

Narrative Report | [Phase 2] Final Project

By: Kein Jake A. Culanggo

This phase centered on gathering and preparing the data required for developing and testing the ASL recognition system. Two main datasets were assembled: a static image dataset for model training and a custom video dataset for initial testing and validation of data suitability.

The primary training dataset was sourced from Kaggle, which contained labeled images of American Sign Language letters organized into class-specific folders. This structure allowed for consistent preprocessing and automatic label assignment. The dataset included minor class imbalances, such as the letter t having fewer samples, but these inconsistencies were small and manageable. Since the images came from varied lighting conditions and hand positions, they provided a diverse base for training and helped support generalization.

Alongside the Kaggle dataset, the group created a supplemental video dataset intended for testing. Each member recorded short clips of themselves spelling their names in ASL, providing a realistic sample of how the system might be used in practice. These videos helped evaluate whether the training data aligned well with real-world samples. Although this process involved observing how the system responded to the recorded gestures, it was conducted only to check the suitability of our collected data. It was not part of Phase 3 model experimentation. This early validation was necessary because even if a dataset performs well on paper, it may not translate effectively to real-life input.

The team then observed that the initial recordings did not match well with the visual characteristics of the Kaggle images. This led us to revisit how the videos were captured. We repeated recordings using clearer backgrounds and minimized arm visibility, since many images in the Kaggle dataset showed only the hand. These adjustments were made to ensure that the testing data was consistent and properly aligned with the visual patterns learned during training. This process confirmed that dataset characteristics played a significant role in model performance, which guided our decision to explore additional datasets later on.

Once the datasets were finalized, a standard preprocessing pipeline was applied. All images and extracted video frames were converted to a consistent size and format to ensure compatibility with the models to be used in the next phase. Frames from the video recordings were extracted at a fixed rate and processed in the same manner as the training images, allowing both datasets to share a uniform structure. Labeling was handled automatically through directory organization, enabling reproducible and efficient preparation.

Throughout this process, individual team roles supported the data pipeline. Dela Cruz handled the implementation of preprocessing procedures and ensured that the datasets followed a consistent structure. Abainza provided equipment and assisted with GPU requirements needed for data preparation tasks. Casino oversaw the correctness of the video recordings and ensured that the testing samples followed the intended guidelines. I contributed by monitoring the coherence between the training and testing datasets, participating in the recording process, and refining documentation to maintain a clear and organized report of the procedures.

Overall, Phase 2 established a complete and well-structured data foundation for the project. The combination of standardized static images and carefully prepared video samples ensured that the

following phase would begin with datasets that were clean, organized, and suitable for systematic experimentation.