Report on Action Recognition System - VCL MiniProject 2

Problem Statement

Design a convolutional neural network for classifying video clips using the given dataset.

The code along with pre-trained models and deployment link can be found here.

Dataset Details

- **Class Names:** 'punch', 'pick', 'laugh', 'pour', 'pullup'.
- **Number of Clips for Training:** 35023
- **Number of Clips for Validation:** 8756

Model Performance

Validation Accuracy

- **Validation Accuracy for Regular Clips:** 97.56
- **Validation Accuracy for Redundant Frame Clips:** 93.42

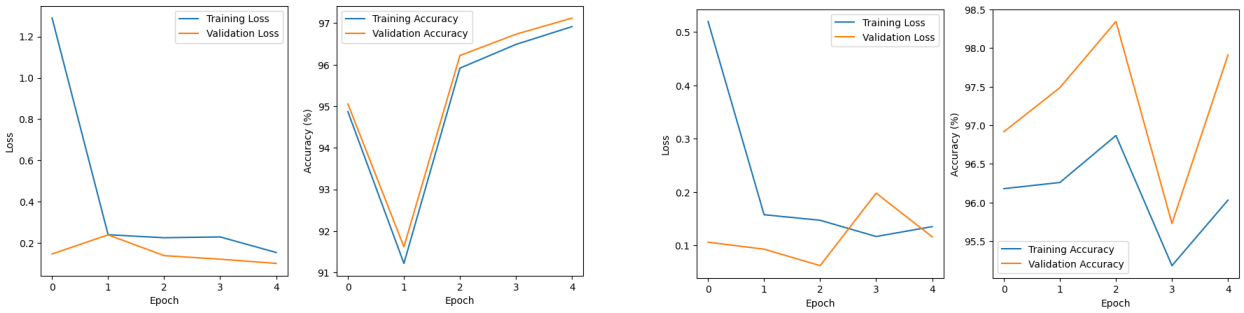
Strategy for Improving Validation Accuracy on Redundant Frame Clips

By Increasing the Depth of the model with an additional 3D-Convolutional layer.

Improved Testing Accuracy for Redundant Frame Clips

- **Improved Testing Accuracy for Regular Frame Clips:** 99.68
- **Improved Testing Accuracy for Redundant Frame Clips:** 98.82

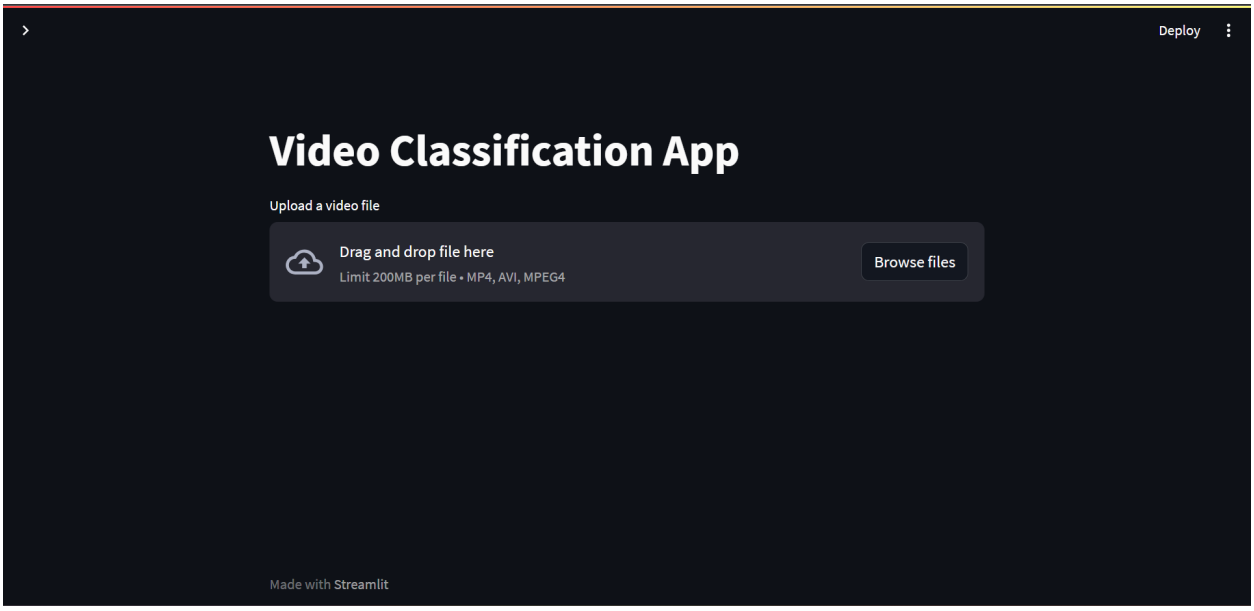
Training and Validation Graphs



- The above graph illustrates the training loss, training accuracy, validation loss, and validation accuracy across epochs for both models - Vanilla(left) and Improved(right)

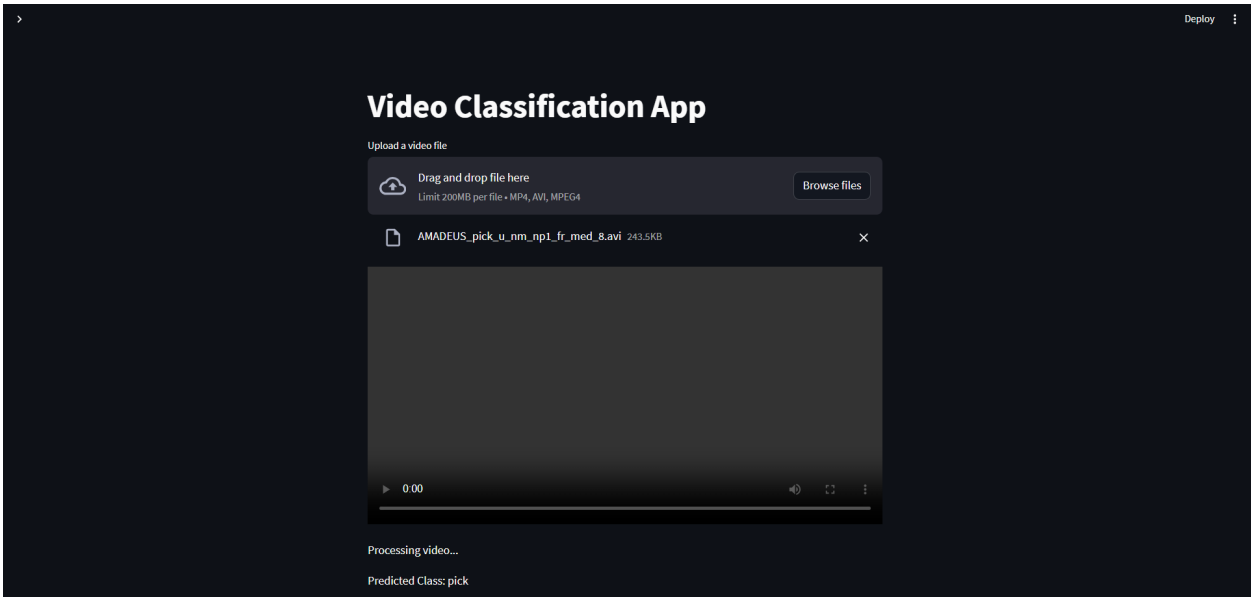
Model Deployment Screenshots

Before Predictions



- This screenshot shows the web-deployed model interface before making predictions.

After Predictions



- This screenshot shows the web-deployed model interface after making predictions.