CS-512 COMPUTER VISION

Anomaly-Driven Video Summarization for Real-Time Surveillance Systems

PROJECT PROPOSAL

VISHNU PRIYAN SELLAM SHANMUGAVEL - A20561323
AKASH THIRUMURUGANANTHAM - A20539883

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.

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.

MAIN & ADDITIONAL PAPERS

Main Paper: Real-world Anomaly Detection in Surveillance Videos (2019)
by Waqas Sultani, Chen Chen, and Mubarak Shah

- The authors propose an innovative anomaly detection approach using weakly labeled data, leveraging a Multiple Instance Learning (MIL) framework. Video-level labels indicate anomaly presence without segment-level annotations. The model assigns high anomaly scores to unusual segments, using a ranking loss function with sparsity and temporal smoothness constraints to improve anomaly localization.

Additional Paper: A Three-Stage Anomaly Detection Framework for Traffic Videos (2022)
by Junzhou Chen, Jiancheng Wang, Jiajun Pu, Ronghui Zhang

- The paper presents a three-stage anomaly detection framework for traffic videos, focusing on feature extraction, anomaly scoring, and decision-making to identify unusual events. It does three stages of processing to enhance real-time traffic monitoring, improving detection accuracy and reducing false positives.

Key Takeaways:

- Surveillance videos are long
- Importance of video summarization for efficient surveillance.

METHODOLOGY

Overview:

- Detect anomalies in surveillance footage using C3D and 3D ResNet.
- Summarize the video by selecting keyframes from ranked segments.
- Improve anomaly localization with feature extraction, modified loss function and ranking mechanism.

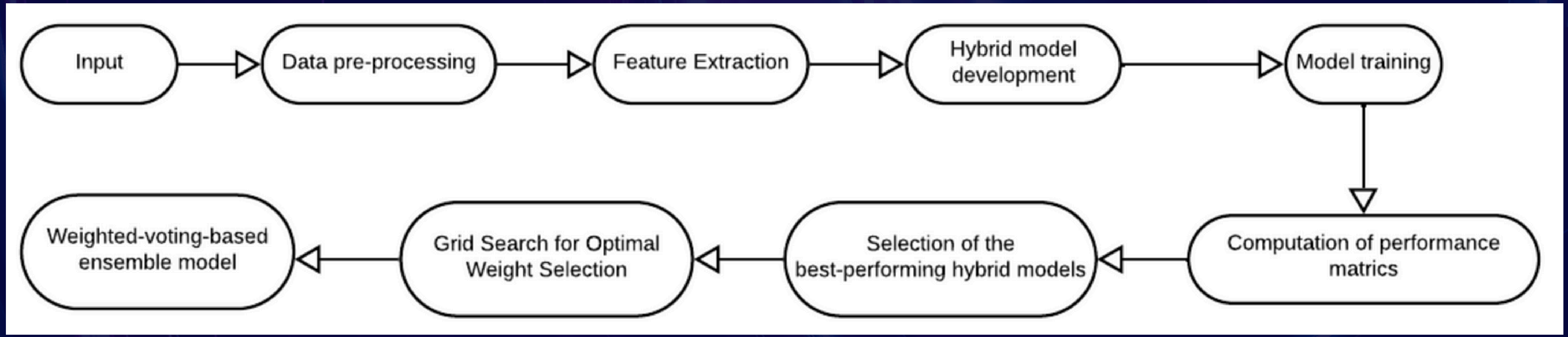
Data Used:

- UFC-Crime: For real-world crime detection.

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.

PLAN


Input
Data pre-processing
Feature Extraction
Hybrid model development
Model training
Computation of performance matrics
Selection of the best-performing hybrid models
Grid Search for Optimal Weight Selection
Weighted-voting-based ensemble model

Preprocessing

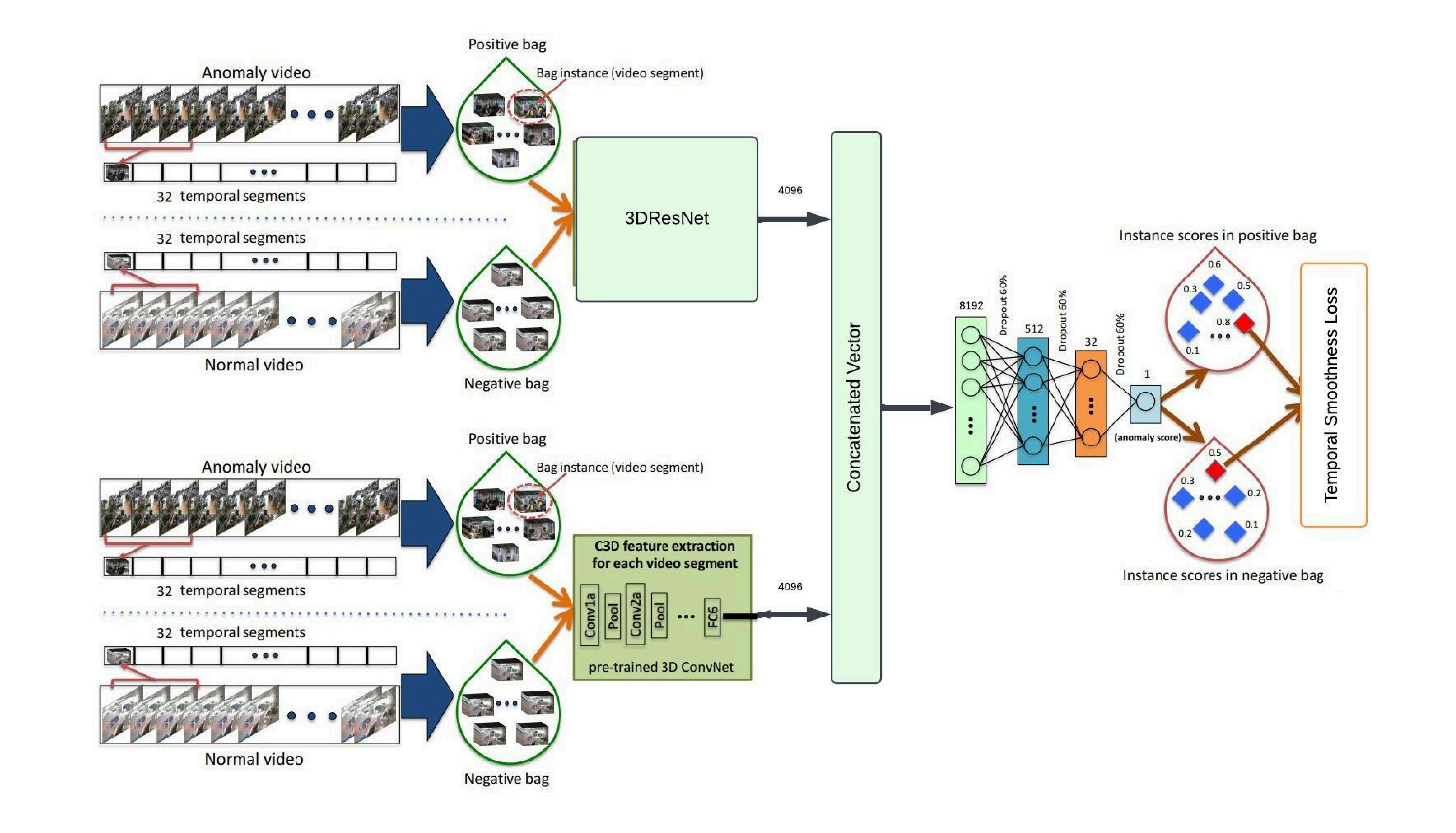
Overview:

- Labeling Segments: Identified and labeled segments where anomalies occur.
- Segmentation: Divided video data into meaningful segments based on anomaly occurrences.
- Efficient Storage: Stored labeled segments as .npy files for efficient data usage and reduced storage requirements.

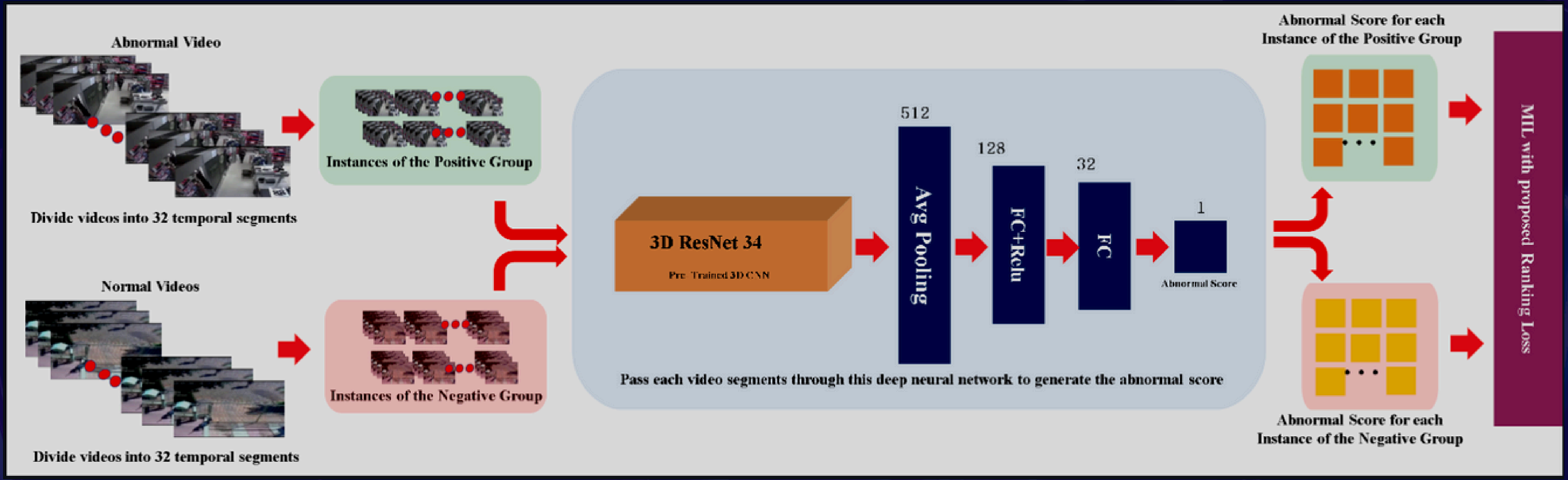
Key Benefits:

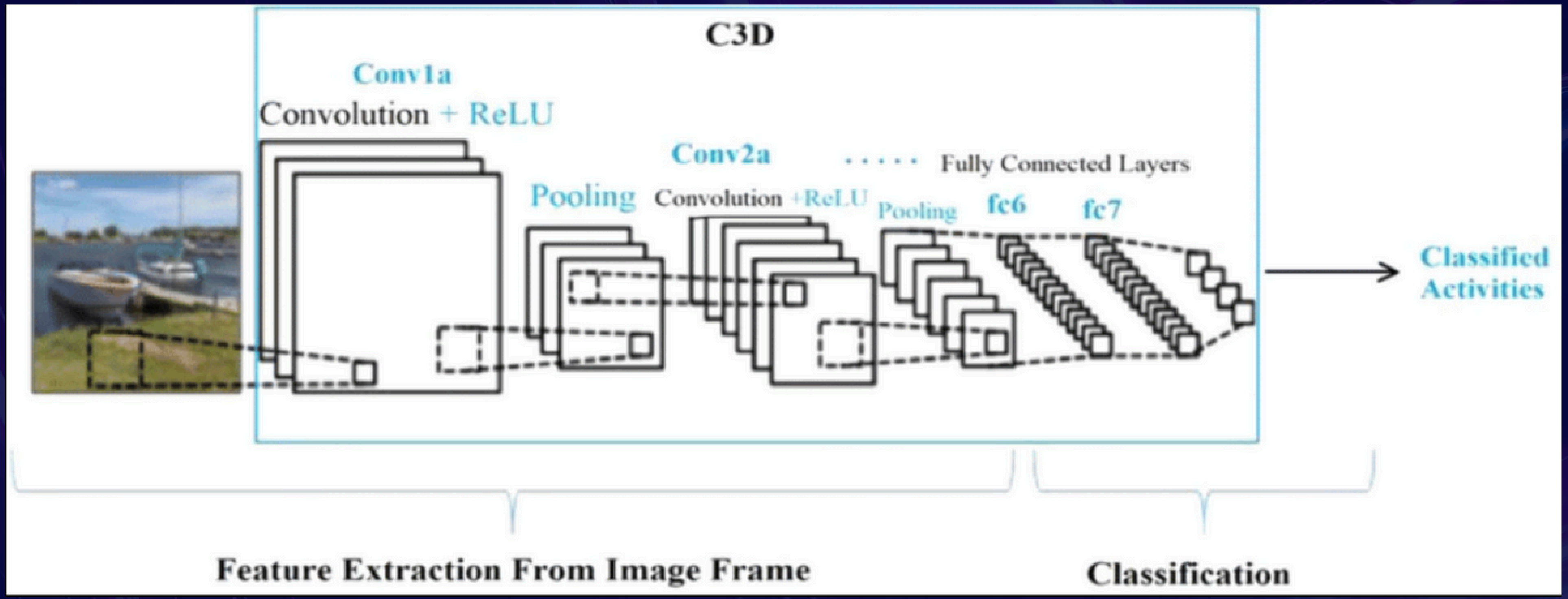
- Faster data loading and retrieval.
- Lower storage consumption by using .npy format.
- Easier management and manipulation of labeled anomaly segments.

ANOMALY DETECTION APPROACH:



Feature extraction Architectures





LOSS FUNCTION

Ranking Loss

- Minimizes distance for positive pairs; maximizes margin for negative pairs.
- Formula: $\text{loss} = y_{\text{true}} \cdot (y_{\text{true}} - y_{\text{pred}})^2 + (1 - y_{\text{true}}) \cdot \max(0, m - (y_{\text{true}} - y_{\text{pred}})^2)$

Temporal Smoothness Loss

- Enforces prediction consistency across frames.
- Formula: $\text{loss} = \text{mean}(|y_{\text{pred}}[i+1] - y_{\text{pred}}[i]|)$

Combined Loss

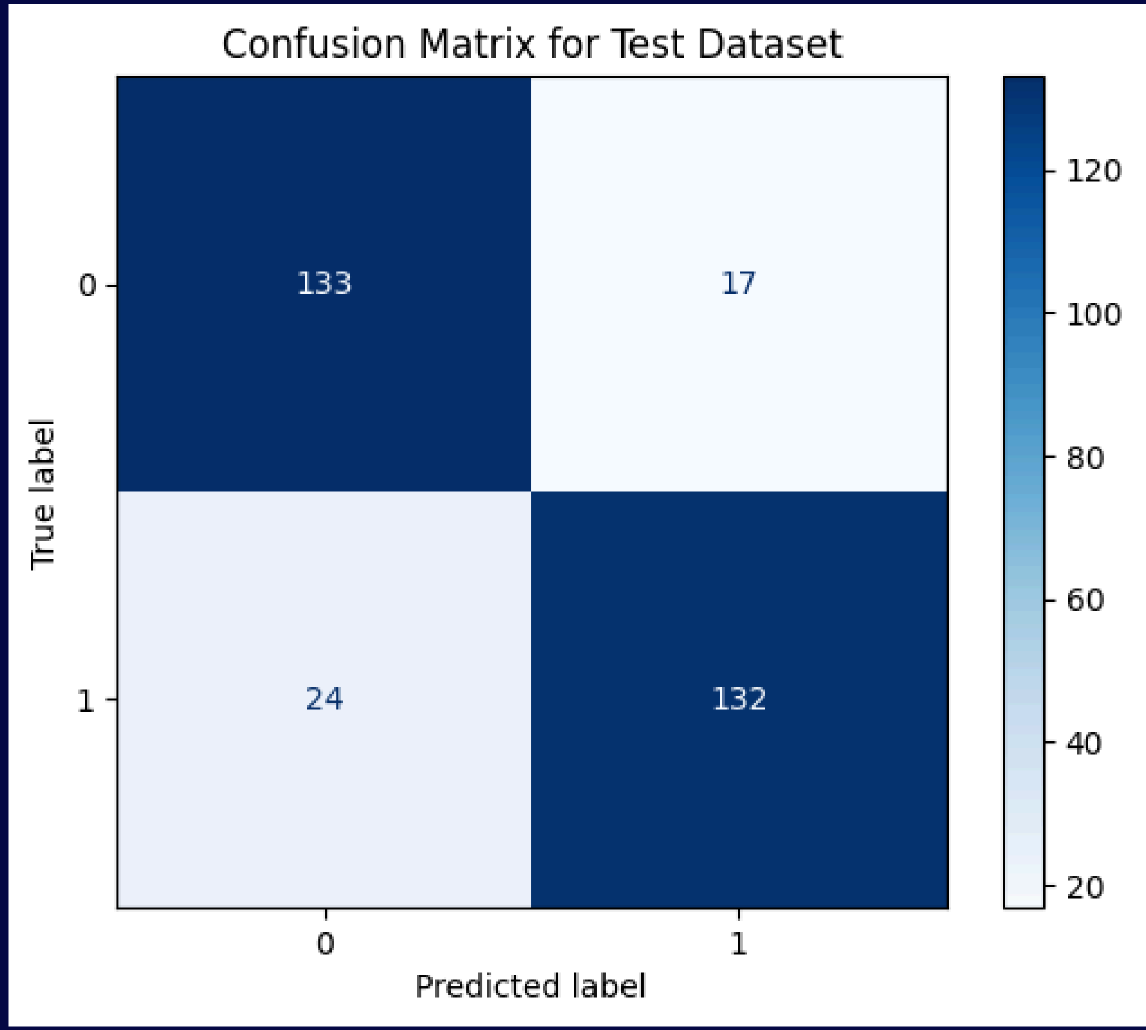
- Balances Binary Cross-Entropy, Ranking Loss, and Temporal Smoothness.
- Formula: $\text{total\_loss} = \text{BCE}(y_{\text{true}}, y_{\text{pred}}) + w_r \cdot \text{ranking\_loss} - w_s \cdot \text{temporal\_loss}$

Test, train and validation reslts COMPARION

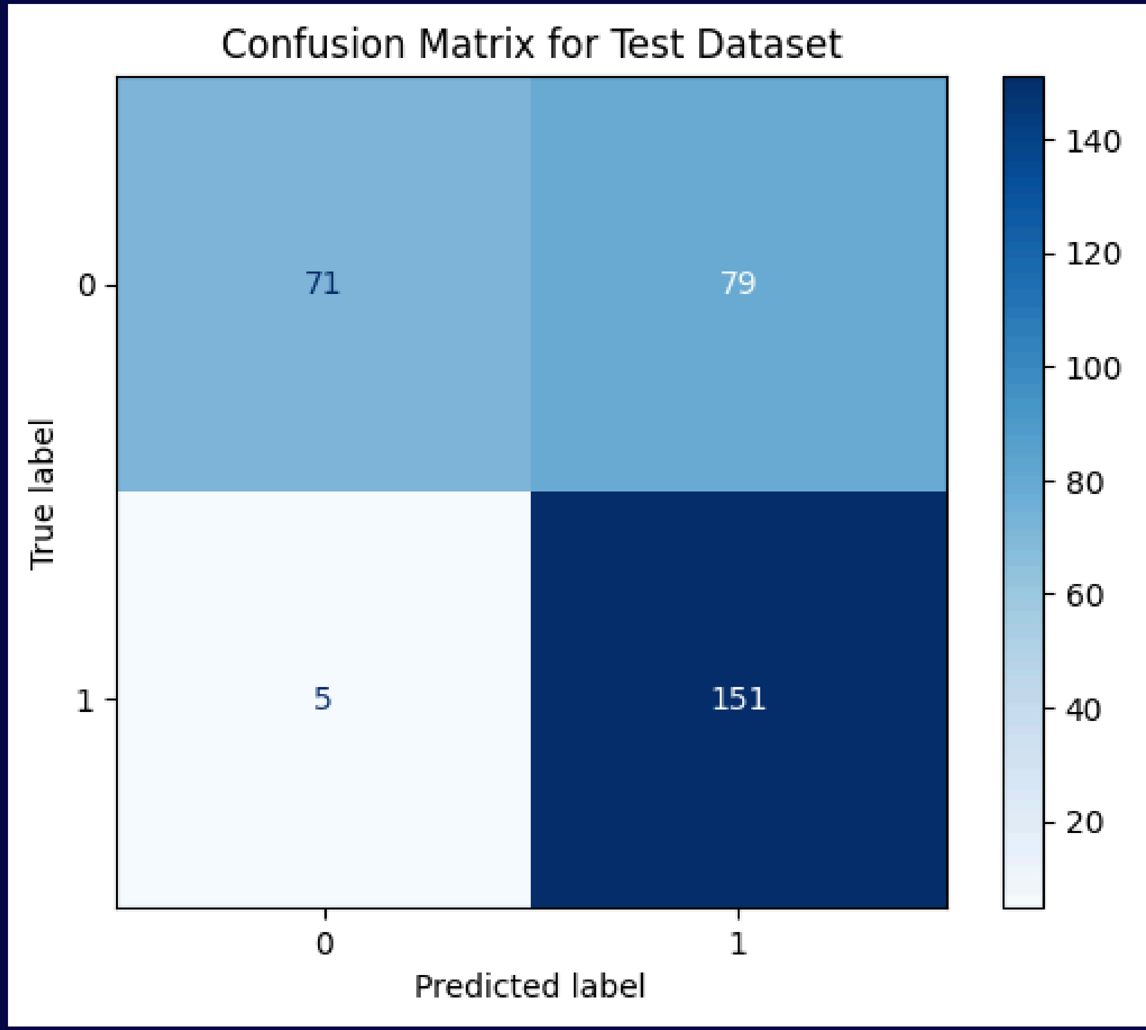
Model	Dataset	Accuracy	Loss	Precision	Recall	F1-Score	AUC	mAP	IoU	Best IoU
c3d	Training	0.904263	0.222517	0.992614	0.835821	0.907495	0.913923	0.992212	0.345592	0.3456
	Validation	0.784916	0.45763	0.850649	0.708108	0.772861	0.78758	0.900692	0.306766	
	Test	0.830065	0.425666	0.933333	0.717949	0.811594	0.832308	0.924393	0.284775	
resnet	Training	0.67645	0.597061	0.700119	0.734336	0.71682	0.668906	0.808579	0.399535	0.3995
	Validation	0.687151	0.593824	0.704663	0.712042	0.708333	0.685362	0.810057	0.366399	
	Test	0.650327	0.613895	0.638418	0.724359	0.678679	0.648846	0.765505	0.371699	
c3d + resnet	Training	0.837876	0.330532	0.998239	0.710526	0.830161	0.854473	0.995612	0.301878	0.3019
	Validation	0.768156	0.611334	0.990909	0.570681	0.724252	0.782346	0.950285	0.242196	
	Test	0.77451	0.56789	0.948454	0.589744	0.727273	0.778205	0.931075	0.242944	
c3d with Custom loss	Training	**0.98812**	0.469773	0.986301	0.992481	0.989382	0.987552	0.999659	0.388304	0.3883
	Validation	**0.837989**	0.816955	0.863388	0.827225	0.84492	0.838762	0.940092	0.353278	
	Test	**0.866013**	0.783398	0.885906	0.846154	0.865574	0.86641	0.94793	0.33165	
resnet with Custom loss	Training	0.620545	0.899452	0.598304	0.972431	0.740811	0.574683	0.871989	0.527274	0.5273
	Validation	0.631285	0.915808	0.594249	0.973822	0.738095	0.606671	0.883627	0.495512	
	Test	0.539216	0.975607	0.526132	0.967949	0.681716	0.530641	0.833888	0.49316	
(c3d + resnet) with custom loss	Training	0.856744	0.626417	0.795613	1	0.886174	0.838073	0.995533	0.45046	0.4505
	Validation	0.790503	0.878001	0.732	0.958115	0.829932	0.778459	0.922001	0.433586	
	Test	0.72549	0.918901	0.656522	0.967949	0.782383	0.720641	0.94572	0.436305	

CONFUSION MATRIX of TEST SET

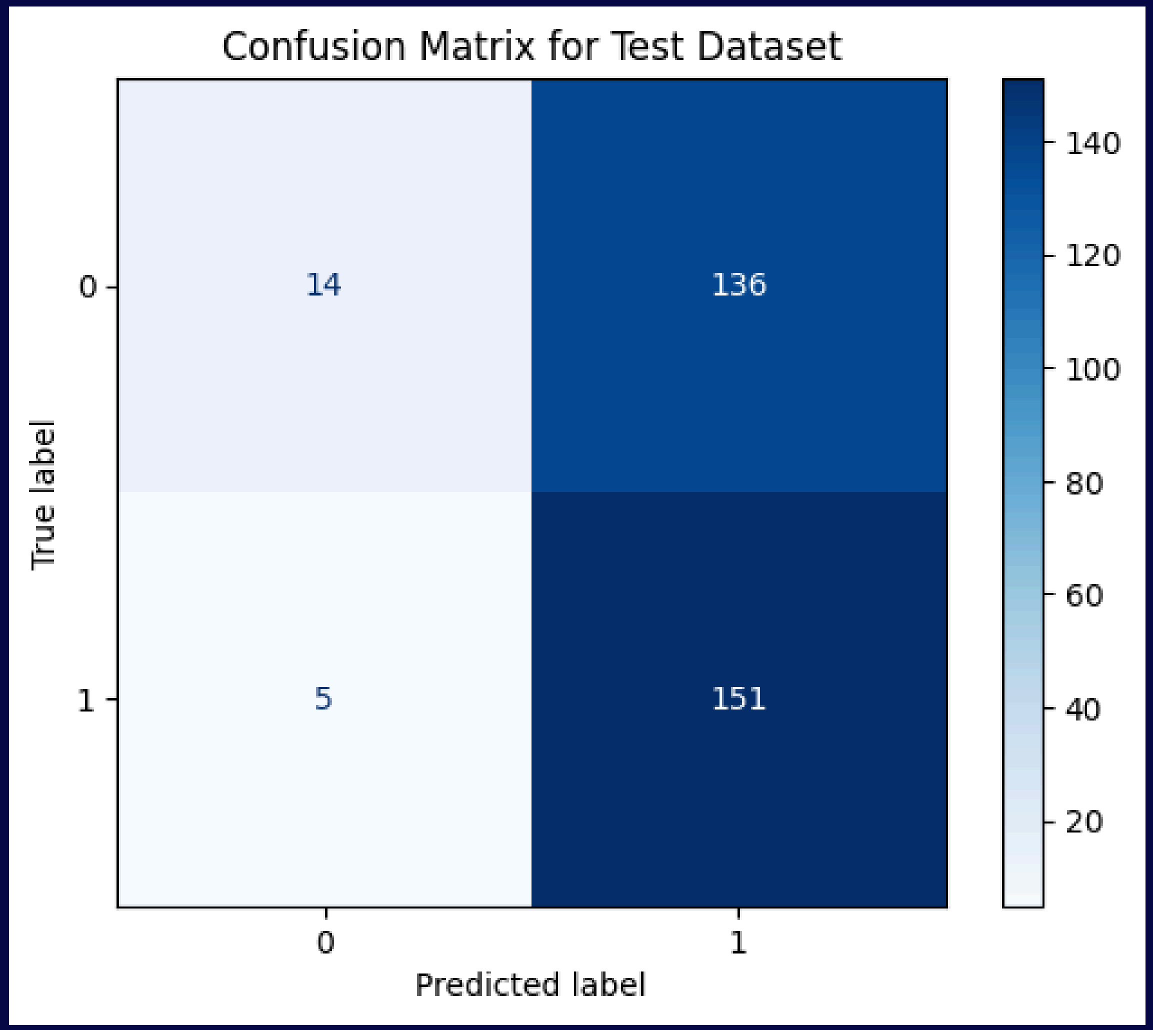
C3D WITH CUSTOM LOSS



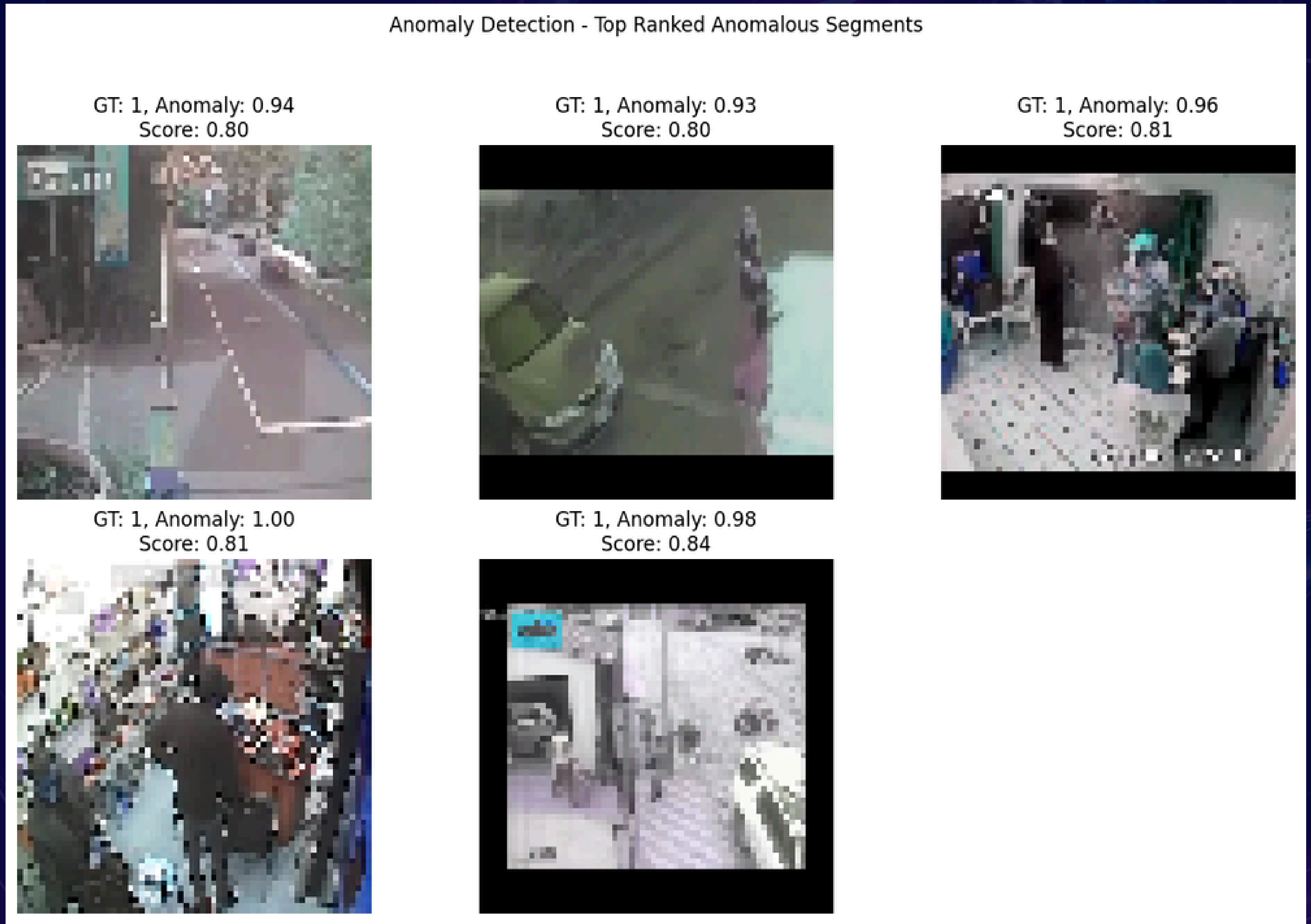
C3D AND 3D RESNET WITH CUSTOM LOSS



3D RESNET WITH CUSTOM LOSS



Ground truth, Prediction and score



CHALLENGES FACED

Overview:

- Labeling Anomalies: Difficulty in accurately labeling segments containing anomalies due to varying patterns and noise.
- Data Imbalance: The imbalance between normal and anomalous segments, leading to potential bias in training.
- Efficient Storage: Ensuring that segment storage in .npy format remains efficient while maintaining high-quality data.
- Model Overfitting: Preventing overfitting due to a limited number of labeled anomaly segments for training.

Key Takeaways:

- Implementing techniques to handle class imbalance, such as oversampling or cost-sensitive learning.
- Optimizing storage formats for easy access without compromising performance.

FUTURE ENHANCEMENTS

Overview:

- Real-Time Detection: Implement a real-time anomaly detection pipeline for live video feeds, making it suitable for monitoring applications.
- Data Augmentation: Augmentation techniques to address data imbalance and improve model robustness.
- Multimodal Data: Integrate additional modalities (e.g., audio or sensor data) to enhance the accuracy of anomaly detection in complex traffic scenarios.

Possible Enhancement:

- Focus on leveraging more advanced techniques for handling complex, real-time anomaly detection tasks.
- Prioritize model interpretability for better integration and trust in practical applications.

THANK YOU