﻿**DSAA 5002 - Data Mining and Knowledge Discovery in Data Science**

**Final Exam Report – Q3 Short Video Classification**

**50015940 Jiaxiang Gao**

1. **Algorithm Overview**

The algorithm employs a 3D Convolutional Neural Network (CNN) to classify videos. The process is divided into several key steps:

1. **Video Frame Extraction:**

Utilizes FFmpeg to extract key frames from videos. The extract\_key\_frames\_ffmpeg function takes a video file and saves its key frames at a specified rate (default is one frame per second). The process\_videos\_with\_ffmpeg function processes all videos in each folder using FFmpeg.

1. **Label Loading and Preprocessing:**

Reads a tag file that contains labels for each video. The load\_labels function creates a DataFrame mapping video names to their corresponding labels.

1. **Frame Loading and Preprocessing:**

Loads and preprocesses video frames. The load\_video\_frames function reads the extracted frames, resizes them, and ensures uniformity in the number of frames per video. Frames are normalized, and labels (if available) are one-hot encoded.

1. **3D CNN Model Building:**

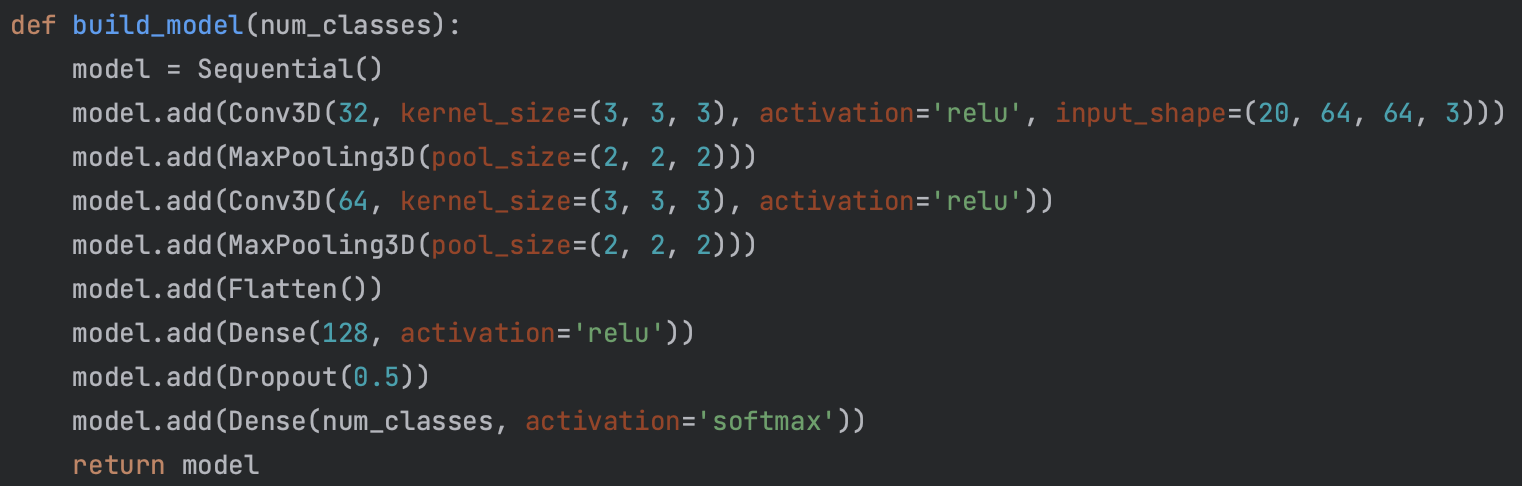
A Sequential model is built using Keras. The model contains layers of 3D convolutions, max pooling, flattening, dense layers, and dropout for regularization. The final layer uses a SoftMax activation function for classification.

1. **Model Training and Evaluation:**

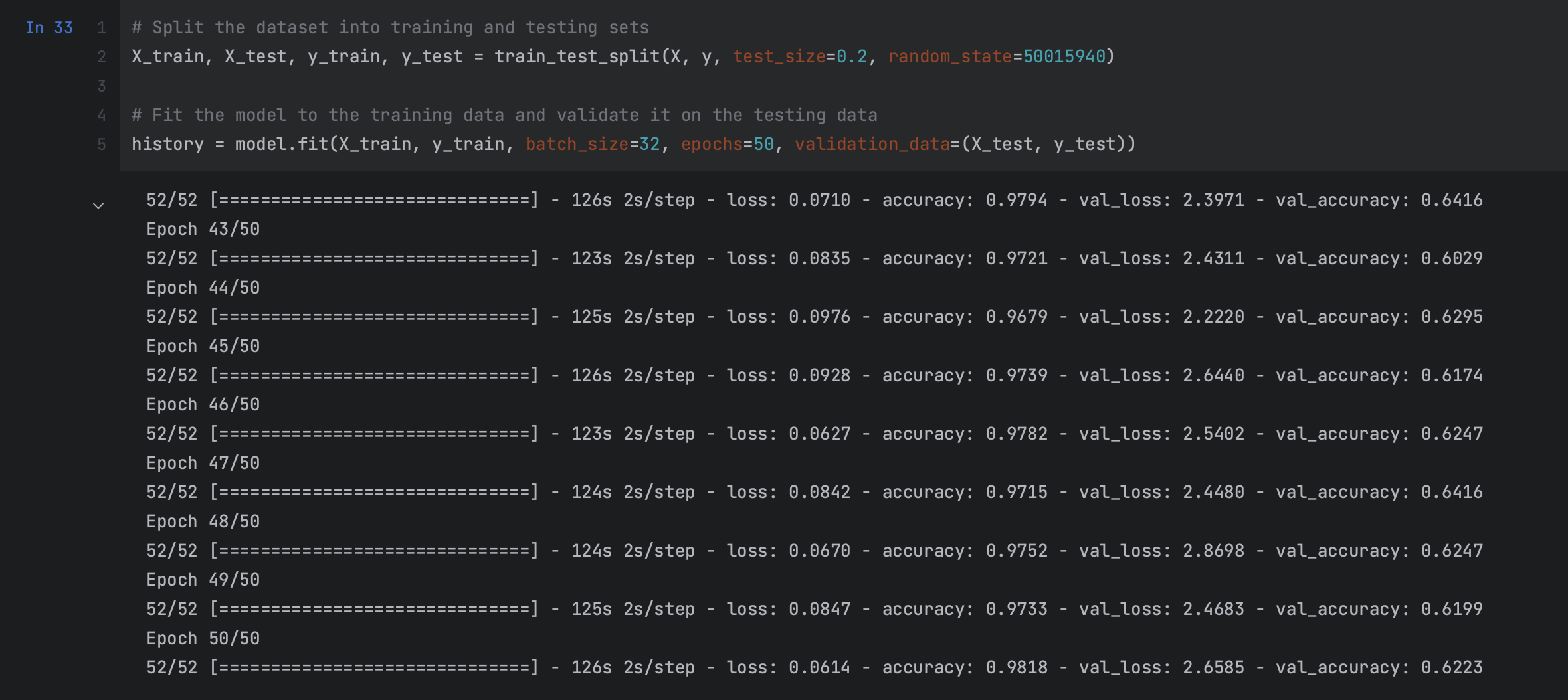
The dataset is split into training and testing sets. The model is compiled with categorical cross-entropy loss and the Adam optimizer. Training is performed over 50 epochs with batch size 32.

1. **Model Architecture:**

* Input Shape: (20, 64, 64, 3) for 20 frames, each 64x64 in size, with 3 color channels.
* Convolutional Layers: Multiple layers with varying number of filters (32, 64).
* Max Pooling Layers: Reduce spatial dimensions.
* Flatten Layer: Converts 3D feature maps to 1D feature vectors.
* Dense Layers: Fully connected layers for classification.
* Dropout: Regularization to prevent overfitting.
* Output Layer: Softmax activation for multi-class classification



1. **Model Training**



**4. Model Evaluate**

Accuracy in spilt training set: 0.62

