, which play a critical role in generating the final feature representations. The layer names and their corresponding outputs were captured, and a feature map model was constructed using these layers. This allowed us to visualize the activations and understand how the model processed different input images at a deeper level. Analyzing intermediate layers is crucial for understanding the hierarchical feature extraction process in CNNs, especially

in explainable AI contexts where interpreting hidden neuron activations can provide transparency into model decision-making.

5. **Evaluate Model Performance:** The performance of the trained ResNet50V2 model was evaluated using the test dataset. Predictions were generated by passing the test images through the model, and the predicted labels were compared with the true class labels. The final class prediction for each image was determined by selecting the class with the highest predicted probability. The evaluation metrics included precision, recall, and F1-score, calculated for each of the 10 scene categories. These metrics provided a comprehensive assessment of the model's classification accuracy. The evaluation process helped to identify the model's strengths in recognizing certain scenes and potential areas where the model's performance could be improved. Additionally, the predicted labels and corresponding image filenames were stored for further analysis. This allowed for the examination of misclassified instances, which can offer insights into areas for model refinement.
6. **Generate Feature Map Predictions:** To further analyze the model's behavior, feature map predictions were generated for the dense layer of the trained ResNet50V2 model. This analysis focused on extracting activation patterns from the 64 neurons in the dense layer, providing a detailed view of how the model responds to different input images. The test dataset was passed through a feature map model, which captured the activations from the dense layer. These activations were recorded for each of the test images, generating a set of predictions that reflected the neuron-level responses across the entire dataset. The results were saved to a structured file for further examination, allowing for a more granular analysis of the model's internal representations. This step was crucial in

understanding the relationship between specific neurons and their contribution to the final classification decisions.

7. **Categorize Activations:** After generating the feature map predictions, the next step involved categorizing the neuron activations into positive and negative groups. The activations of the 64 neurons in the dense layer were analyzed to determine their response to each input image. For each neuron, the highest activation value was identified, and two thresholds were established: one at 80% of the highest value to categorize positive activations, and one at 20% to categorize negative activations. Images with activation values above the 80% threshold were categorized as positive activations, indicating that the neuron was highly activated by those images. Conversely, images with activation values below the 20% threshold were categorized as negative activations, indicating low or minimal activation. The categorized images were saved in separate files for each neuron, allowing further exploration of neuron-specific behaviors. This step is crucial for understanding which images strongly activate certain neurons and contributes to the overall explainability of the model.
8. **Generate ECII Configuration Files:** In the final step of this stage, neuron-specific configuration files were generated to serve as inputs for the subsequent concept induction process. These configuration files were created based on the positive and negative image activations identified in the previous step. A pre-defined template configuration file was used as the base, which included key parameters such as concept limits, horn clause limits, and reasoner settings. For each neuron, the corresponding positive and negative image activations were converted into uniform resource locators (URLs) pointing to the

associated images in the knowledge base. For each neuron, a configuration file was generated containing the following key components:

- **A knowledge base (or ontology):** The file referenced a pre-built ontology in OWL format.
- **Positive examples (P):** This consisted of URLs representing images that strongly activated the neuron (above the 80% threshold).
- **Negative examples (N):** This consisted of URLs representing images that weakly activated the neuron (below the 20% threshold).

Configuration files were only generated for neurons with at least one positive or negative activation. These files were stored in the output directory and would be used in the next stage for running the ECII tool, enabling concept induction for each neuron.

3.1.4 Outputs

Upon completion of the model training and data configuration process, several key outputs were generated, which were critical for the subsequent stages of the project (Figures 3.9 to 3.11):

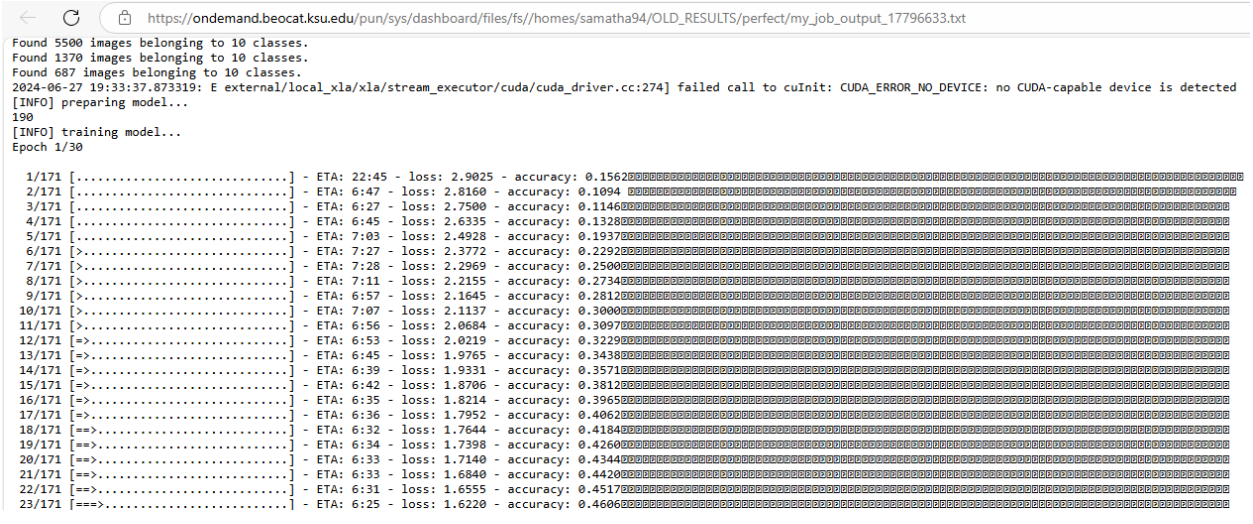
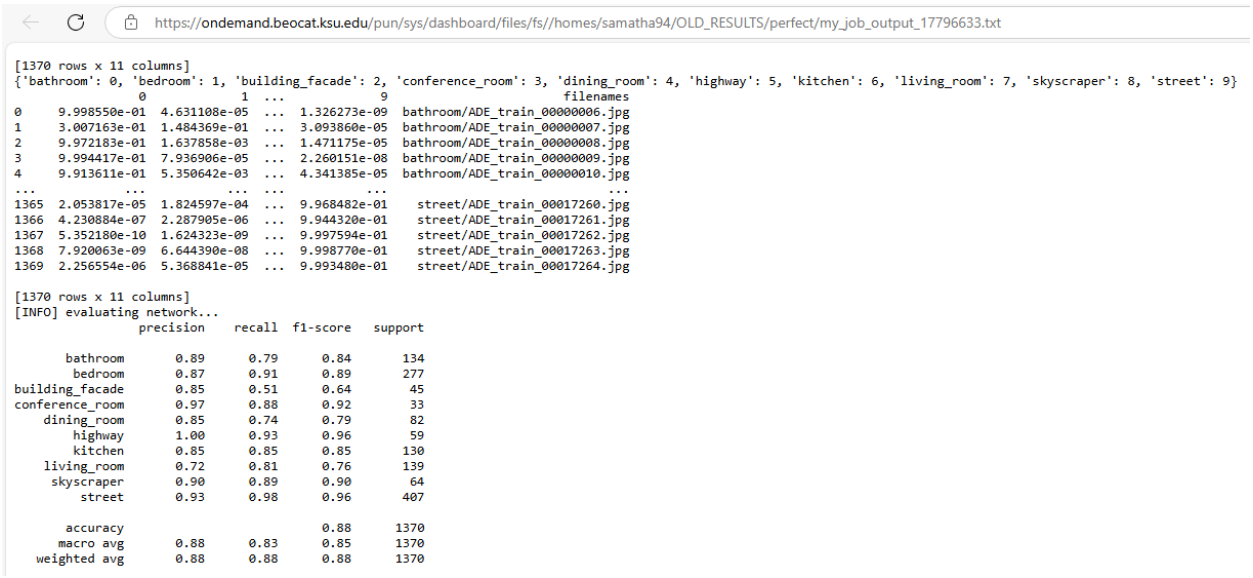


Figure 3.9: Input Dataset Summary: This figure provides a summary of the number of images belonging to each selected class in the ADE20K dataset used for training. It shows the dataset distribution across 10 scene categories, which formed the training set for configuring and fine-tuning the ResNet50V2 model.



https://ondemand.beocat.ksu.edu/pun/sys/dashboard/files/fs//homes/samatha94/OLD_RESULTS/perfect/my_job_output_17796633.txt

```
[1370 rows x 11 columns]
{'bathroom': 0, 'bedroom': 1, 'building_facade': 2, 'conference_room': 3, 'dining_room': 4, 'highway': 5, 'kitchen': 6, 'living_room': 7, 'skyscraper': 8, 'street': 9}
              0             1  ...             9                        filenames
0     9.998550e-01  4.631108e-05  ...  1.326273e-09  bathroom/ADE_train_00000006.jpg
1     3.007163e-01  1.484369e-01  ...  3.093860e-05  bathroom/ADE_train_00000007.jpg
2     9.972183e-01  1.637858e-03  ...  1.471175e-05  bathroom/ADE_train_00000008.jpg
3     9.994417e-01  7.936906e-05  ...  2.260151e-08  bathroom/ADE_train_00000009.jpg
4     9.913611e-01  5.350642e-03  ...  4.341385e-05  bathroom/ADE_train_00000010.jpg
...            ...           ...  ...           ...                            ...
1365  2.053817e-05  1.824597e-04  ...  9.968482e-01    street/ADE_train_00017260.jpg
1366  4.230884e-07  2.287905e-06  ...  9.944320e-01    street/ADE_train_00017261.jpg
1367  5.352180e-10  1.624323e-09  ...  9.997594e-01    street/ADE_train_00017262.jpg
1368  7.920063e-09  6.644390e-08  ...  9.998770e-01    street/ADE_train_00017263.jpg
1369  2.256554e-06  5.368841e-05  ...  9.993480e-01    street/ADE_train_00017264.jpg

[1370 rows x 11 columns]
[INFO] evaluating network...
               precision    recall  f1-score   support

      bathroom       0.89      0.79      0.84       134
       bedroom       0.87      0.91      0.89       277
building_facade       0.85      0.51      0.64        45
conference_room       0.97      0.88      0.92        33
    dining_room       0.85      0.74      0.79        82
        highway       1.00      0.93      0.96        59
        kitchen       0.85      0.85      0.85       130
    living_room       0.72      0.81      0.76       139
     skyscraper       0.90      0.89      0.90        64
         street       0.93      0.98      0.96       407

       accuracy                           0.88      1370
      macro avg       0.88      0.83      0.85      1370
   weighted avg       0.88      0.88      0.88      1370
```

Figure 3.10: Scene Classification Evaluation Report: This figure presents the evaluation metrics for 10 selected scene categories from the ADE20K dataset. The table reports precision, recall, and f1-score for each category, highlighting an overall model accuracy of 88%. These metrics provide a comprehensive assessment of the model's classification performance.

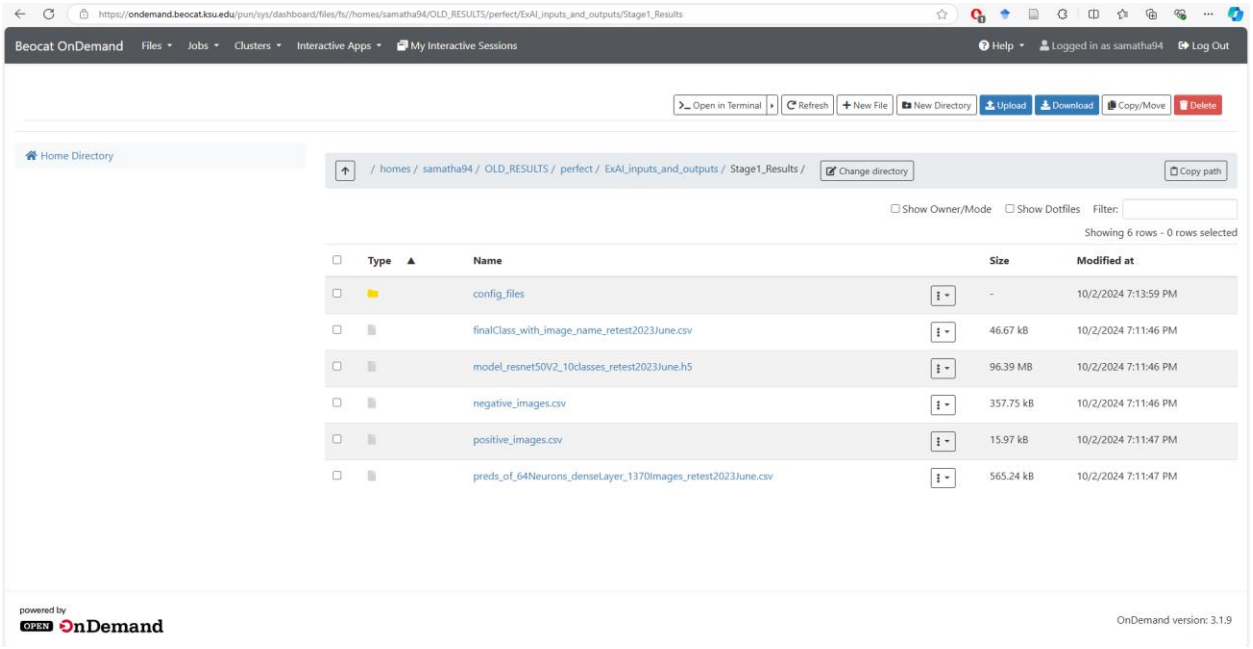


Figure 3.11: Generated Files After Model Training and Prediction: The output directory contains essential files, including the trained ResNet50V2 model file (.h5 format), dense layer predictions, positive and negative activation datasets, and neuron-specific configuration files.

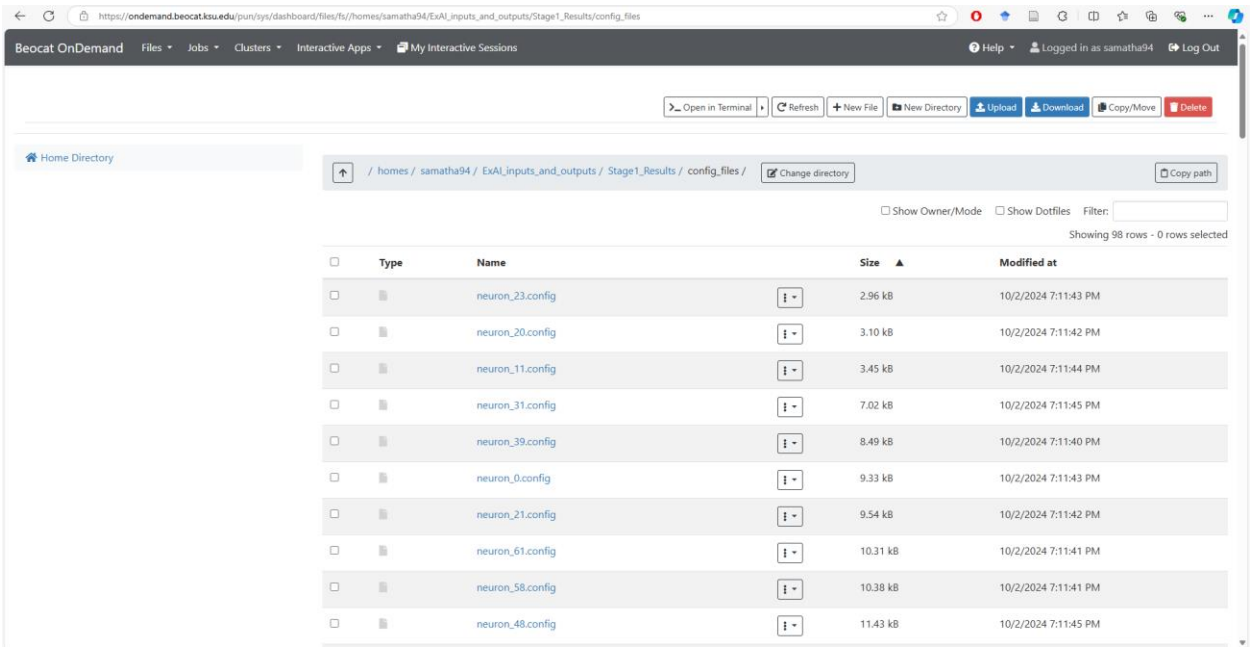


Figure 3.12: Neuron-Specific Configuration Files: This figure displays the configuration files generated for each of the 64 neurons. These files contain essential references and parameters used in the next stage to induce concepts using the ECII tool. The configuration files play a significant role in linking neuron activations to specific concepts.

1. **Trained ResNet50V2 Model:** The final trained model, after 30 epochs, was saved in .h5 format. This model encapsulated the learned weights and configurations specific to the scene classification task, enabling further use in prediction and analysis.
2. **Intermediate Layer Activations:** The activations from the last dense layer of the model were extracted for 64 neurons. These activations provided insight into the internal behavior of the model and were used for further analysis of neuron-specific responses. These activations were stored in a CSV file for future reference and exploration.
3. **Feature Map Predictions:** Dense layer activation predictions were generated for the test dataset, capturing the response of each neuron to the input images. This output served as a critical input for understanding neuron-specific activation patterns and for categorizing positive and negative activations.

4. **Positive and Negative Activations:** Neuron-specific activations were categorized into positive and negative sets. Images that triggered strong activations (above 80%) were categorized as positive, while those with weak activations (below 20%) were categorized as negative. These results were stored as CSV files for each neuron.
5. **ECII Configuration Files:** Neuron-specific configuration files were generated based on the positive and negative activations**.** These files (Figure 3.12) included references to the ontology knowledge base and served as the input for the next stage, where the ECII tool would be used for concept induction. These outputs were systematically saved in the output directory, providing the necessary data for subsequent stages of analysis, including neuron interpretation and concept induction.

3.2 Stage 2: Parallelized Concept Induction and Label Hypothesis Generation

3.2.1 Technical Requirements

The execution of the concept induction and label hypothesis generation in Stage 2 required a high-performance computing setup to efficiently process 64 neurons in parallel. The following technical requirements were critical for this stage:

1. **Compute Resources:** The job was executed on 1 node, utilizing 1 task per node with exclusive node access. This allowed for efficient parallelization across neurons by leveraging SLURM's array feature, which assigned each neuron to a unique task. The job was split into 64 tasks, corresponding to the 64 neurons, enabling simultaneous concept induction for each neuron.

2. **Memory Allocation:** Each job array task was allocated 38 GB of memory, ensuring sufficient resources for loading large knowledge bases (OWL files) and handling computationally intensive concept induction operations.
3. **Execution Time:** A maximum time limit of 2 hours was allocated for the job to ensure that each neuron-specific task had ample time to complete without exceeding resource limits.
4. **Software Requirements:** The ECII tool, written in Java, required Java 1.8 and Maven 3.6.3 for execution. The project was built using Maven to manage dependencies and automate execution across tasks. The latest version of the ECII tool was cloned from a private GitHub repository using a secure authentication token. This ensured that the most up-to-date code was used for the concept induction process.
5. **Automated Execution with SLURM:** The process was fully automated using a SLURM-configured Bash script, which managed parallel execution across 64 tasks, handled environment setup, and ensured that each task ran independently while sharing common dependencies like the ontology files.

The technical setup ensured that the concept induction process for each neuron was handled efficiently and consistently, leveraging parallel processing to reduce overall execution time and enhance scalability.

3.2.2 Inputs and Data Setup

The concept induction process in Stage 2 relied on specific inputs generated in Stage 1 and a combined knowledge base (ontology) to perform semantic reasoning. The following inputs were essential:

1. **Neuron-Specific Configuration Files:** Each neuron was assigned to a configuration file, generated in Stage 1, that specified the positive and negative image activations for that neuron. These files contained the necessary parameters for the concept induction process.
2. **Combined Knowledge Source (OWL File):** A combined ontology was used as the background knowledge for the ECII process. This ontology was constructed by merging ADE20K image annotations with the Wikipedia concept hierarchy, providing semantic context for generating meaningful labels.

These inputs were loaded automatically into the pipeline to ensure a smooth and consistent execution of the concept induction process across all neurons.

3.2.3 Process

The concept induction and label hypothesis generation process for each of the 64 neurons in the CNN's dense layer was performed using the ECII (Efficient Concept Induction Implementation) tool, which utilizes heuristic-based deductive reasoning. This stage relied on a large-scale background knowledge source, the Wikipedia concept hierarchy, consisting of approximately 2 million classes, to generate meaningful semantic labels for the neuron activations. The entire process was automated and parallelized to maximize computational efficiency, which shows the Stage 2 workflow (Figure 3.13) for executing the ECII tool.

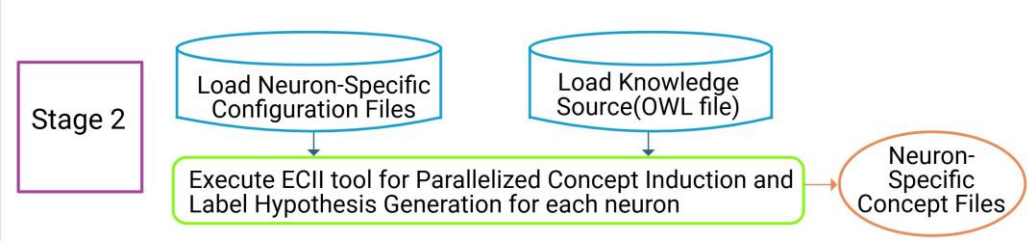


Figure 3.13: Stage 2 workflow for executing the ECII tool to induce neuron-specific concepts and generate label hypotheses.

1. **Loading Neuron-Specific Configuration Files:**

    - The configuration files generated in Stage 1, containing positive and negative image activations for each neuron, were loaded into the ECII tool. These files defined the positive set P (images with activations above 80% of the neuron's maximum) and the negative set N (images with activations below 20%).
    - The configuration files also included links to background knowledge, allowing the ECII system to infer meaningful semantic concepts.

2. **Loading the Combined Knowledge Source (OWL File):**

    - A combined ontology, merging the ADE20K image annotations and the Wikipedia concept hierarchy, was used as the knowledge base (K) for the ECII tool. This ontology provided semantic context, with classes mapped using the Levenshtein string similarity metric to match image annotations to the class hierarchy.
    - The Pellet reasoner was employed for inference over this ontology, enabling the system to deduce high-level semantic concepts from the structured data.

3. **Parallel Execution of ECII Tool:**

    - The ECII tool was executed in parallel across 64 tasks using SLURM's array feature. The parallel processing approach efficiently allocated resources and minimized execution time by allowing all neurons to be processed simultaneously. Each job processed its neuron-specific configuration file and the combined ontology.
    - The ECII tool took the neuron's positive and negative image sets along with the background knowledge to generate label hypotheses. These labels represented the semantic concepts that best described the neuron's activation patterns.

- For each neuron, ECII produced a list of hypothesized labels ranked by coverage scores, representing how well each concept described the neuron's activation patterns. The results also included detailed statistics such as precision, recall, F-measure, and the top solutions, prioritizing the most accurate labels.
- The entire process was automated using a Bash script, which handled job scheduling, environment setup, and repository cloning via SLURM. This ensured consistency and reproducibility while reducing manual intervention. The outputs for each neuron were saved as text files, containing the semantic labels and performance metrics, ready for further analysis and interpretation in subsequent stages.

3.2.4 Outputs

Upon completing the parallelized concept induction process, several key outputs were generated for each of the 64 neurons in the CNN's dense layer:

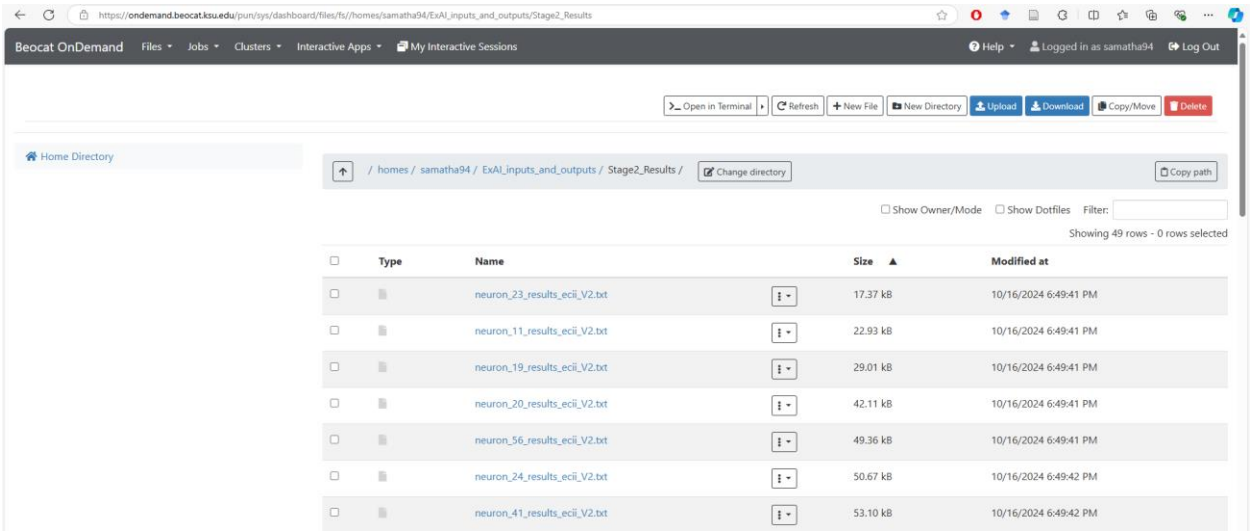


Figure 3.14: Neuron-Specific Concept Files Generated in Stage 2: This figure presents the structured output files generated as part of the parallelized concept induction process. Each file corresponds to a specific neuron and contains the top hypothesized semantic concepts. These concepts are derived from positive and negative individual types, highlighting the most relevant solutions with the highest coverage and precision in identifying key image features.

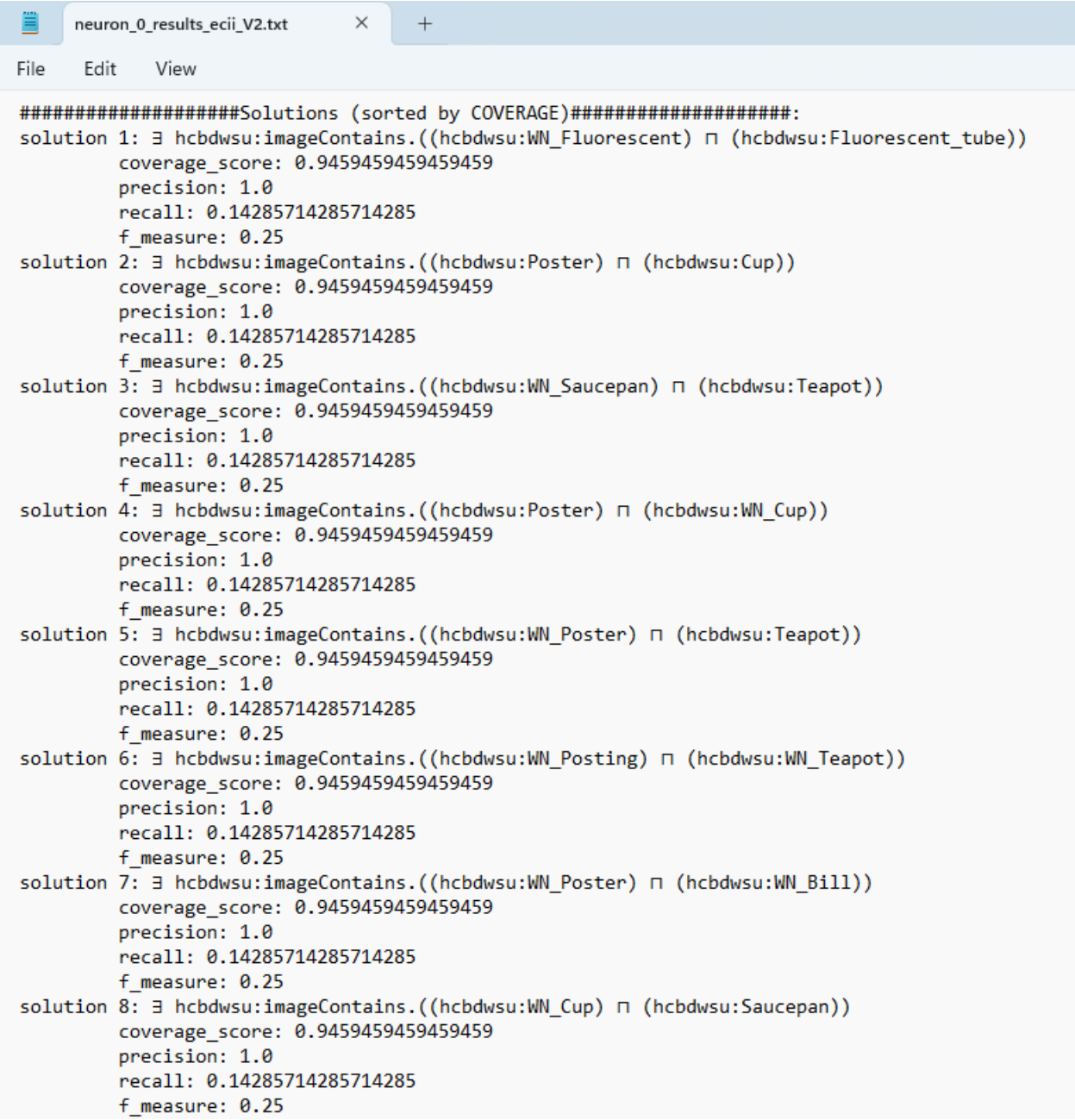


Figure 3.15: Neuron-Specific Concept Solutions Generated in Stage 2 – This figure displays the solutions generated by the ECII tool for Neuron 0, ranked by coverage score. Each solution provides semantic concepts associated with the neuron's activations, along with key performance metrics such as precision, recall, and F-measure, offering insight into how the neuron responds to different image features from the ADE20K dataset.

1. **Neuron-Specific Concept Files**: Each neuron produced a set of hypothesized semantic concepts, which represent the neuron's activation patterns in relation to the ADE20K dataset and the Wikipedia background knowledge. These concepts were derived using deductive reasoning from the positive and negative image sets and were stored in structured text files. The files contained the following key elements:
    - **Top Hypothesized Concepts:** Each neuron had multiple hypothesized semantic labels (Figure 3.15) with the highest coverage scores, reflecting the extent to which the concept explained the neuron's activations.
    - **Performance Metrics:** Precision, recall, and F-measure were reported for each hypothesized label, indicating the quality of the predictions for positive and negative activations. Coverage scores were particularly significant, ranking the solutions based on their relevance.
        - **Positive and Negative Sets:** The files also contained the lists of images from the ADE20K dataset that strongly activated (positive examples) or weakly activated (negative examples) each neuron.
2. **Coverage Score Rankings:** The hypothesized concepts for each neuron were ranked by their coverage score, which is a metric that quantifies how well a specific semantic label fits the neuron's activations. Coverage was computed for both positive and negative sets, with higher scores indicating better fits to the neuron's behavior. This ranking helped prioritize the most relevant and meaningful labels for further analysis.
3. **Detailed Solutions for Each Neuron:** Each neuron-specific result file contained up to 200 possible solutions, with each solution representing a potential explanation of the neuron's activations. For each solution, key metrics such as coverage score, precision,

recall, and F-measure were recorded. The top-ranked solutions were the ones with the highest coverage scores and strong precision, providing a deep understanding of the neuron's role in image classification.

4. **Structured Text Files:** All results were systematically saved in structured text files (Figure 3.14), with clear segmentation for each neuron. These files included the positive and negative examples, the hypothesized concepts, and the corresponding performance metrics, enabling detailed post-analysis of the neurons and their activations. This structured storage also allows for reproducibility and traceability for future evaluations.

These outputs are critical for understanding the internal representations learned by the CNN model's neurons and for linking their activations to high-level semantic concepts. They also serve as the foundation for future stages, including model interpretation and validation, enhancing transparency in AI decision-making.

3.3 Stage 3: Parallelized Image Retrieval and Classification

3.3.1 Technical Requirements

The execution of image retrieval and classification tasks in Stage 3 required a high-performance computational setup to manage large-scale data processing efficiently. The following technical requirements were critical for this stage:

1. **Compute Resources:** Each neuron-specific task was handled using one CPU core, with a maximum of 1 task per node. The job utilized multiple nodes to ensure parallel execution of tasks across all neurons. SLURM's job array feature allowed for each neuron to be assigned to a unique task, enabling concurrent retrieval and classification processes.

2. **Memory Allocation:** Each task was allocated 38 GB of memory. This provided sufficient resources to handle the image retrieval, data processing, and classification tasks, especially for larger datasets or image-heavy tasks.
3. **Execution Time:** A maximum time limit of 2 hours was set for each task, ensuring adequate time for image retrieval, processing, and classification while staying within the compute resource limits of the Beocat cluster.
4. **Software Requirements:** Python 3.11.5 was used for executing the image retrieval and classification tasks. A virtual environment was set up to manage dependencies such as pygoogle_image (for retrieving images from Google) and TensorFlow (for classification). SLURM was used to manage task scheduling, resource allocation, and parallel execution. The process was automated using a SLURM-configured Bash script, which handled environment setup, dependency installations, neuron-specific task execution, and data processing without requiring manual intervention.
5. **Automated Execution with SLURM:** The entire workflow was automated, from neuron-specific configuration file handling to image retrieval and classification. A pre-processing step was performed using an additional Bash script to generate neuron indices based on the results of Stage 1, ensuring that only neurons with valid concept files were processed in Stage 3. SLURM managed parallel execution across multiple nodes, optimizing the workflow for scalability and efficiency, thereby minimizing overall execution time.

3.3.2 Inputs and Data Setup

In Stage 3, the inputs required were drawn from the outputs of the previous stages, as well as additional resources for new data collection and processing. The key inputs include:

1. **Neuron-Specific Concept Files:** These files, generated in Stage 2, contain the hypothesized semantic labels (concepts) for each of the 64 neurons. The top three solutions for each neuron are extracted from these files to guide the image retrieval process.
2. **Pre-Trained Model:** The ResNet50V2 model, pre-trained during Stage 1, serves as the classifier to evaluate the images retrieved during this stage. This model is used to determine the activation patterns of each neuron for the newly retrieved images.
3. **Google Image Search:** For each neuron, images are retrieved from Google based on the extracted keywords. These images serve as the test set to confirm or reject the label hypotheses for each neuron.

The data is automatically prepared and handled by a Bash script that ensures the proper loading and processing of neuron indices, concept files, and model weights. All inputs were efficiently managed through automated parallelized scripts running on Beocat's high-performance computing environment, ensuring consistency and scalability across all neurons.

3.3.3 Process

In Stage 3, the core objective is to validate the neuron-specific label hypotheses by retrieving and classifying images based on the extracted concepts. The process is automated, leveraging parallel execution and Beocat's high-performance computing resources(Figure 3.16).

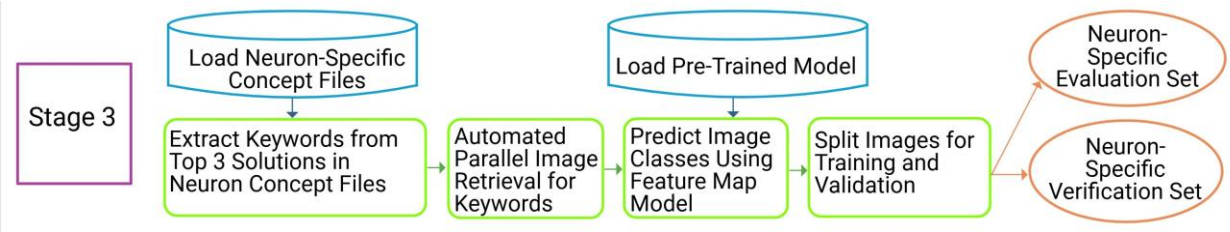


Figure 3.16: Stage 3 workflow for automated image retrieval and classification using neuron-specific concept files and pre-trained models.

1. **Neuron Index Generation:** Prior to initiating the main pipeline, a script generates neuron indices based on the availability of valid configuration files (those with non-empty positive and negative activation sets). These indices dictate which neurons proceed to the image retrieval and classification stages.
2. **Keyword Extraction:** The system extracts the top three semantic concepts from the neuron-specific concept files generated in Stage 2. These concepts are parsed into relevant keywords using a custom keyword extraction module, with each set of keywords representing the hypothesized label for the neuron.
3. **Automated Image Retrieval:** The extracted keywords are fed into the pygoogle_image library, which retrieves up to 100 images per keyword from Google Images. These images are filtered based on size and format, ensuring that they are compatible with the pre-trained ResNet50V2 model (JPEG format, minimum size of 224x224 pixels). The retrieval is fully parallelized, with each neuron being processed independently using SLURM's job array functionality, drastically reducing execution time.
4. **Image Classification:** Once images are retrieved, they are passed through the pre-trained CNN model from Stage 1. The process_and_classify_images function processes the images, classifies them using the feature maps extracted from the ResNet50V2 model, and assigns predicted classes. For each neuron, the system predicts whether the retrieved

images activate the corresponding neuron, thus confirming or rejecting the neuron-specific hypotheses.

5. **Data Splitting:** The retrieved and classified images are divided into two sets:
    - **Evaluation Set (80% of the images):** Used for initial neuron activation testing.
    - **Verification Set (20%):** Reserved for more detailed statistical evaluation in later stages.

        This split ensures that the results are statistically valid and that further analysis can be performed with robust data.
6. **Output Generation:** The results, including the neuron-specific image classifications, are saved as structured CSV files. Evaluation and verification datasets are stored in separate directories to facilitate downstream statistical analysis.

By fully automating the process through SLURM and parallelizing the execution across all neurons, Stage 3 ensures that the validation of neuron-specific hypotheses is both efficient and scalable. This parallelized approach reduces the overall computational time, ensuring that the process is completed in a fraction of the time compared to sequential execution.

3.3.4 Outputs

The outputs from Stage 3 consist of several key elements, organized into structured datasets for further analysis and validation (Figure 3.17). These outputs include:

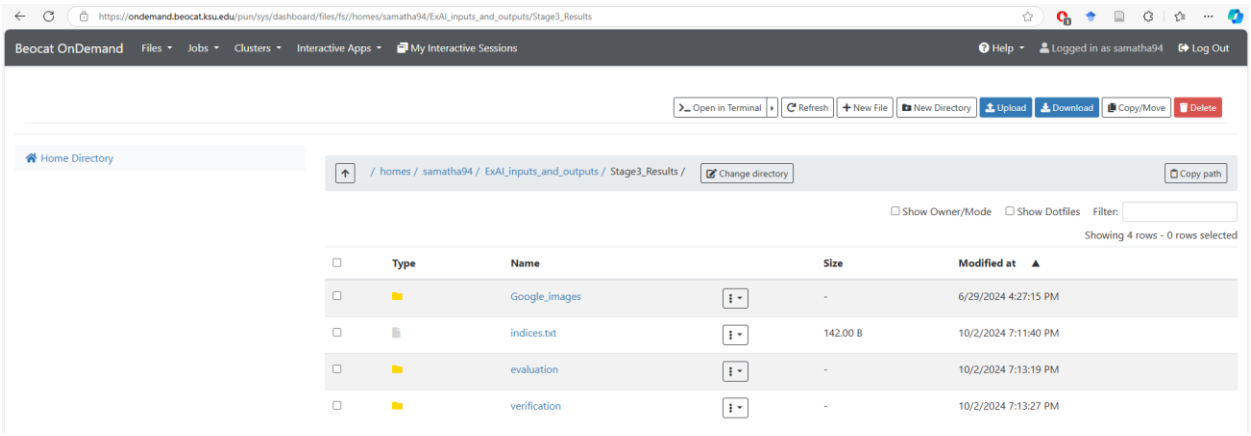


Figure 3.17: Outputs from Stage 3: This figure displays the directory structure for Stage 3 outputs. It highlights key folders and files, including 'Google Images' containing retrieved image datasets, 'indices.txt' listing the neurons selected for image retrieval, and two folders 'evaluation' and 'verification' for organizing results from the classification model.

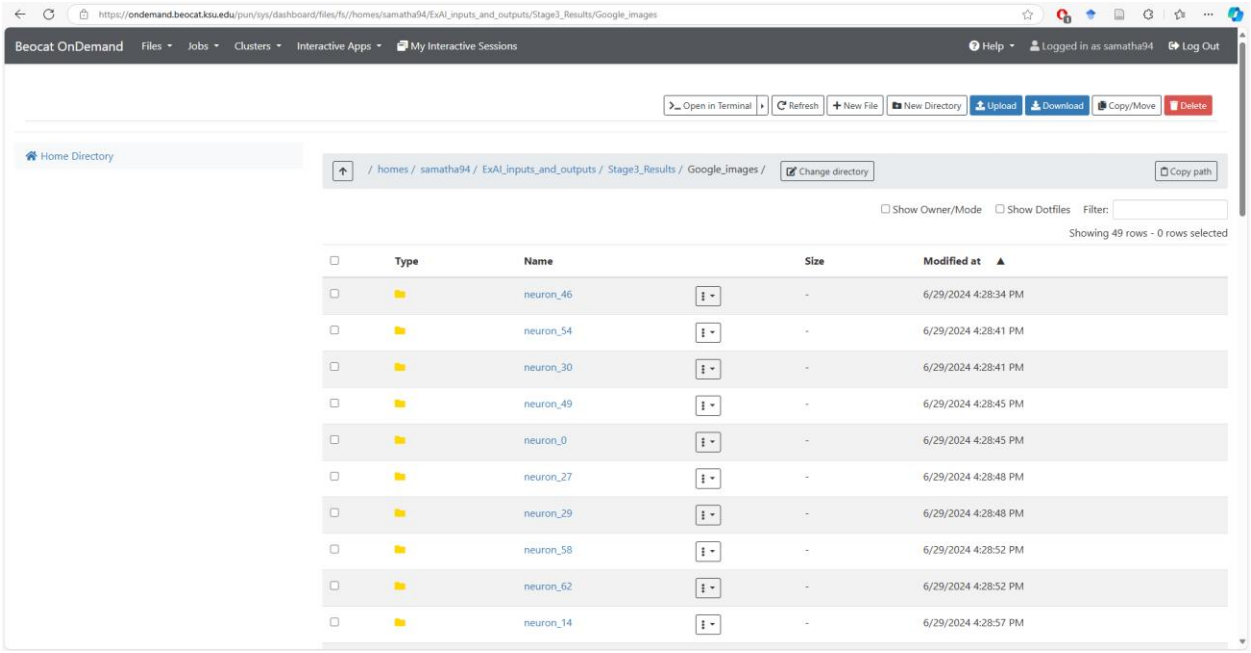


Figure 3.18: Neuron-Specific Folders in Google Images Directory – This figure presents the organization within the "Google Images" folder. It contains folders named after specific neurons, indicating that each neuron has its own dedicated subdirectory for storing retrieved images.

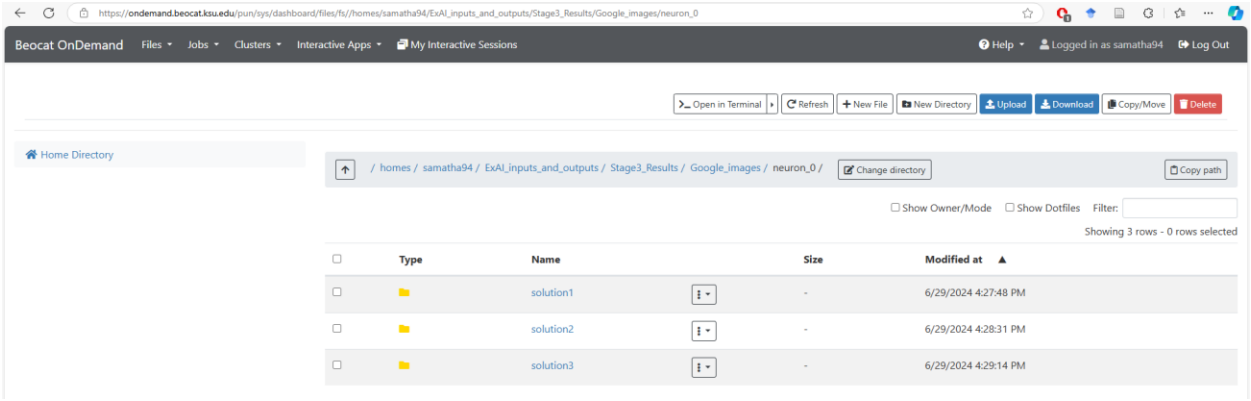


Figure 3.19: Solution Folders within a Neuron Folder – This figure shows the structure within a specific neuron's folder, which includes subfolders named after the top three solutions.

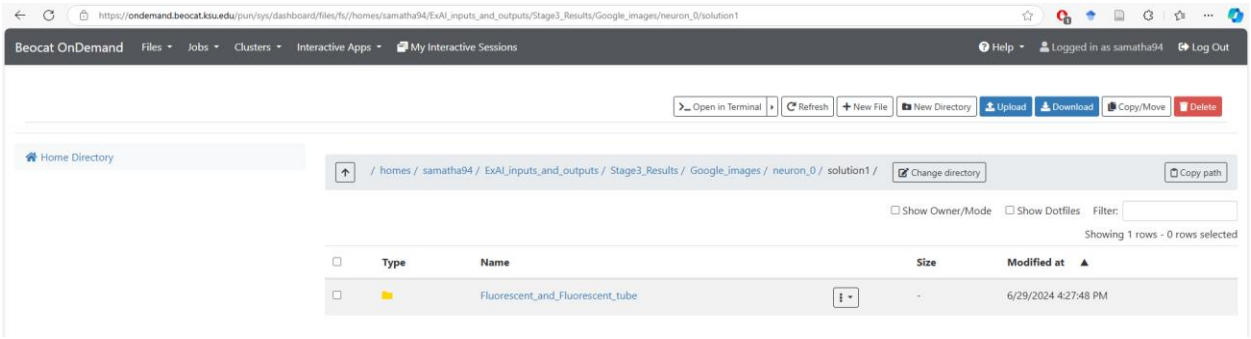


Figure 3.20: Concept Folder within a Solution Folder: This figure illustrates the organization of images within a specific solution folder. The images are stored with filenames corresponding to the concept names or keywords extracted from the solutions.

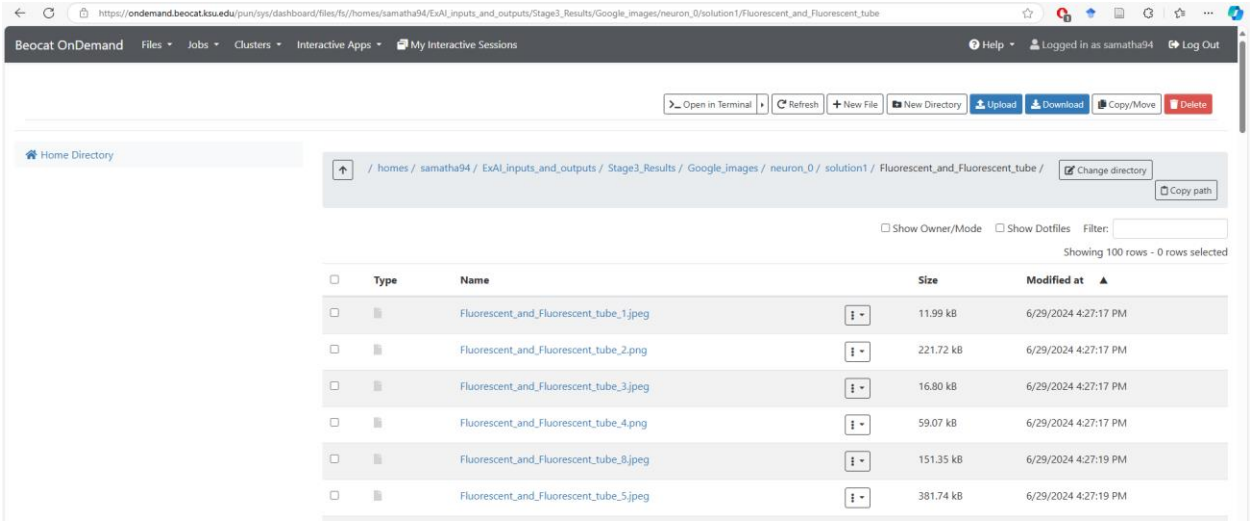


Figure 3.21: Example Images within the Concept Folder: This figure illustrates the images stored within a specific solution folder associated with the label hypothesis "Fluorescent_and_Fluorescent_tube." Each image is named according to the extracted keywords relevant to the hypothesized concept. Up to 100 images per keyword were retrieved.

Figure 3.22: Example of a Retrieved Image – This figure provides a sample image retrieved and stored within a solution folder. The image is a visual representation of the specific hypothesized label for which the neuron showed activation.

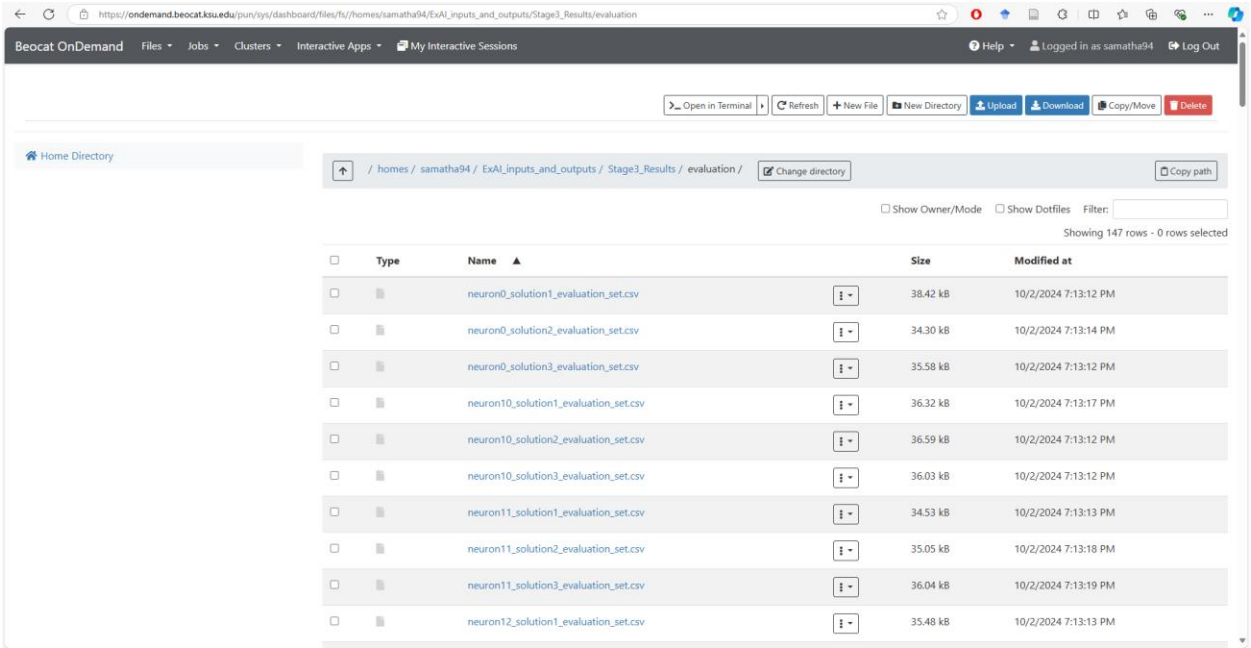


Figure 3.23: Organization of Evaluation Dataset – This figure displays the structured organization of the evaluation dataset for each neuron and its corresponding solutions. Each CSV file contains essential details, such as image filenames, predicted class labels, and neuron activation values.

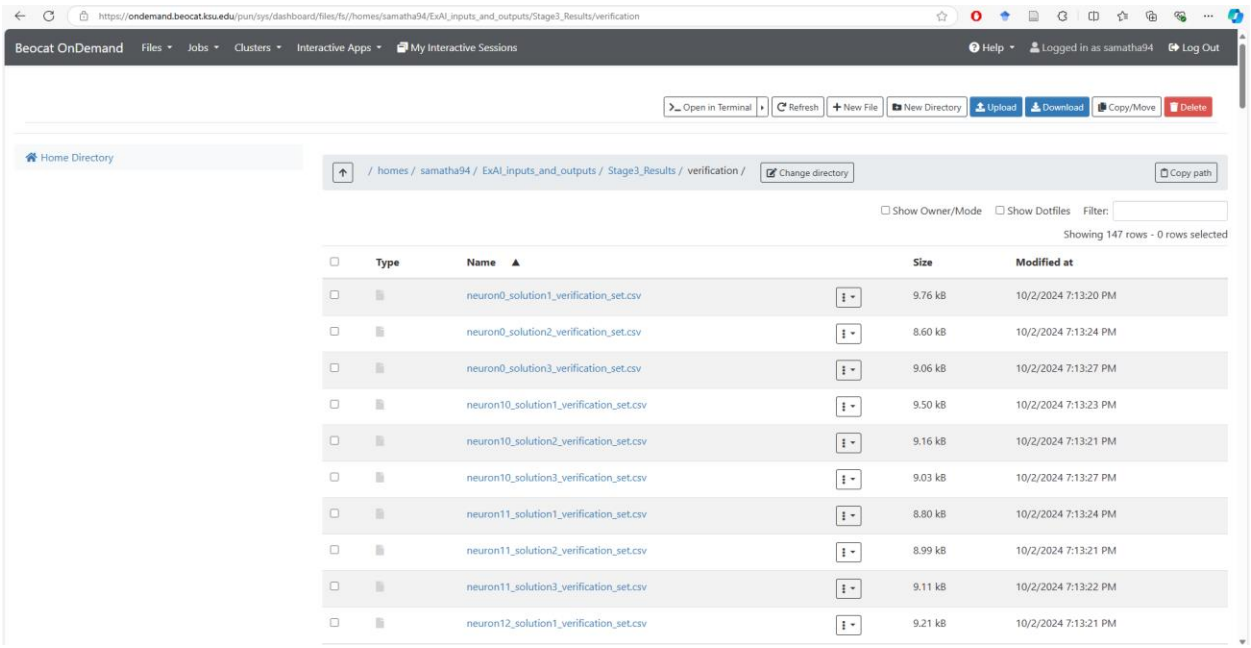


Figure 3.24: Organization of Verification Dataset – This figure illustrates the verification dataset structure, which holds 20% of the retrieved images not used in the evaluation set. Similar to the evaluation set, each CSV file contains neuron activations and predicted class labels.

1. **Indices File:** This file lists the indices of neurons for which valid configuration files were generated in Stage 1. It serves as the basis for identifying neurons to be processed in Stage 3, ensuring that only those neurons with non-empty configuration files undergo the image retrieval and classification procedures.
2. **Google Images:** For each neuron identified in the indices file, a set of images was automatically retrieved from Google Images (Figure 3.18). Using the top three label hypotheses (solutions) for each neuron, keywords were extracted, and up to 100 images per solution were retrieved. The images were stored in a structured folder hierarchy(Figure 3.19), with a folder for each neuron, further subdivided into folders for each solution. Each solution folder contains the images related to the specific label hypothesis, retrieved based on the extracted keywords. The images were saved in JPEG

format, with a minimum resolution of 224x224 pixels, ensuring compatibility with the classification model used in this stage (Figure 3.20 to Figure 3.22).

3. **Evaluation Dataset:** The evaluation set was generated for each neuron and its corresponding solutions using the retrieved images. The evaluation set consists of 80% of the total images, which were classified using the trained CNN model from Stage 1. The evaluation dataset is organized in CSV files(Figure 3.23), containing details such as image filenames, predicted class labels, and neuron activation values. These sets are used to initially assess the activation of the neuron for the specific labels.
4. **Verification Dataset:** The remaining 20% of the retrieved images were allocated to the verification set, which is used for final validation of the neuron activations. Similar to the evaluation set, these images were classified using the trained CNN model, and the results were stored in CSV files(Figure 3.24). The verification dataset includes neuron activations and predicted class labels, allowing for a robust analysis of whether the neuron consistently activates for the hypothesized labels.

Each of these outputs, retrieved images, evaluation sets, and verification sets plays a critical role in confirming or rejecting the label hypotheses generated in Stage 2. The structured outputs enable reproducible, automated analysis, allowing for a comprehensive evaluation of the model's performance in linking neuron activations to high-level semantic concepts.

3.4 Stage 4: Parallelized Image Retrieval and Classification

3.4.1 Technical Requirements

The final stage of the pipeline involves statistical analysis and verification of neuron activations to confirm the hypotheses generated in previous stages. The following technical requirements were established for the successful execution of Stage 4:

1. **Compute Resources:** The job was executed on 1 node with 1 task per node, utilizing a single processing core for each neuron-specific task. This setup provided ample computational power for the statistical analysis of neuron activation metrics.
2. **Memory Allocation:** Each task was allocated 4 GB of memory to handle the loading and processing of neuron-specific activation datasets, ensuring sufficient resources for statistical computations and data handling.
3. **Execution Time:** A job time limit of 1 hour was set for each task, which was more than adequate for completing the statistical analysis tasks for each neuron, including the combination of activation data and the performance of statistical tests.
4. **Software Requirements:** The statistical analysis was performed using Python 3.11.5. A virtual environment was created to manage dependencies, including libraries such as Pandas for data processing, SciPy for statistical tests, and OpenPyXL for Excel file management.
5. **Automated Execution with SLURM:** SLURM workload manager was used to automate the execution process, including environment setup and task execution. A Bash script handled all the steps, from setting up the environment to executing the Python script that performed the statistical analysis on neuron-specific activation data, ensuring the entire process was reproducible and efficient.

3.4.2 Inputs and Data Setup

In Stage 4, statistical analysis and verification of neuron activations rely on organized datasets generated in Stage 3. The input data for this stage includes evaluation and verification sets, each containing neuron-specific activation data. The process begins by loading these datasets, which form the foundation for computing various statistical metrics and validating the label hypotheses. The key components of the input and data setup are outlined below:

1. **Neuron-Specific Evaluation Sets:** The evaluation sets are comprised of activation data collected in Stage 3, representing images that significantly activate each neuron. These datasets, organized in neuron-specific CSV files, include activation values for both target and non-target images. The input process loads these evaluation sets for each neuron and solution into a combined format, ensuring that all relevant data is available for calculating initial metrics such as target activation rates and coverage scores.
2. **Neuron-Specific Verification Sets:** Verification sets contain a separate set of images that undergo statistical testing to confirm the robustness of the label hypotheses. These sets are loaded for neurons identified as having high target activation during the evaluation phase, ensuring that only neurons with significant activation are subjected to further verification.
3. **Neuron ID and Configuration Files:** Neuron IDs are extracted from an indices file that specifies the neurons of interest for the analysis. Additionally, configuration files are referenced to retrieve coverage scores for each neuron and solution. These scores are crucial for evaluating how well the neuron responds to specific labels derived from the ECII tool in previous stages.

4. **Activation Data:** The core input consists of neuron activation values for target and non-target images. These values are aggregated across neurons and solutions, enabling the calculation of activation metrics. The activation data provides a quantitative basis for identifying neurons with meaningful activations, guiding the selection of neurons for further statistical analysis.

This setup ensures that all required inputs, including the activation datasets and neuron-specific evaluation and verification sets, are organized and ready for processing in the subsequent analysis steps. The data setup emphasizes the reproducibility and accuracy of the activation metrics, laying a strong foundation for statistical verification.

3.4.3 Process

In Stage 4, the focus is on analyzing and verifying neuron activations from the evaluation and verification sets (Figure 3.25). The process is divided into two phases: evaluation and verification. This stage ensures that neurons with high target activations are identified and statistically verified through non-parametric tests, providing robust conclusions about label activations.

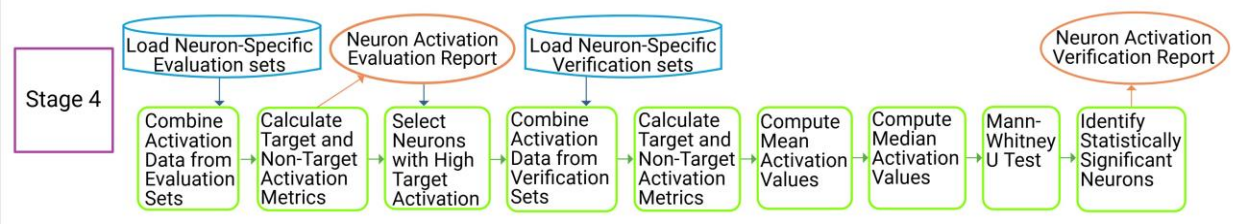


Figure 3.25: Stage 4 workflow for statistical analysis and verification of neuron activations.

1. **Combining Activation Data in the Evaluation Phase:** The process starts by loading the neuron-specific evaluation sets and combining the activation data across neurons for each

solution. Neuron IDs are read from the indices file, and evaluation data from CSV files are consolidated into an Excel workbook for each neuron and solution, with the combined data stored in sheets named solution*_all_activations, where * represents the solution number.

2. **Calculating Activation Metrics in the Evaluation Phase:** For each neuron, the script calculates key metrics such as target and non-target activation percentages. These metrics help identify neurons that strongly activate for target images. Additionally, metrics like activation percentages above 20%, 40%, and 60% of maximum activation are calculated for further insight. These metrics are consolidated into a summary sheet named summary_*. The summary sheets includes columns for neuron ID, target and non-target activation, percentages for activations above the threshold, and the coverage score for each neuron.
3. **Filtering Neurons Based on Target Activation in the Evaluation Phase:** Neurons with high target activation (i.e., above a pre-set threshold of 80%) are selected for further analysis in the verification phase. These neurons are listed in a final_summary_* sheet, which includes only neurons with high target activations, making it easy to identify neurons for the next phase of statistical analysis.
4. **Combining Activation Data in the Verification Phase:** After selecting neurons from the evaluation phase, the script loads neuron-specific verification sets. It then combines the activation data for target and non-target images across neurons, similar to the evaluation phase. The data is stored in solution*_all_activations sheets for each solution in the verification phase.

5. **Calculating Activation Metrics in the Verification Phase:** The script calculates the percentage of target and non-target activations for each neuron, providing an initial indication of neuron performance in the verification phase. The calculated activation percentages are stored in summary sheets named summary_*, with columns for neuron ID, target and non-target activation, and other metrics required for statistical analysis.
6. **Performing Statistical Analysis in the Verification Phase:** For each neuron, the script computes key statistical measures, including the mean and median activation values for both target and non-target images. These metrics provide a detailed view of the distribution of activation values for both categories.

    The Mann-Whitney U test is used to statistically compare target and non-target activations, given its suitability for non-parametric data. The test generates z-scores and p-values, which are recorded in the summary_* sheets, helping to determine whether the target activations are significantly different from non-target activations. Neurons with p-values less than 0.05 are highlighted as statistically significant.

3.4.4 Outputs

The outputs generated in Stage 4 are stored in two main Excel workbooks: evaluation_combined.xlsx and verification_combined.xlsx. Each workbook contains a set of structured sheets summarizing the neuron activation data and statistical analysis results (Figure 3.26).

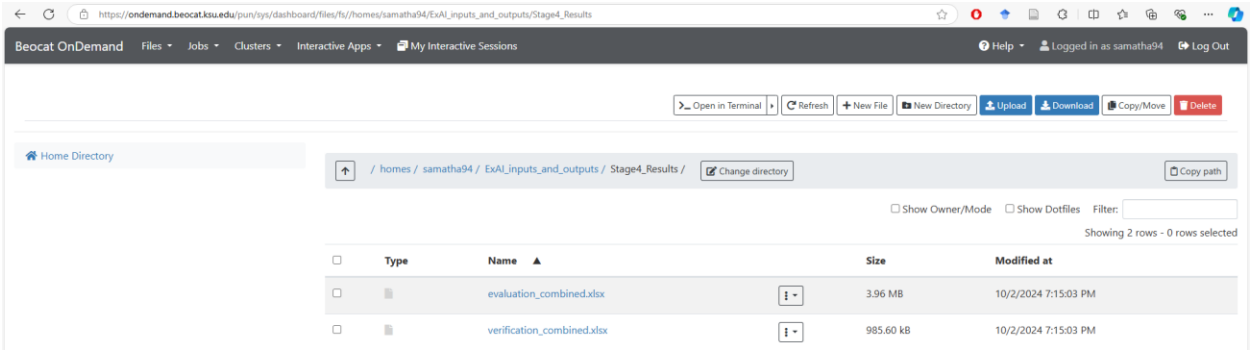


Figure 3.26: Outputs from Stage 4: This figure displays the primary output files generated during Stage 4, organized into two main Excel workbooks: evaluation_combined.xlsx and verification_combined.xlsx. Each workbook consolidates the neuron activation data and statistical analysis results for evaluation and verification phases respectively.

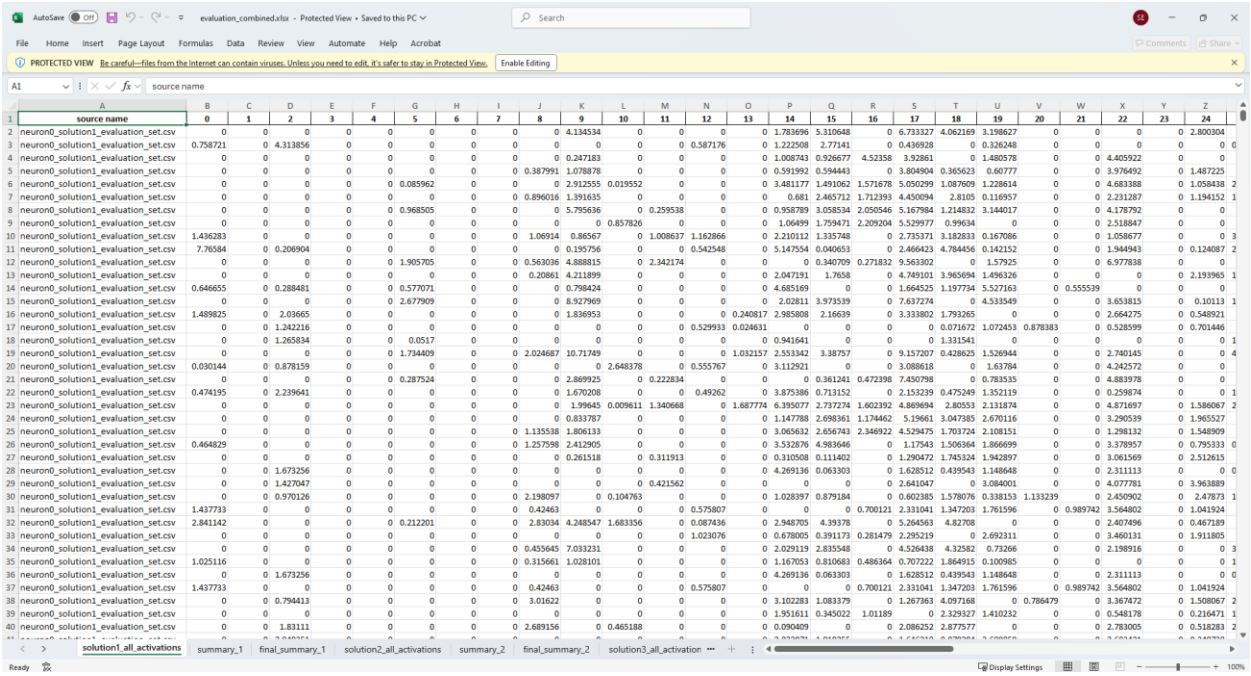
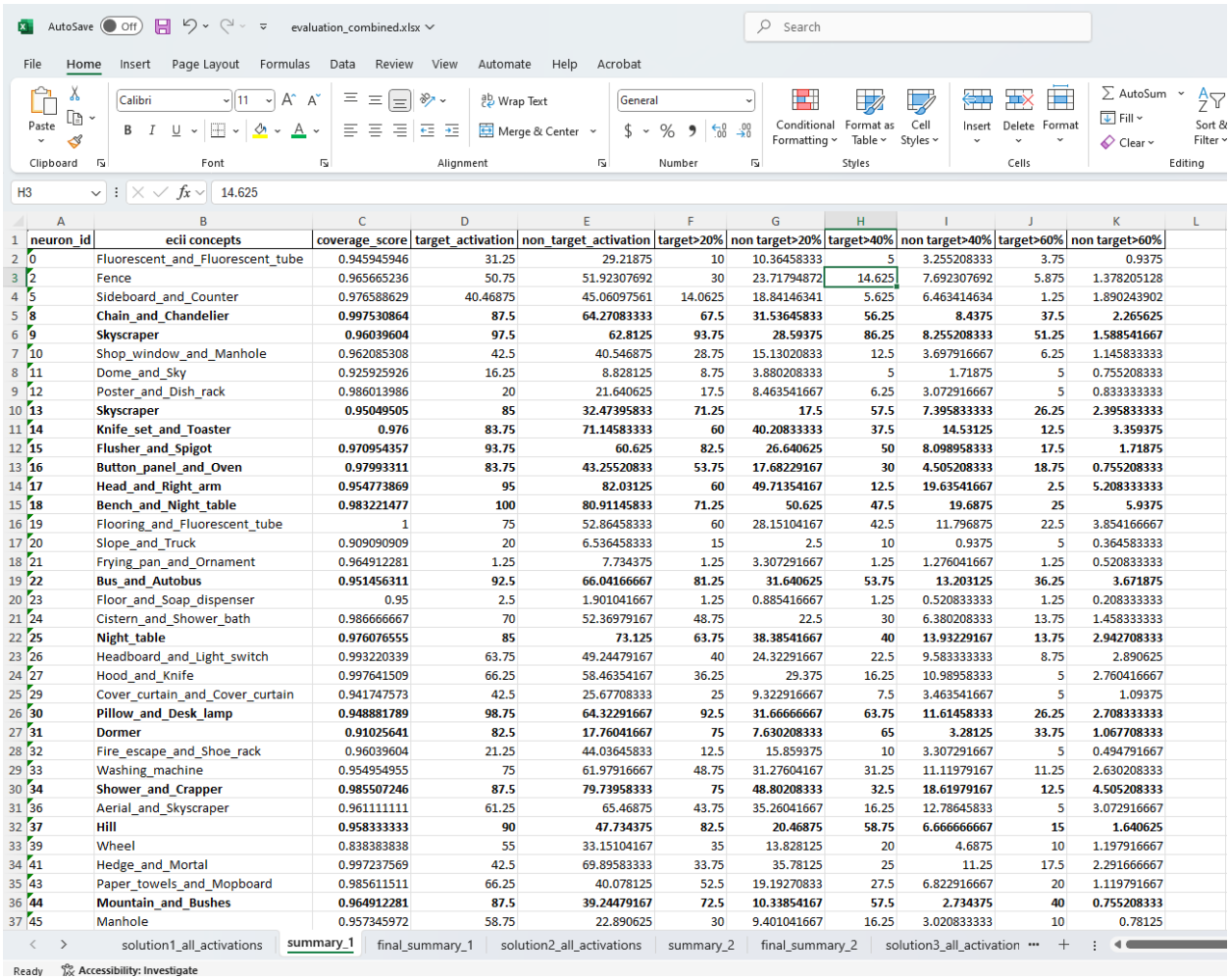
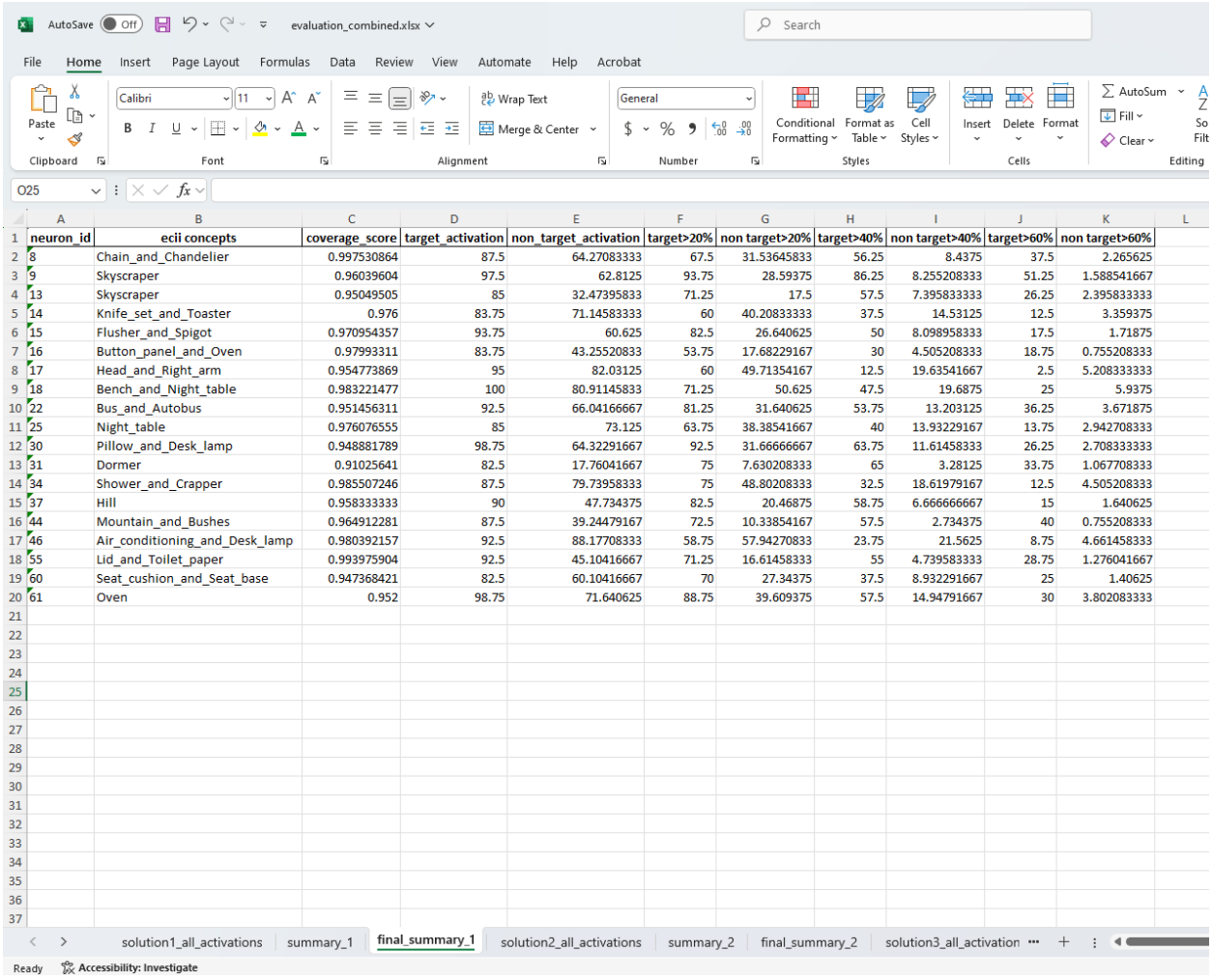


Figure 3.27: Solution-Level Neuron Activation Data in Evaluation Report: This figure illustrates the consolidated neuron activation data in the evaluation_combined.xlsx workbook.



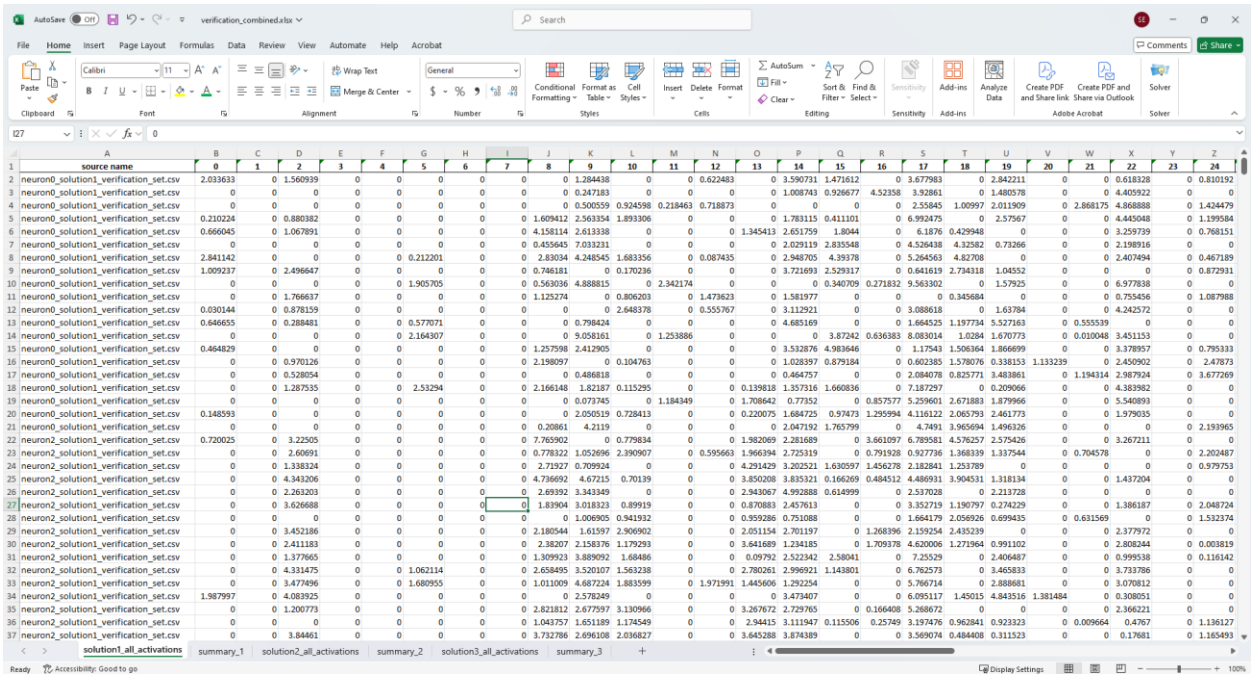
neuron_id	ecii concepts	coverage_score	target_activation	non_target_activation	target>20%	non target>20%	target>40%	non target>40%	target>60%	non target>60%
0	Fluorescent_and_Fluorescent_tube	0.945945946	31.25	29.21875	10	10.36458333	5	3.255208333	3.75	0.9375
2	Fence	0.965665236	50.75	51.92307692	30	23.71794872	14.625	7.692307692	5.875	1.378205128
5	Sideboard_and_Counter	0.976588629	40.46875	45.06097561	14.0625	18.84146341	5.625	6.463414634	1.25	1.890243902
8	**Chain_and_Chandelier**	**0.997530864**	**87.5**	**64.27083333**	**67.5**	**31.53645833**	**56.25**	**8.4375**	**37.5**	**2.265625**
9	**Skyscraper**	**0.96039604**	**97.5**	**62.8125**	**93.75**	**28.59375**	**86.25**	**8.255208333**	**51.25**	**1.588541667**
10	Shop_window_and_Manhole	0.962085308	42.5	40.546875	28.75	15.13020833	12.5	3.697916667	6.25	1.145833333
11	Dome_and_Sky	0.925925926	16.25	8.828125	8.75	3.880208333	5	1.71875	5	0.755208333
12	Poster_and_Dish_rack	0.986013986	20	21.640625	17.5	8.463541667	6.25	3.072916667	5	0.833333333
13	**Skyscraper**	**0.95049505**	**85**	**32.47395833**	**71.25**	**17.5**	**57.5**	**7.395833333**	**26.25**	**2.395833333**
14	**Knife_set_and_Toaster**	**0.976**	**83.75**	**71.14583333**	**60**	**40.20833333**	**37.5**	**14.53125**	**12.5**	**3.359375**
15	**Flusher_and_Spigot**	**0.970954357**	**93.75**	**60.625**	**82.5**	**26.640625**	**50**	**8.098958333**	**17.5**	**1.71875**
16	**Button_panel_and_Oven**	**0.97993311**	**83.75**	**43.25520833**	**53.75**	**17.68229167**	**30**	**4.505208333**	**18.75**	**0.755208333**
17	**Head_and_Right_arm**	**0.954773869**	**95**	**82.03125**	**60**	**49.71354167**	**12.5**	**19.63541667**	**2.5**	**5.208333333**
18	**Bench_and_Night_table**	**0.983221477**	**100**	**80.91145833**	**71.25**	**50.625**	**47.5**	**19.6875**	**25**	**5.9375**
19	Flooring_and_Fluorescent_tube	1	75	52.86458333	60	28.15104167	42.5	11.796875	22.5	3.854166667
20	Slope_and_Truck	0.909090909	20	6.536458333	15	2.5	10	0.9375	5	0.364583333
21	Frying_pan_and_Ornament	0.964912281	1.25	7.734375	1.25	3.307291667	1.25	1.276041667	1.25	0.520833333
22	**Bus_and_Autobus**	**0.951456311**	**92.5**	**66.04166667**	**81.25**	**31.640625**	**53.75**	**13.203125**	**36.25**	**3.671875**
23	Floor_and_Soap_dispenser	0.95	2.5	1.901041667	1.25	0.885416667	1.25	0.520833333	1.25	0.208333333
24	Cistern_and_Shower_bath	0.986666667	70	52.36979167	48.75	22.5	30	6.380208333	13.75	1.458333333
25	**Night_table**	**0.976076555**	**85**	**73.125**	**63.75**	**38.38541667**	**40**	**13.93229167**	**13.75**	**2.942708333**
26	Headboard_and_Light_switch	0.993220339	63.75	49.24479167	40	24.32291667	22.5	9.583333333	8.75	2.890625
27	Hood_and_Knife	0.997641509	66.25	58.46354167	36.25	29.375	16.25	10.98958333	5	2.760416667
29	Cover_curtain_and_Cover_curtain	0.941747573	42.5	25.67708333	25	9.322916667	7.5	3.463541667	5	1.09375
30	**Pillow_and_Desk_lamp**	**0.948881789**	**98.75**	**64.32291667**	**92.5**	**31.66666667**	**63.75**	**11.61458333**	**26.25**	**2.708333333**
31	**Dormer**	**0.91025641**	**82.5**	**17.76041667**	**75**	**7.630208333**	**65**	**3.28125**	**33.75**	**1.067708333**
32	Fire_escape_and_Shoe_rack	0.96039604	21.25	44.03645833	12.5	15.859375	10	3.307291667	5	0.494791667
33	Washing_machine	0.954954955	75	61.97916667	48.75	31.27604167	31.25	11.11979167	11.25	2.630208333
34	**Shower_and_Crapper**	**0.985507246**	**87.5**	**79.73958333**	**75**	**48.80208333**	**32.5**	**18.61979167**	**12.5**	**4.505208333**
36	Aerial_and_Skyscraper	0.961111111	61.25	65.46875	43.75	35.26041667	16.25	12.78645833	5	3.072916667
37	**Hill**	**0.958333333**	**90**	**47.734375**	**82.5**	**20.46875**	**58.75**	**6.666666667**	**15**	**1.640625**
39	Wheel	0.838383838	55	33.15104167	35	13.828125	20	4.6875	10	1.197916667
41	Hedge_and_Mortal	0.997237569	42.5	69.89583333	33.75	35.78125	25	11.25	17.5	2.291666667
43	Paper_towels_and_Mopboard	0.985611511	66.25	40.078125	52.5	19.19270833	27.5	6.822916667	20	1.119791667
44	**Mountain_and_Bushes**	**0.964912281**	**87.5**	**39.24479167**	**72.5**	**10.33854167**	**57.5**	**2.734375**	**40**	**0.755208333**
45	Manhole	0.957345972	58.75	22.890625	30	9.401041667	16.25	3.020833333	10	0.78125

Figure 3.28: Summary of Neuron Activation Metrics in Evaluation Report: This figure presents the summary_1 sheet within the evaluation_combined.xlsx workbook. It provides key metrics for each neuron, such as the coverage score, target and non-target activations, and activation percentages at different thresholds. This summary serves as an initial filter for identifying neurons with high target activations.



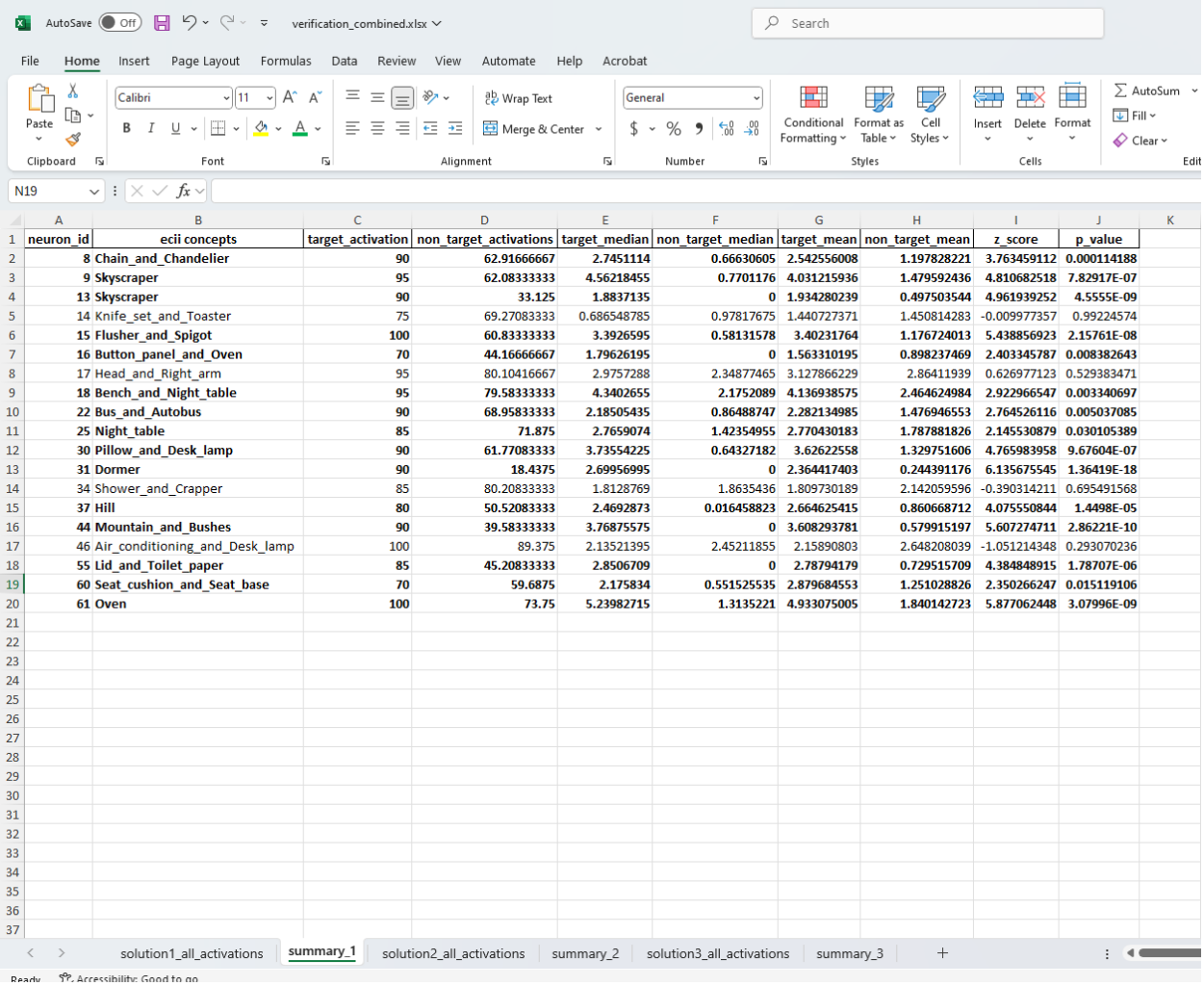
neuron_id	ecii concepts	coverage_score	target_activation	non_target_activation	target>20%	non target>20%	target>40%	non target>40%	target>60%	non target>60%
8	Chain_and_Chandelier	0.997530864	87.5	64.27083333	67.5	31.53645833	56.25	8.4375	37.5	2.265625
9	Skyscraper	0.96039604	97.5	62.8125	93.75	28.59375	86.25	8.255208333	51.25	1.588541667
13	Skyscraper	0.95049505	85	32.47395833	71.25	17.5	57.5	7.395833333	26.25	2.395833333
14	Knife_set_and_Toaster	0.976	83.75	71.14583333	60	40.20833333	37.5	14.53125	12.5	3.359375
15	Flusher_and_Spigot	0.970954357	93.75	60.625	82.5	26.640625	50	8.098958333	17.5	1.71875
16	Button_panel_and_Oven	0.97993311	83.75	43.25520833	53.75	17.68229167	30	4.505208333	18.75	0.755208333
17	Head_and_Right_arm	0.954773869	95	82.03125	60	49.71354167	12.5	19.63541667	2.5	5.208333333
18	Bench_and_Night_table	0.983221477	100	80.91145833	71.25	50.625	47.5	19.6875	25	5.9375
22	Bus_and_Autobus	0.951456311	92.5	66.04166667	81.25	31.640625	53.75	13.203125	36.25	3.671875
25	Night_table	0.976076555	85	73.125	63.75	38.38541667	40	13.93229167	13.75	2.942708333
30	Pillow_and_Desk_lamp	0.948881789	98.75	64.32291667	92.5	31.66666667	63.75	11.61458333	26.25	2.708333333
31	Dormer	0.91025641	82.5	17.76041667	75	7.630208333	65	3.28125	33.75	1.067708333
34	Shower_and_Crapper	0.985507246	87.5	79.73958333	75	48.80208333	32.5	18.61979167	12.5	4.505208333
37	Hill	0.958333333	90	47.734375	82.5	20.46875	58.75	6.666666667	15	1.640625
44	Mountain_and_Bushes	0.964912281	87.5	39.24479167	72.5	10.33854167	57.5	2.734375	40	0.755208333
46	Air_conditioning_and_Desk_lamp	0.980392157	92.5	88.17708333	58.75	57.94270833	23.75	21.5625	8.75	4.661458333
55	Lid_and_Toilet_paper	0.993975904	92.5	45.10416667	71.25	16.61458333	55	4.739583333	28.75	1.276041667
60	Seat_cushion_and_Seat_base	0.947368421	82.5	60.10416667	70	27.34375	37.5	8.932291667	25	1.40625
61	Oven	0.952	98.75	71.640625	88.75	39.609375	57.5	14.94791667	30	3.802083333

Figure 3.29: Final Summary of Neurons with High Target Activations in Evaluation Report: This figure shows the final_summary_1 sheet in the evaluation_combined.xlsx workbook. It highlights neurons that have surpassed the 80% target activation threshold, marking them as candidates for further validation during the verification phase.



source name	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
neuron0_solution1_verification_set.csv	2.033633	0	1.560939	0	0	0	0	0	0	1.284438	0	0	0.622483	0	3.590731	1.471612	0	3.677983	0	2.842211	0	0	0.618328	0	0.810192
neuron0_solution1_verification_set.csv	0	0	0	0	0	0	0	0	0	0.247183	0	0	0	0	1.008743	0.926677	4.52358	3.92861	0	1.480578	0	0	4.405922	0	0
neuron0_solution1_verification_set.csv	0	0	0	0	0	0	0	0	0	0.500559	0.924598	0.218463	0.718873	0	0	0	0	2.55845	1.00997	2.011909	0	2.868175	4.868888	0	1.424479
neuron0_solution1_verification_set.csv	0.210224	0	0.880382	0	0	0	0	0	1.609412	2.563354	1.893306	0	0	0	1.783115	0.411101	0	6.992475	0	2.57567	0	0	4.445048	0	1.199584
neuron0_solution1_verification_set.csv	0.666045	0	1.067891	0	0	0	0	0	4.158114	2.613338	0	0	0	1.345413	2.651759	1.8044	0	6.1876	0.429948	0	0	0	3.259739	0	0.768151
neuron0_solution1_verification_set.csv	0	0	0	0	0	0	0	0	0.455645	7.033231	0	0	0	0	2.029119	2.835548	0	4.526438	4.32582	0.73266	0	0	2.198916	0	0
neuron0_solution1_verification_set.csv	2.841142	0	0	0	0	0.212201	0	0	2.83034	4.248545	1.683356	0	0.087435	0	2.948705	4.39378	0	5.264563	4.82708	0	0	0	2.407494	0	0.467189
neuron0_solution1_verification_set.csv	1.009237	0	2.496647	0	0	0	0	0	0.746181	0	0.170236	0	0	0	3.721693	2.529317	0	0.641619	2.734318	1.04552	0	0	0	0	0.872931
neuron0_solution1_verification_set.csv	0	0	0	0	0	1.905705	0	0	0.563036	4.888815	0	2.342174	0	0	0	0.340709	0.271832	9.563302	0	1.57925	0	0	6.977838	0	0
neuron0_solution1_verification_set.csv	0	0	1.766637	0	0	0	0	0	1.125274	0	0.806203	0	1.473623	0	1.581977	0	0	0	0.345684	0	0	0	0.755456	0	1.087988
neuron0_solution1_verification_set.csv	0.030144	0	0.878159	0	0	0	0	0	0	0	2.648378	0	0.555767	0	3.112921	0	0	3.088618	0	1.63784	0	0	4.242572	0	0
neuron0_solution1_verification_set.csv	0.646655	0	0.288481	0	0	0.577071	0	0	0	0.798424	0	0	0	0	4.685169	0	0	1.664525	1.197734	5.527163	0	0.555539	0	0	0
neuron0_solution1_verification_set.csv	0	0	0	0	0	2.164307	0	0	0	9.058161	0	1.253886	0	0	0	3.87242	0.636383	8.083014	1.0284	1.670773	0	0.010048	3.451153	0	0
neuron0_solution1_verification_set.csv	0.464829	0	0	0	0	0	0	0	1.257598	2.412905	0	0	0	0	3.532876	4.983646	0	1.17543	1.506364	1.866699	0	0	3.378957	0	0.795333
neuron0_solution1_verification_set.csv	0	0	0.970126	0	0	0	0	0	2.198097	0	0.104763	0	0	0	1.028397	0.879184	0	0.602385	1.578076	0.338153	1.133239	0	2.450902	0	2.47873
neuron0_solution1_verification_set.csv	0	0	0.528054	0	0	0	0	0	0	0.486818	0	0	0	0	0.464757	0	0	2.084078	0.825771	3.483861	0	1.194314	2.987924	0	3.677269
neuron0_solution1_verification_set.csv	0	0	1.287535	0	0	2.53294	0	0	2.166148	1.82187	0.115295	0	0	0.139818	1.357316	1.660836	0	7.187297	0	0.209066	0	0	4.383982	0	0
neuron0_solution1_verification_set.csv	0	0	0	0	0	0	0	0	0	0.073745	0	1.184349	0	1.708642	0.77352	0	0.857577	5.259601	2.671883	1.879966	0	0	5.540893	0	0
neuron0_solution1_verification_set.csv	0.148593	0	0	0	0	0	0	0	0	2.050519	0.728413	0	0	0.220075	1.684725	0.97473	1.295994	4.116122	2.065793	2.461773	0	0	1.979035	0	0
neuron0_solution1_verification_set.csv	0	0	0	0	0	0	0	0	0.20861	4.2119	0	0	0	0	2.047192	1.765799	0	4.7491	3.965694	1.496326	0	0	0	0	2.193965
neuron2_solution1_verification_set.csv	0.720025	0	3.22505	0	0	0	0	0	7.765902	0	0.779834	0	0	1.982069	2.281689	0	3.661097	6.789581	4.576257	2.575426	0	0	3.267211	0	0
neuron2_solution1_verification_set.csv	0	0	2.60691	0	0	0	0	0	0.778322	1.052696	2.390907	0	0.595663	1.966394	2.725319	0	0.791928	0.927736	1.368339	1.337544	0	0.704578	0	0	2.202487
neuron2_solution1_verification_set.csv	0	0	1.338324	0	0	0	0	0	2.71927	0.709924	0	0	0	4.291429	3.202521	1.630597	1.456278	2.182841	1.253789	0	0	0	0	0	0.979753
neuron2_solution1_verification_set.csv	0	0	4.343206	0	0	0	0	0	4.736692	4.67215	0.70139	0	0	3.850208	3.835321	0.166269	0.484512	4.486931	3.904531	1.318134	0	0	1.437204	0	0
neuron2_solution1_verification_set.csv	0	0	2.263203	0	0	0	0	0	2.69392	3.343349	0	0	0	2.943067	4.992888	0.614999	0	2.537028	0	2.213728	0	0	0	0	0
neuron2_solution1_verification_set.csv	0	0	3.626688	0	0	0	0	0	1.83904	3.018323	0.89919	0	0	0.870883	2.457613	0	0	3.352719	1.190797	0.274229	0	0	1.386187	0	2.048724
neuron2_solution1_verification_set.csv	0	0	0	0	0	0	0	0	0	1.006905	0.941932	0	0	0.959286	0.751088	0	0	1.664179	2.056926	0.699435	0	0.631569	0	0	1.532374
neuron2_solution1_verification_set.csv	0	0	3.452186	0	0	0	0	0	2.180544	1.61597	2.906902	0	0	2.051154	2.701197	0	1.268396	2.159254	2.435239	0	0	0	2.377972	0	0
neuron2_solution1_verification_set.csv	0	0	2.411183	0	0	0	0	0	2.38207	2.158376	1.179293	0	0	3.641689	1.234185	0	1.709378	4.620006	1.271964	0.991102	0	0	2.808244	0	0.003819
neuron2_solution1_verification_set.csv	0	0	1.377665	0	0	0	0	0	1.309923	3.889092	1.68486	0	0	0.09792	2.522342	2.58041	0	7.25529	0	2.406487	0	0	0.999538	0	0.116142
neuron2_solution1_verification_set.csv	0	0	4.331475	0	0	1.062114	0	0	2.658495	3.520107	1.563238	0	0	2.780261	2.996921	1.143801	0	6.762573	0	3.465833	0	0	3.733786	0	0
neuron2_solution1_verification_set.csv	0	0	3.477496	0	0	1.680955	0	0	1.011009	4.687224	1.883599	0	1.971991	1.445606	1.292254	0	0	5.766714	0	2.888681	0	0	3.070812	0	0
neuron2_solution1_verification_set.csv	1.987997	0	4.083925	0	0	0	0	0	0	2.578249	0	0	0	0	3.473407	0	0	6.095117	1.45015	4.843516	1.381484	0	0.308051	0	0
neuron2_solution1_verification_set.csv	0	0	1.200773	0	0	0	0	0	2.821812	2.677597	3.130966	0	0	3.267672	2.729765	0	0.166408	5.268672	0	0	0	0	2.366221	0	0
neuron2_solution1_verification_set.csv	0	0	0	0	0	0	0	0	1.043757	1.651189	1.174549	0	0	2.94415	3.111947	0.115506	0.25749	3.197476	0.962841	0.923323	0	0.009664	0.4767	0	1.136127
neuron2_solution1_verification_set.csv	0	0	3.84461	0	0	0	0	0	3.732786	2.696108	2.036827	0	0	3.645288	3.874389	0	0	3.569074	0.484408	0.311523	0	0	0.17681	0	1.165493

Figure 3.30: Solution-Level Neuron Activation Data in Verification Report: This figure illustrates the consolidated neuron activation data in the verification_combined.xlsx workbook.



neuron_id	ecii concepts	target_activation	non_target_activations	target_median	non_target_median	target_mean	non_target_mean	z_score	p_value
8	Chain_and_Chandelier	90	62.91666667	2.7451114	0.66630605	2.542556008	1.197828221	3.763459112	0.000114188
9	Skyscraper	95	62.08333333	4.56218455	0.7701176	4.031215936	1.479592436	4.810682518	7.82917E-07
13	Skyscraper	90	33.125	1.8837135	0	1.934280239	0.497503544	4.961939252	4.5555E-09
14	Knife_set_and_Toaster	75	69.27083333	0.686548785	0.97817675	1.440727371	1.450814283	-0.009977357	0.99224574
15	Flusher_and_Spigot	100	60.83333333	3.3926595	0.58131578	3.40231764	1.176724013	5.438856923	2.15761E-08
16	Button_panel_and_Oven	70	44.16666667	1.79626195	0	1.563310195	0.898237469	2.403345787	0.008382643
17	Head_and_Right_arm	95	80.10416667	2.9757288	2.34877465	3.127866229	2.86411939	0.626977123	0.529383471
18	Bench_and_Night_table	95	79.58333333	4.3402655	2.1752089	4.136938575	2.464624984	2.922966547	0.003340697
22	Bus_and_Autobus	90	68.95833333	2.18505435	0.86488747	2.282134985	1.476946553	2.764526116	0.005037085
25	Night_table	85	71.875	2.7659074	1.42354955	2.770430183	1.787881826	2.145530879	0.030105389
30	Pillow_and_Desk_lamp	90	61.77083333	3.73554225	0.64327182	3.62622558	1.329751606	4.765983958	9.67604E-07
31	Dormer	90	18.4375	2.69956995	0	2.364417403	0.244391176	6.135675545	1.36419E-18
34	Shower_and_Crapper	85	80.20833333	1.8128769	1.8635436	1.809730189	2.142059596	-0.390314211	0.695491568
37	Hill	80	50.52083333	2.4692873	0.016458823	2.664625415	0.860668712	4.075550844	1.4498E-05
44	Mountain_and_Bushes	90	39.58333333	3.76875575	0	3.608293781	0.579915197	5.607274711	2.86221E-10
46	Air_conditioning_and_Desk_lamp	100	89.375	2.13521395	2.45211855	2.15890803	2.648208039	-1.051214348	0.293070236
55	Lid_and_Toilet_paper	85	45.20833333	2.8506709	0	2.78794179	0.729515709	4.384848915	1.78707E-06
60	Seat_cushion_and_Seat_base	70	59.6875	2.175834	0.551525535	2.879684553	1.251028826	2.350266247	0.015119106
61	Oven	100	73.75	5.23982715	1.3135221	4.933075005	1.840142723	5.877062448	3.07996E-09

Figure 3.31: Statistical Summary of Neuron Activations in Verification Report: This figure presents the summary_1 sheet within the verification_combined.xlsx workbook. It includes key statistical measures such as target and non-target activation means, medians, z-scores, and p-values, highlighting neurons with statistically significant differences in activations for robust validation of target labels.

1. **Neuron Activation Evaluation Report:**

   Located in the evaluation_combined.xlsx file, this report includes:

   - The **solution*_all_activations sheets** consolidate the activation data for each neuron across different solutions, showing activations for both target and non-

target images (Figure 3.27). These sheets serve as the foundation for understanding neuron-specific activation patterns in the evaluation phase.

- The **summary_ sheets*** present a detailed overview of neuron-specific metrics such as neuron_id, ecii concepts, coverage_score, and percentages of target and non-target activations across various thresholds (>20%, >40%, and >60%) (Figure 3.28). These sheets highlight neurons with high target activations.
- The **final_summary_ sheets*** filter neurons with over 80% target activation from the summary sheets, indicating neurons that proceed to the verification phase (Figure 3.29). These filtered neurons are prime candidates for further validation.

2. **Neuron Activation Verification Report:**

Stored in the verification_combined.xlsx file, this report includes:

- The **solution*_all_activations sheets** for the verification phase, which compile neuron-specific activation data, similar to the evaluation phase, but with a focus on images used for statistical verification (Figure 3.30).
- The **summary_ sheets*** in this workbook present a detailed statistical analysis, including columns for neuron_id, ecii concepts, target_activation, non_target_activations, as well as key statistical measures like target_median, non_target_median, target_mean, non_target_mean, z_score, and p_value. Neurons with statistically significant differences (p-value < 0.05) are highlighted as those showing robust and validated label activations (Figure 3.31).

Both reports offer insights into neuron activation behavior during the evaluation and verification phases, facilitating the final validation of neuron-label hypotheses through detailed statistical analysis.

Chapter 4 - Results

4.1 Stage 1 Results

The completion of the automated model training and data configuration process was monitored through the SLURM workload manager. The job was successfully completed within the allocated time and resources (Figure 4.1). The key metrics obtained from the job execution are summarized below:

1. **Job Completion Status:** Completed successfully with JobID 17796633.
2. **Total CPU Time:** 2 hours, 34 minutes, and 12 seconds, with the job running across 24 CPUs.
3. **Elapsed Time:** The total elapsed time for the job was 38 minutes and 50 seconds.
4. **Memory Usage:** The maximum resident set size (MaxRSS) during execution was 4.87 GB, which was well within the requested memory allocation of 20 GB(i.e. 5GB per node).
5. **Node Allocation:** The job was distributed across 4 nodes, utilizing 6 cores per node for parallel execution.

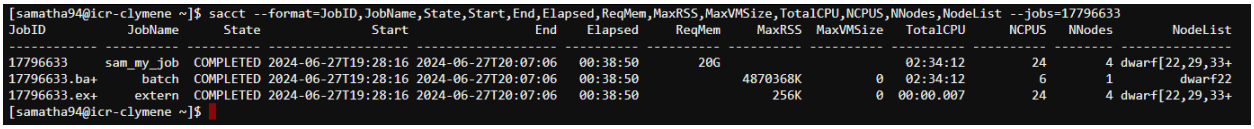


Figure 4.1: SLURM Job Execution Summary for Stage 1: This figure displays the SLURM workload manager's summary of job execution for the automated model training and data configuration process.

These results demonstrate the efficient use of computational resources, with the automated pipeline completing the model training and configuration well within the expected time limits and resource allocations.

4.2 Stage 2 Results

The parallelized concept induction and label hypothesis generation for Stage 2 were executed efficiently using the SLURM workload manager on the Beocat high-performance computing cluster (Figure 4.2). The automation pipeline, designed to analyze all 64 neurons simultaneously, significantly reduced the overall processing time, demonstrating the scalability and efficiency of the pipeline. The key metrics from the job execution are summarized below:

1. **Job Completion Status:** All neuron-specific concept induction tasks were successfully completed with JobID 17820917.
2. **Total CPU Time:** The total CPU time for individual tasks ranged from approximately 24 seconds to 16 minutes, depending on the complexity of the neuron-specific analysis.
3. **Elapsed Time:** The total elapsed time for processing all 64 neurons was approximately 17 minutes. This represents a significant reduction in time compared to sequential execution, which would have taken over 10 hours to complete.
4. **Memory Usage:** The maximum resident set size (MaxRSS) for individual tasks ranged from approximately 800 MB to 22 GB, with a maximum requested memory of 38 GB for each neuron. The memory usage was efficiently distributed across tasks, ensuring no bottlenecks or resource shortages during execution.
5. **Node Allocation:** Each neuron-specific task was allocated to a separate node, utilizing one core per task to prevent resource contention. In total, 64 tasks were distributed across various nodes in the cluster (e.g., hero, dwarf, warlock, and wizard nodes), ensuring parallel execution and optimal resource utilization.
6. **Efficiency Gains:** The use of parallel processing reduced the total execution time from over 10 hours (in a sequential setup) to under 17 minutes. This drastic improvement

highlights the efficiency and scalability of the automated pipeline, allowing the concept induction process to be completed in a fraction of the time.

Host: headnode.beocat.ksu.edu

```
[samatha94@icr-clymene ~]$ sacct --format=JobID,JobName,State,Start,End,Elapsed,ReqMem,MaxRSS,MaxVMSize,TotalCPU,NCPUS,NNodes,NodeList --jobs=17820917

JobID           JobName      State               Start                 End    Elapsed     ReqMem     MaxRSS  MaxVMSize   TotalCPU      NCPUS   NNodes        NodeList
------------ ---------- ---------- ------------------- ------------------- ---------- ---------- ---------- ---------- ---------- ---------- -------- ---------------
17820917_0   sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:51   00:09:34        38G                    07:57.619          1        1          hero53
17820917_0.+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:51   00:09:34             20784304K         0  07:57.618          1        1          hero53
17820917_0.+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:51   00:09:34               188K         0   00:00:00          1        1          hero53
17820917_1   sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:25   00:02:08        38G                    00:40.453          1        1         dwarf50
17820917_1.+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:25   00:02:08              886744K         0  00:40.452          1        1         dwarf50
17820917_1.+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:25   00:02:08               188K         0  00:00.001          1        1         dwarf50
17820917_2   sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:29   00:09:12        38G                    07:36.984          1        1          hero17
17820917_2.+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:29   00:09:12             20569236K         0  07:36.983          1        1          hero17
17820917_2.+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:29   00:09:12               192K         0   00:00:00          1        1          hero17
17820917_3   sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:20   00:02:03        38G                    00:30.330          1        1          hero22
17820917_3.+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:20   00:02:03              892264K         0  00:30.329          1        1          hero22
17820917_3.+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:20   00:02:03               192K         0   00:00:00          1        1          hero22
17820917_4   sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:16   00:01:59        38G                    00:27.381          1        1          hero23
17820917_4.+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:16   00:01:59              846688K         0  00:27.380          1        1          hero23
17820917_4.+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:16   00:01:59               192K         0   00:00:00          1        1          hero23
17820917_5   sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:07   00:09:50        38G                    07:45.076          1        1          hero25
17820917_5.+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:07   00:09:50             20319004K         0  07:45.075          1        1          hero25
17820917_5.+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:07   00:09:50               184K         0   00:00:00          1        1          hero25
17820917_6   sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:49   00:02:32        38G                    00:28.771          1        1          hero25
17820917_6.+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:49   00:02:32              865196K         0  00:28.770          1        1          hero25
17820917_6.+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:49   00:02:32               188K         0   00:00:00          1        1          hero25
17820917_7   sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:17   00:02:00        38G                    00:28.991          1        1          hero27
17820917_7.+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:17   00:02:00               836M         0  00:28.991          1        1          hero27
17820917_7.+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:17   00:02:00               192K         0   00:00:00          1        1          hero27
17820917_8   sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:27   00:10:10        38G                    08:01.663          1        1          hero29
17820917_8.+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:27   00:10:10             18203052K         0  08:01.662          1        1          hero29
17820917_8.+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:27   00:10:10               184K         0   00:00:00          1        1          hero29
17820917_9   sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:48   00:09:31        38G                    07:24.893          1        1          hero29
17820917_9.+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:48   00:09:31             18761104K         0  07:24.892          1        1          hero29
17820917_9.+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:48   00:09:31                 0             0   00:00:00          1        1          hero29
17820917_10  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:35   00:10:18        38G                    08:13.541          1        1          hero30
17820917_10+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:35   00:10:18             18801708K         0  08:13.540          1        1          hero30
17820917_10+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:35   00:10:18                 0             0   00:00:00          1        1          hero30
17820917_11  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:28   00:10:11        38G                    07:58.194          1        1          hero30
17820917_11+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:28   00:10:11             20131632K         0  07:58.194          1        1          hero30
17820917_11+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:28   00:10:11               196K         0   00:00:00          1        1          hero30
17820917_12  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:21   00:12:04        38G                    09:50.814          1        1       warlock12
17820917_12+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:21   00:12:04             16224880K         0  09:50.813          1        1       warlock12
17820917_12+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:21   00:12:04                 0             0  00:00.001          1        1       warlock12
17820917_13  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:00   00:11:43        38G                    09:46.306          1        1       warlock12
17820917_13+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:00   00:11:43             20379096K         0  09:46.305          1        1       warlock12
17820917_13+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:00   00:11:43               196K         0  00:00.001          1        1       warlock12
17820917_16  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:04   00:09:47        38G                    07:54.320          1        1       warlock14
17820917_16+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:04   00:09:47             19148608K         0  07:54.318          1        1       warlock14
17820917_16+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:04   00:09:47                 0             0  00:00.001          1        1       warlock14
```

Host: headnode.beocat.ksu.edu

```
17820917_17  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:19:41   00:13:24        38G                      08:34.747         1         1      warlock14
17820917_17+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:19:41   00:13:24           15840836K         0  08:34.746         1         1      warlock14
17820917_17+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:19:41   00:13:24                   0         0   00:00:00         1         1      warlock14
17820917_18  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:54   00:09:37        38G                      07:43.776         1         1      warlock14
17820917_18+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:54   00:09:37           17513348K         0  07:43.775         1         1      warlock14
17820917_18+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:54   00:09:37                   0         0   00:00.001         1         1      warlock14
17820917_19  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:20:13   00:13:56        38G                      08:34.366         1         1      warlock14
17820917_19+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:20:13   00:13:56           15537112K         0  08:34.365         1         1      warlock14
17820917_19+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:20:13   00:13:56                   0         0   00:00:00         1         1      warlock14
17820917_20  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:38   00:10:21        38G                      08:29.572         1         1      warlock14
17820917_20+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:38   00:10:21           15836224K         0  08:29.571         1         1      warlock14
17820917_20+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:38   00:10:21                   0         0   00:00:00         1         1      warlock14
17820917_21  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:34   00:10:17        38G                      08:23.334         1         1      warlock14
17820917_21+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:34   00:10:17           16262848K         0  08:23.333         1         1      warlock14
17820917_21+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:34   00:10:17                   0         0   00:00:00         1         1      warlock14
17820917_22  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:41   00:10:24        38G                      08:31.016         1         1      warlock14
17820917_22+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:41   00:10:24           16836132K         0  08:31.015         1         1      warlock14
17820917_22+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:41   00:10:24                   0         0   00:00:00         1         1      warlock14
17820917_23  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:29   00:10:12        38G                      08:18.697         1         1      warlock14
17820917_23+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:29   00:10:12           15874496K         0  08:18.696         1         1      warlock14
17820917_23+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:29   00:10:12                   0         0   00:00:00         1         1      warlock14
17820917_24  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:20:39   00:14:22        38G                      08:28.601         1         1      warlock14
17820917_24+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:20:39   00:14:22           16160344K         0  08:28.600         1         1      warlock14
17820917_24+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:20:39   00:14:22                   0         0   00:00:00         1         1      warlock14
17820917_25  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:51   00:09:34        38G                      07:42.157         1         1      warlock14
17820917_25+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:51   00:09:34           16606564K         0  07:42.156         1         1      warlock14
17820917_25+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:15:51   00:09:34                   0         0   00:00:00         1         1      warlock14
17820917_26  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:20:39   00:14:22        38G                      08:37.180         1         1      warlock14
17820917_26+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:20:39   00:14:22             15436M         0  08:37.179         1         1      warlock14
17820917_26+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:20:39   00:14:22                   0         0   00:00:00         1         1      warlock14
17820917_27  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:07   00:09:50        38G                      07:58.155         1         1      warlock14
17820917_27+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:07   00:09:50           15220172K         0  07:58.154         1         1      warlock14
17820917_27+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:16:07   00:09:50                   0         0   00:00.001         1         1      warlock14
17820917_28  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:11:03   00:04:46        38G                      00:24.490         1         1      warlock14
17820917_28+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:11:03   00:04:46             878816K         0  00:24.489         1         1      warlock14
17820917_28+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:11:03   00:04:46                   0         0   00:00:00         1         1      warlock14
17820917_29  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:48   00:16:31        38G                      09:28.439         1         1      warlock15
17820917_29+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:48   00:16:31           15746236K         0  09:28.438         1         1      warlock15
17820917_29+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:48   00:16:31                   0         0   00:00:00         1         1      warlock15
17820917_30  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:58   00:16:41        38G                      09:38.991         1         1      warlock15
17820917_30+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:58   00:16:41           16265272K         0  09:38.990         1         1      warlock15
17820917_30+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:58   00:16:41                   0         0   00:00:00         1         1      warlock15
17820917_31  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:50   00:16:33        38G                      09:27.718         1         1      warlock15
17820917_31+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:50   00:16:33           16199180K         0  09:27.717         1         1      warlock15
17820917_31+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:50   00:16:33                   0         0   00:00:00         1         1      warlock15
17820917_32  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:32   00:16:15        38G                      09:11.840         1         1      warlock15
17820917_32+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:32   00:16:15           16134364K         0  09:11.839         1         1      warlock15
17820917_32+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:32   00:16:15                   0         0   00:00:00         1         1      warlock15
```

Host: headnode.beocat.ksu.edu

```
17820917_33  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:49   00:16:32        38G                      09:31.396         1         1      warlock15
17820917_33+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:49   00:16:32           16957104K         0  09:31.395         1         1      warlock15
17820917_33+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:49   00:16:32                   0         0   00:00:00         1         1      warlock15
17820917_34  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:00   00:11:43        38G                      09:29.910         1         1      warlock15
17820917_34+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:00   00:11:43           15944852K         0  09:29.909         1         1      warlock15
17820917_34+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:03   00:11:46                   0         0   00:00:00         1         1      warlock15
17820917_35  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:58   00:02:41        38G                      00:27.521         1         1      warlock15
17820917_35+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:58   00:02:41             808968K         0  00:27.520         1         1      warlock15
17820917_35+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:58   00:02:41                   0         0   00:00:00         1         1      warlock15
17820917_36  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:17:49   00:11:32        38G                      09:18.501         1         1      warlock15
17820917_36+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:17:49   00:11:32           16025180K         0  09:18.500         1         1      warlock15
17820917_36+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:17:49   00:11:32                   0         0   00:00:00         1         1      warlock15
17820917_37  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:27   00:12:10        38G                      09:55.760         1         1      warlock15
17820917_37+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:27   00:12:10           17785836K         0  09:55.759         1         1      warlock15
17820917_37+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:27   00:12:10                   0         0   00:00:00         1         1      warlock15
17820917_38  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:13:43   00:07:26        38G                      00:29.827         1         1      warlock15
17820917_38+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:13:43   00:07:26             862500K         0  00:29.826         1         1      warlock15
17820917_38+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:13:43   00:07:26                   0         0   00:00:00         1         1      warlock15
17820917_39  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:23   00:16:06        38G                      09:04.839         1         1      warlock15
17820917_39+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:23   00:16:06           15585404K         0  09:04.838         1         1      warlock15
17820917_39+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:22:23   00:16:06                   0         0   00:00:00         1         1      warlock15
17820917_40  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:13:43   00:07:26        38G                      00:28.225         1         1      warlock15
17820917_40+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:13:43   00:07:26             862864K         0  00:28.224         1         1      warlock15
17820917_40+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:13:43   00:07:26                   0         0   00:00:00         1         1      warlock15
17820917_41  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:16   00:16:59        38G                      09:55.565         1         1      warlock15
17820917_41+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:16   00:16:59           15807612K         0  09:55.564         1         1      warlock15
17820917_41+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:16   00:16:59                   0         0   00:00:00         1         1      warlock15
17820917_42  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:54   00:02:37        38G                      00:27.653         1         1      warlock15
17820917_42+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:54   00:02:37             815844K         0  00:27.652         1         1      warlock15
17820917_42+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:08:54   00:02:37                   0         0   00:00:00         1         1      warlock15
17820917_43  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:23   00:17:06        38G                      10:03.313         1         1      warlock15
17820917_43+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:23   00:17:06           16231976K         0  10:03.312         1         1      warlock15
17820917_43+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:24   00:17:07                   0         0   00:00:00         1         1      warlock15
17820917_44  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:17   00:17:00        38G                      09:56.833         1         1      warlock15
17820917_44+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:17   00:17:00           15577172K         0  09:56.832         1         1      warlock15
17820917_44+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:17   00:17:00                   0         0   00:00:00         1         1      warlock15
17820917_45  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:30   00:17:13        38G                      10:10.778         1         1      warlock15
17820917_45+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:30   00:17:13           17089976K         0  10:10.777         1         1      warlock15
17820917_45+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:30   00:17:13                   0         0   00:00:00         1         1      warlock15
17820917_46  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:33   00:17:16        38G                      10:16.926         1         1      warlock15
17820917_46+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:33   00:17:16           16828312K         0  10:16.925         1         1      warlock15
17820917_46+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:23:33   00:17:16                   0         0   00:00:00         1         1      warlock15
17820917_47  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:13:43   00:07:26        38G                      00:29.969         1         1      warlock15
17820917_47+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:13:43   00:07:26             873188K         0  00:29.968         1         1      warlock15
17820917_47+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:13:43   00:07:26                   0         0   00:00:00         1         1      warlock15
17820917_48  sam_my_job  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:32   00:12:15        38G                      10:01.983         1         1      warlock15
17820917_48+      batch  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:32   00:12:15           16081744K         0  10:01.982         1         1      warlock15
17820917_48+     extern  COMPLETED 2024-06-29T15:06:17 2024-06-29T15:18:33   00:12:16                   0         0   00:00:00         1         1      warlock15
```

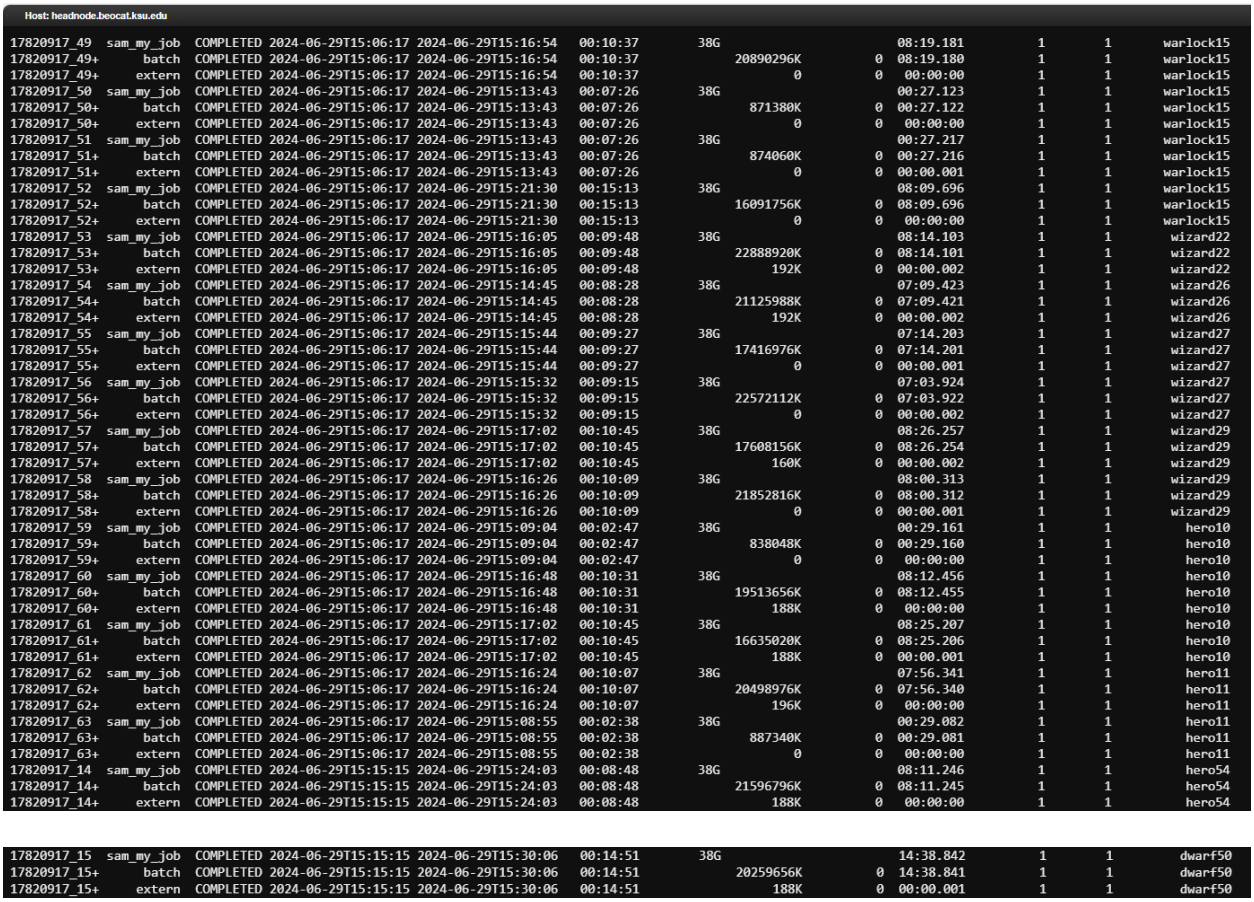


Figure 4.2: Job Execution Summary for Stage 2: Above 5 screenshots shows the SLURM job completion details for parallelized concept induction, where all 64 neuron-specific tasks were successfully executed.

These results underscore the benefits of automation and parallel processing in large-scale concept induction tasks. The efficient use of resources, along with the automation of repetitive tasks, not only accelerated the process but also ensured reproducibility and accuracy in generating neuron-specific label hypotheses. The outputs generated in this stage provide a solid foundation for the next stages of analysis, focusing on model interpretation and validation.

4.3 Stage 3 Results

The completion of the parallelized image retrieval and classification process for Stage 3 was efficiently managed through the SLURM workload manager. The process involved

generating neuron indices, retrieving images based on neuron-specific keywords, and performing image classification using the model trained in Stage 1. The key metrics for the job execution are summarized below:

1. **Job Completion Status:**
    - Neuron index generation was successfully completed with JobID 17821519 (Figure 4.3).
    - The main tasks of image retrieval and classification were completed successfully with JobID 17821523 (Figure 4.4).
2. **Total CPU Time:**
    - Neuron index generation (JobID 17821519): The total CPU time was 0.011 seconds.
    - For image retrieval and classification (JobID 17821523), individual neuron tasks ranged from approximately 1:17.866 to 1:30.920 minutes.
3. **Elapsed Time:**
    - Neuron index generation (JobID 17821519) completed in less than 1 second (elapsed time: 00:00:00).
    - The total elapsed time for individual neuron image retrieval and classification tasks (JobID 17821523) varied between 6 minutes and 8 minutes. All tasks were completed in parallel, ensuring efficient processing.
4. **Memory Usage:**
    - For neuron index generation (JobID 17821519), the maximum resident set size (MaxRSS) was 60 KB, well within the requested memory of 1 GB.

- For image retrieval and classification (JobID 17821523), memory usage per neuron ranged from approximately 900 MB to 1.7 GB. The requested memory for each task was 2 GB, ensuring adequate resource allocation.

5. **Node Allocation:**
   - Neuron index generation was completed on node hero54 using 1 core.
   - Image retrieval and classification tasks were distributed across multiple nodes in the cluster, including hero, mole, dwarf, and wizard. Each task was executed on 1 core, with 64 tasks processed in parallel, utilizing separate nodes to maximize resource efficiency.
6. **Efficiency Gains:** The parallelized approach to image retrieval and classification significantly reduced the overall processing time from an estimated 16 hours to approximately 10 minutes. Each neuron task was processed concurrently, allowing the system to handle the computational load more effectively. The efficient resource utilization, including memory management and node allocation, ensured that all tasks completed within their respective time limits and resource requests, without causing any bottlenecks. By leveraging Beocat's high-performance computing cluster, the pipeline achieved substantial reductions in processing time compared to sequential execution.

```
[samatha94@icr-clymene ~]$ sacct --format=JobID,JobName,State,Start,End,Elapsed,ReqMem,MaxRSS,MaxVMSize,TotalCPU,NCPUS,NNodes,NodeList --jobs=17821519

JobID           JobName      State               Start                End    Elapsed     ReqMem     MaxRSS  MaxVMSize   TotalCPU      NCPUS   NNodes        NodeList
------------ ---------- ---------- ------------------- ------------------- ---------- ---------- ---------- ---------- ---------- ---------- -------- ---------------
17821519     stage3_ge+  COMPLETED 2024-06-29T16:22:12 2024-06-29T16:22:12   00:00:00         1G                    00:00.011          1        1          hero54
17821519.ba+      batch  COMPLETED 2024-06-29T16:22:12 2024-06-29T16:22:12   00:00:00                60K          0  00:00.010          1        1          hero54
17821519.ex+     extern  COMPLETED 2024-06-29T16:22:12 2024-06-29T16:22:12   00:00:00               184K          0   00:00:00          1        1          hero54
```

Figure 4.3: Job Execution Summary for Neuron Index Generation in Stage 3 – This figure displays the SLURM job completion details for the neuron index generation task in Stage 3.

```
[samatha94@icr-clymene ~]$ sacct --format=JobID,JobName,State,Start,End,Elapsed,ReqMem,MaxRSS,MaxVMSize,TotalCPU,NCPUS,NNodes,NodeList --jobs=17821523

JobID           JobName      State               Start                 End    Elapsed     ReqMem     MaxRSS  MaxVMSize   TotalCPU      NCPUS   NNodes        NodeList
------------ ---------- ---------- ------------------- ------------------- ---------- ---------- ---------- ---------- ---------- --------- -------- ---------------
17821523_0   sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:29   00:06:43         2G                    01:17.867          1        1         hero54
17821523_0.+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:29   00:06:43              1467920K         0  01:17.866          1        1         hero54
17821523_0.+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:29   00:06:43               192K         0  00:00.001          1        1         hero54
17821523_1   sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:55   00:07:09         2G                    01:18.721          1        1         hero54
17821523_1.+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:55   00:07:09              1555120K         0  01:18.720          1        1         hero54
17821523_1.+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:55   00:07:09               196K         0  00:00.001          1        1         hero54
17821523_2   sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:19   00:07:33         2G                    01:53.581          1        1        dwarf50
17821523_2.+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:19   00:07:33              1616952K         0  01:53.580          1        1        dwarf50
17821523_2.+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:19   00:07:33                 0         0  00:00.001          1        1        dwarf50
17821523_3   sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:10   00:07:24         2G                    01:54.326          1        1        dwarf50
17821523_3.+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:10   00:07:24              1733336K         0  01:54.325          1        1        dwarf50
17821523_3.+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:10   00:07:24               192K         0  00:00.001          1        1        dwarf50
17821523_4   sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:06   00:07:20         2G                    01:27.101          1        1         mole041
17821523_4.+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:06   00:07:20              1065004K         0  01:27.100          1        1         mole041
17821523_4.+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:06   00:07:20                 0         0  00:00:00           1        1         mole041
17821523_5   sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:07   00:07:21         2G                    01:25.384          1        1         mole041
17821523_5.+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:07   00:07:21              1212452K         0  01:25.383          1        1         mole041
17821523_5.+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:07   00:07:21                 0         0  00:00:00           1        1         mole041
17821523_6   sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:21   00:07:35         2G                    01:29.044          1        1         mole041
17821523_6.+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:21   00:07:35              1184248K         0  01:29.043          1        1         mole041
17821523_6.+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:21   00:07:35               192K         0  00:00:00           1        1         mole041
17821523_7   sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:38   00:06:52         2G                    01:25.546          1        1         mole041
17821523_7.+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:38   00:06:52              1161176K         0  01:25.545          1        1         mole041
17821523_7.+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:38   00:06:52                 0         0  00:00:00           1        1         mole041
17821523_8   sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:18   00:07:32         2G                    01:27.583          1        1         mole041
17821523_8.+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:18   00:07:32              1278260K         0  01:27.582          1        1         mole041
17821523_8.+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:18   00:07:32                 0         0  00:00:00           1        1         mole041
17821523_9   sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:15   00:07:29         2G                    01:25.363          1        1         mole041
17821523_9.+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:15   00:07:29              1096908K         0  01:25.362          1        1         mole041
17821523_9.+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:15   00:07:29                188K         0  00:00:00           1        1         mole041
17821523_10  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:45   00:06:59         2G                    01:20.984          1        1         mole041
17821523_10+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:45   00:06:59              1045908K         0  01:20.983          1        1         mole041
17821523_10+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:45   00:06:59                 0         0  00:00.001          1        1         mole041
17821523_11  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:31   00:07:45         2G                    01:26.336          1        1         mole042
17821523_11+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:31   00:07:45              1113612K         0  01:26.336          1        1         mole042
17821523_11+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:31   00:07:45                 0         0  00:00:00           1        1         mole042
17821523_12  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:00   00:07:14         2G                    01:32.120          1        1         mole042
17821523_12+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:00   00:07:14              1227752K         0  01:32.119          1        1         mole042
17821523_12+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:01   00:07:15                 0         0  00:00.001          1        1         mole042
17821523_13  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:59   00:07:13         2G                    01:33.229          1        1         mole042
17821523_13+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:59   00:07:13              1277476K         0  01:33.228          1        1         mole042
17821523_13+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:59   00:07:13                 0         0  00:00.001          1        1         mole042
```

Host: headnode.beocat.ksu.edu

```
17821523_14  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:27   00:07:41         2G                    01:27.354          1        1         mole042
17821523_14+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:27   00:07:41              1153396K         0  01:27.353          1        1         mole042
17821523_14+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:27   00:07:41                 0         0  00:00:00           1        1         mole042
17821523_15  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:35   00:07:49         2G                    01:30.920          1        1         mole042
17821523_15+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:35   00:07:49               917548K         0  01:30.919          1        1         mole042
17821523_15+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:35   00:07:49                 0         0  00:00.001          1        1         mole042
17821523_16  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:56   00:07:10         2G                    01:36.496          1        1         mole042
17821523_16+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:56   00:07:10              1156856K         0  01:36.495          1        1         mole042
17821523_16+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:56   00:07:10                 0         0  00:00.001          1        1         mole042
17821523_17  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:31:31   00:08:45         2G                    01:27.614          1        1         mole042
17821523_17+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:31:31   00:08:45               975164K         0  01:27.613          1        1         mole042
17821523_17+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:31:31   00:08:45                 0         0  00:00.001          1        1         mole042
17821523_18  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:03   00:07:17         2G                    01:30.745          1        1         mole042
17821523_18+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:03   00:07:17              1074668K         0  01:30.744          1        1         mole042
17821523_18+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:03   00:07:17                 0         0  00:00.001          1        1         mole042
17821523_19  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:41   00:06:55         2G                    01:30.211          1        1         mole042
17821523_19+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:41   00:06:55               940412K         0  01:30.210          1        1         mole042
17821523_19+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:41   00:06:55                 0         0  00:00.001          1        1         mole042
17821523_20  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:38   00:06:52         2G                    01:32.019          1        1         mole042
17821523_20+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:38   00:06:52               907376K         0  01:32.018          1        1         mole042
17821523_20+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:38   00:06:52                 0         0  00:00.001          1        1         mole042
17821523_21  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:35   00:06:49         2G                    01:19.707          1        1         mole042
17821523_21+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:35   00:06:49              1001312K         0  01:19.706          1        1         mole042
17821523_21+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:35   00:06:49                 0         0  00:00.001          1        1         mole042
17821523_22  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:40   00:06:54         2G                    01:24.742          1        1         mole042
17821523_22+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:40   00:06:54              1040140K         0  01:24.741          1        1         mole042
17821523_22+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:40   00:06:54                 0         0  00:00:00           1        1         mole042
17821523_23  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:48   00:07:02         2G                    01:22.150          1        1         mole042
17821523_23+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:48   00:07:02              1094668K         0  01:22.149          1        1         mole042
17821523_23+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:48   00:07:02                 0         0  00:00.001          1        1         mole042
17821523_24  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:49   00:07:03         2G                    01:23.646          1        1         mole042
17821523_24+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:49   00:07:03              1142276K         0  01:23.645          1        1         mole042
17821523_24+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:49   00:07:03                 0         0  00:00.001          1        1         mole042
17821523_25  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:41   00:06:55         2G                    01:22.009          1        1         mole042
17821523_25+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:41   00:06:55               895480K         0  01:22.008          1        1         mole042
17821523_25+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:41   00:06:55                 0         0  00:00.001          1        1         mole042
17821523_26  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:11   00:07:25         2G                    01:21.034          1        1         mole043
17821523_26+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:11   00:07:25               953548K         0  01:21.033          1        1         mole043
17821523_26+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:11   00:07:25                 0         0  00:00:00           1        1         mole043
17821523_27  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:29   00:07:43         2G                    01:22.304          1        1         mole043
17821523_27+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:29   00:07:43              1076732K         0  01:22.303          1        1         mole043
17821523_27+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:29   00:07:43                188K         0  00:00:00           1        1         mole043
17821523_28  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:55   00:07:09         2G                    01:20.845          1        1         mole043
17821523_28+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:55   00:07:09              1158304K         0  01:20.844          1        1         mole043
17821523_28+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:29:55   00:07:09               192K         0  00:00.001          1        1         mole043
17821523_29  sam_my_job  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:13   00:07:27         2G                    01:29.322          1        1         mole043
17821523_29+      batch  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:13   00:07:27              1219656K         0  01:29.322          1        1         mole043
17821523_29+     extern  COMPLETED 2024-06-29T16:22:46 2024-06-29T16:30:13   00:07:27                 0         0  00:00:00           1        1         mole043
```

Host: headnode.beocat.ksu.edu

17821523_30	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:59	00:08:13	2G			01:24.159	1	1	mole043
17821523_30+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:59	00:08:13		1132152K	0	01:24.158	1	1	mole043
17821523_30+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:59	00:08:13		0	0	00:00:00	1	1	mole043
17821523_31	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:24	00:07:38	2G			01:22.317	1	1	mole043
17821523_31+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:24	00:07:38		1148544K	0	01:22.316	1	1	mole043
17821523_31+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:25	00:07:39		0	0	00:00.001	1	1	mole043
17821523_32	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:19	00:07:33	2G			01:22.528	1	1	mole043
17821523_32+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:19	00:07:33		1223720K	0	01:22.527	1	1	mole043
17821523_32+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:19	00:07:33		0	0	00:00:00	1	1	mole043
17821523_33	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:58	00:07:12	2G			01:30.543	1	1	mole045
17821523_33+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:58	00:07:12		1196772K	0	01:30.542	1	1	mole045
17821523_33+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:58	00:07:12		0	0	00:00:00	1	1	mole045
17821523_34	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:03	00:07:17	2G			01:29.027	1	1	mole045
17821523_34+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:03	00:07:17		1218884K	0	01:29.026	1	1	mole045
17821523_34+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:03	00:07:17		0	0	00:00:00	1	1	mole045
17821523_35	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:42	00:07:56	2G			01:24.019	1	1	mole045
17821523_35+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:42	00:07:56		1140092K	0	01:24.018	1	1	mole045
17821523_35+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:42	00:07:56		188K	0	00:00:00	1	1	mole045
17821523_36	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:23	00:06:37	2G			01:22.051	1	1	mole045
17821523_36+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:23	00:06:37		1234332K	0	01:22.050	1	1	mole045
17821523_36+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:23	00:06:37		0	0	00:00:00	1	1	mole045
17821523_37	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:36	00:06:50	2G			01:25.102	1	1	mole045
17821523_37+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:36	00:06:50		1079536K	0	01:25.101	1	1	mole045
17821523_37+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:36	00:06:50		184K	0	00:00.001	1	1	mole045
17821523_38	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:33	00:06:47	2G			01:24.196	1	1	mole045
17821523_38+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:33	00:06:47		998772K	0	01:24.195	1	1	mole045
17821523_38+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:33	00:06:47		184K	0	00:00.001	1	1	mole045
17821523_39	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:29	00:07:43	2G			01:24.002	1	1	mole045
17821523_39+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:29	00:07:43		1097520K	0	01:24.001	1	1	mole045
17821523_39+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:29	00:07:43		0	0	00:00:00	1	1	mole045
17821523_40	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:13	00:07:27	2G			01:18.769	1	1	mole046
17821523_40+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:13	00:07:27		1108752K	0	01:18.769	1	1	mole046
17821523_40+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:13	00:07:27		204K	0	00:00:00	1	1	mole046
17821523_41	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:41	00:06:55	2G			01:19.968	1	1	mole046
17821523_41+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:41	00:06:55		973828K	0	01:19.967	1	1	mole046
17821523_41+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:41	00:06:55		0	0	00:00:00	1	1	mole046
17821523_42	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:55	00:07:09	2G			01:22.498	1	1	mole046
17821523_42+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:55	00:07:09		1301744K	0	01:22.497	1	1	mole046
17821523_42+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:55	00:07:09		192K	0	00:00:00	1	1	mole046
17821523_43	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:18	00:07:32	2G			01:21.728	1	1	mole046
17821523_43+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:18	00:07:32		1283740K	0	01:21.727	1	1	mole046
17821523_43+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:18	00:07:32		188K	0	00:00:00	1	1	mole046
17821523_44	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:08	00:07:22	2G			01:26.948	1	1	mole046
17821523_44+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:08	00:07:22		1328180K	0	01:26.947	1	1	mole046
17821523_44+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:08	00:07:22		0	0	00:00:00	1	1	mole046
17821523_45	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:53	00:07:07	2G			01:26.092	1	1	mole046
17821523_45+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:53	00:07:07		1217452K	0	01:26.091	1	1	mole046
17821523_45+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:53	00:07:07		184K	0	00:00:00	1	1	mole046

Host: headnode.beocat.ksu.edu

17821523_46	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:14	00:07:28	2G			01:23.860	1	1	mole046
17821523_46+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:14	00:07:28		1106992K	0	01:23.859	1	1	mole046
17821523_46+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:14	00:07:28		0	0	00:00.001	1	1	mole046
17821523_47	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:23	00:07:37	2G			01:23.006	1	1	mole047
17821523_47+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:23	00:07:37		1386196K	0	01:23.005	1	1	mole047
17821523_47+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:30:23	00:07:37		0	0	00:00:00	1	1	mole047
17821523_48	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:35	00:06:49	2G			01:18.312	1	1	mole047
17821523_48+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:35	00:06:49		1455400K	0	01:18.311	1	1	mole047
17821523_48+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:29:35	00:06:49		0	0	00:00.001	1	1	mole047
17821523_49	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.574	1	1	mole047
17821523_49+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		32732K	0	00:09.573	1	1	mole047
17821523_49+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		0	0	00:00.001	1	1	mole047
17821523_50	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.569	1	1	mole047
17821523_50+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		32556K	0	00:09.568	1	1	mole047
17821523_50+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		0	0	00:00.001	1	1	mole047
17821523_51	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.547	1	1	mole047
17821523_51+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		32892K	0	00:09.546	1	1	mole047
17821523_51+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		188K	0	00:00.001	1	1	mole047
17821523_52	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.609	1	1	mole047
17821523_52+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		32332K	0	00:09.608	1	1	mole047
17821523_52+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		0	0	00:00:00	1	1	mole047
17821523_53	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.519	1	1	mole047
17821523_53+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		32876K	0	00:09.518	1	1	mole047
17821523_53+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		0	0	00:00:00	1	1	mole047
17821523_54	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.449	1	1	mole048
17821523_54+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		32180K	0	00:09.448	1	1	mole048
17821523_54+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		0	0	00:00.001	1	1	mole048
17821523_55	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.543	1	1	mole048
17821523_55+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		34196K	0	00:09.542	1	1	mole048
17821523_55+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		0	0	00:00:00	1	1	mole048
17821523_56	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.532	1	1	mole048
17821523_56+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		33544K	0	00:09.531	1	1	mole048
17821523_56+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		0	0	00:00.001	1	1	mole048
17821523_57	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.430	1	1	mole048
17821523_57+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		35476K	0	00:09.429	1	1	mole048
17821523_57+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		192K	0	00:00.001	1	1	mole048
17821523_58	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.556	1	1	mole048
17821523_58+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		36400K	0	00:09.555	1	1	mole048
17821523_58+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		0	0	00:00:00	1	1	mole048
17821523_59	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.548	1	1	mole048
17821523_59+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		36072K	0	00:09.547	1	1	mole048
17821523_59+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		188K	0	00:00.001	1	1	mole048
17821523_60	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:09.429	1	1	mole048
17821523_60+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		33568K	0	00:09.428	1	1	mole048
17821523_60+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		0	0	00:00:00	1	1	mole048
17821523_61	sam_my_job	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24	2G			00:08.263	1	1	mole049
17821523_61+	batch	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		39720K	0	00:08.263	1	1	mole049
17821523_61+	extern	COMPLETED	2024-06-29T16:22:46	2024-06-29T16:24:10	00:01:24		188K	0	00:00:00	1	1	mole049

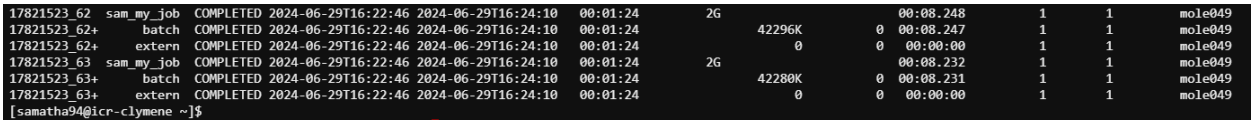


Figure 4.4: Parallelized Execution Summary for Stage 3 Image Retrieval and Classification: This figure showcases the SLURM workload manager's execution of Stage 3, where image retrieval and classification tasks were performed concurrently.

These results underscore the efficiency and scalability of the Stage 3 pipeline, with all tasks completed within the requested resources and time. This provides a solid foundation for the subsequent stages of analysis and model validation.

4.4 Stage 4 Results

The statistical analysis and verification of neuron activations in Stage 4 were efficiently managed through the Beocat high-performance computing environment. This stage was aimed at verifying the label hypotheses generated in the earlier stages by calculating and statistically comparing neuron activations for target and non-target images. The following are key metrics obtained during the job execution:

1. **Job Completion Status:** The job was successfully completed with JobID 17821704 (Figure 4.5), confirming that the neuron-specific statistical analysis tasks were executed without errors.
2. **Total CPU Time:** The total CPU time used for the job was approximately 49.88 seconds, indicating efficient processing of neuron activation data and statistical calculations.
3. **Elapsed Time:** The job’s total elapsed time was 2 minutes and 41 seconds, reflecting the quick turnaround time required for executing the entire statistical analysis and verification of neuron activations.

4. **Memory Usage:** The maximum resident set size (MaxRSS) during the job execution was 840 MB, while the total requested memory for the job was 4 GB. This memory usage was well within the allocated resources, ensuring smooth execution without resource bottlenecks.
5. **Node Allocation:** The job was executed on a single node, utilizing one CPU core (mole041), ensuring exclusive access to the node for running the statistical analysis without interference from other processes.
6. **Efficiency Gains:** The high-performance computing resources of Beocat facilitated a rapid and automated verification process in Stage 4, significantly reducing the time required for neuron-specific statistical analysis. What would typically be a resource-intensive task was completed in just under 3 minutes. The automation efficiently handled neuron activations from both the evaluation and verification phases, ensuring that statistical tests, including the Mann-Whitney U test, were performed for every neuron.

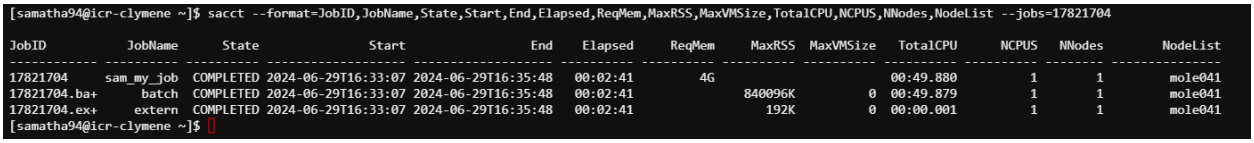


Figure 4.5: SLURM Workload Summary for Stage 4 Statistical Analysis: This figure presents the job details of the statistical analysis and verification of neuron activations in Stage 4.

The statistical analysis in Stage 4 validated neuron activations and confirmed that concept induction with structured background knowledge yielded meaningful and consistent labels. The target activation metrics for each neuron were aligned with the hypothesized labels, and the statistical tests provided robust confirmation of these relationships. This streamlined approach allowed for the rapid verification of neuron activations, highlighting the robustness and relevance of the label hypotheses in explaining neuron activation patterns.

These results underline the system's capability to handle large-scale, neuron-specific statistical analysis with minimal resource consumption and within a short execution window, enabling the validation of neuron-label hypotheses through robust statistical evaluation.

Chapter 5 - Scalability and Performance Evaluation of the Automation Pipeline on New Scene Categories

In addition to the initial set of scene categories, we extended the analysis by selecting 20 additional classes from the ADE20K dataset, each based on the number of available images in each class. This allowed us to assess the scalability and performance of the automation pipeline when applied to a wider variety of image categories. The execution was performed in two stages, each with 10 new classes.

For the first set of classes: “airport terminal”, “art studio”, “attic”, “corridor”, “game room”, “home office”, “hotel room”, “mountain snowy”, “office”, and “waiting room”—the total number of images was 1,197(Figure 5.1 to Figure 5.4). The automation pipeline was executed across all four stages with the following times:

1. **Stage 1 (Model Training and Data Configuration):** 12 minutes and 52 seconds (Figure 5.5)
2. **Stage 2 (Concept Induction and Label Hypothesis Generation):** 18 minutes and 10 seconds (Figure 5.6)
3. **Stage 3 (Image Retrieval and Classification):** 3 minutes and 16 seconds (Figure 5.7 and Figure 5.8)
4. **Stage 4 (Statistical Analysis and Verification):** 2 minutes and 25 seconds (Figure 5.9)

The total time taken for all stages to process this set of 1,197 images was 36 minutes and 43 seconds.

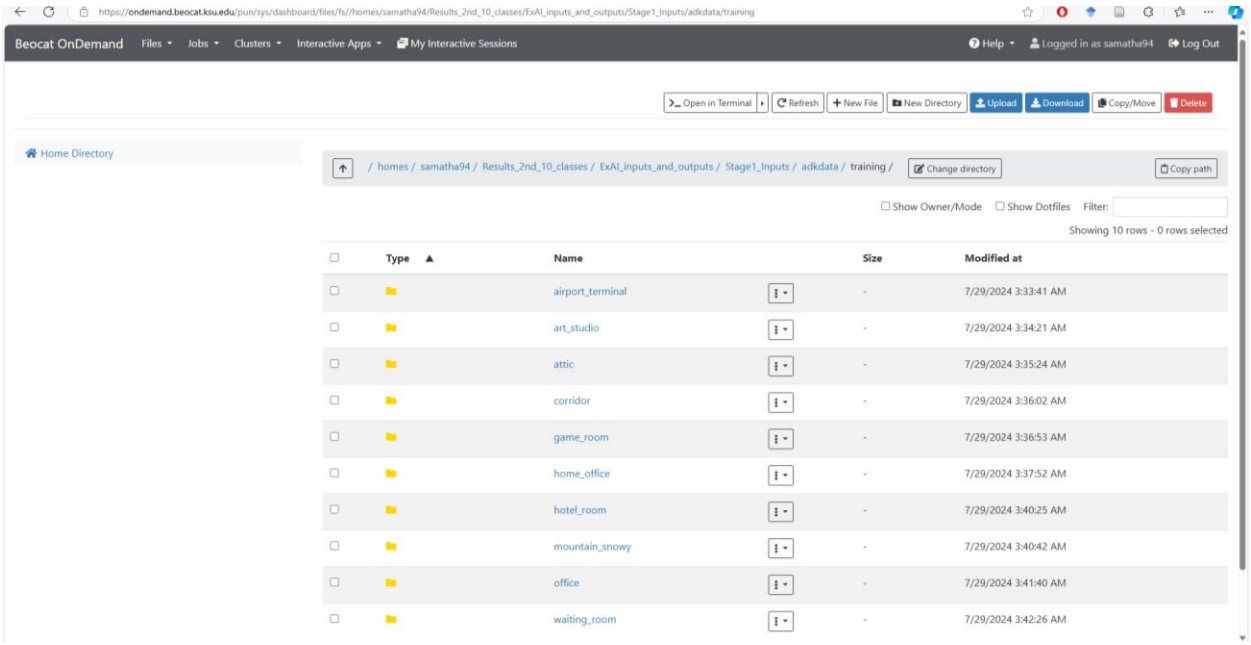


Figure 5.1: Training Folder Organization for New Scene Categories: This figure displays the folder structure used to organize training images for the second set of 10 new scene categories from the ADE20K dataset.

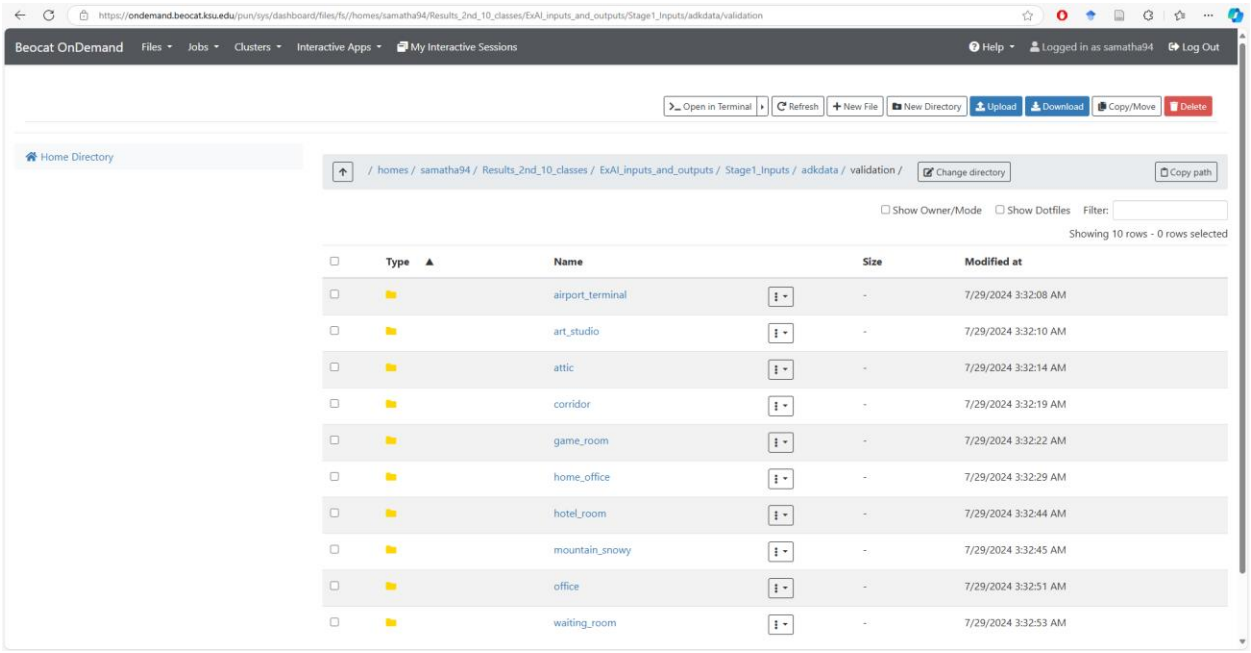


Figure 5.2: Validation Folder Organization for New Scene Categories: This figure highlights the structured folder layout for the validation dataset containing the new set of 10 scene categories.

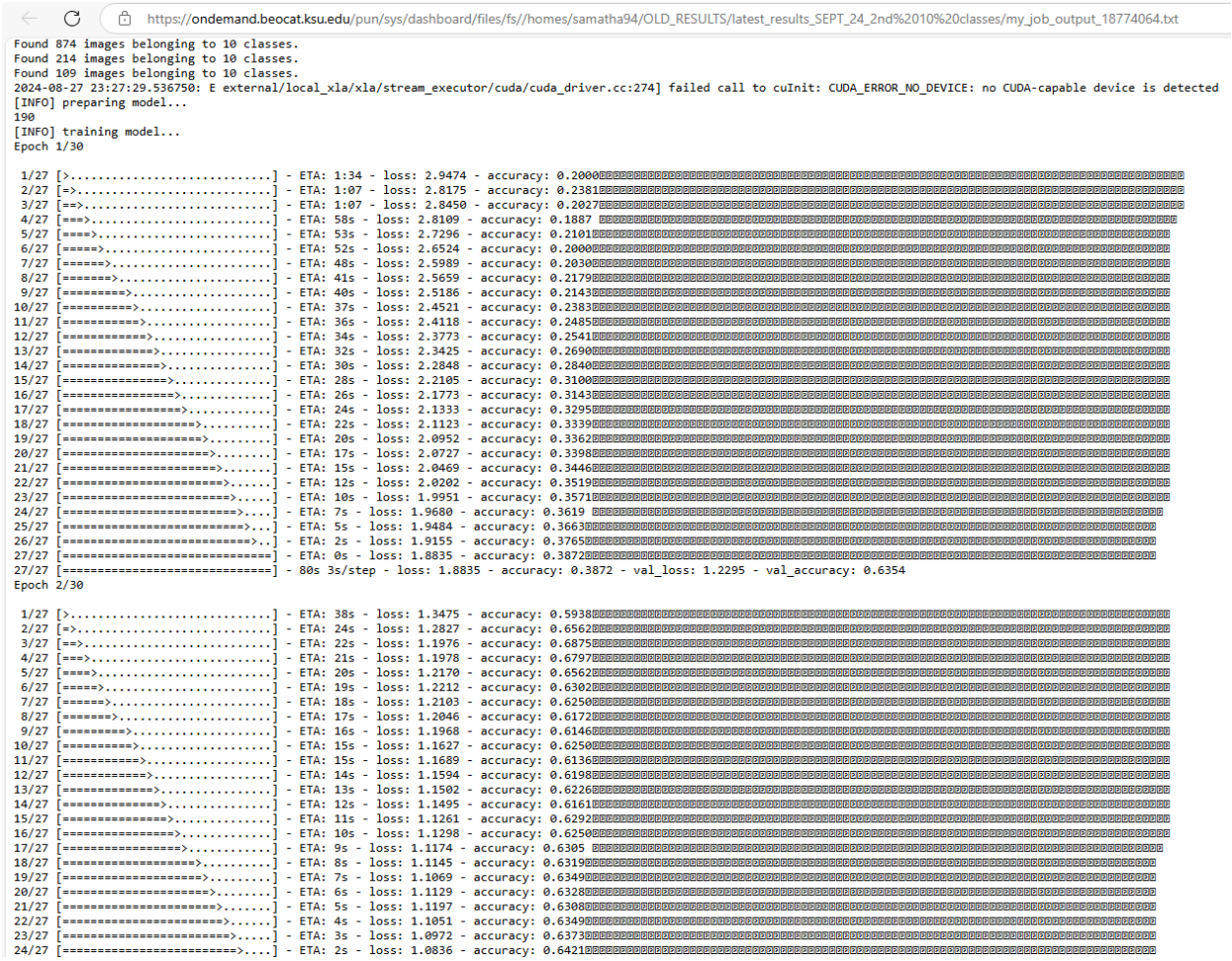


Figure 5.3: Input Dataset Summary for New Scene Categories – This figure presents a summary of the number of images in each selected class from the ADE20K dataset used in training the ResNet50V2 model. It displays the dataset distribution across the new set of 10 scene categories.

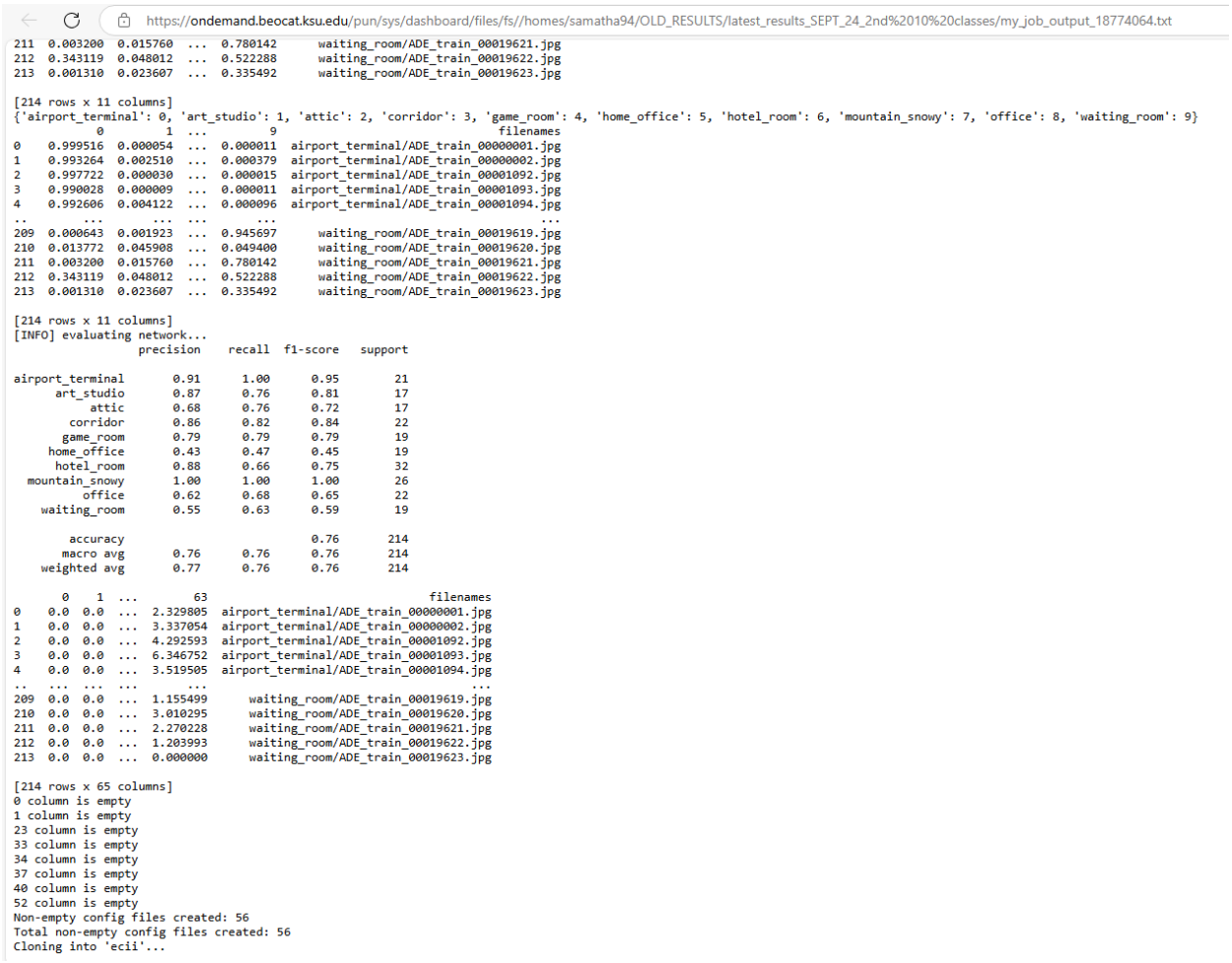


Figure 5.4: Scene Classification Evaluation Report for New Scene Categories – This figure provides the evaluation metrics for the new set of 10 scene categories. It presents the precision, recall, and F1-score for each category, reporting an overall model accuracy of 76% and highlighting the model's ability to differentiate between the new categories effectively.

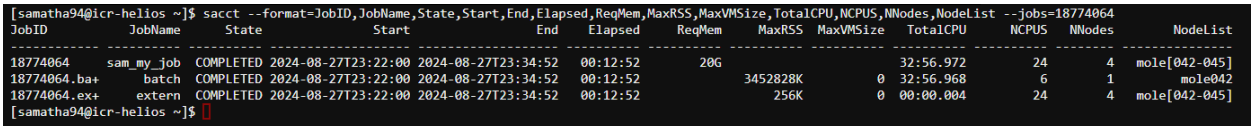


Figure 5.5: SLURM Workload Summary for Stage 1 Training Execution – This figure displays the SLURM workload summary for the successful execution of the model training and data configuration process in Stage 1 for the second set of 10 scene categories.

Host: headnode.beocat.ksu.edu

```
[samatha94@icr-helios ~]$ sacct --format=JobID,JobName,State,Start,End,Elapsed,ReqMem,MaxRSS,MaxVMSize,TotalCPU,NCPUS,NNodes,NodeList --jobs=18774756
JobID           JobName      State               Start                End    Elapsed     ReqMem     MaxRSS  MaxVMSize   TotalCPU      NCPUS   NNodes        NodeList
------------ ---------- ---------- ------------------- ------------------- ---------- ---------- ---------- ---------- ---------- ---------- -------- ---------------
18774756_0   sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:39:30   00:00:43        38G                    00:26.361          1        1          hero29
18774756_0.+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:39:30   00:00:43             677916K          0  00:26.360          1        1          hero29
18774756_0.+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:39:30   00:00:43                188K          0   00:00:00          1        1          hero29
18774756_1   sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:39:33   00:00:46        38G                    00:27.083          1        1          hero30
18774756_1.+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:39:33   00:00:46             886496K          0  00:27.082          1        1          hero30
18774756_1.+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:39:33   00:00:46                192K          0   00:00:00          1        1          hero30
18774756_2   sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:09   00:09:22        38G                    07:43.714          1        1          hero13
18774756_2.+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:09   00:09:22           20217040K          0  07:43.713          1        1          hero13
18774756_2.+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:09   00:09:22                  0          0  00:00.001          1        1          hero13
18774756_3   sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:34   00:09:47        38G                    07:51.805          1        1          hero13
18774756_3.+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:34   00:09:47           14162720K          0  07:51.804          1        1          hero13
18774756_3.+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:34   00:09:47                  0          0  00:00.001          1        1          hero13
18774756_4   sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:29   00:09:42        38G                    07:48.335          1        1          hero13
18774756_4.+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:29   00:09:42           14128380K          0  07:48.334          1        1          hero13
18774756_4.+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:30   00:09:43                  0          0   00:00:00          1        1          hero13
18774756_5   sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:36   00:09:49        38G                    07:40.684          1        1          hero14
18774756_5.+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:36   00:09:49           18953024K          0  07:40.683          1        1          hero14
18774756_5.+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:36   00:09:49                192K          0   00:00:00          1        1          hero14
18774756_6   sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:48   00:10:01        38G                    07:52.607          1        1          hero14
18774756_6.+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:48   00:10:01           15071212K          0  07:52.606          1        1          hero14
18774756_6.+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:48   00:10:01                196K          0  00:00.001          1        1          hero14
18774756_7   sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:46   00:09:59        38G                    07:52.129          1        1          hero14
18774756_7.+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:46   00:09:59           14397992K          0  07:52.128          1        1          hero14
18774756_7.+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:46   00:09:59                184K          0   00:00:00          1        1          hero14
18774756_8   sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:49:36   00:10:49        38G                    07:44.227          1        1          hero15
18774756_8.+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:49:36   00:10:49           15657152K          0  07:44.226          1        1          hero15
18774756_8.+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:49:36   00:10:49                  0          0   00:00:00          1        1          hero15
18774756_9   sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:49:32   00:10:45        38G                    07:40.363          1        1          hero15
18774756_9.+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:49:32   00:10:45           15145548K          0  07:40.362          1        1          hero15
18774756_9.+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:49:32   00:10:45                  0          0  00:00.001          1        1          hero15
18774756_10  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:43   00:09:56        38G                    07:45.353          1        1          hero15
18774756_10+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:43   00:09:56           18617548K          0  07:45.352          1        1          hero15
18774756_10+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:43   00:09:56                184K          0   00:00:00          1        1          hero15
18774756_11  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:15   00:09:28        38G                    07:50.388          1        1          hero16
18774756_11+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:15   00:09:28           19764260K          0  07:50.387          1        1          hero16
18774756_11+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:15   00:09:28                  0          0  00:00.001          1        1          hero16
18774756_12  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:23   00:09:36        38G                    07:37.538          1        1          hero16
18774756_12+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:23   00:09:36           15102812K          0  07:37.537          1        1          hero16
18774756_12+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:23   00:09:36                  0          0  00:00.001          1        1          hero16
18774756_13  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:28   00:09:41        38G                    07:43.182          1        1          hero16
18774756_13+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:28   00:09:41           15293192K          0  07:43.181          1        1          hero16
18774756_13+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:28   00:09:41                  0          0  00:00.001          1        1          hero16
18774756_14  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:49:49   00:11:02        38G                    09:41.655          1        1      warlock12
18774756_14+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:49:49   00:11:02           19612920K          0  09:41.654          1        1      warlock12
18774756_14+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:49:49   00:11:02                188K          0  00:00.001          1        1      warlock12
```

Host: headnode.beocat.ksu.edu

```
18774756_15  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:33   00:14:46        38G                    07:23.986          1        1      warlock14
18774756_15+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:33   00:14:46           14920188K          0  07:23.984          1        1      warlock14
18774756_15+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:33   00:14:46                  0          0  00:00.001          1        1      warlock14
18774756_16  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:39   00:14:52        38G                    07:29.936          1        1      warlock14
18774756_16+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:39   00:14:52           15746744K          0  07:29.935          1        1      warlock14
18774756_16+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:39   00:14:52                  0          0  00:00.001          1        1      warlock14
18774756_17  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:37   00:14:50        38G                    07:27.916          1        1      warlock14
18774756_17+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:37   00:14:50           14663172K          0  07:27.915          1        1      warlock14
18774756_17+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:37   00:14:50                  0          0   00:00:00          1        1      warlock14
18774756_18  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:37   00:14:50        38G                    07:28.206          1        1      warlock14
18774756_18+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:37   00:14:50           14316368K          0  07:28.204          1        1      warlock14
18774756_18+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:38   00:14:51                  0          0  00:00.001          1        1      warlock14
18774756_19  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:34   00:14:47        38G                    07:25.551          1        1      warlock14
18774756_19+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:34   00:14:47           13951220K          0  07:25.550          1        1      warlock14
18774756_19+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:34   00:14:47                  0          0   00:00:00          1        1      warlock14
18774756_20  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:37   00:14:50        38G                    07:28.641          1        1      warlock14
18774756_20+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:37   00:14:50           14478580K          0  07:28.640          1        1      warlock14
18774756_20+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:37   00:14:50                  0          0   00:00:00          1        1      warlock14
18774756_21  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:38   00:14:51        38G                    07:28.273          1        1      warlock14
18774756_21+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:38   00:14:51           14885136K          0  07:28.273          1        1      warlock14
18774756_21+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:38   00:14:51                  0          0   00:00:00          1        1      warlock14
18774756_22  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:40   00:14:53        38G                    07:31.536          1        1      warlock14
18774756_22+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:40   00:14:53           15449828K          0  07:31.535          1        1      warlock14
18774756_22+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:40   00:14:53                  0          0   00:00:00          1        1      warlock14
18774756_23  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:32   00:07:45        38G                    00:25.061          1        1      warlock14
18774756_23+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:32   00:07:45             873468K          0  00:25.060          1        1      warlock14
18774756_23+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:32   00:07:45                  0          0  00:00.001          1        1      warlock14
18774756_24  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:02   00:14:15        38G                    06:54.099          1        1      warlock14
18774756_24+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:02   00:14:15           19129324K          0  06:54.098          1        1      warlock14
18774756_24+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:02   00:14:15                  0          0  00:00.001          1        1      warlock14
18774756_25  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:10   00:15:23        38G                    10:07.510          1        1      warlock15
18774756_25+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:10   00:15:23           14582324K          0  10:07.509          1        1      warlock15
18774756_25+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:10   00:15:23                192K          0   00:00:00          1        1      warlock15
18774756_26  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:14   00:15:27        38G                    10:09.613          1        1      warlock15
18774756_26+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:14   00:15:27           14714536K          0  10:09.612          1        1      warlock15
18774756_26+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:14   00:15:27                188K          0  00:00.001          1        1      warlock15
18774756_27  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:20   00:15:33        38G                    10:15.749          1        1      warlock15
18774756_27+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:20   00:15:33           15207984K          0  10:15.749          1        1      warlock15
18774756_27+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:20   00:15:33                  0          0   00:00:00          1        1      warlock15
18774756_28  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:27   00:15:40        38G                    10:22.447          1        1      warlock15
18774756_28+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:27   00:15:40           14852024K          0  10:22.446          1        1      warlock15
18774756_28+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:27   00:15:40                188K          0   00:00:00          1        1      warlock15
18774756_29  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:04   00:15:17        38G                    10:01.679          1        1      warlock15
18774756_29+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:04   00:15:17           16206948K          0  10:01.678          1        1      warlock15
18774756_29+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:05   00:15:18                  0          0   00:00:00          1        1      warlock15
18774756_30  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:53   00:15:06        38G                    09:49.224          1        1      warlock15
18774756_30+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:53   00:15:06           17251656K          0  09:49.223          1        1      warlock15
18774756_30+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:53   00:15:06                  0          0   00:00:00          1        1      warlock15
```

Host: headnode.beocat.ksu.edu

```
18774756_31  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:07   00:15:20        38G                    10:03.392         1         1       warlock15
18774756_31+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:07   00:15:20          14681432K         0  10:03.391         1         1       warlock15
18774756_31+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:54:07   00:15:20                0         0   00:00:00         1         1       warlock15
18774756_32  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:52:46   00:13:59        38G                    06:05.849         1         1       warlock31
18774756_32+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:52:46   00:13:59          14402992K         0  06:05.848         1         1       warlock31
18774756_32+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:52:46   00:13:59                0         0   00:00:00         1         1       warlock31
18774756_33  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:27   00:08:40        38G                    00:24.859         1         1       warlock31
18774756_33+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:27   00:08:40            876980K         0  00:24.858         1         1       warlock31
18774756_33+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:27   00:08:40                0         0  00:00.001         1         1       warlock31
18774756_34  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:05   00:07:18        38G                    00:24.419         1         1       warlock31
18774756_34+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:05   00:07:18            868216K         0  00:24.417         1         1       warlock31
18774756_34+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:05   00:07:18                0         0  00:00.001         1         1       warlock31
18774756_35  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:11   00:14:24        38G                    06:06.894         1         1       warlock31
18774756_35+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:11   00:14:24          15202404K         0  06:06.893         1         1       warlock31
18774756_35+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:11   00:14:24                0         0  00:00.001         1         1       warlock31
18774756_36  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:18   00:07:31        38G                    05:47.409         1         1       warlock31
18774756_36+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:18   00:07:31          20015492K         0  05:47.408         1         1       warlock31
18774756_36+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:18   00:07:31                0         0  00:00.001         1         1       warlock31
18774756_37  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:23   00:08:36        38G                    00:24.841         1         1       warlock31
18774756_37+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:23   00:08:36            866604K         0  00:24.840         1         1       warlock31
18774756_37+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:23   00:08:36                0         0  00:00.001         1         1       warlock31
18774756_38  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:07   00:14:20        38G                    06:05.675         1         1       warlock31
18774756_38+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:07   00:14:20          14018268K         0  06:05.674         1         1       warlock31
18774756_38+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:07   00:14:20                0         0  00:00.001         1         1       warlock31
18774756_39  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:07   00:14:20        38G                    06:05.251         1         1       warlock31
18774756_39+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:07   00:14:20          13940368K         0  06:05.250         1         1       warlock31
18774756_39+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:53:07   00:14:20                0         0  00:00.001         1         1       warlock31
18774756_40  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:14   00:07:27        38G                    00:22.982         1         1       warlock31
18774756_40+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:14   00:07:27            896924K         0  00:22.981         1         1       warlock31
18774756_40+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:14   00:07:27                0         0  00:00.001         1         1       warlock31
18774756_41  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:52   00:08:05        38G                    05:40.328         1         1       warlock31
18774756_41+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:52   00:08:05          15719940K         0  05:40.327         1         1       warlock31
18774756_41+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:46:52   00:08:05                0         0  00:00.001         1         1       warlock31
18774756_42  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:43   00:08:56        38G                    07:50.171         1         1        wizard22
18774756_42+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:43   00:08:56          16765812K         0  07:50.169         1         1        wizard22
18774756_42+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:43   00:08:56                0         0  00:00.002         1         1        wizard22
18774756_43  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:37   00:08:50        38G                    07:43.867         1         1        wizard22
18774756_43+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:37   00:08:50          18865744K         0  07:43.864         1         1        wizard22
18774756_43+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:37   00:08:50                0         0  00:00.002         1         1        wizard22
18774756_44  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:50:32   00:11:45        38G                    07:14.978         1         1        wizard24
18774756_44+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:50:32   00:11:45          14918896K         0  07:14.977         1         1        wizard24
18774756_44+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:50:32   00:11:45                0         0  00:00.001         1         1        wizard24
18774756_45  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:42   00:08:55        38G                    06:51.011         1         1        wizard24
18774756_45+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:42   00:08:55          13915524K         0  06:51.010         1         1        wizard24
18774756_45+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:47:42   00:08:55                0         0  00:00.001         1         1        wizard24
18774756_46  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:55:04   00:16:17        38G                    06:52.379         1         1        wizard24
18774756_46+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:55:04   00:16:17          14612080K         0  06:52.377         1         1        wizard24
18774756_46+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:55:04   00:16:17                0         0  00:00.001         1         1        wizard24
```

Host: headnode.beocat.ksu.edu

```
18774756_47  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:50:16   00:11:29        38G                    07:09.826         1         1        wizard24
18774756_47+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:50:16   00:11:29          13988020K         0  07:09.825         1         1        wizard24
18774756_47+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:50:16   00:11:29                0         0  00:00.001         1         1        wizard24
18774756_48  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:57:00   00:18:13        38G                    08:51.386         1         1        wizard24
18774756_48+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:57:00   00:18:13          14394752K         0  08:51.385         1         1        wizard24
18774756_48+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:57:00   00:18:13                0         0  00:00.001         1         1        wizard24
18774756_49  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:49   00:18:02        38G                    08:42.045         1         1        wizard24
18774756_49+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:49   00:18:02          15111004K         0  08:42.044         1         1        wizard24
18774756_49+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:49   00:18:02                0         0  00:00.001         1         1        wizard24
18774756_50  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:54   00:18:07        38G                    08:50.566         1         1        wizard24
18774756_50+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:54   00:18:07          15516924K         0  08:50.565         1         1        wizard24
18774756_50+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:54   00:18:07               188K         0  00:00.001         1         1        wizard24
18774756_51  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:46   00:17:59        38G                    08:37.390         1         1        wizard24
18774756_51+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:46   00:17:59          15264180K         0  08:37.389         1         1        wizard24
18774756_51+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:46   00:17:59                0         0  00:00.001         1         1        wizard24
18774756_52  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:35   00:09:48        38G                    00:30.400         1         1        wizard24
18774756_52+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:35   00:09:48            865396K         0  00:30.399         1         1        wizard24
18774756_52+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:48:35   00:09:48                0         0  00:00.001         1         1        wizard24
18774756_53  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:45   00:17:58        38G                    08:39.904         1         1        wizard24
18774756_53+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:45   00:17:58          14423084K         0  08:39.903         1         1        wizard24
18774756_53+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:45   00:17:58                0         0  00:00.001         1         1        wizard24
18774756_54  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:49   00:18:02        38G                    08:40.897         1         1        wizard24
18774756_54+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:49   00:18:02          14695588K         0  08:40.896         1         1        wizard24
18774756_54+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:49   00:18:02                0         0  00:00.001         1         1        wizard24
18774756_55  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:52   00:18:05        38G                    08:44.780         1         1        wizard24
18774756_55+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:52   00:18:05          15051700K         0  08:44.779         1         1        wizard24
18774756_55+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:52   00:18:05                0         0  00:00.001         1         1        wizard24
18774756_56  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:30   00:17:43        38G                    08:24.030         1         1        wizard24
18774756_56+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:30   00:17:43          14470484K         0  08:24.029         1         1        wizard24
18774756_56+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:30   00:17:43                0         0  00:00.001         1         1        wizard24
18774756_57  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:50   00:18:03        38G                    08:44.313         1         1        wizard24
18774756_57+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:50   00:18:03          15477788K         0  08:44.312         1         1        wizard24
18774756_57+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:50   00:18:03                0         0  00:00.001         1         1        wizard24
18774756_58  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:39   00:17:52        38G                    08:33.319         1         1        wizard24
18774756_58+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:39   00:17:52          14786156K         0  08:33.318         1         1        wizard24
18774756_58+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:39   00:17:52                0         0  00:00.001         1         1        wizard24
18774756_59  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:48   00:18:01        38G                    08:41.202         1         1        wizard24
18774756_59+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:48   00:18:01          15592380K         0  08:41.201         1         1        wizard24
18774756_59+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:48   00:18:01                0         0  00:00.001         1         1        wizard24
18774756_60  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:57   00:18:10        38G                    08:49.975         1         1        wizard24
18774756_60+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:57   00:18:10          15668376K         0  08:49.974         1         1        wizard24
18774756_60+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:57   00:18:10                0         0  00:00.001         1         1        wizard24
18774756_61  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:46   00:17:59        38G                    08:40.252         1         1        wizard24
18774756_61+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:46   00:17:59          15200996K         0  08:40.251         1         1        wizard24
18774756_61+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:46   00:17:59                0         0  00:00.001         1         1        wizard24
18774756_62  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:43   00:17:56        38G                    08:36.635         1         1        wizard24
18774756_62+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:43   00:17:56          14092292K         0  08:36.634         1         1        wizard24
18774756_62+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:43   00:17:56                0         0   00:00:00         1         1        wizard24
```

```
18774756_63  sam_my_job  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:40   00:17:53        38G                    08:31.537         1       1        wizard24
18774756_63+      batch  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:40   00:17:53            14407756K        0  08:31.536         1       1        wizard24
18774756_63+     extern  COMPLETED 2024-08-28T01:38:47 2024-08-28T01:56:40   00:17:53                  0        0  00:00.001         1       1        wizard24
[samatha94@icr-helios ~]$
```

Figure 5.6: Job Execution Summary for Stage 2 – This figure displays five screenshots illustrating the SLURM job completion details for parallelized concept induction in Stage 2 for the second set of 10 scene categories, where all 64 neuron-specific tasks were successfully executed.

```
[samatha94@icr-helios ~]$ sacct --format=JobID,JobName,State,Start,End,Elapsed,ReqMem,MaxRSS,MaxVMSize,TotalCPU,NCPUS,NNodes,NodeList --jobs=18775050
JobID           JobName      State               Start                 End    Elapsed     ReqMem     MaxRSS  MaxVMSize   TotalCPU      NCPUS   NNodes        NodeList
------------ ---------- ---------- ------------------- ------------------- ---------- ---------- ---------- ---------- ---------- ---------- -------- ---------------
18775050     job_stage+  COMPLETED 2024-08-28T02:00:29 2024-08-28T02:00:29   00:00:00         1G                    00:00.009          1        1          hero26
18775050.ba+      batch  COMPLETED 2024-08-28T02:00:29 2024-08-28T02:00:29   00:00:00                 64K        0  00:00.008          1        1          hero26
18775050.ex+     extern  COMPLETED 2024-08-28T02:00:29 2024-08-28T02:00:29   00:00:00                188K        0   00:00:00          1        1          hero26
```

Figure 5.7: Job Execution Summary for Neuron Index Generation in Stage 3 – This figure displays the SLURM job completion details for the neuron index generation task in Stage 3 for the second set of 10 scene categories.

```
Host: headnode.beocat.ksu.edu

[samatha94@icr-helios ~]$ sacct --format=JobID,JobName,State,Start,End,Elapsed,ReqMem,MaxRSS,MaxVMSize,TotalCPU,NCPUS,NNodes,NodeList --jobs=18775116
JobID           JobName      State               Start                 End    Elapsed     ReqMem     MaxRSS  MaxVMSize   TotalCPU      NCPUS   NNodes        NodeList
------------ ---------- ---------- ------------------- ------------------- ---------- ---------- ---------- ---------- ---------- ---------- -------- ---------------
18775116_0   sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:32:53   00:01:41         2G                    01:06.966          1        1          hero26
18775116_0.+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:32:53   00:01:41            1145528K        0  01:06.965          1        1          hero26
18775116_0.+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:32:53   00:01:41                188K        0   00:00:00          1        1          hero26
18775116_1   sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:03   00:01:51         2G                    01:05.736          1        1          hero26
18775116_1.+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:03   00:01:51            1020624K        0  01:05.735          1        1          hero26
18775116_1.+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:03   00:01:51                  0        0   00:00:00          1        1          hero26
18775116_2   sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:12   00:02:00         2G                    01:07.349          1        1          hero26
18775116_2.+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:12   00:02:00             766460K        0  01:07.348          1        1          hero26
18775116_2.+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:12   00:02:00                  0        0   00:00:00          1        1          hero26
18775116_3   sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:34:15   00:03:03         2G                    01:12.012          1        1          hero29
18775116_3.+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:34:15   00:03:03            1120436K        0  01:12.011          1        1          hero29
18775116_3.+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:34:15   00:03:03                  0        0   00:00:00          1        1          hero29
18775116_4   sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:32   00:02:20         2G                    01:15.228          1        1          hero29
18775116_4.+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:32   00:02:20             960332K        0  01:15.227          1        1          hero29
18775116_4.+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:32   00:02:20                  0        0   00:00:00          1        1          hero29
18775116_5   sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:34:03   00:02:51         2G                    01:21.043          1        1          hero29
18775116_5.+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:34:03   00:02:51             940500K        0  01:21.043          1        1          hero29
18775116_5.+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:34:03   00:02:51                192K        0   00:00:00          1        1          hero29
18775116_6   sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:16   00:02:04         2G                    01:20.089          1        1          hero30
18775116_6.+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:16   00:02:04            1237596K        0  01:20.088          1        1          hero30
18775116_6.+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:16   00:02:04                  0        0   00:00:00          1        1          hero30
18775116_7   sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:06   00:01:54         2G                    01:10.812          1        1          hero30
18775116_7.+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:06   00:01:54             778244K        0  01:10.811          1        1          hero30
18775116_7.+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:06   00:01:54                188K        0   00:00:00          1        1          hero30
18775116_8   sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:17   00:02:05         2G                    01:20.473          1        1          hero30
18775116_8.+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:17   00:02:05            1128832K        0  01:20.472          1        1          hero30
18775116_8.+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:17   00:02:05                  0        0   00:00:00          1        1          hero30
18775116_9   sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:14   00:02:02         2G                    01:06.665          1        1          mole042
18775116_9.+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:14   00:02:02            1005240K        0  01:06.664          1        1          mole042
18775116_9.+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:14   00:02:02                  0        0   00:00:00          1        1          mole042
18775116_10  sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:04   00:01:52         2G                    01:07.646          1        1          mole042
18775116_10+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:04   00:01:52            1069560K        0  01:07.645          1        1          mole042
18775116_10+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:04   00:01:52                  0        0   00:00:00          1        1          mole042
18775116_11  sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:08   00:01:56         2G                    01:10.038          1        1          mole042
18775116_11+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:08   00:01:56             953272K        0  01:10.038          1        1          mole042
18775116_11+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:08   00:01:56                  0        0   00:00:00          1        1          mole042
18775116_12  sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:34:10   00:02:58         2G                    01:09.703          1        1          mole043
18775116_12+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:34:10   00:02:58            1009268K        0  01:09.702          1        1          mole043
18775116_12+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:34:10   00:02:58                  0        0   00:00:00          1        1          mole043
18775116_13  sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:19   00:02:07         2G                    01:19.846          1        1          mole043
18775116_13+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:19   00:02:07            1193300K        0  01:19.845          1        1          mole043
18775116_13+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:19   00:02:07                  0        0   00:00:00          1        1          mole043
18775116_14  sam_my_job  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:23   00:02:11         2G                    01:09.338          1        1          mole043
18775116_14+      batch  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:23   00:02:11            1060484K        0  01:09.337          1        1          mole043
18775116_14+     extern  COMPLETED 2024-08-28T02:31:12 2024-08-28T02:33:23   00:02:11                188K        0   00:00:00          1        1          mole043
```

Host: headnode.beocat.ksu.edu

18775116_15	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:34:28	00:03:16	2G			01:09.224	1	1	mole044
18775116_15+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:34:28	00:03:16		1004552K	0	01:09.223	1	1	mole044
18775116_15+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:34:28	00:03:16		0	0	00:00:00	1	1	mole044
18775116_16	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:06	00:01:54	2G			01:10.379	1	1	mole044
18775116_16+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:06	00:01:54		902104K	0	01:10.378	1	1	mole044
18775116_16+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:06	00:01:54		0	0	00:00:00	1	1	mole044
18775116_17	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:32:52	00:01:40	2G			01:08.232	1	1	mole044
18775116_17+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:32:52	00:01:40		1097096K	0	01:08.231	1	1	mole044
18775116_17+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:32:52	00:01:40		192K	0	00:00:00	1	1	mole044
18775116_18	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:21	00:02:09	2G			01:20.432	1	1	mole047
18775116_18+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:21	00:02:09		1160184K	0	01:20.431	1	1	mole047
18775116_18+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:21	00:02:09		0	0	00:00:00	1	1	mole047
18775116_19	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:24	00:02:12	2G			01:08.919	1	1	mole047
18775116_19+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:24	00:02:12		1275620K	0	01:08.918	1	1	mole047
18775116_19+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:24	00:02:12		0	0	00:00:00	1	1	mole047
18775116_20	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:11	00:01:59	2G			01:11.664	1	1	mole047
18775116_20+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:11	00:01:59		858356K	0	01:11.663	1	1	mole047
18775116_20+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:11	00:01:59		196K	0	00:00:00	1	1	mole047
18775116_21	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:33	00:02:21	2G			01:04.967	1	1	mole048
18775116_21+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:33	00:02:21		902420K	0	01:04.966	1	1	mole048
18775116_21+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:33	00:02:21		192K	0	00:00:00	1	1	mole048
18775116_22	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:34:23	00:03:11	2G			01:09.906	1	1	mole048
18775116_22+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:34:23	00:03:11		1315620K	0	01:09.906	1	1	mole048
18775116_22+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:34:23	00:03:11		0	0	00:00:00	1	1	mole048
18775116_23	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:19	00:02:07	2G			01:09.642	1	1	mole048
18775116_23+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:19	00:02:07		1017132K	0	01:09.641	1	1	mole048
18775116_23+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:19	00:02:07		0	0	00:00:00	1	1	mole048
18775116_24	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:07	00:01:55	2G			01:10.857	1	1	mole080
18775116_24+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:07	00:01:55		882932K	0	01:10.856	1	1	mole080
18775116_24+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:07	00:01:55		0	0	00:00:00	1	1	mole080
18775116_25	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:30	00:02:18	2G			01:11.442	1	1	mole080
18775116_25+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:30	00:02:18		902468K	0	01:11.441	1	1	mole080
18775116_25+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:30	00:02:18		0	0	00:00:00	1	1	mole080
18775116_26	sam_my_job	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:41	00:02:29	2G			01:10.609	1	1	mole080
18775116_26+	batch	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:41	00:02:29		884460K	0	01:10.608	1	1	mole080
18775116_26+	extern	COMPLETED	2024-08-28T02:31:12	2024-08-28T02:33:41	00:02:29		192K	0	00:00:00	1	1	mole080
18775116_27	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:38	00:01:55	2G			01:14.394	1	1	hero13
18775116_27+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:38	00:01:55		788560K	0	01:14.393	1	1	hero13
18775116_27+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:38	00:01:55		0	0	00:00:00	1	1	hero13
18775116_28	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:22	00:02:39	2G			01:13.275	1	1	hero13
18775116_28+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:22	00:02:39		1026000K	0	01:13.274	1	1	hero13
18775116_28+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:22	00:02:39		0	0	00:00:00	1	1	hero13
18775116_29	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:20	00:02:37	2G			01:17.450	1	1	hero13
18775116_29+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:20	00:02:37		1090324K	0	01:17.449	1	1	hero13
18775116_29+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:20	00:02:37		0	0	00:00:00	1	1	hero13
18775116_30	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:59	00:02:16	2G			01:25.204	1	1	hero13
18775116_30+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:59	00:02:16		1188012K	0	01:25.203	1	1	hero13
18775116_30+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:59	00:02:16		0	0	00:00:00	1	1	hero13

Host: headnode.beocat.ksu.edu

18775116_31	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:24	00:02:41	2G			01:16.640	1	1	hero13
18775116_31+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:24	00:02:41		1143580K	0	01:16.639	1	1	hero13
18775116_31+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:24	00:02:41		188K	0	00:00:00	1	1	hero13
18775116_32	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:58	00:02:15	2G			01:24.068	1	1	hero13
18775116_32+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:58	00:02:15		1081372K	0	01:24.067	1	1	hero13
18775116_32+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:58	00:02:15		0	0	00:00.001	1	1	hero13
18775116_33	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:59	00:02:16	2G			01:23.807	1	1	hero13
18775116_33+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:59	00:02:16		1127232K	0	01:23.806	1	1	hero13
18775116_33+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:59	00:02:16		0	0	00:00:00	1	1	hero13
18775116_34	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:42	00:01:59	2G			01:10.703	1	1	hero13
18775116_34+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:42	00:01:59		910324K	0	01:10.702	1	1	hero13
18775116_34+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:42	00:01:59		0	0	00:00.001	1	1	hero13
18775116_35	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:52	00:02:09	2G			01:13.106	1	1	hero13
18775116_35+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:52	00:02:09		1189096K	0	01:13.105	1	1	hero13
18775116_35+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:52	00:02:09		192K	0	00:00:00	1	1	hero13
18775116_36	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:52	00:03:09	2G			01:12.317	1	1	hero13
18775116_36+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:52	00:03:09		951532K	0	01:12.317	1	1	hero13
18775116_36+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:52	00:03:09		196K	0	00:00:00	1	1	hero13
18775116_37	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:16	00:02:33	2G			01:11.571	1	1	hero13
18775116_37+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:16	00:02:33		926M	0	01:11.570	1	1	hero13
18775116_37+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:16	00:02:33		0	0	00:00:00	1	1	hero13
18775116_38	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:06	00:02:23	2G			01:15.849	1	1	hero13
18775116_38+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:06	00:02:23		935188K	0	01:15.848	1	1	hero13
18775116_38+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:07	00:02:24		0	0	00:00:00	1	1	hero13
18775116_39	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:16	00:02:33	2G			01:14.297	1	1	hero13
18775116_39+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:16	00:02:33		1092196K	0	01:14.296	1	1	hero13
18775116_39+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:16	00:02:33		0	0	00:00:00	1	1	hero13
18775116_40	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:09	00:02:26	2G			01:21.428	1	1	hero13
18775116_40+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:09	00:02:26		1000736K	0	01:21.427	1	1	hero13
18775116_40+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:09	00:02:26		0	0	00:00:00	1	1	hero13
18775116_41	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:11	00:02:28	2G			01:24.426	1	1	hero13
18775116_41+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:11	00:02:28		1140980K	0	01:24.425	1	1	hero13
18775116_41+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:11	00:02:28		0	0	00:00:00	1	1	hero13
18775116_42	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:36	00:01:53	2G			01:10.834	1	1	hero13
18775116_42+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:36	00:01:53		882292K	0	01:10.833	1	1	hero13
18775116_42+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:36	00:01:53		0	0	00:00:00	1	1	hero13
18775116_43	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:11	00:02:28	2G			01:15.565	1	1	hero13
18775116_43+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:11	00:02:28		911660K	0	01:15.564	1	1	hero13
18775116_43+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:11	00:02:28		0	0	00:00:00	1	1	hero13
18775116_44	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:23	00:02:40	2G			01:21.903	1	1	hero13
18775116_44+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:23	00:02:40		1200876K	0	01:21.902	1	1	hero13
18775116_44+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:23	00:02:40		0	0	00:00:00	1	1	hero13
18775116_45	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:04	00:02:21	2G			01:14.708	1	1	hero13
18775116_45+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:04	00:02:21		900980K	0	01:14.707	1	1	hero13
18775116_45+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:34:04	00:02:21		0	0	00:00:00	1	1	hero13
18775116_46	sam_my_job	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:51	00:02:08	2G			01:22.824	1	1	hero13
18775116_46+	batch	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:51	00:02:08		1333320K	0	01:22.823	1	1	hero13
18775116_46+	extern	COMPLETED	2024-08-28T02:31:43	2024-08-28T02:33:51	00:02:08		0	0	00:00:00	1	1	hero13

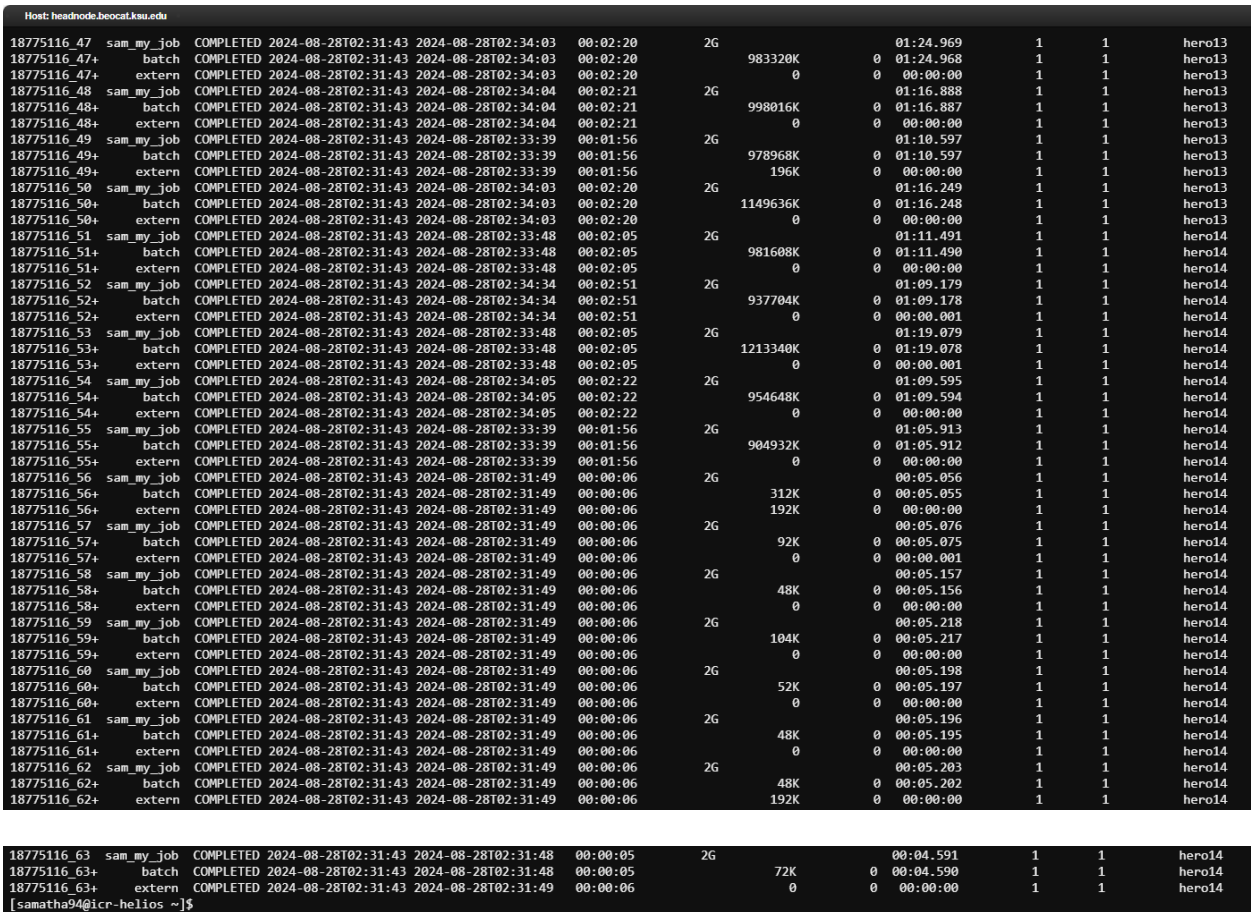


Figure 5.8: Job Execution Summary for Stage 3 – This figure displays five screenshots illustrating the SLURM job completion details for parallelized image retrieval and classification in Stage 3 for the second set of 10 scene categories.

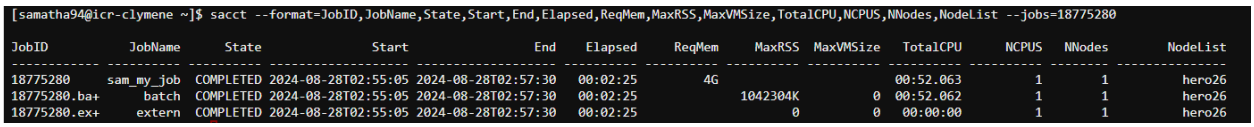


Figure 5.9: Job Execution Summary for Stage 4 – This figure displays the SLURM job completion details for statistical analysis and verification of neuron activations in Stage 4 for the second set of 10 scene categories.

For the second set of classes— “alley”, “art gallery”, “beach”, “castle”, “child’s room”, “closet”, “coast”, “mountain”, and “park”—we processed 822 images(Figure 5.10 to Figure 5.13). The automation pipeline was executed as follows:

1. **Stage 1 (Model Training and Data Configuration):** 12 minutes and 33 seconds (Figure 5.14)
2. **Stage 2 (Concept Induction and Label Hypothesis Generation):** 15 minutes (Figure 5.15)
3. **Stage 3 (Image Retrieval and Classification):** 7 minutes and 47 seconds (Figure 5.16 and Figure 5.17)
4. **Stage 4 (Statistical Analysis and Verification):** 3 minutes and 1 second (Figure 5.18)

The total time taken for all stages to process this set of 822 images was 38 minutes and 21 seconds.

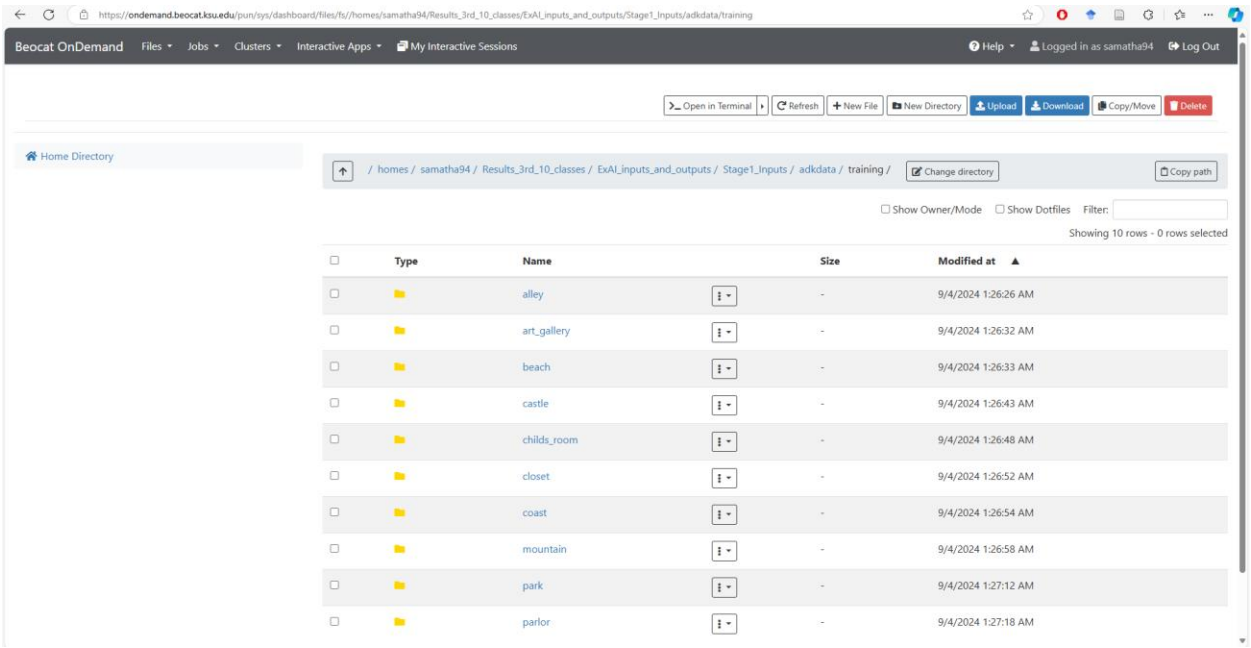


Figure 5.10: Training Folder Organization for New Scene Categories – This figure displays the folder structure used to organize training images for the third set of 10 new scene categories from the ADE20K dataset.

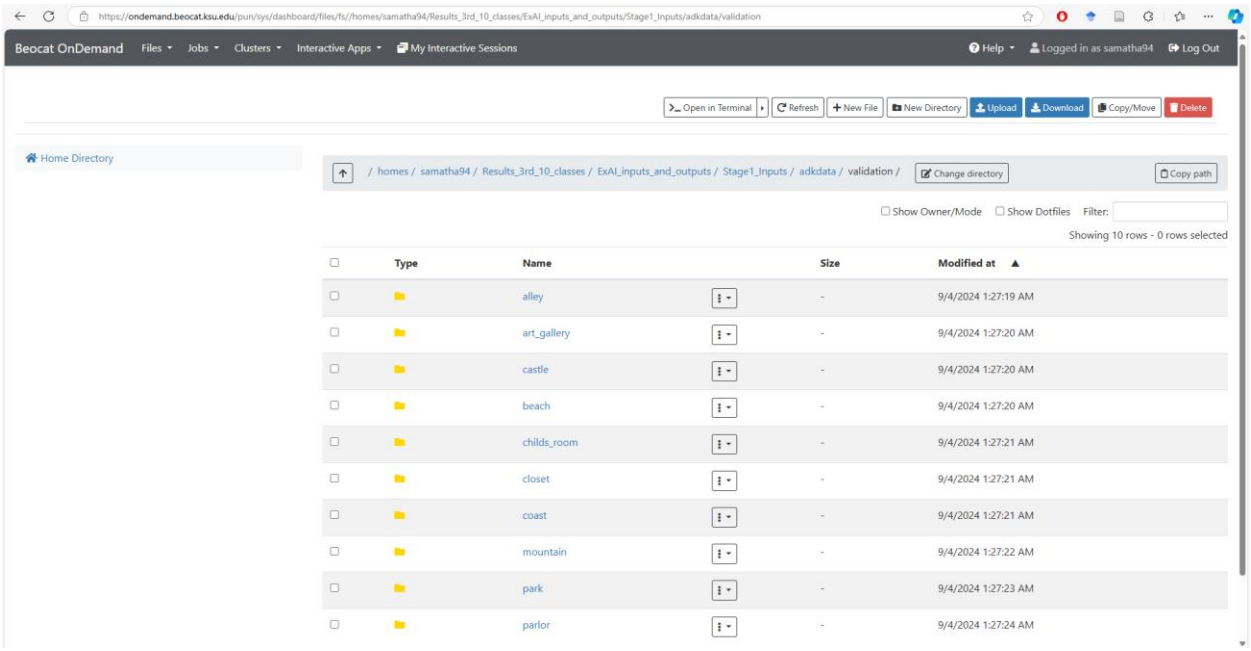


Figure 5.11: Validation Folder Organization for New Scene Categories – This figure highlights the structured folder layout for the validation dataset containing the new set of 10 scene categories.

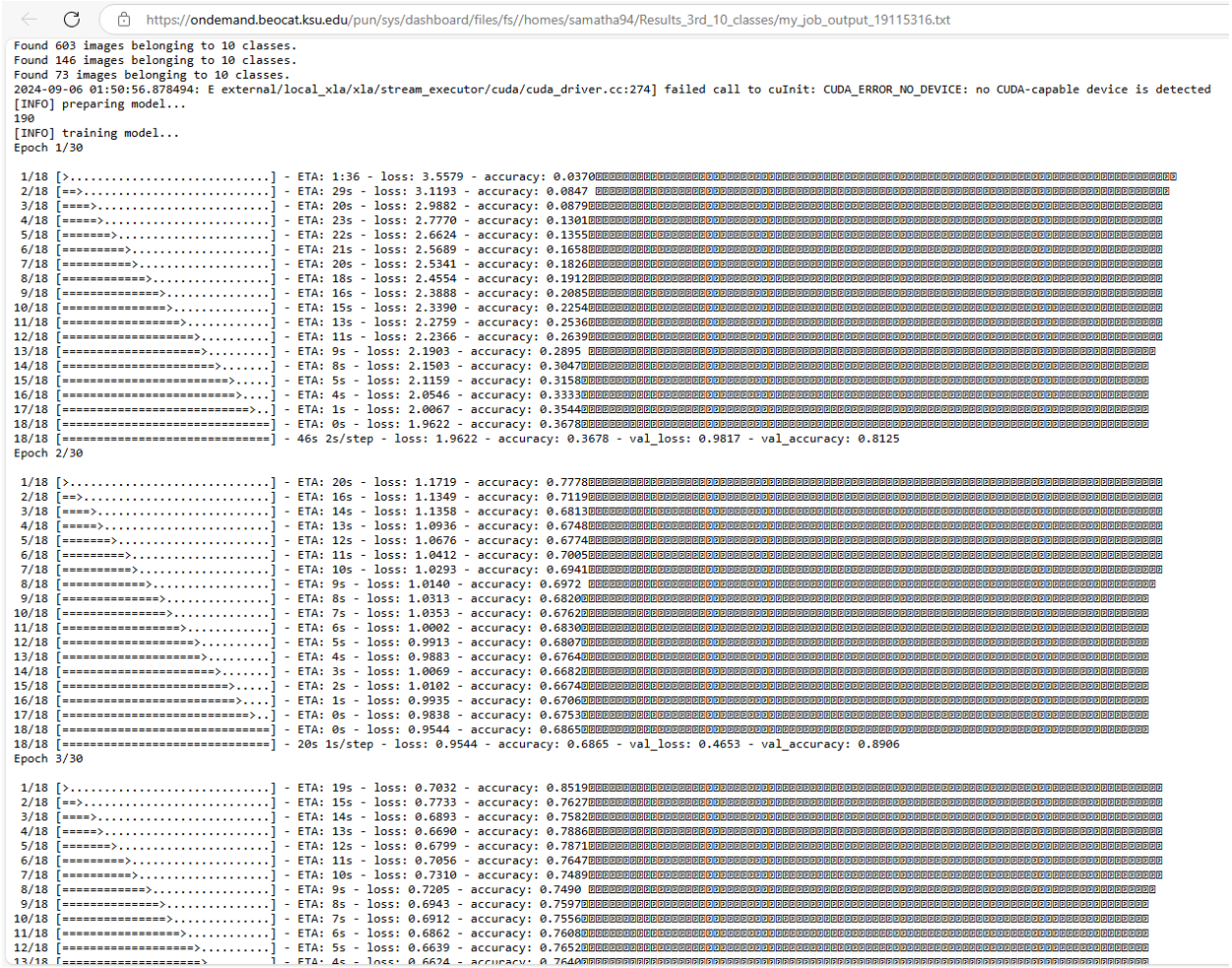


Figure 5.12: Input Dataset Summary for New Scene Categories: This figure presents a summary of the number of images in each selected class from the ADE20K dataset used in training the ResNet50V2 model

https://ondemand.beocat.ksu.edu/pun/sys/dashboard/files/fs//homes/samatha94/Results_3rd_10_classes/my_job_output_19115316.txt

```
[146 rows x 11 columns]
{'alley': 0, 'art_gallery': 1, 'beach': 2, 'castle': 3, 'childs_room': 4, 'closet': 5, 'coast': 6, 'mountain': 7, 'park': 8, 'parlor': 9}
          0         1  ...         9                    filenames
0    0.259941  0.004283  ...  0.001661   alley/ADE_train_00001233.jpg
1    0.988199  0.000680  ...  0.000150   alley/ADE_train_00001234.jpg
2    0.599735  0.027857  ...  0.041258   alley/ADE_train_00001235.jpg
3    0.963947  0.013735  ...  0.001161   alley/ADE_train_00001236.jpg
4    0.884039  0.047463  ...  0.013188   alley/ADE_train_00001237.jpg
..        ...       ...  ...       ...                          ...
141  0.009883  0.015030  ...  0.890530  parlor/ADE_train_00015159.jpg
142  0.917528  0.029106  ...  0.014117  parlor/ADE_train_00015160.jpg
143  0.014007  0.029851  ...  0.909471  parlor/ADE_train_00015161.jpg
144  0.008809  0.005195  ...  0.963505  parlor/ADE_train_00015162.jpg
145  0.009681  0.000637  ...  0.974037  parlor/ADE_train_00015163.jpg

[146 rows x 11 columns]
[INFO] evaluating network...
              precision    recall  f1-score   support

       alley       0.88      0.93      0.90        15
 art_gallery       0.87      1.00      0.93        13
       beach       0.83      0.71      0.77        14
      castle       0.80      1.00      0.89        12
 childs_room       0.86      0.86      0.86        14
      closet       1.00      0.86      0.92        14
       coast       0.76      0.81      0.79        16
    mountain       0.94      1.00      0.97        17
        park       1.00      0.76      0.87        17
      parlor       0.86      0.86      0.86        14

    accuracy                           0.88       146
   macro avg       0.88      0.88      0.88       146
weighted avg       0.88      0.88      0.88       146

           0         1    2  ...        62        63                    filenames
0    1.739340  0.000000  0.0  ...  1.211792  0.211209   alley/ADE_train_00001233.jpg
1    2.367265  0.000000  0.0  ...  2.007598  0.000000   alley/ADE_train_00001234.jpg
2    2.371900  0.000000  0.0  ...  1.624250  0.000000   alley/ADE_train_00001235.jpg
3    2.379179  0.000000  0.0  ...  3.096490  0.000000   alley/ADE_train_00001236.jpg
4    3.251593  0.000000  0.0  ...  1.929230  0.000000   alley/ADE_train_00001237.jpg
..        ...       ...  ...  ...       ...       ...                          ...
141  0.438810  0.000000  0.0  ...  1.393316  2.971103  parlor/ADE_train_00015159.jpg
142  1.843267  0.000000  0.0  ...  1.951641  1.043936  parlor/ADE_train_00015160.jpg
143  2.693723  0.308063  0.0  ...  1.934529  3.223967  parlor/ADE_train_00015161.jpg
144  2.360412  0.124546  0.0  ...  1.118846  1.517442  parlor/ADE_train_00015162.jpg
145  2.170304  1.245035  0.0  ...  1.430789  1.343011  parlor/ADE_train_00015163.jpg

[146 rows x 65 columns]
9 column is empty
14 column is empty
27 column is empty
29 column is empty
32 column is empty
33 column is empty
40 column is empty
44 column is empty
46 column is empty
53 column is empty
59 column is empty
Non-empty config files created: 53
Total non-empty config files created: 53
Cloning into 'ecii'...
```

Figure 5.13: Scene Classification Evaluation Report for New Scene Categories: This figure provides the evaluation metrics for the new set of 10 scene categories. It presents the precision, recall, and F1-score for each category, reporting an overall model accuracy of 88%

```
[samatha94@icr-clymene ~]$ sacct --format=JobID,JobName,State,Start,End,Elapsed,ReqMem,MaxRSS,MaxVMSize,TotalCPU,NCPUS,NNodes,NodeList --jobs=19115316
JobID           JobName      State               Start                 End    Elapsed     ReqMem     MaxRSS  MaxVMSize   TotalCPU      NCPUS   NNodes        NodeList
------------ ---------- ---------- ------------------- ------------------- ---------- ---------- ---------- ---------- ---------- ---------- -------- ---------------
19115316     sam_my_job  COMPLETED 2024-09-06T01:41:38 2024-09-06T01:54:11   00:12:33        20G                          13:45.806         24        4 mole[042-043,0+
19115316.ba+      batch  COMPLETED 2024-09-06T01:41:38 2024-09-06T01:54:11   00:12:33             2826932K          0  13:45.802          6        1          mole042
19115316.ex+     extern  COMPLETED 2024-09-06T01:41:38 2024-09-06T01:54:11   00:12:33                256K          0  00:00.003         24        4 mole[042-043,0+
```

Figure 5.14: SLURM Workload Summary for Stage 1 Training Execution – This figure displays the SLURM workload summary for the successful execution of the model training and data configuration process in Stage 1 for the thrid set of 10 scene categories.

Host: headnode.beocat.ksu.edu

```
[samatha94@icr-clymene ~]$ sacct --format=JobID,JobName,State,Start,End,Elapsed,ReqMem,MaxRSS,MaxVMSize,TotalCPU,NCPUS,NNodes,NodeList --jobs=19115601
JobID           JobName      State               Start                 End    Elapsed     ReqMem     MaxRSS  MaxVMSize   TotalCPU      NCPUS   NNodes        NodeList
------------ ---------- ---------- ------------------- ------------------- ---------- ---------- ---------- ---------- ---------- ---------- -------- ---------------
19115601_0   sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:15:08   00:11:31        38G                    09:25.121          1        1        wizard22
19115601_0.+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:15:08   00:11:31             21660172K          0  09:25.118          1        1        wizard22
19115601_0.+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:15:08   00:11:31                192K          0  00:00.002          1        1        wizard22
19115601_1   sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:27   00:10:50        38G                    07:23.093          1        1          hero13
19115601_1.+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:27   00:10:50             16616968K          0  07:23.092          1        1          hero13
19115601_1.+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:27   00:10:50                192K          0  00:00.001          1        1          hero13
19115601_2   sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:11   00:10:34        38G                    07:08.173          1        1          hero13
19115601_2.+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:11   00:10:34             16323236K          0  07:08.172          1        1          hero13
19115601_2.+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:11   00:10:34                196K          0  00:00.001          1        1          hero13
19115601_3   sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:20   00:10:43        38G                    07:58.468          1        1          hero02
19115601_3.+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:20   00:10:43             15307004K          0  07:58.466          1        1          hero02
19115601_3.+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:20   00:10:43                  0          0  00:00.001          1        1          hero02
19115601_4   sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:09   00:10:32        38G                    07:45.151          1        1          hero02
19115601_4.+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:09   00:10:32             19434620K          0  07:45.150          1        1          hero02
19115601_4.+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:09   00:10:32                188K          0  00:00.001          1        1          hero02
19115601_5   sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:10:39   00:07:02        38G                    05:34.704          1        1        wizard39
19115601_5.+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:10:39   00:07:02             19164288K          0  05:34.702          1        1        wizard39
19115601_5.+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:10:39   00:07:02                188K          0  00:00.001          1        1        wizard39
19115601_6   sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:56   00:11:19        38G                    09:17.494          1        1        warlock12
19115601_6.+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:56   00:11:19             20280496K          0  09:17.492          1        1        warlock12
19115601_6.+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:14:56   00:11:19                188K          0  00:00.001          1        1        warlock12
19115601_7   sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:35   00:08:58        38G                    05:47.193          1        1        warlock31
19115601_7.+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:35   00:08:58             20117672K          0  05:47.192          1        1        warlock31
19115601_7.+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:35   00:08:58                  0          0  00:00.001          1        1        warlock31
19115601_8   sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:49   00:09:12        38G                    05:53.901          1        1        warlock31
19115601_8.+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:49   00:09:12             14856212K          0  05:53.900          1        1        warlock31
19115601_8.+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:49   00:09:12                  0          0  00:00.001          1        1        warlock31
19115601_9   sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:24   00:03:47        38G                    00:27.044          1        1        wizard08
19115601_9.+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:24   00:03:47              818944K          0  00:27.042          1        1        wizard08
19115601_9.+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:24   00:03:47                188K          0  00:00.001          1        1        wizard08
19115601_10  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:41   00:10:04        38G                    06:34.816          1        1        wizard08
19115601_10+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:41   00:10:04             20600016K          0  06:34.815          1        1        wizard08
19115601_10+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:41   00:10:04                  0          0  00:00.001          1        1        wizard08
19115601_11  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:21   00:09:44        38G                    07:21.773          1        1        wizard14
19115601_11+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:21   00:09:44             19293324K          0  07:21.772          1        1        wizard14
19115601_11+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:21   00:09:44                192K          0  00:00.001          1        1        wizard14
19115601_12  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:16   00:09:39        38G                    07:13.639          1        1        wizard16
19115601_12+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:16   00:09:39             20682960K          0  07:13.638          1        1        wizard16
19115601_12+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:16   00:09:39                192K          0  00:00.001          1        1        wizard16
19115601_13  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:15:08   00:11:31        38G                    08:46.379          1        1         dwarf54
19115601_13+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:15:08   00:11:31             19532136K          0  08:46.378          1        1         dwarf54
19115601_13+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:15:08   00:11:31                  0          0  00:00.001          1        1         dwarf54
19115601_14  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:06:59   00:03:22        38G                    00:40.099          1        1         dwarf54
19115601_14+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:06:59   00:03:22              943124K          0  00:40.098          1        1         dwarf54
19115601_14+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:06:59   00:03:22                192K          0  00:00:00          1        1         dwarf54
```

Host: headnode.beocat.ksu.edu

```
19115601_15  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:12   00:08:35        38G                    06:30.292          1        1        warlock16
19115601_15+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:12   00:08:35             19531744K          0  06:30.291          1        1        warlock16
19115601_15+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:12   00:08:35                188K          0  00:00:00          1        1        warlock16
19115601_16  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:16:23   00:12:46        38G                    06:29.352          1        1        warlock24
19115601_16+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:16:23   00:12:46             16434376K          0  06:29.351          1        1        warlock24
19115601_16+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:16:23   00:12:46                188K          0  00:00.001          1        1        warlock24
19115601_17  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:17:04   00:13:27        38G                    06:11.459          1        1        warlock24
19115601_17+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:17:04   00:13:27             14431572K          0  06:11.457          1        1        warlock24
19115601_17+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:17:06   00:13:29                192K          0  00:00.001          1        1        warlock24
19115601_18  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:36   00:09:59        38G                    06:37.295          1        1        warlock05
19115601_18+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:36   00:09:59             18334880K          0  06:37.294          1        1        warlock05
19115601_18+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:36   00:09:59                192K          0  00:00:00          1        1        warlock05
19115601_19  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:29   00:09:52        38G                    06:28.547          1        1        warlock05
19115601_19+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:29   00:09:52             15226872K          0  06:28.546          1        1        warlock05
19115601_19+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:29   00:09:52                  0          0  00:00:00          1        1        warlock05
19115601_20  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:14   00:09:37        38G                    06:41.683          1        1        warlock07
19115601_20+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:14   00:09:37             15225820K          0  06:41.682          1        1        warlock07
19115601_20+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:14   00:09:37                192K          0  00:00:00          1        1        warlock07
19115601_21  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:03   00:09:26        38G                    06:30.221          1        1        warlock07
19115601_21+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:03   00:09:26             20160192K          0  06:30.220          1        1        warlock07
19115601_21+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:06   00:09:29                184K          0  00:00:00          1        1        warlock07
19115601_22  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:39   00:10:02        38G                    06:43.419          1        1        warlock08
19115601_22+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:39   00:10:02             18847296K          0  06:43.418          1        1        warlock08
19115601_22+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:39   00:10:02                  0          0  00:00.001          1        1        warlock08
19115601_23  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:34   00:09:57        38G                    06:33.463          1        1        warlock08
19115601_23+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:34   00:09:57             16159732K          0  06:33.462          1        1        warlock08
19115601_23+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:34   00:09:57                  0          0  00:00.001          1        1        warlock08
19115601_24  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:18:15   00:14:38        38G                    07:25.102          1        1        warlock09
19115601_24+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:18:15   00:14:38             14203480K          0  07:25.102          1        1        warlock09
19115601_24+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:18:15   00:14:38                  0          0  00:00:00          1        1        warlock09
19115601_25  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:18:10   00:14:33        38G                    07:21.156          1        1        warlock09
19115601_25+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:18:10   00:14:33             15004980K          0  07:21.155          1        1        warlock09
19115601_25+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:18:10   00:14:33                  0          0  00:00:00          1        1        warlock09
19115601_26  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:18:17   00:14:40        38G                    07:27.146          1        1        warlock09
19115601_26+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:18:17   00:14:40             14700500K          0  07:27.145          1        1        warlock09
19115601_26+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:18:17   00:14:40                  0          0  00:00:00          1        1        warlock09
19115601_27  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:06   00:03:29        38G                    00:18.188          1        1        warlock25
19115601_27+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:06   00:03:29              843504K          0  00:18.187          1        1        warlock25
19115601_27+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:06   00:03:29                192K          0  00:00:00          1        1        warlock25
19115601_28  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:11:51   00:08:14        38G                    05:02.809          1        1        warlock25
19115601_28+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:11:51   00:08:14             19678824K          0  05:02.808          1        1        warlock25
19115601_28+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:11:51   00:08:14                  0          0  00:00:00          1        1        warlock25
19115601_29  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:16   00:03:39        38G                    00:18.852          1        1        warlock26
19115601_29+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:16   00:03:39              871656K          0  00:18.851          1        1        warlock26
19115601_29+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:16   00:03:39                188K          0  00:00:00          1        1        warlock26
19115601_30  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:16   00:08:39        38G                    05:11.942          1        1        warlock26
19115601_30+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:16   00:08:39             19693296K          0  05:11.941          1        1        warlock26
19115601_30+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:16   00:08:39                  0          0  00:00:00          1        1        warlock26
```

Host: headnode.beocat.ksu.edu

```
19115601_31  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:33  00:09:56        38G                      06:29.291         1         1      warlock27
19115601_31+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:33  00:09:56          19062540K          0  06:29.290         1         1      warlock27
19115601_31+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:33  00:09:56                 0          0  00:00.001         1         1      warlock27
19115601_32  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:28  00:03:51        38G                      00:22.038         1         1      warlock27
19115601_32+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:28  00:03:51            945564K          0  00:22.036         1         1      warlock27
19115601_32+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:07:28  00:03:51               188K          0  00:00.001         1         1      warlock27
19115601_33  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:06:10  00:02:33        38G                      00:28.386         1         1       wizard31
19115601_33+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:06:10  00:02:33            911608K          0  00:28.385         1         1       wizard31
19115601_33+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:06:10  00:02:33               192K          0  00:00.001         1         1       wizard31
19115601_34  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:54  00:09:17        38G                      06:01.256         1         1      wizard36
19115601_34+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:54  00:09:17          18071104K          0  06:01.255         1         1      wizard36
19115601_34+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:54  00:09:17               188K          0  00:00:00          1         1      wizard36
19115601_35  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:51  00:09:14        38G                      05:58.270         1         1      wizard36
19115601_35+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:51  00:09:14          17253284K          0  05:58.269         1         1      wizard36
19115601_35+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:51  00:09:14               188K          0  00:00:00          1         1      wizard36
19115601_36  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:06  00:08:29        38G                      06:11.683         1         1      wizard37
19115601_36+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:06  00:08:29          20487636K          0  06:11.682         1         1      wizard37
19115601_36+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:06  00:08:29               192K          0  00:00:00          1         1      wizard37
19115601_37  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:56  00:09:19        38G                      05:56.356         1         1      wizard38
19115601_37+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:56  00:09:19          18321656K          0  05:56.355         1         1      wizard38
19115601_37+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:12:56  00:09:19                 0          0  00:00:00          1         1      wizard38
19115601_38  sam_my_job  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:13  00:09:36        38G                      06:10.534         1         1      wizard38
19115601_38+      batch  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:13  00:09:36          15954964K          0  06:10.534         1         1      wizard38
19115601_38+     extern  COMPLETED 2024-09-06T02:03:37 2024-09-06T02:13:13  00:09:36               192K          0  00:00:00          1         1      wizard38
19115601_39  sam_my_job  COMPLETED 2024-09-06T02:03:38 2024-09-06T02:13:23  00:09:45        38G                      07:17.677         1         1      warlock18
19115601_39+      batch  COMPLETED 2024-09-06T02:03:38 2024-09-06T02:13:23  00:09:45          20454856K          0  07:17.676         1         1      warlock18
19115601_39+     extern  COMPLETED 2024-09-06T02:03:38 2024-09-06T02:13:23  00:09:45               196K          0  00:00.001         1         1      warlock18
19115601_40  sam_my_job  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:07:42  00:03:05        38G                      00:35.283         1         1         hero09
19115601_40+      batch  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:07:42  00:03:05            897328K          0  00:35.282         1         1         hero09
19115601_40+     extern  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:07:42  00:03:05                 0          0  00:00.001         1         1         hero09
19115601_41  sam_my_job  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:16:06  00:11:29        38G                      08:54.091         1         1         hero09
19115601_41+      batch  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:16:06  00:11:29          17843028K          0  08:54.090         1         1         hero09
19115601_41+     extern  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:16:06  00:11:29                 0          0  00:00.001         1         1         hero09
19115601_42  sam_my_job  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:16:06  00:11:29        38G                      08:55.029         1         1         hero09
19115601_42+      batch  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:16:06  00:11:29          15381544K          0  08:55.028         1         1         hero09
19115601_42+     extern  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:16:06  00:11:29               188K          0  00:00.001         1         1         hero09
19115601_43  sam_my_job  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:15:14  00:10:37        38G                      08:35.004         1         1         hero42
19115601_43+      batch  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:15:14  00:10:37          18993760K          0  08:35.003         1         1         hero42
19115601_43+     extern  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:15:14  00:10:37               192K          0  00:00.001         1         1         hero42
19115601_44  sam_my_job  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:07:12  00:02:35        38G                      00:35.864         1         1         hero42
19115601_44+      batch  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:07:12  00:02:35            908536K          0  00:35.863         1         1         hero42
19115601_44+     extern  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:07:12  00:02:35                 0          0  00:00.001         1         1         hero42
19115601_45  sam_my_job  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:13:51  00:09:14        38G                      06:49.988         1         1      warlock14
19115601_45+      batch  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:13:51  00:09:14          18939204K          0  06:49.987         1         1      warlock14
19115601_45+     extern  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:13:51  00:09:14                 0          0  00:00:00          1         1      warlock14
19115601_46  sam_my_job  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:07:27  00:02:50        38G                      00:22.707         1         1      warlock14
19115601_46+      batch  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:07:27  00:02:50            939828K          0  00:22.705         1         1      warlock14
19115601_46+     extern  COMPLETED 2024-09-06T02:04:37 2024-09-06T02:07:27  00:02:50               184K          0  00:00.001         1         1      warlock14
```

Host: headnode.beocat.ksu.edu

```
19115601_47  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:15:49  00:10:11        38G                      07:21.272         1         1      warlock20
19115601_47+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:15:49  00:10:11          18306636K          0  07:21.270         1         1      warlock20
19115601_47+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:15:49  00:10:11               192K          0  00:00.002         1         1      warlock20
19115601_48  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:15:49  00:10:11        38G                      07:25.504         1         1      warlock20
19115601_48+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:15:49  00:10:11          16261400K          0  07:25.501         1         1      warlock20
19115601_48+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:15:49  00:10:11               192K          0  00:00.002         1         1      warlock20
19115601_49  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:16:51  00:11:13        38G                      06:00.166         1         1      warlock24
19115601_49+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:16:51  00:11:13          15087336K          0  06:00.165         1         1      warlock24
19115601_49+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:16:51  00:11:13                 0          0  00:00.001         1         1      warlock24
19115601_50  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:17:17  00:11:39        38G                      06:04.377         1         1      warlock24
19115601_50+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:17:17  00:11:39          14654572K          0  06:04.375         1         1      warlock24
19115601_50+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:17:17  00:11:39                 0          0  00:00.001         1         1      warlock24
19115601_51  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:17:11  00:11:33        38G                      05:55.679         1         1      warlock24
19115601_51+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:17:11  00:11:33          14858404K          0  05:55.678         1         1      warlock24
19115601_51+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:17:11  00:11:33               192K          0  00:00.001         1         1      warlock24
19115601_52  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:11:46  00:06:08        38G                      05:56.025         1         1      warlock24
19115601_52+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:11:46  00:06:08          19734988K          0  05:56.024         1         1      warlock24
19115601_52+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:11:46  00:06:08                 0          0  00:00.001         1         1      warlock24
19115601_53  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:11:21  00:05:43        38G                      00:18.458         1         1      warlock24
19115601_53+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:11:21  00:05:43            858892K          0  00:18.457         1         1      warlock24
19115601_53+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:11:21  00:05:43                 0          0  00:00.001         1         1      warlock24
19115601_54  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:16:06  00:10:28        38G                      07:55.091         1         1      warlock02
19115601_54+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:16:06  00:10:28          16605164K          0  07:55.089         1         1      warlock02
19115601_54+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:16:06  00:10:28               192K          0  00:00.001         1         1      warlock02
19115601_55  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:16:29  00:10:51        38G                      08:22.333         1         1      warlock02
19115601_55+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:16:29  00:10:51          16095420K          0  08:22.331         1         1      warlock02
19115601_55+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:16:29  00:10:51               236K          0  00:00.001         1         1      warlock02
19115601_56  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:18:56  00:13:18        38G                      07:12.143         1         1      warlock09
19115601_56+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:18:56  00:13:18          14055064K          0  07:12.142         1         1      warlock09
19115601_56+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:18:56  00:13:18                 0          0  00:00:00          1         1      warlock09
19115601_57  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:18:56  00:13:18        38G                      07:15.014         1         1      warlock09
19115601_57+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:18:56  00:13:18          15072796K          0  07:15.013         1         1      warlock09
19115601_57+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:18:56  00:13:18                 0          0  00:00:00          1         1      warlock09
19115601_58  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:18:56  00:13:18        38G                      07:13.799         1         1      warlock09
19115601_58+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:18:56  00:13:18          15393936K          0  07:13.798         1         1      warlock09
19115601_58+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:18:56  00:13:18                 0          0  00:00:00          1         1      warlock09
19115601_59  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:11:59  00:06:21        38G                      00:22.099         1         1      warlock09
19115601_59+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:11:59  00:06:21            861556K          0  00:22.098         1         1      warlock09
19115601_59+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:11:59  00:06:21                 0          0  00:00:00          1         1      warlock09
19115601_60  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:19:04  00:13:26        38G                      07:20.884         1         1      warlock09
19115601_60+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:19:04  00:13:26          15036168K          0  07:20.884         1         1      warlock09
19115601_60+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:19:06  00:13:28                 0          0  00:00:00          1         1      warlock09
19115601_61  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:13:15  00:07:37        38G                      06:40.192         1         1      warlock09
19115601_61+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:13:15  00:07:37          18790672K          0  06:40.191         1         1      warlock09
19115601_61+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:13:15  00:07:37               192K          0  00:00:00          1         1      warlock09
19115601_62  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:12:59  00:07:21        38G                      06:59.308         1         1       wizard31
19115601_62+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:12:59  00:07:21          20137604K          0  06:59.307         1         1       wizard31
19115601_62+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:12:59  00:07:21                 0          0  00:00.001         1         1       wizard31
```

```
19115601_63  sam_my_job  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:13:14   00:07:36        38G                    06:47.199         1         1        wizard31
19115601_63+      batch  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:13:14   00:07:36           14305588K         0  06:47.198         1         1        wizard31
19115601_63+     extern  COMPLETED 2024-09-06T02:05:38 2024-09-06T02:13:14   00:07:36               196K         0   00:00:00         1         1        wizard31
[samatha94@icr-clymene ~]$
```

Figure 5.15: Job Execution Summary for Stage 2 – This figure displays five screenshots illustrating the SLURM job completion details for parallelized concept induction in Stage 2 for the third set of 10 scene categories, where all 64 neuron-specific tasks were successfully executed.

```
[samatha94@icr-helios ~]$ sacct --format=JobID,JobName,State,Start,End,Elapsed,ReqMem,MaxRSS,MaxVMSize,TotalCPU,NCPUS,NNodes,NodeList --jobs=19115819
JobID           JobName      State               Start                 End    Elapsed     ReqMem     MaxRSS  MaxVMSize   TotalCPU      NCPUS   NNodes        NodeList
------------ ---------- ---------- ------------------- ------------------- ---------- ---------- ---------- ---------- ---------- --------- -------- ---------------
19115819     job_stage+  COMPLETED 2024-09-06T02:22:38 2024-09-06T02:22:39   00:00:01         1G                    00:00.011         1        1         dwarf52
19115819.ba+      batch  COMPLETED 2024-09-06T02:22:38 2024-09-06T02:22:39   00:00:01                64K         0  00:00.010         1        1         dwarf52
19115819.ex+     extern  COMPLETED 2024-09-06T02:22:38 2024-09-06T02:22:39   00:00:01               196K         0  00:00.001         1        1         dwarf52
```

Figure 5.16: Job Execution Summary for Neuron Index Generation in Stage 3 – This figure displays the SLURM job completion details for the neuron index generation task in Stage 3 for the third set of 10 scene categories.

```
Host: headnode.beocat.ksu.edu

[samatha94@icr-helios ~]$ sacct --format=JobID,JobName,State,Start,End,Elapsed,ReqMem,MaxRSS,MaxVMSize,TotalCPU,NCPUS,NNodes,NodeList --jobs=19115845
JobID           JobName      State               Start                 End    Elapsed     ReqMem     MaxRSS  MaxVMSize   TotalCPU      NCPUS   NNodes        NodeList
------------ ---------- ---------- ------------------- ------------------- ---------- ---------- ---------- ---------- ---------- --------- -------- ---------------
19115845_0   sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:31:14   00:07:36         2G                    01:36.828         1        1         dwarf52
19115845_0.+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:31:14   00:07:36           1199236K         0  01:36.827         1        1         dwarf52
19115845_0.+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:31:14   00:07:36                  0         0  00:00.001         1        1         dwarf52
19115845_1   sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:00   00:06:22         2G                    01:37.043         1        1         dwarf52
19115845_1.+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:00   00:06:22           1143540K         0  01:37.041         1        1         dwarf52
19115845_1.+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:00   00:06:22                  0         0  00:00.001         1        1         dwarf52
19115845_2   sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:15   00:06:37         2G                    01:36.347         1        1         dwarf52
19115845_2.+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:15   00:06:37           1049712K         0  01:36.346         1        1         dwarf52
19115845_2.+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:15   00:06:37                  0         0   00:00.00         1        1         dwarf52
19115845_3   sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:31:15   00:07:37         2G                    01:32.351         1        1         dwarf52
19115845_3.+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:31:15   00:07:37           1080572K         0  01:32.349         1        1         dwarf52
19115845_3.+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:31:15   00:07:37                  0         0  00:00.001         1        1         dwarf52
19115845_4   sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:06   00:06:28         2G                    01:30.226         1        1         dwarf52
19115845_4.+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:06   00:06:28            905904K         0  01:30.225         1        1         dwarf52
19115845_4.+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:07   00:06:29                  0         0  00:00.001         1        1         dwarf52
19115845_5   sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:35   00:06:57         2G                    01:32.149         1        1         dwarf52
19115845_5.+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:35   00:06:57           1065788K         0  01:32.148         1        1         dwarf52
19115845_5.+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:35   00:06:57                  0         0   00:00.00         1        1         dwarf52
19115845_6   sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:20   00:06:42         2G                    01:43.078         1        1         dwarf52
19115845_6.+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:20   00:06:42           1160816K         0  01:43.077         1        1         dwarf52
19115845_6.+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:20   00:06:42                  0         0  00:00.001         1        1         dwarf52
19115845_7   sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:40   00:07:02         2G                    01:31.432         1        1         dwarf52
19115845_7.+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:40   00:07:02            872124K         0  01:31.431         1        1         dwarf52
19115845_7.+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:40   00:07:02                  0         0  00:00.001         1        1         dwarf52
19115845_8   sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:29   00:06:51         2G                    01:29.098         1        1         dwarf52
19115845_8.+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:29   00:06:51           1107576K         0  01:29.097         1        1         dwarf52
19115845_8.+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:29   00:06:51                  0         0   00:00.00         1        1         dwarf52
19115845_9   sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:34   00:06:56         2G                    01:31.484         1        1         dwarf52
19115845_9.+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:34   00:06:56           1130396K         0  01:31.483         1        1         dwarf52
19115845_9.+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:34   00:06:56               192K         0   00:00.00         1        1         dwarf52
19115845_10  sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:25   00:06:47         2G                    01:32.640         1        1         dwarf52
19115845_10+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:25   00:06:47           1287808K         0  01:32.639         1        1         dwarf52
19115845_10+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:25   00:06:47               184K         0  00:00.001         1        1         dwarf52
19115845_11  sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:07   00:06:29         2G                    01:32.512         1        1         dwarf52
19115845_11+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:07   00:06:29           1017568K         0  01:32.511         1        1         dwarf52
19115845_11+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:07   00:06:29                  0         0  00:00.001         1        1         dwarf52
19115845_12  sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:29:22   00:05:44         2G                    01:12.481         1        1          hero53
19115845_12+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:29:22   00:05:44           1031260K         0  01:12.480         1        1          hero53
19115845_12+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:29:22   00:05:44               192K         0   00:00.00         1        1          hero53
19115845_13  sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:28   00:06:50         2G                    01:13.861         1        1          hero53
19115845_13+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:28   00:06:50            897776K         0  01:13.861         1        1          hero53
19115845_13+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:28   00:06:50                  0         0   00:00.00         1        1          hero53
19115845_14  sam_my_job  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:11   00:06:33         2G                    01:16.688         1        1          hero53
19115845_14+      batch  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:11   00:06:33           1150128K         0  01:16.687         1        1          hero53
19115845_14+     extern  COMPLETED 2024-09-06T02:23:38 2024-09-06T02:30:11   00:06:33                  0         0   00:00.00         1        1          hero53
```

Host: headnode.beocat.ksu.edu

19115845_15	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:30	00:05:52	2G			01:12.293	1	1	hero53
19115845_15+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:30	00:05:52		943392K	0	01:12.292	1	1	hero53
19115845_15+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:30	00:05:52		0	0	00:00:00	1	1	hero53
19115845_16	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:18	00:06:40	2G			01:14.092	1	1	hero53
19115845_16+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:18	00:06:40		1147312K	0	01:14.091	1	1	hero53
19115845_16+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:18	00:06:40		192K	0	00:00:00	1	1	hero53
19115845_17	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:13	00:06:35	2G			01:11.164	1	1	hero53
19115845_17+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:13	00:06:35		1112072K	0	01:11.163	1	1	hero53
19115845_17+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:13	00:06:35		0	0	00:00:00	1	1	hero53
19115845_18	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:29	00:05:51	2G			01:12.055	1	1	hero53
19115845_18+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:29	00:05:51		991780K	0	01:12.054	1	1	hero53
19115845_18+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:29	00:05:51		0	0	00:00:00	1	1	hero53
19115845_19	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:24	00:06:46	2G			01:17.606	1	1	hero53
19115845_19+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:24	00:06:46		911120K	0	01:17.605	1	1	hero53
19115845_19+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:24	00:06:46		0	0	00:00:00	1	1	hero53
19115845_20	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:13	00:06:35	2G			01:18.892	1	1	hero53
19115845_20+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:13	00:06:35		1083236K	0	01:18.891	1	1	hero53
19115845_20+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:13	00:06:35		0	0	00:00:00	1	1	hero53
19115845_21	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:34	00:05:56	2G			01:14.203	1	1	hero53
19115845_21+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:34	00:05:56		1089624K	0	01:14.202	1	1	hero53
19115845_21+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:34	00:05:56		0	0	00:00:00	1	1	hero53
19115845_22	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:04	00:06:26	2G			01:22.594	1	1	hero53
19115845_22+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:04	00:06:26		1007252K	0	01:22.593	1	1	hero53
19115845_22+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:06	00:06:28		0	0	00:00:00	1	1	hero53
19115845_23	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:20	00:05:42	2G			01:04.583	1	1	hero53
19115845_23+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:20	00:05:42		1291424K	0	01:04.582	1	1	hero53
19115845_23+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:20	00:05:42		0	0	00:00:00	1	1	hero53
19115845_24	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:06	00:06:28	2G			01:16.799	1	1	hero54
19115845_24+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:06	00:06:28		1066836K	0	01:16.798	1	1	hero54
19115845_24+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:07	00:06:29		0	0	00:00:00	1	1	hero54
19115845_25	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:31:25	00:07:47	2G			01:13.778	1	1	hero54
19115845_25+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:31:25	00:07:47		931772K	0	01:13.777	1	1	hero54
19115845_25+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:31:25	00:07:47		188K	0	00:00:00	1	1	hero54
19115845_26	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:47	00:06:09	2G			01:14.171	1	1	hero54
19115845_26+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:47	00:06:09		1088476K	0	01:14.170	1	1	hero54
19115845_26+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:47	00:06:09		0	0	00:00:00	1	1	hero54
19115845_27	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:33	00:05:55	2G			01:10.638	1	1	hero54
19115845_27+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:33	00:05:55		1070464K	0	01:10.637	1	1	hero54
19115845_27+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:33	00:05:55		0	0	00:00:00	1	1	hero54
19115845_28	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:24	00:06:46	2G			01:12.311	1	1	hero54
19115845_28+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:24	00:06:46		997284K	0	01:12.310	1	1	hero54
19115845_28+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:24	00:06:46		0	0	00:00:00	1	1	hero54
19115845_29	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:52	00:06:14	2G			01:13.589	1	1	hero54
19115845_29+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:52	00:06:14		949420K	0	01:13.588	1	1	hero54
19115845_29+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:52	00:06:14		0	0	00:00:00	1	1	hero54
19115845_30	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:06	00:06:28	2G			01:13.948	1	1	hero54
19115845_30+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:06	00:06:28		884180K	0	01:13.947	1	1	hero54
19115845_30+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:07	00:06:29		0	0	00:00:00	1	1	hero54

Host: headnode.beocat.ksu.edu

19115845_31	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:08	00:06:30	2G			01:12.446	1	1	hero54
19115845_31+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:08	00:06:30		923704K	0	01:12.445	1	1	hero54
19115845_31+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:08	00:06:30		0	0	00:00:00	1	1	hero54
19115845_32	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:47	00:06:09	2G			01:15.330	1	1	hero54
19115845_32+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:47	00:06:09		1067452K	0	01:15.329	1	1	hero54
19115845_32+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:47	00:06:09		0	0	00:00:00	1	1	hero54
19115845_33	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:00	00:06:22	2G			01:14.044	1	1	hero54
19115845_33+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:00	00:06:22		930212K	0	01:14.043	1	1	hero54
19115845_33+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:00	00:06:22		0	0	00:00:00	1	1	hero54
19115845_34	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:52	00:06:14	2G			01:15.079	1	1	hero54
19115845_34+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:52	00:06:14		954096K	0	01:15.079	1	1	hero54
19115845_34+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:52	00:06:14		0	0	00:00:00	1	1	hero54
19115845_35	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:30	00:06:52	2G			01:12.350	1	1	hero54
19115845_35+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:30	00:06:52		1176404K	0	01:12.349	1	1	hero54
19115845_35+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:30	00:06:52		0	0	00:00.001	1	1	hero54
19115845_36	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:26:44	00:03:06	2G			01:09.587	1	1	mole042
19115845_36+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:26:44	00:03:06		1000116K	0	01:09.586	1	1	mole042
19115845_36+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:26:44	00:03:06		188K	0	00:00:00	1	1	mole042
19115845_37	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:27:16	00:03:38	2G			01:10.936	1	1	mole042
19115845_37+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:27:16	00:03:38		1093096K	0	01:10.935	1	1	mole042
19115845_37+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:27:16	00:03:38		0	0	00:00:00	1	1	mole042
19115845_38	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:27:46	00:04:08	2G			01:10.897	1	1	mole042
19115845_38+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:27:46	00:04:08		1123052K	0	01:10.896	1	1	mole042
19115845_38+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:27:47	00:04:09		0	0	00:00:00	1	1	mole042
19115845_39	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:55	00:06:17	2G			01:21.747	1	1	mole043
19115845_39+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:55	00:06:17		1277004K	0	01:21.746	1	1	mole043
19115845_39+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:55	00:06:17		188K	0	00:00:00	1	1	mole043
19115845_40	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:23	00:06:45	2G			01:12.896	1	1	mole043
19115845_40+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:23	00:06:45		1081968K	0	01:12.895	1	1	mole043
19115845_40+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:23	00:06:45		188K	0	00:00:00	1	1	mole043
19115845_41	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:45	00:07:07	2G			01:16.999	1	1	mole043
19115845_41+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:45	00:07:07		1420688K	0	01:16.999	1	1	mole043
19115845_41+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:45	00:07:07		188K	0	00:00:00	1	1	mole043
19115845_42	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:57	00:06:19	2G			01:05.682	1	1	mole046
19115845_42+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:57	00:06:19		1240884K	0	01:05.681	1	1	mole046
19115845_42+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:58	00:06:20		0	0	00:00:00	1	1	mole046
19115845_43	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:23	00:06:45	2G			01:10.807	1	1	mole046
19115845_43+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:23	00:06:45		1194460K	0	01:10.807	1	1	mole046
19115845_43+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:23	00:06:45		0	0	00:00:00	1	1	mole046
19115845_44	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:21	00:06:43	2G			01:10.636	1	1	mole046
19115845_44+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:21	00:06:43		1299860K	0	01:10.636	1	1	mole046
19115845_44+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:21	00:06:43		0	0	00:00:00	1	1	mole046
19115845_45	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:57	00:07:19	2G			01:20.911	1	1	mole047
19115845_45+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:57	00:07:19		1819544K	0	01:20.910	1	1	mole047
19115845_45+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:30:57	00:07:19		192K	0	00:00:00	1	1	mole047
19115845_46	sam_my_job	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:47	00:06:09	2G			01:10.442	1	1	mole051
19115845_46+	batch	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:47	00:06:09		1356368K	0	01:10.441	1	1	mole051
19115845_46+	extern	COMPLETED	2024-09-06T02:23:38	2024-09-06T02:29:47	00:06:09		0	0	00:00:00	1	1	mole051

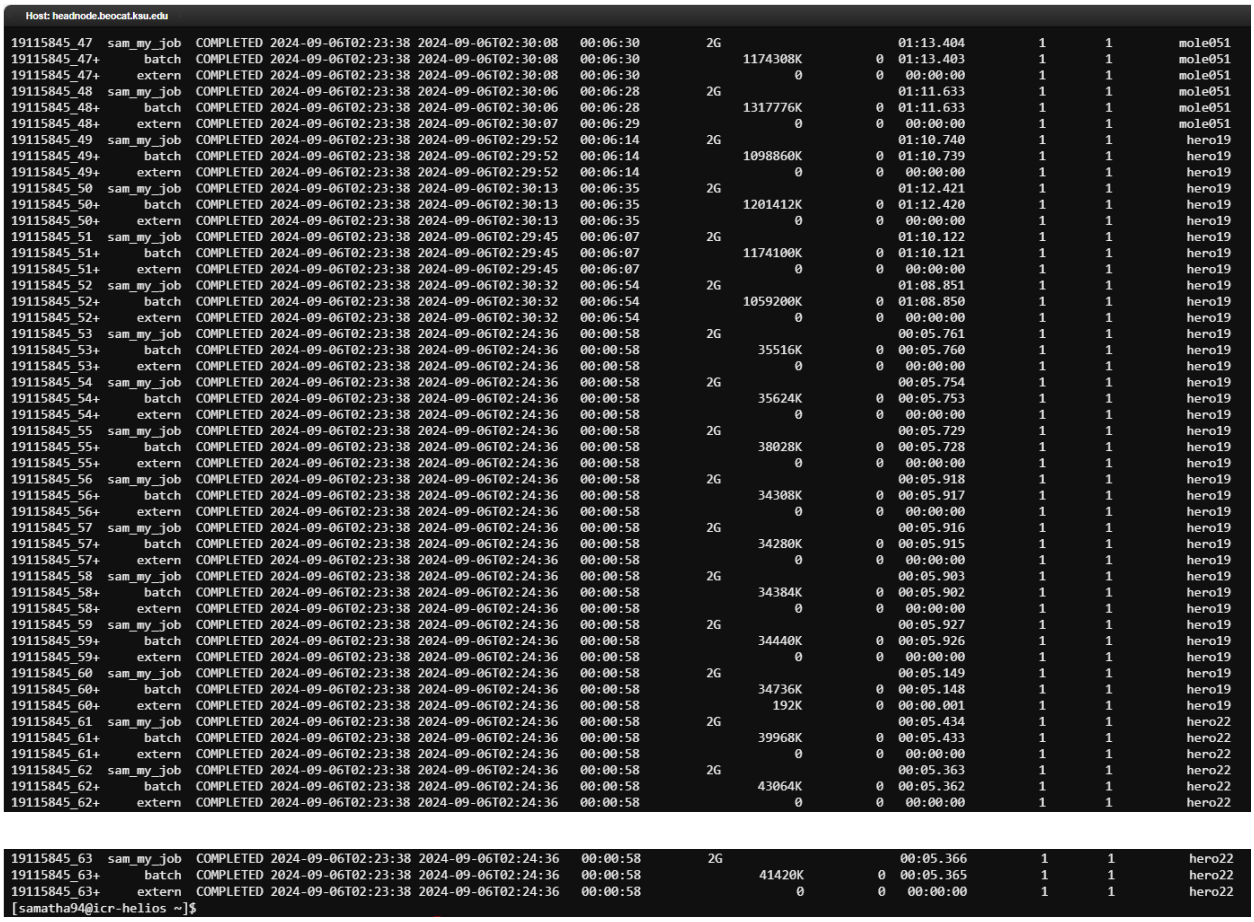


Figure 5.17: Job Execution Summary for Stage 3 – This figure displays five screenshots illustrating the SLURM job completion details for parallelized image retrieval and classification in Stage 3 for the third set of 10 scene categories.

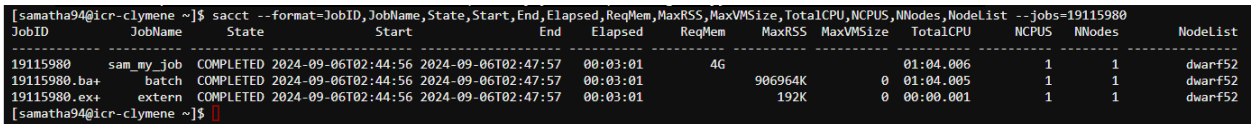


Figure 5.18: Job Execution Summary for Stage 4 – This figure displays the SLURM job completion details for statistical analysis and verification of neuron activations in Stage 4 for the third set of 10 scene categories.

The results show that the automation pipeline scales efficiently even with different sets of classes and varying image counts. The execution times for each stage remained consistent, even when processing different volumes of data. Stage 1 and Stage 2 saw slightly longer processing times due to the training and concept induction tasks, which are more dependent on the number

of neurons and available images. In contrast, Stages 3 and 4 demonstrated highly efficient performance, as these stages involved retrieving and classifying images, followed by statistical validation, tasks that benefit from parallelized execution.

The results confirm that the automation pipeline is capable of handling new image categories while maintaining computational efficiency across all stages. The accuracy results—76% for the first set and 88% for the second set—demonstrate the model's robustness and adaptability to new scene categories, ensuring that the system can be extended to additional image classes without significant increases in resource consumption or processing time.

Chapter 6 - Analysis of Multiple Neurons Activated by the Same ECII Concept

In the original Stage 4 process, statistical analysis was performed on a per-neuron basis, calculating activation metrics for individual neurons. This analysis represents a modification of the Stage 4 pipeline that extends its capabilities to accommodate the statistical analysis of combined neuron activations triggered by the same ECII concept. The core motivation behind this modification is to explore how multiple neurons, activated by the same ECII concept, collectively influence the label hypotheses generated in earlier stages. This extension allows for a more comprehensive evaluation of the consistency and robustness of neuron activations related to specific image categories or features.

6.1 Technical Requirements

The technical requirements for this analysis remain the same as described in Stage 4, involving the use of high-performance computing resources for parallelized processing. This includes 1 node per task, 4 GB of memory per task, and Python 3.11.5, with dependencies managed via a virtual environment. The job scheduling and resource management are handled by the SLURM workload manager.

6.2 Inputs and Data Setup

The inputs and data setup for this analysis are the same as in Stage 4. Neuron-specific evaluation and verification sets are used to calculate activation metrics for target and non-target images. Neuron IDs are loaded from the indices.txt file, and configuration files provide the

coverage scores. The input datasets are structured the same way, allowing for direct reuse of the neuron activation data and configuration files.

6.3 Process

The process in this modified analysis focuses on grouping neuron activations by ECII concepts and performing combined statistical analysis. The key steps are as follows:
Loading Neuron-Specific Activation Data: The pipeline begins by reading neuron-specific evaluation sets, and the activation data for each neuron is loaded separately. Neuron IDs are read from the indices.txt file, and individual neuron activation data is stored in the solution*_all_activations sheets, just like in the original Stage 4 process.

1. **Calculating Activation Metrics in the Evaluation Phase:** The neurons that are associated with the same ECII concept are grouped for the purpose of calculating activation metrics. The activation values for neurons that belong to the same ECII concept are pooled only during the metric calculation process, which computes combined target and non-target activation percentages for these groups. Additional metrics, such as activations exceeding thresholds of 20%, 40%, and 60%, are calculated for the pooled neuron activations.
2. **Filtering ECII Concepts Based on Combined Target Activation:** The ECII concepts with combined neuron target activation percentages exceeding 80% are selected for further analysis in the verification phase. The filtered neurons are saved in final_summary_* sheets, marking them as significant for the next stage of analysis.
3. **Statistical Analysis in the Verification Phase:** In the verification phase, the combined activations of neuron groups are subjected to statistical tests. The Mann-Whitney U test is

used to compare the combined target and non-target activations for each ECII concept group. The test produces z-scores and p-values, which are recorded in summary_* sheets. ECII concepts with p-values below 0.05 are marked as statistically significant, validating the combined activations of neurons for the given concept.

6.4 Outputs

The outputs generated in Stage 4 are stored in two main Excel workbooks: evaluation_combined.xlsx and verification_combined.xlsx (Figure 6.1). Each workbook contains a set of structured sheets summarizing the neuron activation data and statistical analysis results.

1. **Neuron Activation Evaluation Report:**
    - **The solution*_all_activations:** These sheets consolidate the activation data for each neuron across different solutions, showing activations for both target and non-target images. These sheets serve as the foundation for understanding neuron-specific activation patterns in the evaluation phase.
    - **summary_ sheets*:** These sheets display combined metrics for multiple neurons, including target and non-target activation percentages, coverage scores, and activation percentages above 20%, 40%, and 60%. Neurons are grouped by ECII concepts (Figure 6.2).
    - **The final_summary_ sheets*:** ECII concepts where combined target activations exceed 80% are highlighted, marking them for further statistical verification (Figure 6.3).
2. **Neuron Activation Verification Report:**

- **solution*_all_activations:** These sheets compile neuron-specific activation data from the verification phase, grouped by ECII concept.
    - **The summary_ sheets*:** These sheets include statistical metrics for the combined neurons, such as target and non-target means, medians, z-scores, and p-values. ECII concepts with statistically significant differences (p-value < 0.05) are highlighted to indicate robust activations (Figure 6.4).

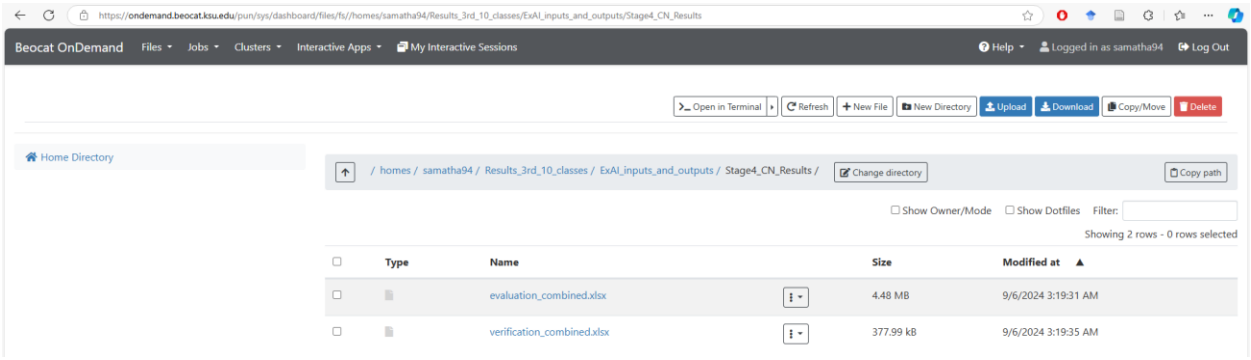
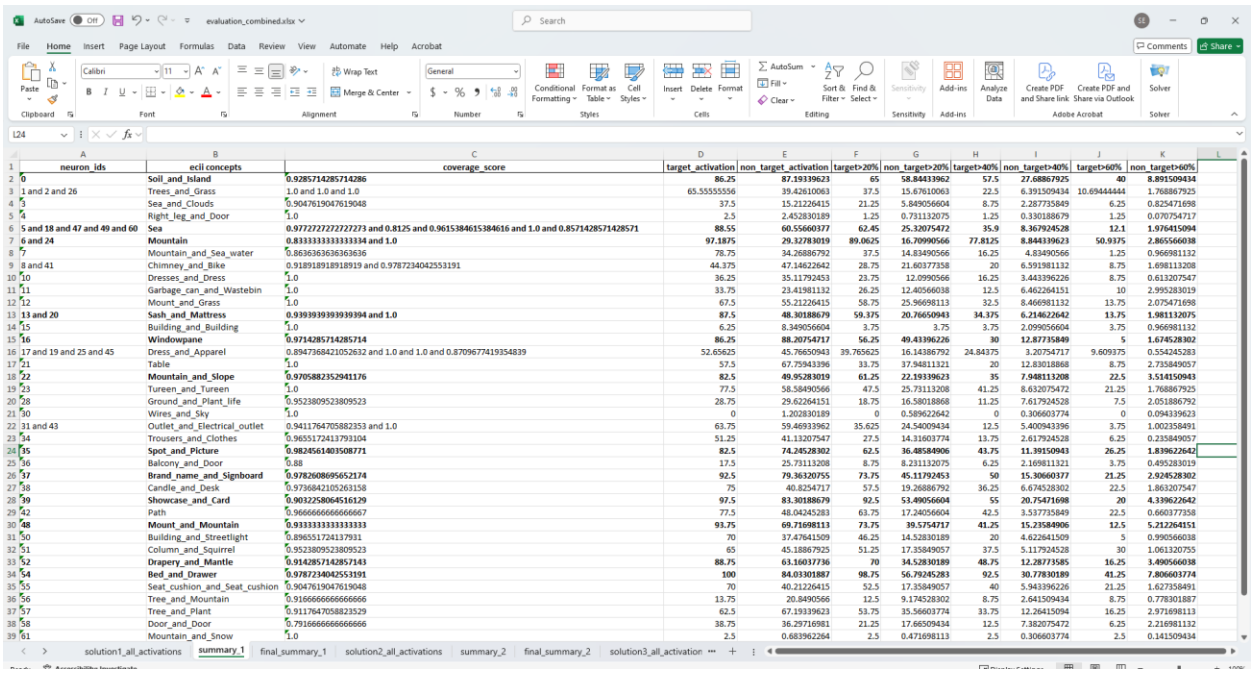
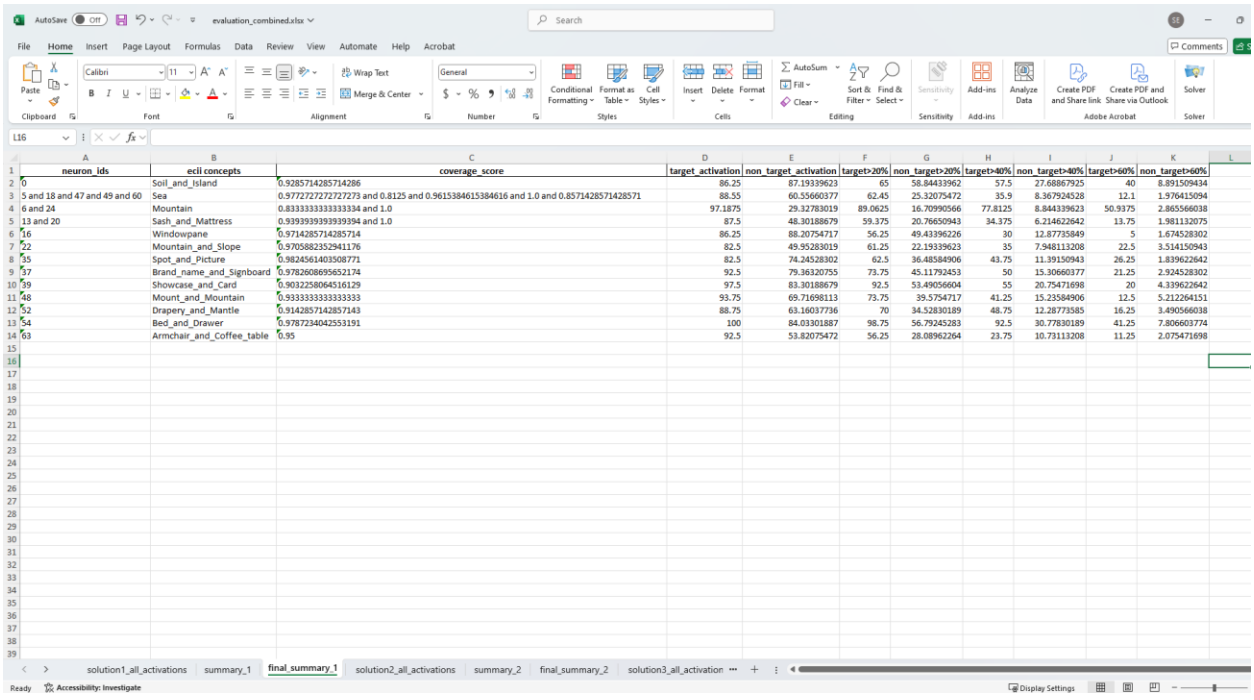


Figure 6.1: Results of Combined Neuron Activation analysis: This figure shows the two primary Excel workbooks generated during the combined neuron activation analysis: evaluation_combined.xlsx and verification_combined.xlsx. Each workbook contains structured sheets summarizing the consolidated activation metrics, grouped by ECII concepts, and the statistical analysis results, providing insights into the combined influence of multiple neurons on specific label hypotheses.



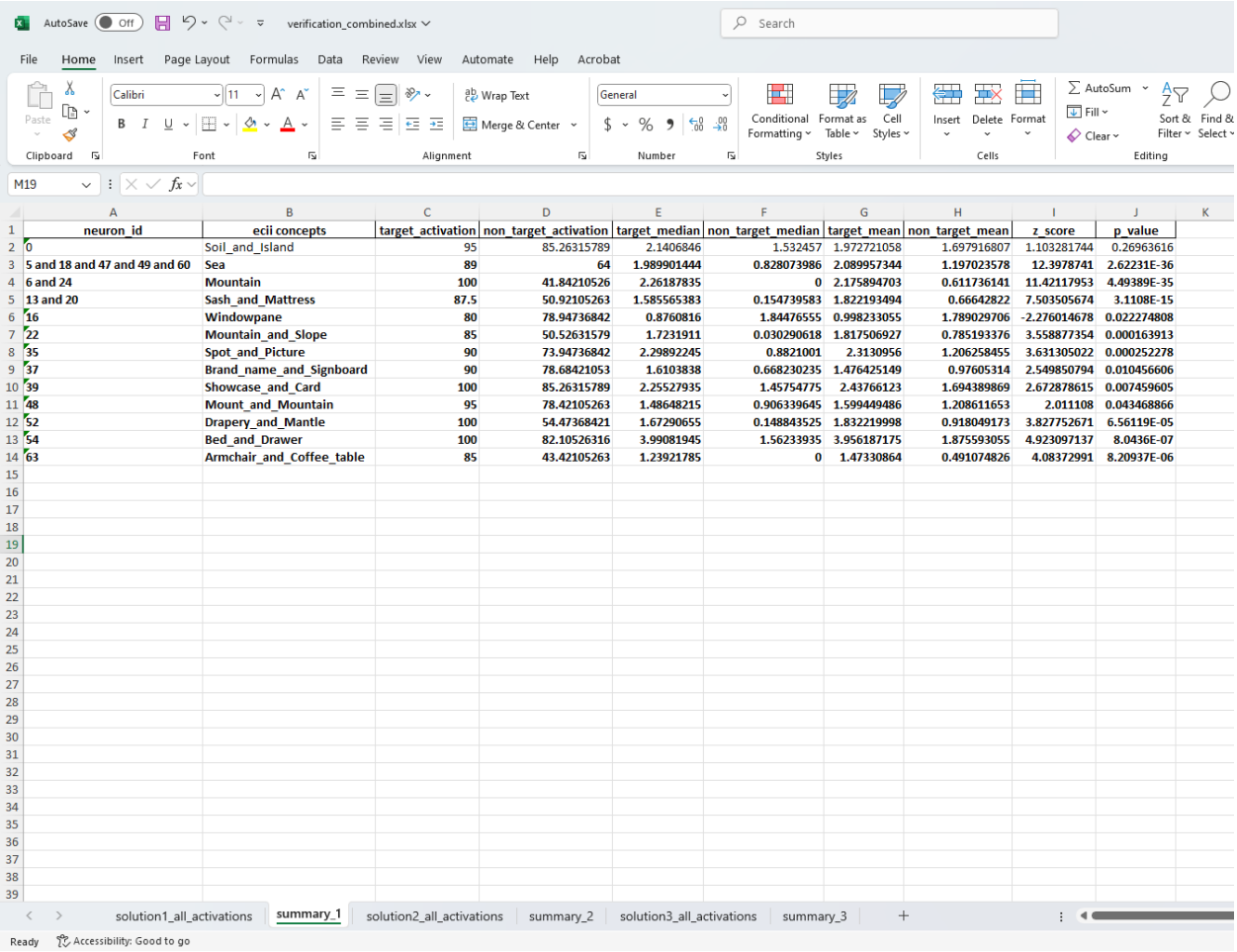
neuron_ids	ecii concepts	coverage_score	target_activation	non_target_activation	target>20%	non_target>20%	target>40%	non_target>40%	target>60%	non_target>60%
0	Soil_and_Island	0.928571428571486	86.25	87.19339623	65	58.84433962	57.5	27.68867925	40	8.891509434
1 and 2 and 26	Trees_and_Grass	1.0 and 1.0 and 1.0	65.55555556	39.42610063	37.5	15.67610063	22.5	6.391509434	10.69444444	1.768867925
5	Sea_and_Clouds	0.9047619047619048	37.5	15.21226415	21.25	5.849056604	8.75	2.287735849	6.25	0.825471698
4	Right_leg_and_Door	1.0	2.5	2.452830189	1.25	0.731132075	1.25	0.330188679	1.25	0.070754717
5 and 18 and 47 and 49 and 60	Sea	0.9772727272727273 and 0.8125 and 0.9615384615384616 and 1.0 and 0.8571428571428571	88.55	60.55660377	62.45	25.32075472	35.9	8.367924528	12.1	1.976415094
6 and 24	Mountain	0.8333333333333334 and 1.0	97.1875	29.32783019	89.0625	16.70990566	77.8125	8.844339623	50.9375	2.865566038
7	Mountain_and_Sea_water	0.8636363636363636	78.75	34.26886792	37.5	14.83490566	16.25	4.83490566	1.25	0.966981132
8 and 41	Chimney_and_Bike	0.9189189189189 and 0.9787234042553191	44.375	47.14622642	28.75	21.60377358	20	6.591981132	8.75	1.698113208
10	Dresses_and_Dress	1.0	36.25	35.11792453	23.75	12.0990566	16.25	3.443396226	8.75	0.613207547
11	Garbage_can_and_Wastebin	1.0	33.75	23.41981132	26.25	12.40566038	12.5	6.462264151	10	2.995283019
12	Mount_and_Grass	1.0	67.5	55.21226415	58.75	25.96698113	32.5	8.466981132	13.75	2.075471698
13 and 20	Sash_and_Mattress	0.9393939393939394 and 1.0	87.5	48.30188679	59.375	20.76650943	34.375	6.214622642	13.75	1.981132075
15	Building_and_Building	1.0	6.25	8.349056604	3.75	3.75	3.75	2.099056604	3.75	0.966981132
16	Windowpane	0.971428571428574	86.25	88.20754717	56.25	49.43396226	30	12.87735849	5	1.674528302
17 and 19 and 25 and 45	Dress_and_Apparel	0.8947368421052632 and 1.0 and 1.0 and 0.8709677419354839	52.65625	45.76650943	39.765625	16.14386792	24.84375	3.20754717	9.609375	0.554245283
21	Table	1.0	57.5	67.75943396	33.75	37.94811321	20	12.83018868	8.75	2.735849057
22	Mountain_and_Slope	0.9705882352941176	82.5	49.95283019	61.25	22.19339623	35	7.948113208	22.5	3.514150943
23	Tureen_and_Tureen	1.0	77.5	58.58490566	47.5	25.73113208	41.25	8.632075472	21.25	1.768867925
28	Ground_and_Plant_life	0.9523809523809523	28.75	29.62264151	18.75	16.58018868	11.25	7.617924528	7.5	2.051886792
30	Wires_and_Sky	1.0	0	1.202830189	0	0.589622642	0	0.306603774	0	0.094339623
31 and 43	Outlet_and_Electrical_outlet	0.9411764705882353 and 1.0	63.75	59.46933962	35.625	24.54009434	12.5	5.400943396	3.75	1.002358491
54	Trousers_and_Clothes	0.9655172413793104	51.25	41.13207547	27.5	14.31603774	13.75	2.617924528	6.25	0.235849057
35	Spot_and_Picture	0.9824561403508771	82.5	74.24528302	62.5	36.48584906	43.75	11.39150943	26.25	1.839622642
36	Balcony_and_Door	0.88	17.5	25.73113208	8.75	8.231132075	6.25	2.169811321	3.75	0.495283019
37	Brand_name_and_Signboard	0.9782608695652174	92.5	79.36320755	73.75	45.11792453	50	15.30660377	21.25	2.924528302
38	Candle_and_Desk	0.9736842105263158	75	40.8254717	57.5	19.26886792	36.25	6.674528302	22.5	1.863207547
39	Showcase_and_Card	0.9032258064516129	97.5	83.30188679	92.5	53.49056604	55	20.75471698	20	4.339622642
42	Path	0.9666666666666667	77.5	48.04245283	63.75	17.24056604	42.5	3.537735849	22.5	0.660377358
48	Mount_and_Mountain	0.9333333333333333	93.75	69.71698113	73.75	39.5754717	41.25	15.23584906	12.5	5.212264151
50	Building_and_Streetlight	0.896551724137931	70	37.47641509	46.25	14.52830189	20	4.622641509	5	0.990566038
51	Column_and_Squirrel	0.9523809523809523	65	45.18867925	51.25	17.35849057	37.5	5.117924528	30	1.061320755
52	Drapery_and_Mantle	0.9142857142857143	88.75	63.16037736	70	34.52830189	48.75	12.28773585	16.25	3.490566038
54	Bed_and_Drawer	0.9787234042553191	100	84.03301887	98.75	56.79245283	92.5	30.77830189	41.25	7.806603774
55	Seat_cushion_and_Seat_cushion	0.9047619047619048	70	40.21226415	52.5	17.35849057	40	5.943396226	21.25	1.627358491
56	Tree_and_Mountain	0.9166666666666666	13.75	20.8490566	12.5	9.174528302	8.75	2.641509434	8.75	0.778301887
57	Tree_and_Plant	0.9117647058823529	62.5	67.19339623	53.75	35.56603774	33.75	12.26415094	16.25	2.971698113
58	Door_and_Door	0.7916666666666666	38.75	36.2971698	21.25	17.66509434	12.5	7.382075472	6.25	2.216981132
61	Mountain_and_Snow	1.0	2.5	0.683962264	2.5	0.471698113	2.5	0.306603774	2.5	0.141509434

Figure 6.2: Combined Neuron Activation Metrics in the Evaluation Phase: This figure displays the calculated activation metrics for groups of neurons associated with the same ECII concept during the evaluation phase.



neuron_ids	ecii concepts	coverage_score	target_activation	non_target_activation	target>20%	non_target>20%	target>40%	non_target>40%	target>60%	non_target>60%
0	Soil_and_Island	0.928571428571486	86.25	87.19339623	65	58.84433962	57.5	27.68867925	40	8.891509434
5 and 18 and 47 and 49 and 60	Sea	0.9772727272727273 and 0.8125 and 0.9615384615384616 and 1.0 and 0.8571428571428571	88.55	60.55660377	62.45	25.32075472	35.9	8.367924528	12.1	1.976415094
6 and 24	Mountain	0.8333333333333334 and 1.0	97.1875	29.32783019	89.0625	16.70990566	77.8125	8.844339623	50.9375	2.865566038
13 and 20	Sash_and_Mattress	0.9393939393939394 and 1.0	87.5	48.30188679	59.375	20.76650943	34.375	6.214622642	13.75	1.981132075
16	Windowpane	0.971428571428574	86.25	88.20754717	56.25	49.43396226	30	12.87735849	5	1.674528302
22	Mountain_and_Slope	0.9705882352941176	82.5	49.95283019	61.25	22.19339623	35	7.948113208	22.5	3.514150943
35	Spot_and_Picture	0.9824561403508771	82.5	74.24528302	62.5	36.48584906	43.75	11.39150943	26.25	1.839622642
37	Brand_name_and_Signboard	0.9782608695652174	92.5	79.36320755	73.75	45.11792453	50	15.30660377	21.25	2.924528302
39	Showcase_and_Card	0.9032258064516129	97.5	83.30188679	92.5	53.49056604	55	20.75471698	20	4.339622642
48	Mount_and_Mountain	0.9333333333333333	93.75	69.71698113	73.75	39.5754717	41.25	15.23584906	12.5	5.212264151
52	Drapery_and_Mantle	0.9142857142857143	88.75	63.16037736	70	34.52830189	48.75	12.28773585	16.25	3.490566038
54	Bed_and_Drawer	0.9787234042553191	100	84.03301887	98.75	56.79245283	92.5	30.77830189	41.25	7.806603774
63	Armchair_and_Coffee_table	0.95	92.5	53.82075472	56.25	28.08962264	23.75	10.73113208	11.25	2.075471698

Figure 6.3: Selection of ECII Concepts for Statistical Verification: This figure shows the process of filtering ECII concepts based on combined target activation exceeding 80%. The concepts are highlighted for further statistical verification, indicating neurons with robust target activations.



neuron_id	ecii concepts	target_activation	non_target_activation	target_median	non_target_median	target_mean	non_target_mean	z_score	p_value
0	Soil_and_Island	95	85.26315789	2.1406846	1.532457	1.972721058	1.697916807	1.103281744	0.26963616
5 and 18 and 47 and 49 and 60	Sea	89	64	1.989901444	0.828073986	2.089957344	1.197023578	12.3978741	2.62231E-36
6 and 24	Mountain	100	41.84210526	2.26187835	0	2.175894703	0.611736141	11.42117953	4.49389E-35
13 and 20	Sash_and_Mattress	87.5	50.92105263	1.585565383	0.154739583	1.822193494	0.66642822	7.503505674	3.1108E-15
16	Windowpane	80	78.94736842	0.8760816	1.84476555	0.998233055	1.789029706	-2.276014678	0.022274808
22	Mountain_and_Slope	85	50.52631579	1.7231911	0.030290618	1.817506927	0.785193376	3.558877354	0.000163913
35	Spot_and_Picture	90	73.94736842	2.29892245	0.8821001	2.3130956	1.206258455	3.631305022	0.000252278
37	Brand_name_and_Signboard	90	78.68421053	1.6103838	0.668230235	1.476425149	0.97605314	2.549850794	0.010456606
39	Showcase_and_Card	100	85.26315789	2.25527935	1.45754775	2.43766123	1.694389869	2.672878615	0.007459605
48	Mount_and_Mountain	95	78.42105263	1.48648215	0.906339645	1.599449486	1.208611653	2.011108	0.043468866
52	Drapery_and_Mantle	100	54.47368421	1.67290655	0.148843525	1.832219998	0.918049173	3.827752671	6.56119E-05
54	Bed_and_Drawer	100	82.10526316	3.99081945	1.56233935	3.956187175	1.875593055	4.923097137	8.0436E-07
63	Armchair_and_Coffee_table	85	43.42105263	1.23921785	0	1.47330864	0.491074826	4.08372991	8.20937E-06

Figure 6.4: Summary of Statistical Analysis for Combined Neurons in the Verification Phase: This figure summarizes the statistical analysis of combined neuron activations in the verification phase. The results include target and non-target means, medians, z-scores, and p-values, highlighting ECII concepts with statistically significant differences in neuron activations.

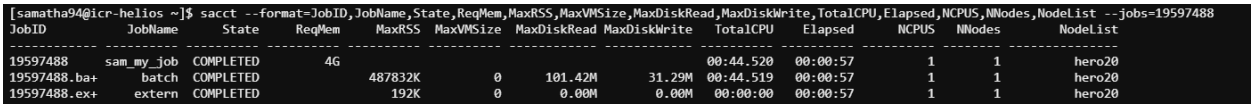
6.5 Results

The statistical analysis of combined neuron activations for the same ECII concept was executed efficiently through the Beocat high-performance computing environment. This stage was focused on verifying label hypotheses by pooling activations from multiple neurons

associated with the same ECII concept and performing statistical comparisons between target and non-target images. The following are key metrics obtained during the job execution:

1. **Job Completion Status:** The job was successfully completed with JobID 19597488 (Figure 6.5), confirming that the combined neuron-specific statistical analysis tasks were executed without any errors.
2. **Total CPU Time:** The total CPU time used for the job was approximately 44.52 seconds, indicating the efficient processing of pooled neuron activation data and statistical calculations for each ECII concept.
3. **Elapsed Time:** The job's total elapsed time was 57 seconds, reflecting the quick turnaround time required for executing the entire statistical analysis and verification of pooled neuron activations across ECII concepts.
4. **Memory Usage:** The maximum resident set size (MaxRSS) during job execution was approximately 487 MB, while the total requested memory for the job was 4 GB.
5. **Node Allocation:** The job was executed on a single node (hero20), utilizing one CPU core. Exclusive access to this node ensured that the statistical analysis was performed without interference from other processes, allowing efficient use of computational resources.
6. **Efficiency Gains:** The use of Beocat's high-performance computing resources enabled the rapid execution of statistical analysis for multiple neurons associated with the same ECII concept. The entire job was completed in under 1 minute, showcasing the efficiency of the automation pipeline in handling large-scale, combined neuron-specific statistical analysis. The Mann-Whitney U test, applied to the pooled activations, validated the

neuron activations for each ECII concept, confirming that the target activation metrics were consistent with the hypothesized labels.



```
[samatha94@icr-helios ~]$ sacct --format=JobID,JobName,State,ReqMem,MaxRSS,MaxVMSize,MaxDiskRead,MaxDiskWrite,TotalCPU,Elapsed,NCPUS,NNodes,Nodelist --jobs=19597488
JobID           JobName      State     ReqMem     MaxRSS  MaxVMSize  MaxDiskRead MaxDiskWrite   TotalCPU    Elapsed      NCPUS   NNodes        NodeList
------------ ---------- ---------- ---------- ---------- ---------- ------------ ------------ ---------- ---------- ---------- -------- ---------------
19597488     sam_my_job  COMPLETED         4G                                                  00:44.520   00:00:57          1        1          hero20
19597488.ba+      batch  COMPLETED               487832K          0      101.42M       31.29M  00:44.519   00:00:57          1        1          hero20
19597488.ex+     extern  COMPLETED                 192K          0        0.00M        0.00M   00:00:00   00:00:57          1        1          hero20
```

Figure 6.5: Job Execution Summary for Combined Neuron Activation Analysis in Stage 4 – This figure displays the SLURM job completion details for the combined neuron activation analysis and verification of activations in Stage 4.

The streamlined process ensured that the label hypotheses for ECII concept groups were verified with minimal resource consumption, validating the robustness of the neuron activations through statistically significant comparisons. These results underscore the system's capability to handle large-scale analysis of pooled neuron activations, providing efficient and accurate verification of neuron-label hypotheses across ECII concepts with minimal resource consumption.

Chapter 7 - Conclusion

This project successfully developed and implemented an automated, scalable pipeline for neuron-specific concept induction, label hypothesis generation, and statistical validation of neuron activations. By leveraging the Beocat high-performance computing cluster, we were able to streamline the entire process, dramatically reducing execution times across all stages without compromising the depth and rigor of the analysis.

Stage 1 focused on training the CNN model and configuring the data, achieving completion in 38 minutes and 50 seconds across 24 CPU cores. This stage laid the groundwork for neuron-specific tasks by providing a well-trained model to interpret neuron activations. The rapid model training demonstrated the efficiency of the parallelized system, which made optimal use of available computational resources.

In Stage 2, we applied parallelized concept induction and label hypothesis generation across 64 neurons, reducing what would have been a 10-hour task to just 17 minutes. This stage illustrated the pipeline's capacity to handle large-scale tasks efficiently, ensuring high throughput without any bottlenecks or memory constraints.

Stage 3 further demonstrated the power of parallelization through automated image retrieval and classification. This stage transformed what would have been a 16-hour sequential process into one that completed in just 10 minutes. The effective distribution of tasks across multiple nodes, combined with efficient memory management, played a critical role in achieving these results. This stage also highlighted the importance of integrating external data sources in the context of neural network interpretability.

Finally, Stage 4 validated the hypotheses generated in the earlier stages by performing statistical analysis on neuron activations. The stage, which completed in just under 3 minutes,

used statistical tests, including the Mann-Whitney U test, to confirm the reliability and consistency of neuron activations for both target and non-target images. The results confirmed that the concept induction with structured background knowledge produced meaningful labels that robustly explained neuron activation patterns. The automated approach ensured both speed and accuracy, allowing for rapid hypothesis testing and validation.

1. **Scalability and Performance Evaluation on New Scene Categories:** To further evaluate the scalability and performance of the automation pipeline, we applied it to 20 new scene categories from the ADE20K dataset, split into two sets of 10 classes each. The pipeline performed consistently across these new categories, handling 1,197 images in the first set and 822 images in the second. The execution times remained efficient, with Stage 1 completing in approximately 12 minutes for each set, Stage 2 ranging from 15 to 18 minutes, Stage 3 varying between 3 and 8 minutes, and Stage 4 consistently under 3 minutes. Additionally, the model's classification accuracy across these new categories was satisfactory, achieving an overall accuracy of 76% for the first set and 88% for the second. These results confirm the robustness of the pipeline when applied to new and diverse image categories, demonstrating its scalability and flexibility in handling varying datasets without significant increases in resource consumption or time.
2. **Analysis of Multiple Neurons Activated by the Same ECII Concept:** In addition to scalability, we extended the analysis by pooling activations from multiple neurons associated with the same ECII concept. By combining neuron activations that correspond to the same concept, we were able to provide a more comprehensive statistical validation. The results of this analysis demonstrated that pooled neuron activations were consistent with the hypothesized labels, as validated through statistical tests such as the Mann-

Whitney U test. The entire analysis was completed in under 1 minute, reflecting the pipeline's efficiency in handling more complex neuron grouping scenarios. This approach enhances the robustness of the label hypotheses by ensuring that multiple neurons can be validated collectively, providing deeper insights into the shared activations of neurons under the same concept.

Overall, the initial four stages of the pipeline were completed in approximately 1 hour and 46 minutes. Including the two additional analyses, the entire pipeline was executed efficiently, demonstrating substantial improvements in performance and scalability. The combination of automation and high-performance computing has proven to be a robust framework for large-scale neuron-specific analyses, reducing execution times across all stages while maintaining computational efficiency and reproducibility. By automating concept induction, image retrieval, and statistical validation, the system ensures scalability and adaptability to new datasets and more complex neuron activation scenarios.

These advances contribute significantly to the broader field of neural network interpretability, providing a solid foundation for future research into neuron activation patterns and their real-world applications in AI-driven systems. The project’s results demonstrate that high-performance computing, combined with structured knowledge and scalable methodologies, can yield meaningful insights into neural network behavior, pushing the boundaries of explainable AI.

Chapter 8 - Limitations

While the automation of neuron-specific concept induction, image classification, and statistical analysis provided considerable efficiency gains, several limitations arise concerning the current level of automation implemented in the pipeline:

1. **Limited Flexibility in Customization:** The automated pipeline is designed around a highly structured workflow with predefined stages and tasks, limiting the flexibility to accommodate more dynamic or exploratory tasks. For example, users must adhere to the predefined neuron sets and labeled datasets. In situations where researchers need to explore alternative hypotheses or inject custom workflows—such as varying model architectures, using different statistical tests, or handling unstructured data—manual intervention is often required.
2. **Dependency on Predefined Workflow:** The pipeline's automation is tailored to a specific sequence of operations: loading neuron-specific datasets, performing model training, conducting neuron activation analysis, and validating results using statistical methods. While this highly structured process benefits efficiency and reproducibility, it limits the pipeline's flexibility. Changes to one stage (e.g., using a different statistical method or varying input formats) require manual adjustment to other stages, reducing the overall adaptability of the system. Automating more dynamic workflows where processes can be updated or replaced seamlessly would provide greater versatility.
3. **Resource Allocation Challenges:** The current automation system assumes static resource allocation for each task based on predefined memory and CPU limits. While this works effectively for the current dataset and task sizes, scaling to more complex tasks, larger datasets, or deeper neural networks could lead to inefficient use of computational

resources. For instance, memory and CPU allocations might either be under-utilized (leading to wasted resources) or insufficient (leading to performance degradation or job failure). Automating dynamic resource scaling based on real-time task demands is an area where the pipeline could be enhanced to improve efficiency further.

Chapter 9 - Future Work

To address these limitations and enhance the pipeline's capabilities, several avenues for future research and development can be explored:

1. **Dynamic Workflow Adaptation:** Implement a more flexible automation framework that allows dynamic adjustments to workflows based on intermediate results, error conditions, or evolving data inputs.
2. **Automated Resource Scaling:** Introduce intelligent resource allocation mechanisms that adjust memory, CPU, and other resources in real time based on the specific needs of each task, optimizing for both performance and efficiency.
3. **Expanded Label Hypothesis Generation:** Currently, the label hypothesis generation process is limited to the top three solutions for each neuron. Future iterations could explore incorporating a variable number of top solutions, allowing for more diverse and comprehensive analyses of neuron activations. Moreover, integrating advanced evaluation metrics—such as precision, recall, and F-measure alongside the coverage score—would provide a more robust measure of label validity, improving the overall quality of the neuron-label hypotheses.
4. **Enhanced Image Retrieval and Classification:** The pipeline's image retrieval process was restricted to 100 images per keyword. Expanding the number of images retrieved per keyword could improve the quality and depth of the validation sets, allowing for a more thorough evaluation of neuron activations. Further, investigating the use of more advanced neural network architectures, such as EfficientNet or Transformer-based models, could significantly improve the performance and accuracy of the image classification task.

5. **Incorporation of Alternative Neural Architectures:** While the ResNet50V2 model demonstrated strong performance in classifying neuron activations, future studies should evaluate the effectiveness of alternative neural architectures. Models such as EfficientNet, Vision Transformers (ViT), or hybrid architectures could be tested to further improve classification accuracy and interpretability of neuron activations.

By addressing these limitations and incorporating the suggested enhancements, the pipeline can evolve into a more powerful, adaptable, and efficient tool for large-scale neuron-specific analysis, expanding its applicability in computational neuroscience and AI interpretability.

References

[1] A. Adadi and M. Berrada, “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI),” *IEEE Access*, vol. 6, pp. 52138–52160, 2018, doi: 10.1109/ACCESS.2018.2870052.

[2] M. van Lent, W. Fisher, and M. Mancuso, “An Explainable Artificial Intelligence System for Small-unit Tactical Behavior”.

[3] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang, “XAI—Explainable artificial intelligence,” *Sci. Robot.*, vol. 4, no. 37, p. eaay7120, Dec. 2019, doi: 10.1126/scirobotics.aay7120.

[4] “Explainable Artificial Intelligence.” Accessed: Sep. 07, 2024. [Online]. Available: https://www.darpa.mil/program/explainable-artificial-intelligence

[5] A. Dalal, M. K. Sarker, A. Barua, E. Vasserman, and P. Hitzler, “Understanding CNN Hidden Neuron Activations Using Structured Background Knowledge and Deductive Reasoning,” Aug. 09, 2023, *arXiv*: arXiv:2308.03999. Accessed: Oct. 21, 2024. [Online]. Available: http://arxiv.org/abs/2308.03999

[6] J. D. Kelleher, *Deep learning*. in The MIT press essential knowledge series. Cambridge, Massachusetts: The MIT Press, 2019.

[7] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[8] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. in Adaptive computation and machine learning. Cambridge, Mass: The MIT press, 2016.

[9] A. Shrestha and A. Mahmood, “Review of Deep Learning Algorithms and Architectures,” *IEEE Access*, vol. 7, pp. 53040–53065, 2019, doi: 10.1109/ACCESS.2019.2912200.

[10] M. Nielsen, “Neural Networks and Deep Learning”.

[11] J. Gu *et al.*, “Recent advances in convolutional neural networks,” *Pattern Recognition*, vol. 77, pp. 354–377, May 2018, doi: 10.1016/j.patcog.2017.10.013.

[12] Y. LeCun *et al.*, “Handwritten Digit Recognition with a Back-Propagation Network,” in *Advances in Neural Information Processing Systems*, Morgan-Kaufmann, 1989. Accessed: Oct. 05, 2024. [Online]. Available: https://proceedings.neurips.cc/paper/1989/hash/53c3bce66e43be4f209556518c2fcb54-Abstract.html

[13] L. Zhao and Z. Zhang, “A improved pooling method for convolutional neural networks,” *Sci Rep*, vol. 14, no. 1, p. 1589, Jan. 2024, doi: 10.1038/s41598-024-51258-6.

[14] “What is Semantic Reasoning?” Accessed: Oct. 19, 2024. [Online]. Available: https://www.oxfordsemantic.tech/faqs/what-is-semantic-reasoning

[15] M. K. Sarker *et al.*, “Wikipedia Knowledge Graph for Explainable AI,” in *Knowledge Graphs and Semantic Web*, vol. 1232, B. Villazón-Terrazas, F. Ortiz-Rodríguez, S. M. Tiwari, and S. K. Shandilya, Eds., in Communications in Computer and Information Science, vol. 1232. , Cham: Springer International Publishing, 2020, pp. 72–87. doi: 10.1007/978-3-030-65384-2_6.

[16] K. Li, Y. Zhang, K. Li, Y. Li, and Y. Fu, “Visual Semantic Reasoning for Image-Text Matching,” in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South): IEEE, Oct. 2019, pp. 4653–4661. doi: 10.1109/ICCV.2019.00475.

[17] A. Dalal, M. K. Sarker, A. Barua, and P. Hitzler, “Explaining Deep Learning Hidden Neuron Activations using Concept Induction,” Jan. 23, 2023, *arXiv*: arXiv:2301.09611. Accessed: Oct. 19, 2024. [Online]. Available: http://arxiv.org/abs/2301.09611

[18] A. Dalal, R. Rayan, A. Barua, E. Y. Vasserman, M. K. Sarker, and P. Hitzler, “On the Value of Labeled Data and Symbolic Methods for Hidden Neuron Activation Analysis,” in *Neural-Symbolic Learning and Reasoning*, vol. 14980, T. R. Besold, A. d’Avila Garcez, E. Jimenez-Ruiz, R. Confalonieri, P. Madhyastha, and B. Wagner, Eds., in Lecture Notes in Computer Science, vol. 14980. , Cham: Springer Nature Switzerland, 2024, pp. 109–131. doi: 10.1007/978-3-031-71170-1_12.

[19] A. Dalal, R. Rayan, and P. Hitzler, “Error-Margin Analysis for Hidden Neuron Activation Labels,” in *Neural-Symbolic Learning and Reasoning*, vol. 14980, T. R. Besold, A. d’Avila Garcez, E. Jimenez-Ruiz, R. Confalonieri, P. Madhyastha, and B. Wagner, Eds., in Lecture Notes in Computer Science, vol. 14980. , Cham: Springer Nature Switzerland, 2024, pp. 149–164. doi: 10.1007/978-3-031-71170-1_14.

[20] B. Zhou, H. Zhao, X. Puig, T. Xiao, S. Fidler, A. Barriuso, A. Torralba, Semantic understanding of scenes through the ade20k dataset, International Journal of Computer Vision 127 (2019) 302–321. doi:10.1007/s11263-018-1140-0.

[21] Sarker MK, Hitzler P. Efficient Concept Induction for Description Logics. AAAI. 2019 Jul 17;33(01):3036–43.

[22] K. Hutson, D. Andresen, A. Tygart, and D. Turner, “Managing a Heterogeneous Cluster,” in Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning), Chicago IL USA: ACM, Jul. 2019, pp. 1–6. doi: 10.1145/3332186.3332251.

[23] S. Hutchison, D. Andresen, M. Neilsen, W. Hsu, and B. Parsons, “High Performance Computing Queue Time Prediction Using Clustering and Regression,” in Parallel Processing and Applied Mathematics, vol. 13827, R. Wyrzykowski, J. Dongarra, E. Deelman, and K. Karczewski, Eds., in Lecture Notes in Computer Science, vol. 13827. , Cham: Springer International Publishing, 2023, pp. 260–272. doi: 10.1007/978-3-031-30445-3_22.