CS-512 COMPUTER VISION

Anomaly-Driven Video Summarization for Real-Time Surveillance Systems

PROJECT PROPOSAL

VISHNU PRIYAN SELLAM SHANMUGAVEL - A20561323
AKASH THIRUMURUGANANTHAM - A20539883

OBJECTIVE

OUR PROJECT FOCUSES ON OPTIMIZING SURVEILLANCE VIDEO PROCESSING BY AUTOMATICALLY DETECTING AND SUMMARIZING ANOMALOUS EVENTS. THIS APPROACH REDUCES THE NEED FOR MANUAL VIDEO REVIEW, ENABLING QUICKER THREAT IDENTIFICATION AND EFFICIENT STORAGE MANAGEMENT.

PROBLEM STATEMENT

SURVEILLANCE SYSTEMS GENERATE MASSIVE AMOUNTS OF FOOTAGE, MOSTLY STORED PASSIVELY, WHICH IS INEFFICIENT FOR MONITORING AND DIFFICULT TO ANALYZE IN REAL TIME. MANUAL REVIEW OF SUCH DATA IS LABOR-INTENSIVE, PRONE TO ERRORS, AND SLOW, PARTICULARLY WHEN SEARCHING FOR SPECIFIC ANOMALOUS EVENTS.

PROJECT OVERVIEW

Our project, "Anomaly-Driven Video Summarization for Real-Time Surveillance Systems," aims to enhance the efficiency of surveillance monitoring by automatically detecting anomalies in video footage and summarizing relevant segments.

Our approach segments videos into smaller portions, extracts spatial and temporal features using deep learning models (C3D and 3D ResNet), and applies a modified loss function to improve anomaly localization.

Using datasets like UFC-Crime, UBNormal, and UBI-Flights, our model is trained to detect and prioritize significant footage, enabling quicker and more reliable monitoring and response.

Create a concise video summary by selecting keyframes from the top-ranked segments, providing a quick overview of anomalous activities.

LITERATURE REVIEW

MAIN PAPER SUMMARY

The paper, “**Real-world Anomaly Detection in Surveillance Videos”** by Waqas Sultani, Chen Chen, and Mubarak Shah, introduces an innovative approach for anomaly detection using weakly labeled data. The authors leverage a Multiple Instance Learning (MIL) framework where video-level labels (indicating whether an anomaly is present) are used without specific segment-level annotations. The model assigns high anomaly scores to segments with unusual behavior, aided by a ranking loss function with sparsity and temporal smoothness constraints, enhancing its ability to localize anomalies within the footage.

MAIN PAPER SUMMARY

Our project addresses key gaps by focusing not only on detecting anomalies but also on summarizing them, enabling efficient video monitoring and retrieval. Additionally, our model incorporates a modified loss function and ranks segments based on factors like memorability and entropy, refining anomaly localization and improving overall detection precision in real-world surveillance environments, which require quick and accurate anomaly identification.

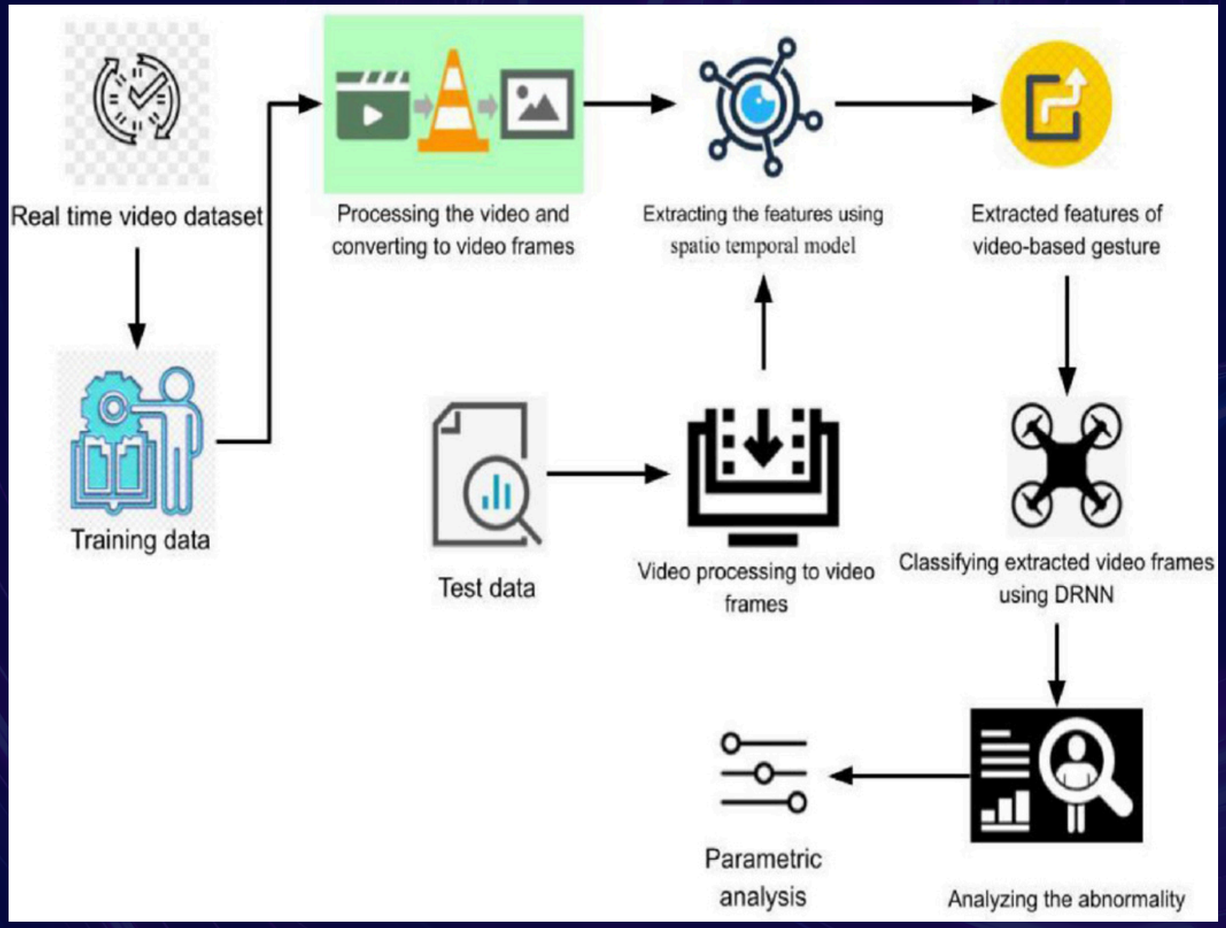
PROPOSED SOLUTION

Our solution automates anomaly detection in surveillance videos by segmenting footage, extracting spatial-temporal features using C3D and 3D ResNet, and applying a modified loss function for better anomaly localization. We rank segments by significance, using scoring metrics, to create a summary that highlights key events for efficient monitoring and threat identification.

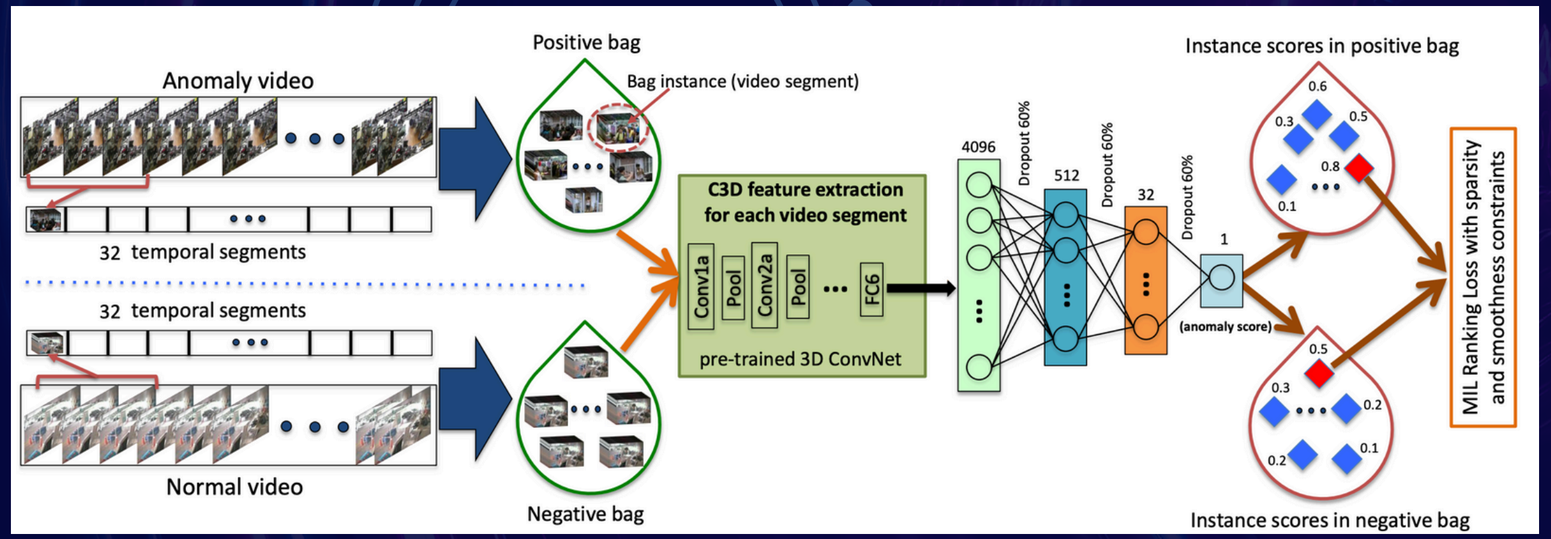
SOLUTION APPROACH

1	**SHOT SEGMENTATION**	Divide videos into smaller temporal segments
2	**FEATURE EXTRACTION**	Use advanced models like C3D (Convolutional 3D) and 3D ResNet to capture both spatial and temporal features
3	**LOSS FUNCTION**	Enhance the ranking loss function by adding Temporal Smoothness and Sparsity Constraints
4	**SCORING AND RANKING**	Score each segment based on memorability, entropy, and temporal dynamics.
5	**VIDEO SUMMARIZATION**	Generate a summary of the footage by selecting keyframes from the highest-ranked segments, ensuring efficient storage and faster retrieval.

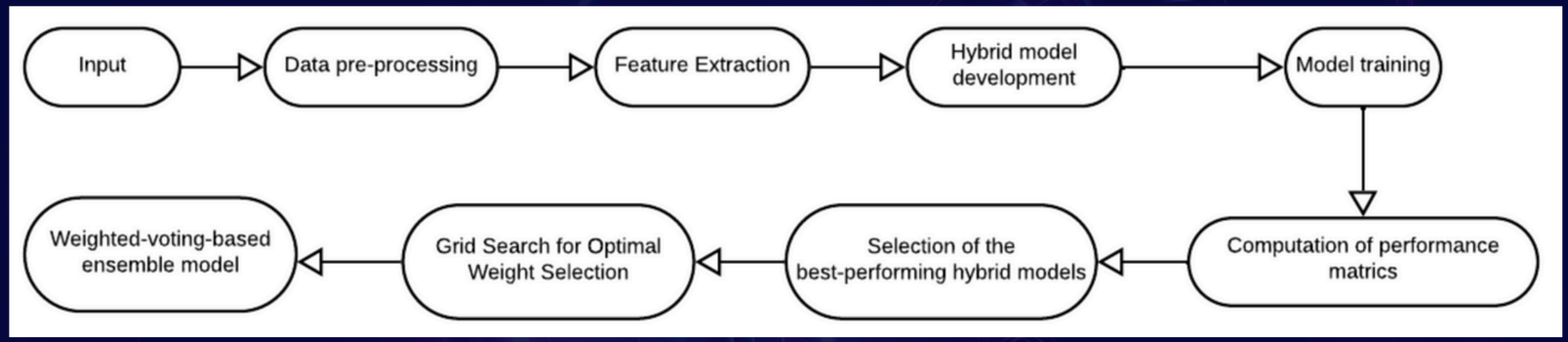
ANOMALY DETECTION APPROACH

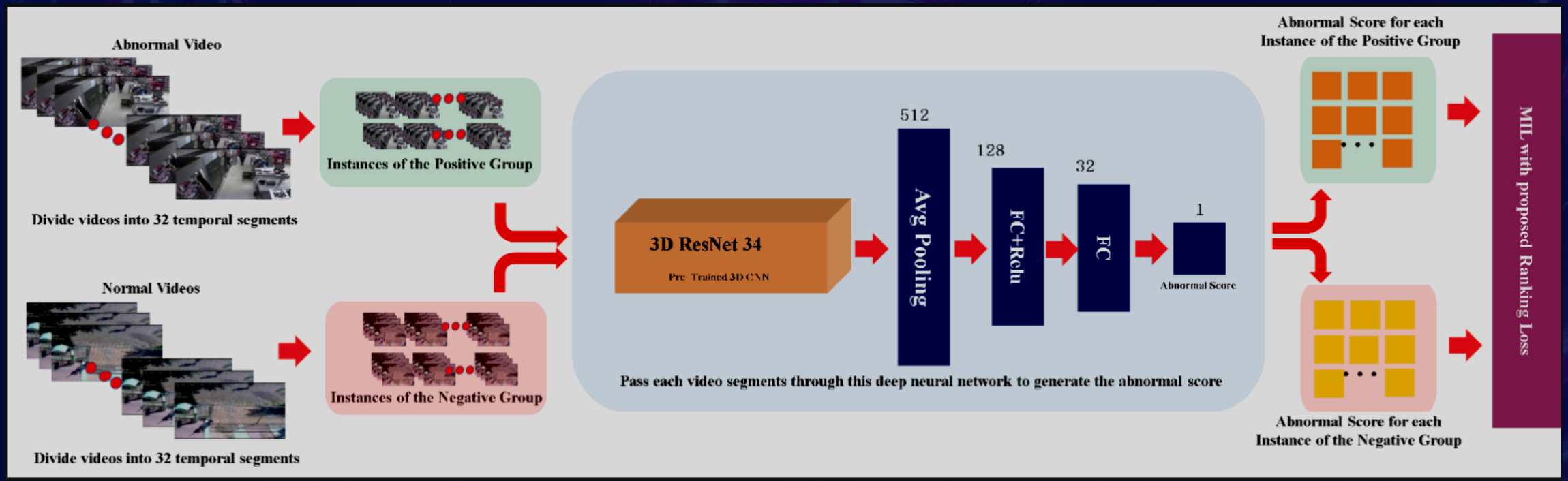


DETECTION EXAMPLES

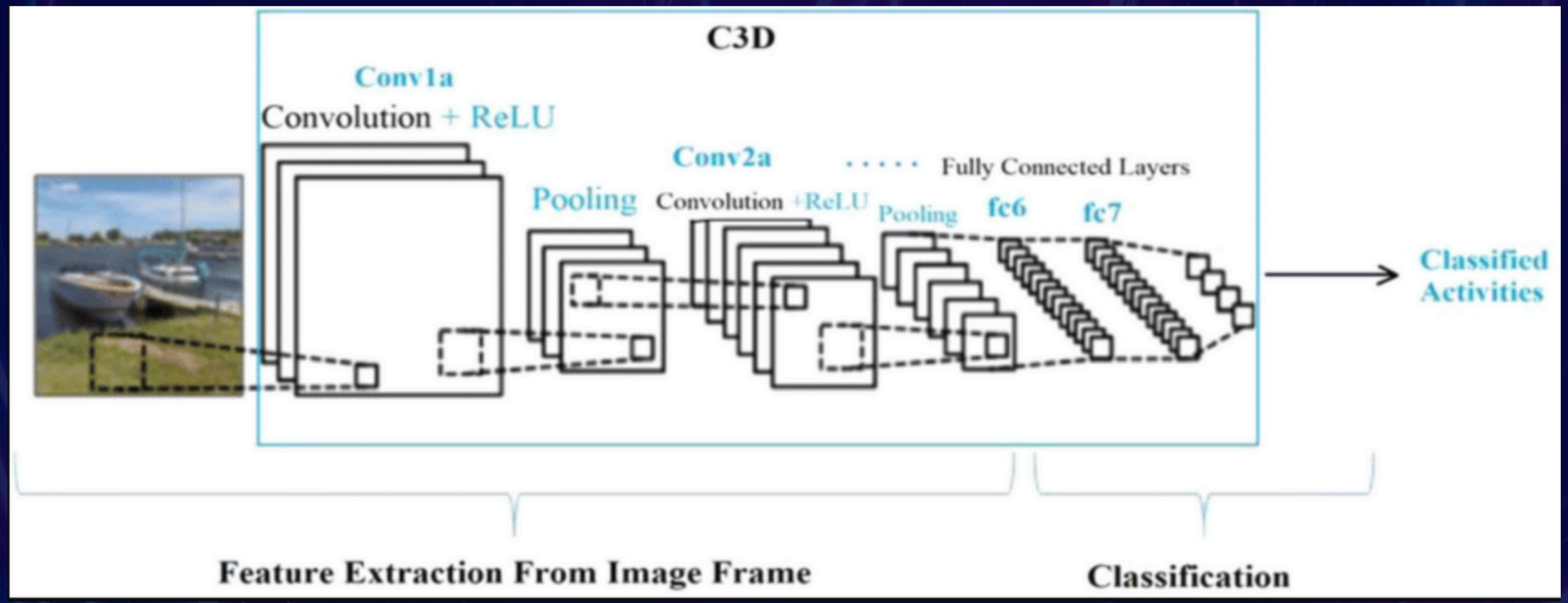


OVERALL PROCESS

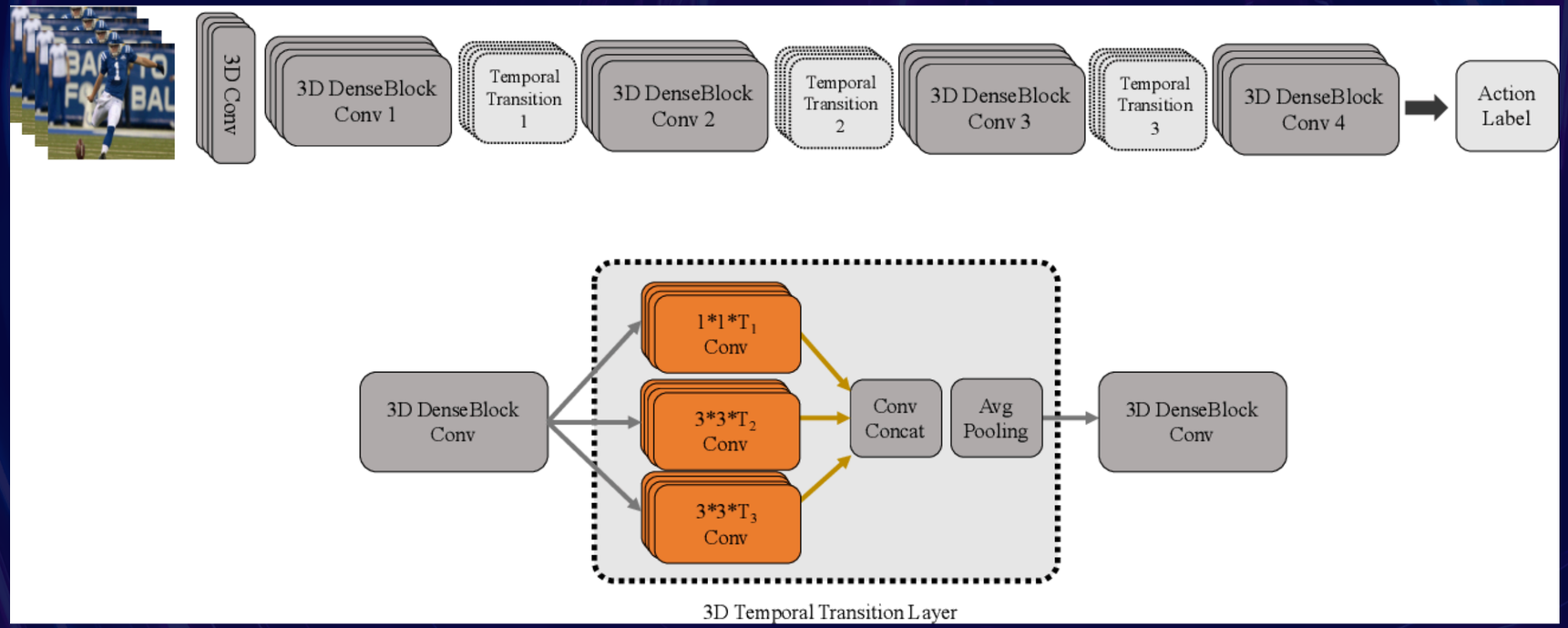




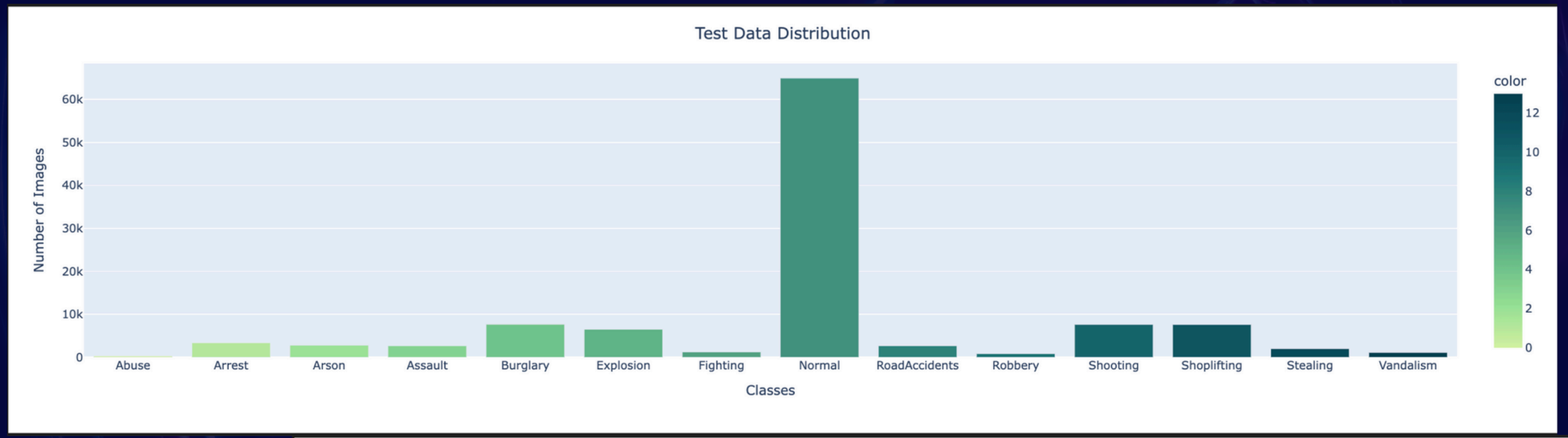
C3D Feature Extraction



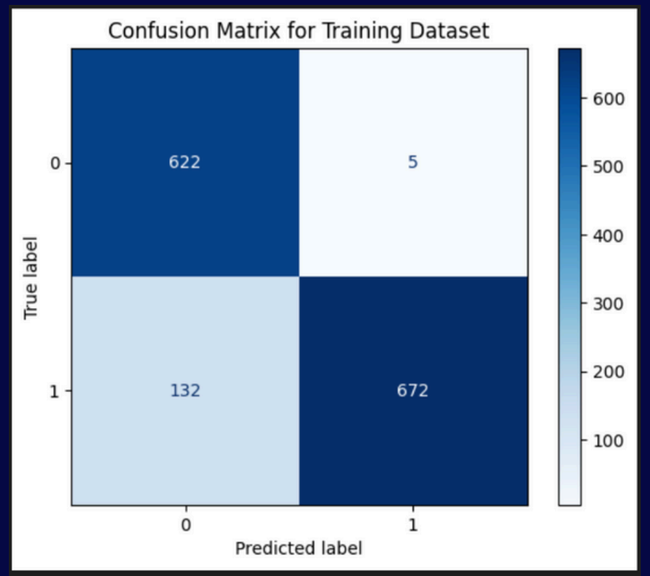
3D RESNET

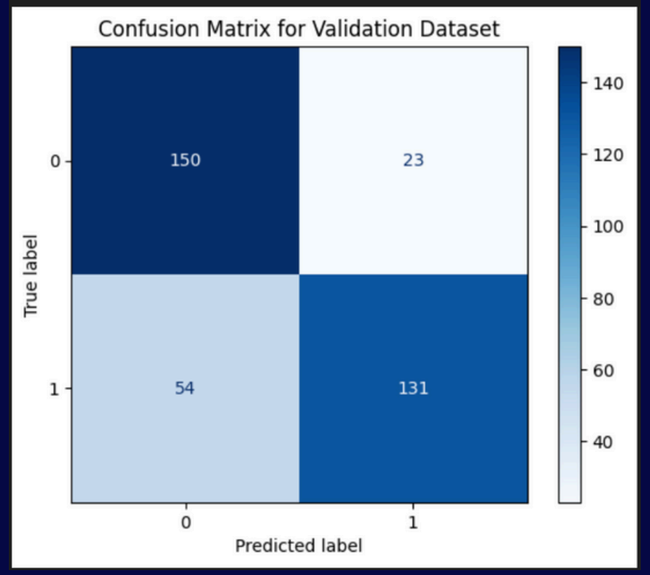

3D Conv
3D DenseBlock Conv 1
Temporal Transition 1
3D DenseBlock Conv 2
Temporal Transition 2
3D DenseBlock Conv 3
Temporal Transition 3
3D DenseBlock Conv 4
Action Label
3D DenseBlock Conv
1*1*T1 Conv
3*3*T2 Conv
3*3*T3 Conv
Conv Concat
Avg Pooling
3D DenseBlock Conv
3D Temporal Transition Layer

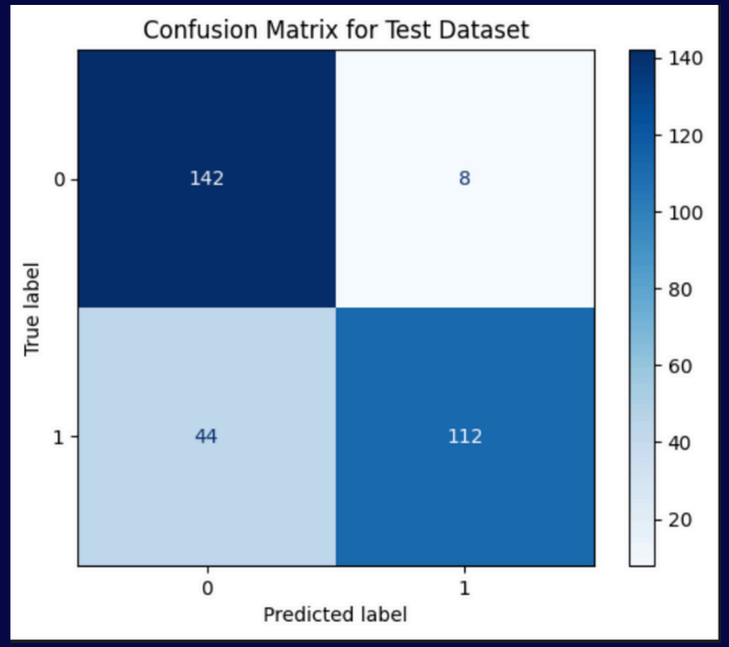
TEST AND TRAIN RESLTS

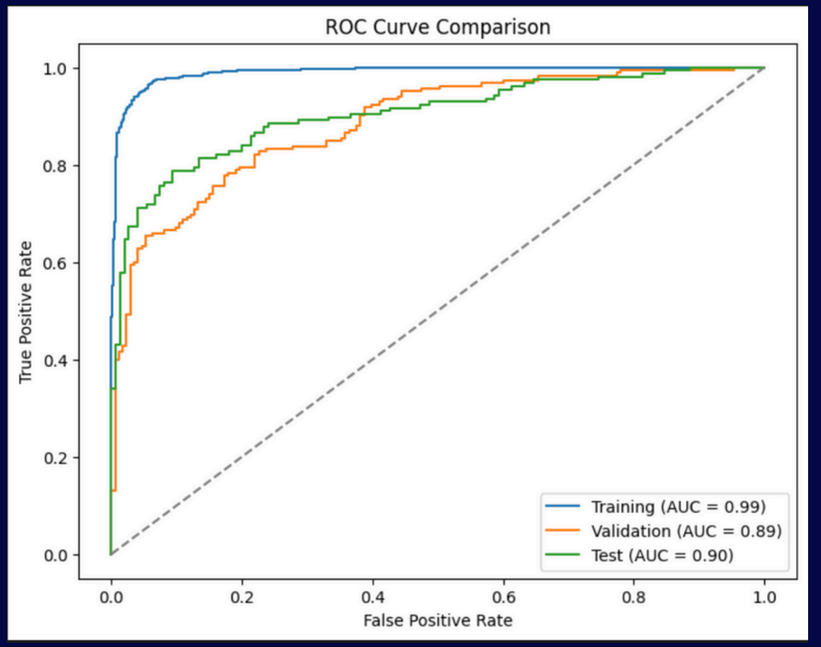


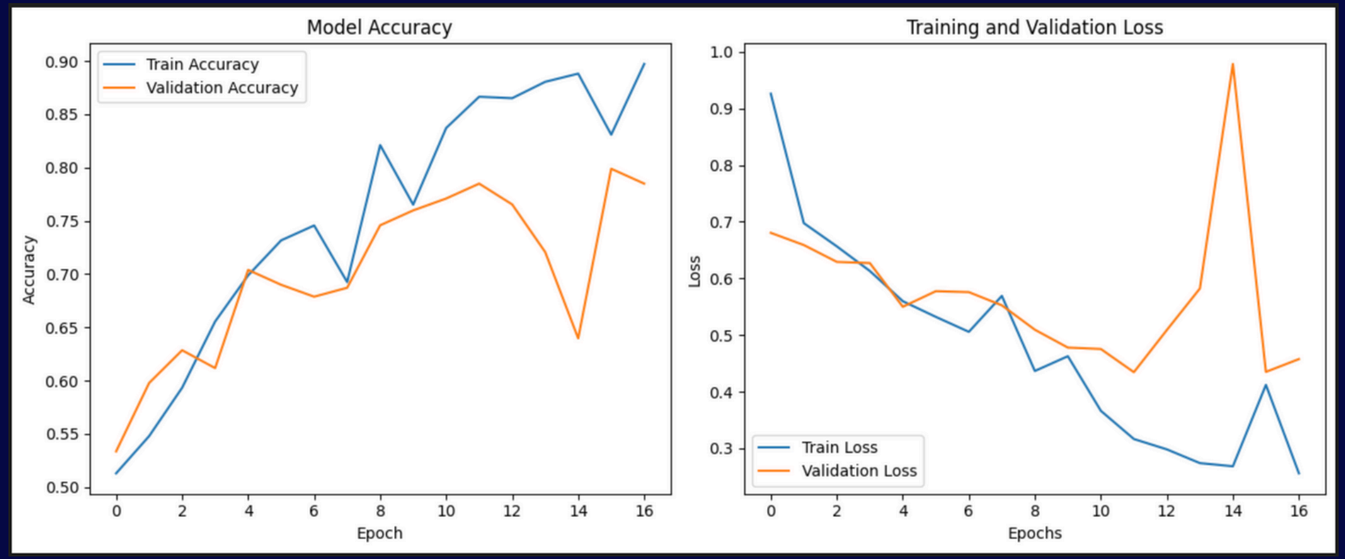
TEST, TRAIN AND VALIDATION RESLTS



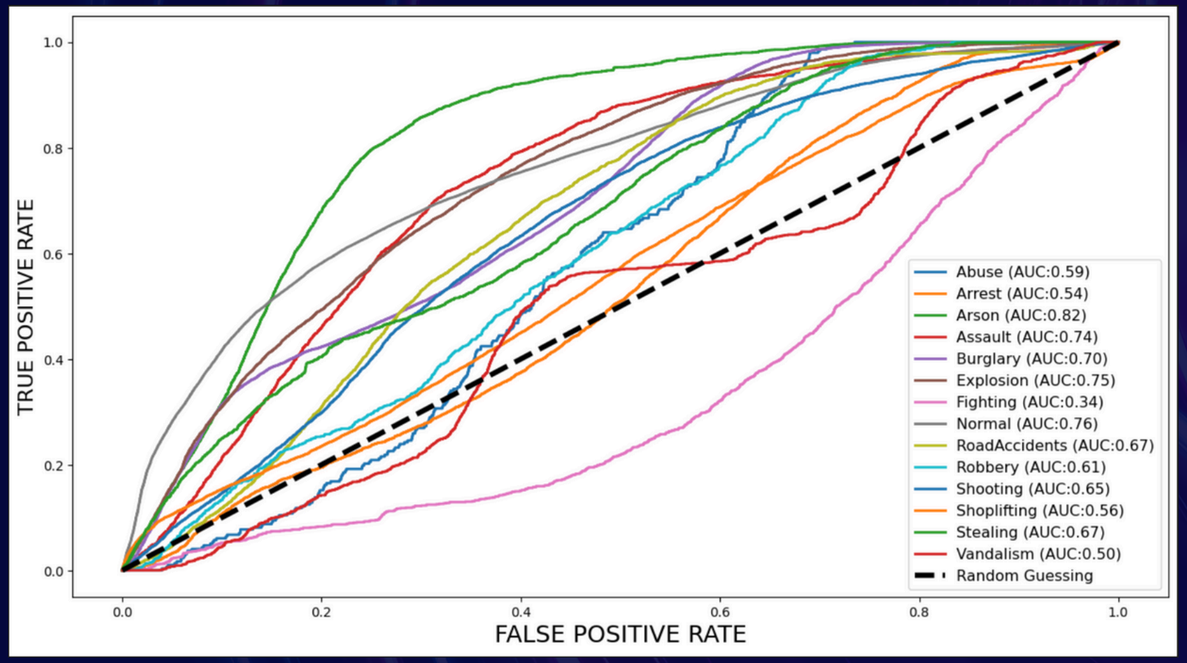


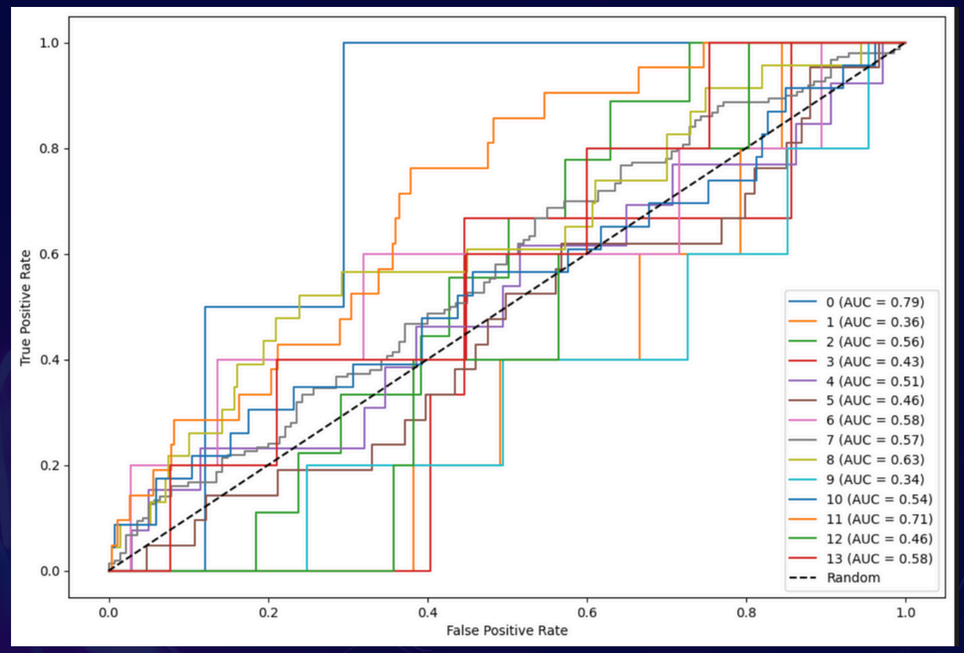


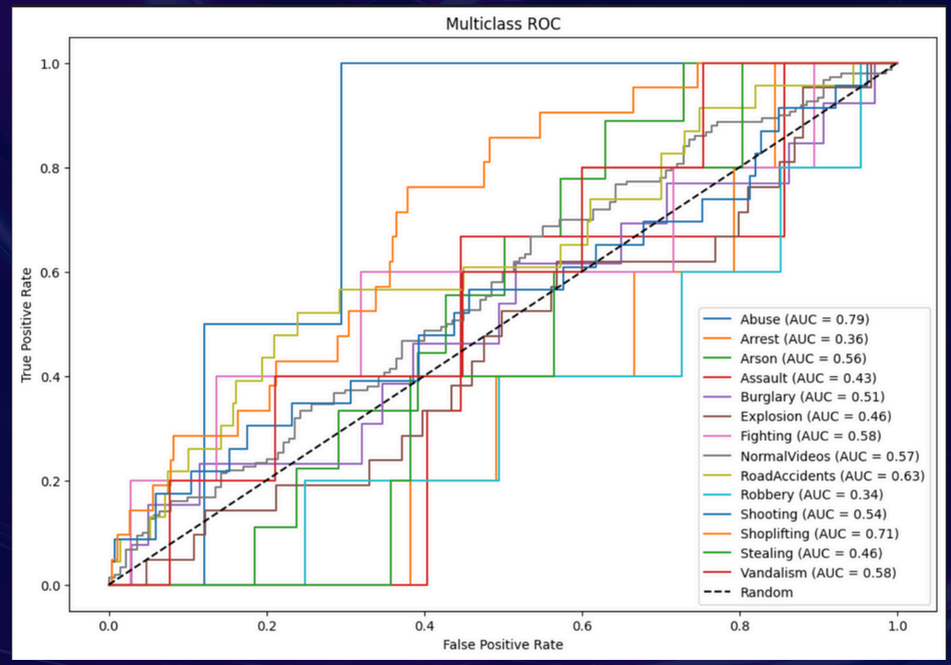




RESULTS AND COMPARISONS







THANK YOU