**ASL Gesture Recognition Using CNNs for Real-Time Translation**

**19.11.2024**

**Applied Machine Learning**

**Aiymbay Sunggat**

**Madiyar Abatov BDA-2206**

**Zhantore Darmenov BDA-2205**

**Astana IT University**

**Overview**

This project involves classifying American Sign Language (ASL) alphabet signs using deep learning models. It utilizes a pretrained MobileNetV2 architecture and includes dataset preparation, training, evaluation, and iterative improvements. The final model is deployed using Streamlit for real-world usability.

**Requirements Addressed**

**Dataset Details:**

* Source: ASL Alphabet Dataset from Kaggle.
* Classes: 29 classes, including letters A-Z and "space," "nothing," "del."
* Preprocessing:
* Images resized to 224x224.
* Normalization applied.
* Data Split:
* Training: 178,447 samples (80%).
* Validation: 44,627 samples (20%).
* Test: 28 samples (1 per class, for simplicity).

**Model Training Process:**

* Base Model: MobileNetV2 pretrained on ImageNet.
* Training Configuration:
* Loss function: Sparse categorical cross-entropy.
* Optimizer: Adam with a learning rate of 0.0001.
* Epochs: 10 for initial training; 2 for fine-tuning.
* Batch size: 32.
* Improvements:
* Fine-tuned last 10 layers of the MobileNetV2 base.
* Added regularization via learning rate reduction.

**Metrics Documented:**

* Training accuracy: 82.16% after initial training.
* Validation accuracy: 93.53% after initial training.
* Test accuracy: 92.86% after fine-tuning.
* Loss values and learning rate reductions documented.

**Deployment:**

* Final model deployed using Streamlit.
* Users can upload and classify hand sign images via an intuitive interface.

**Code Submission:**

Project code hosted on GitHub, including a README file for setup.

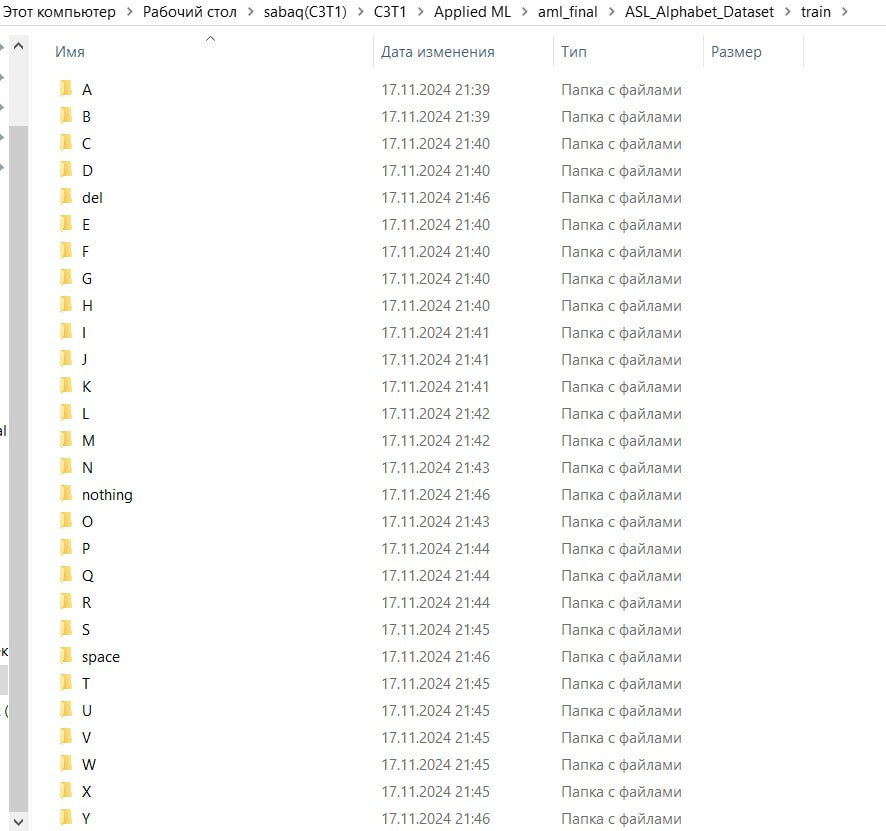
**Dataset Preparation**

**Structure:**

Images organized into folders named after class labels (e.g., "A", "B", ..., "space").

**Visualization:**

Samples visualized for data distribution and integrity.



*Folder structure of data*

**Model Training**

**Base Model:**

* Architecture: MobileNetV2 with modified classification layers for 29 classes.
* Configuration:
* Optimizer: Adam (learning rate = 0.0001).
* Loss Function: Sparse categorical cross-entropy.
* Metrics: Accuracy, loss.

**Fine-Tuning:**

* Adjusted learning rate to 0.00001
* Improved validation accuracy: 95.40%
* Reduced overfitting by freezing fewer layers in the pretrained model.

**Results**

**Classification Metrics:**

* Test Accuracy: 92.86%
* Confusion matrix highlights minor misclassifications in similar gestures.
* Validation loss consistently decreased: 0.22 after 10 epochs.

**Visualization:**

Confusion matrices and random predictions illustrate performance.

**Deployment**

**Streamlit Application:**

* Features:
* Upload images to classify ASL alphabet signs.
* Displays predicted class with confidence scores.
* Usage Instructions:

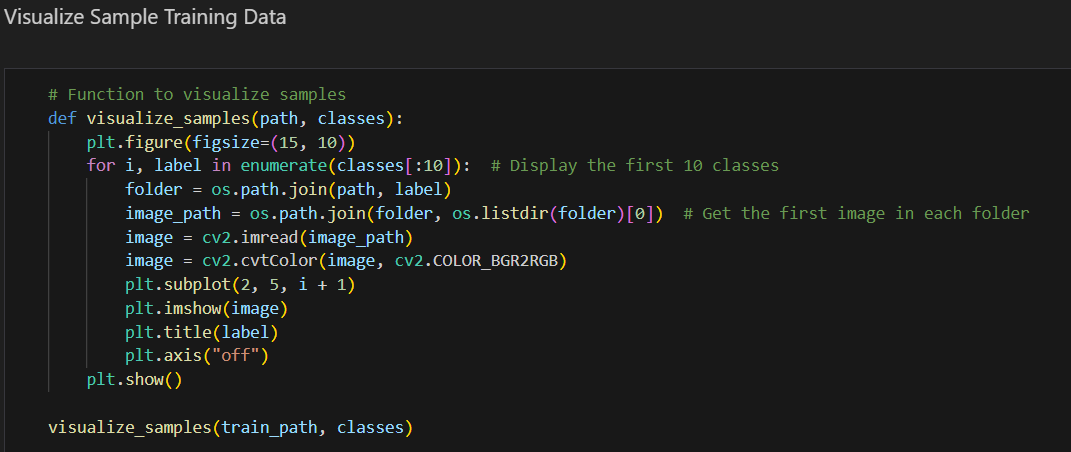
1. Clone the GitHub repository.
2. Install dependencies using `pip install -r requirements.txt`.
3. Run the app with `streamlit run app.py`.

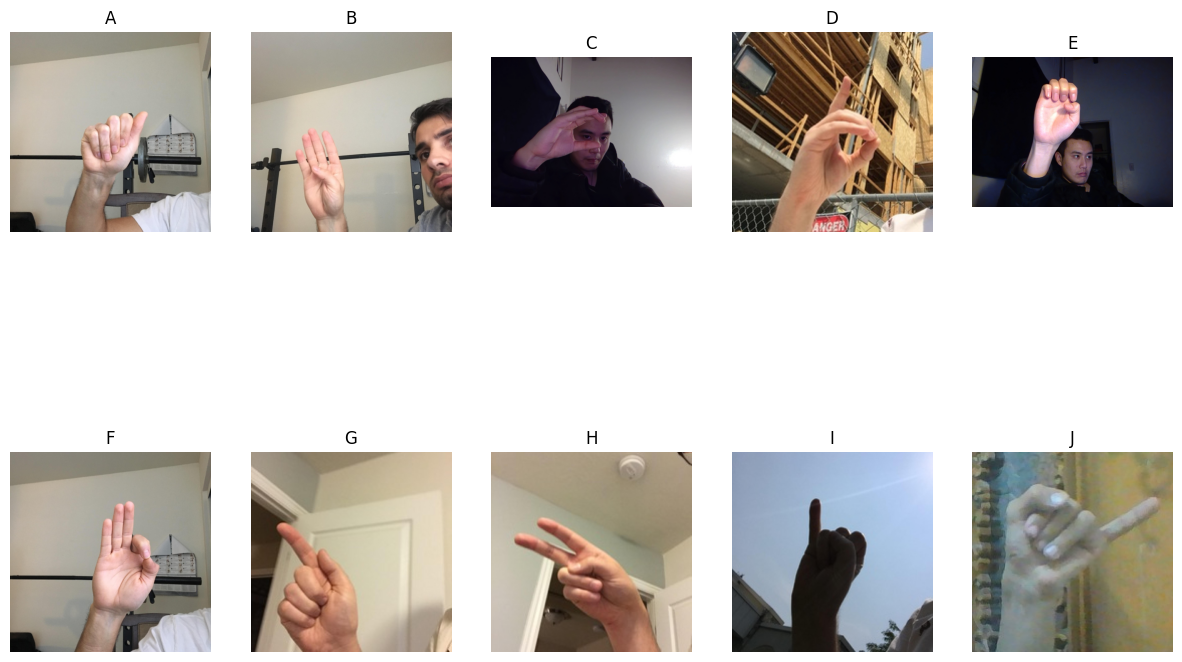
**GitHub Repository:**

Contains training scripts, model weights, and deployment files.

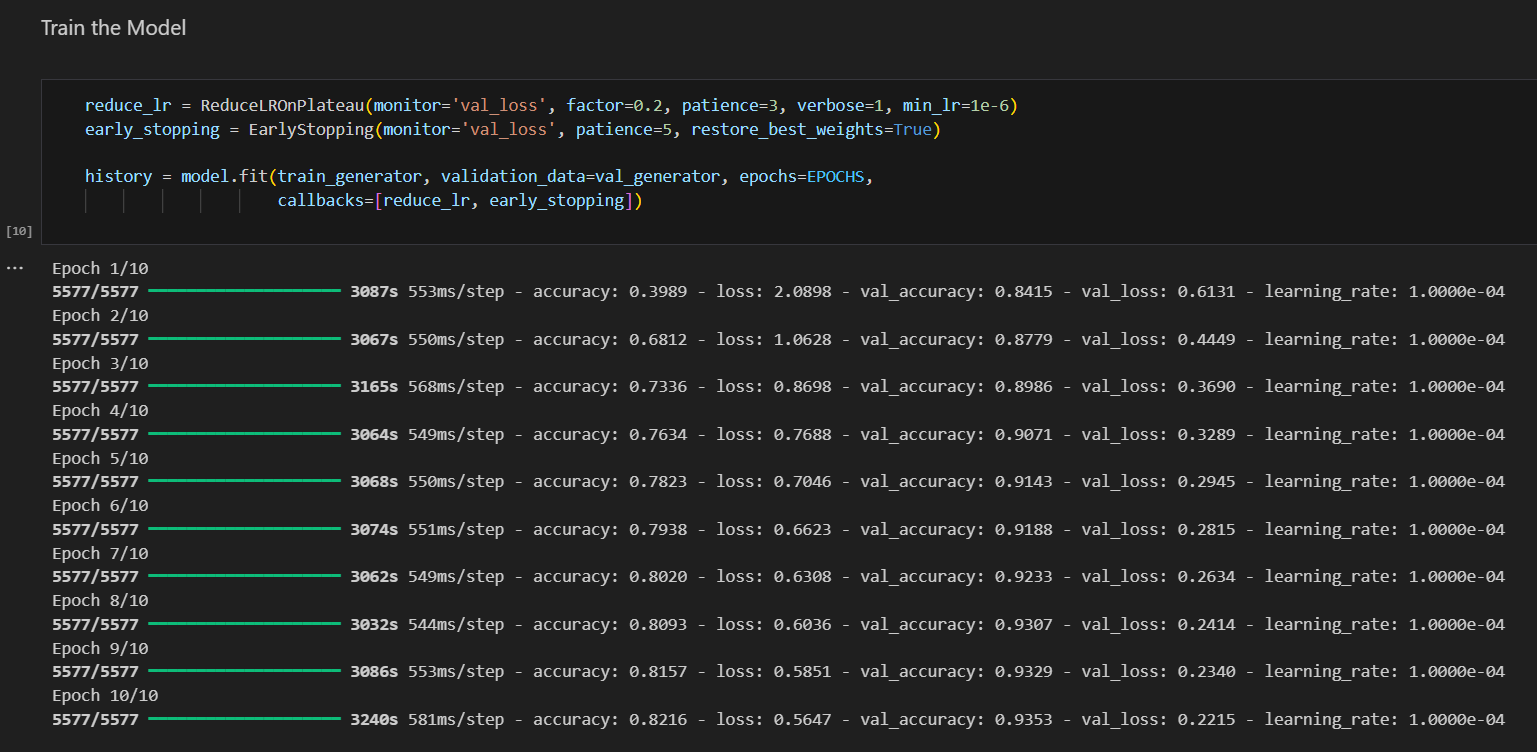
**Screenshots**

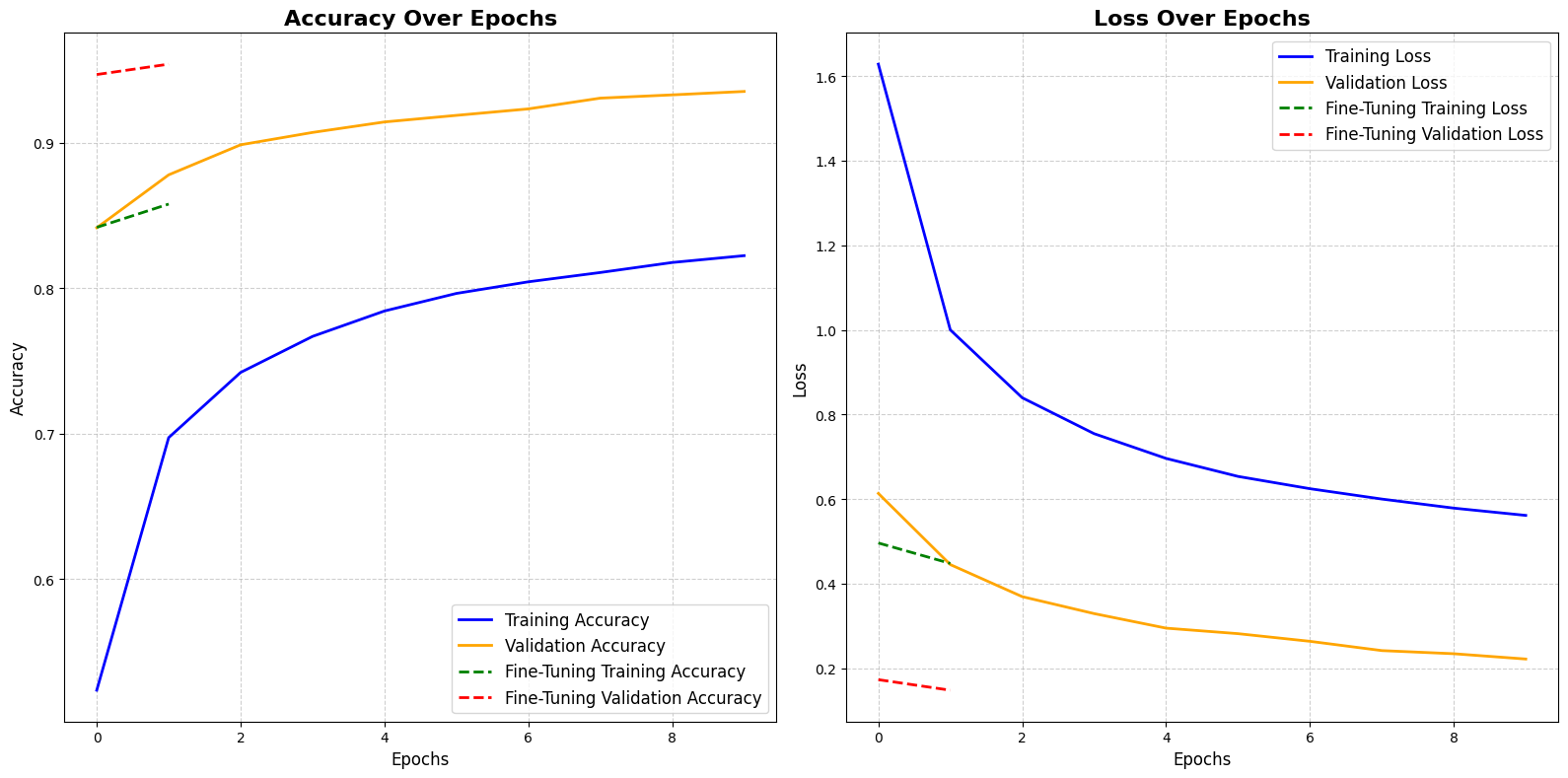
Dataset Exploration

****

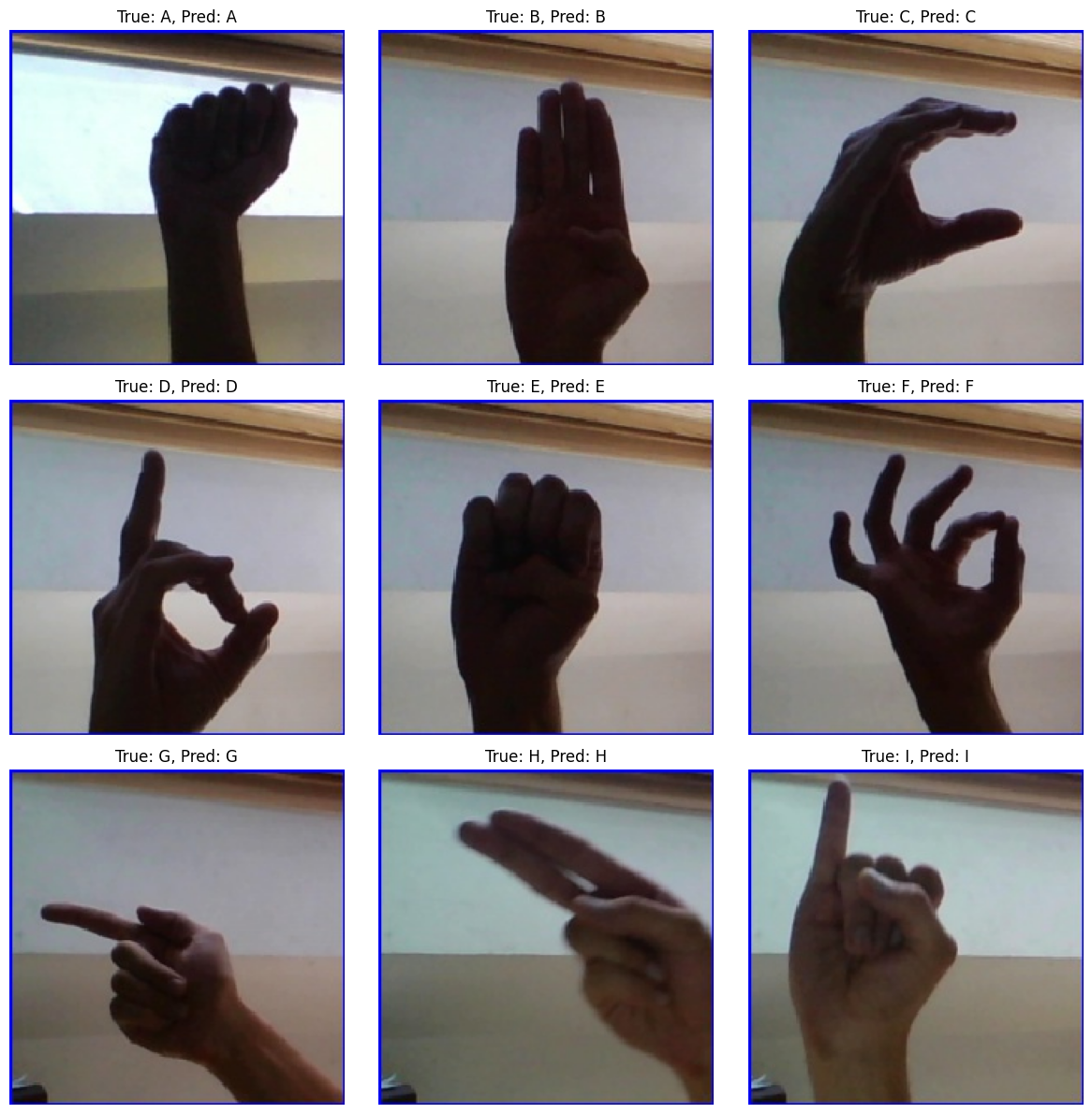
****

Training process

****

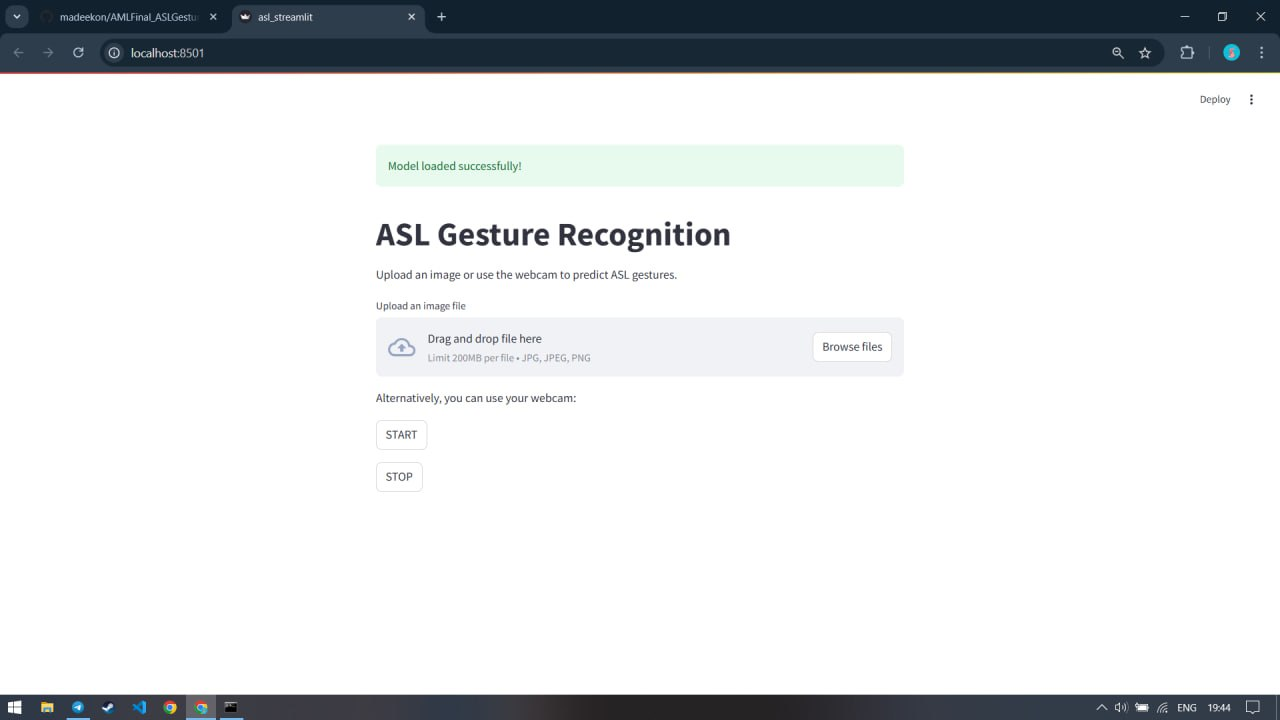
****

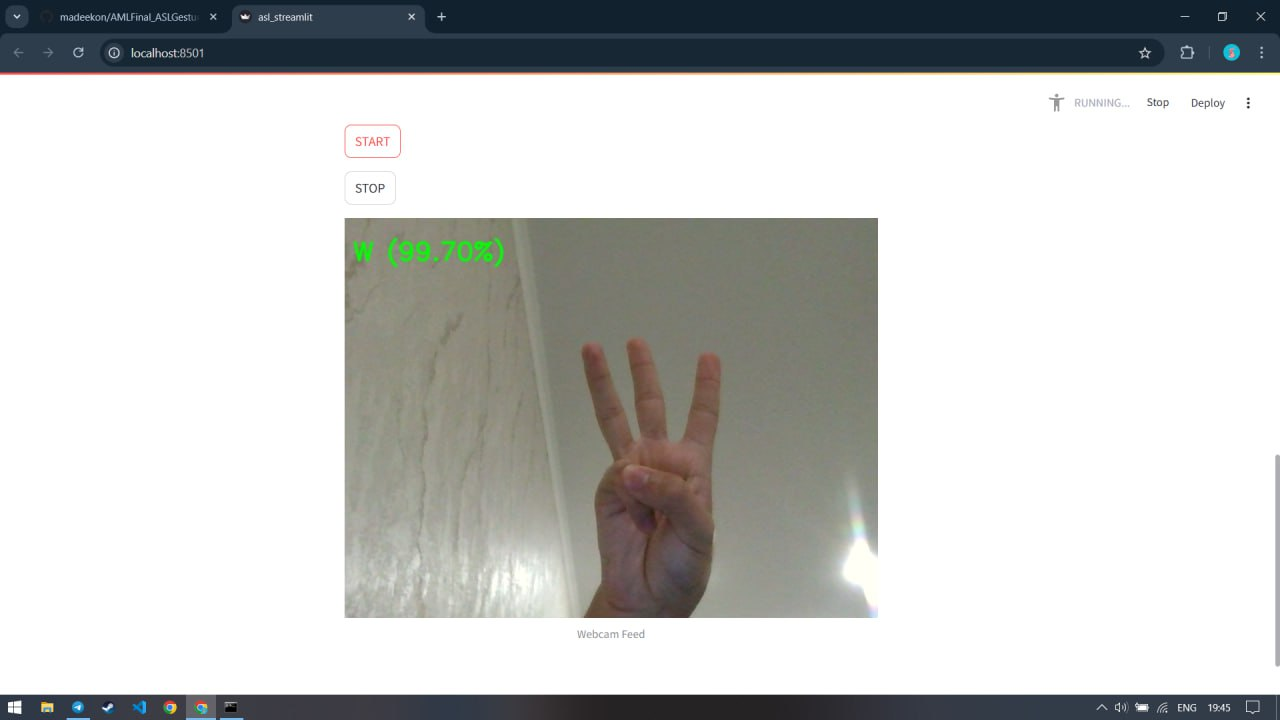
*Visualization of Training Performance*

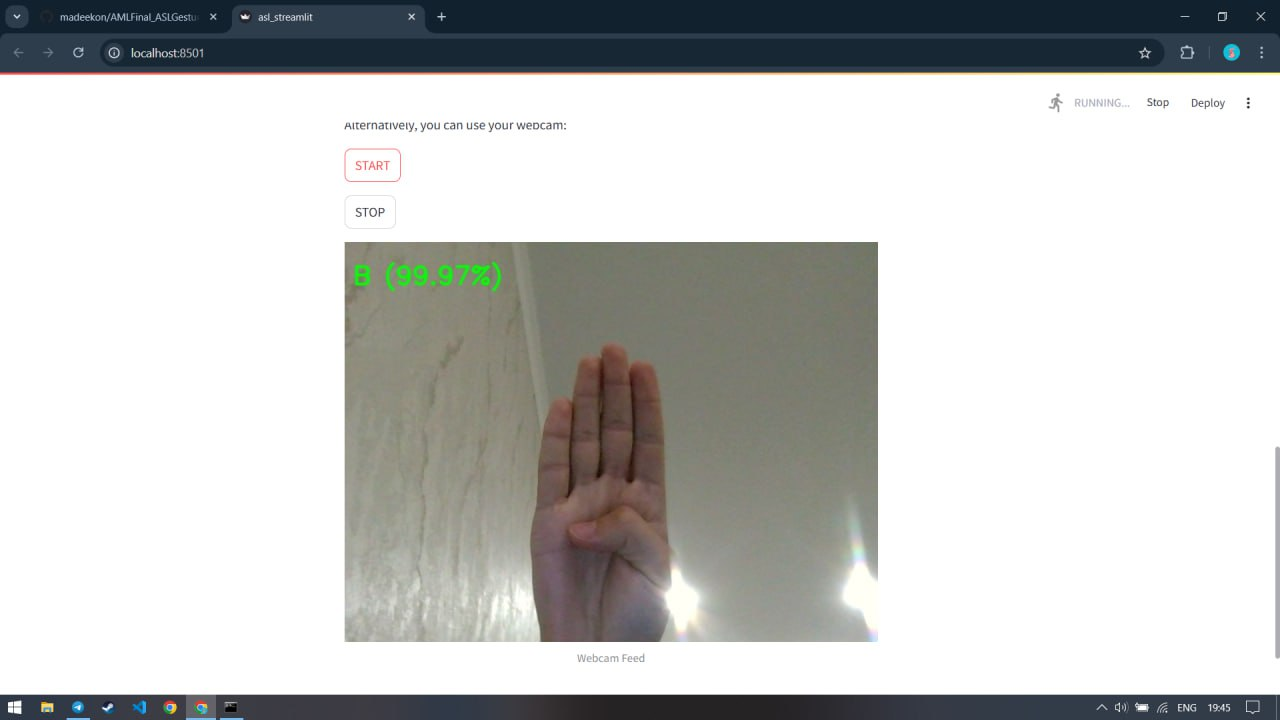


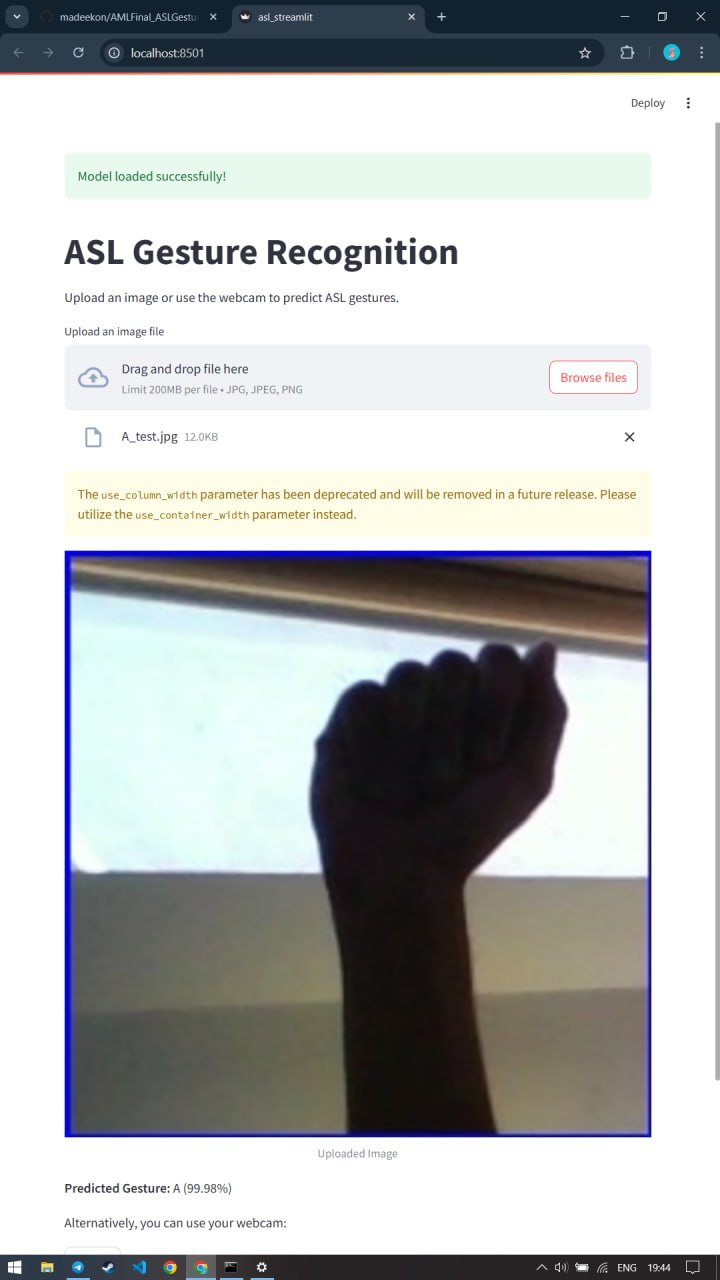
*Predictions*

**Streamlit deployment**









**Conclusion**

This project demonstrates effective classification of ASL hand signs using deep learning. The fine-tuned model achieves high accuracy, and the deployment via Streamlit ensures practical applicability. Future improvements could include dynamic gesture recognition and multilingual support.