

Gaps and Results

Lightweight and Explainable **Criminal Activity Detection**

Muhammad Mohib Qureshi

25K2617

Identified Research Gaps

Our literature review reveals four critical gaps:

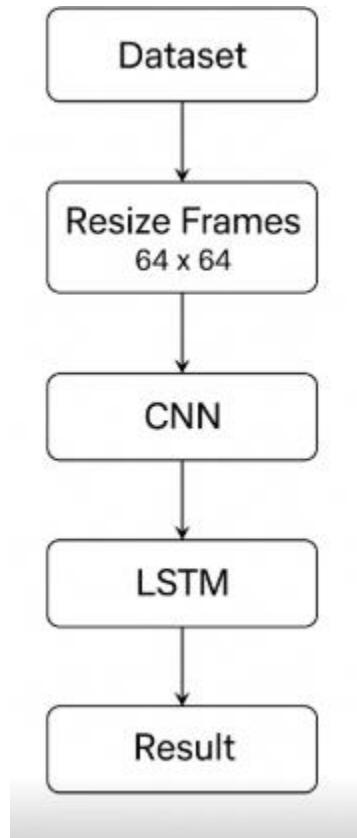
1. **Computational Efficiency:** Most state-of-the-art models employ heavy architectures (ResNet, DenseNet, YOLO) that are impractical for real-time edge deployment. There is a clear need for lightweight alternatives that maintain competitive accuracy.
2. **Limited Generalization:** The overwhelming focus on UCF-Crime for evaluation raises concerns about model generalization. Cross-dataset testing is rarely performed, leaving questions about real-world applicability unanswered.
3. **Lack of Explainability:** The absence of interpretable explanations in current systems limits their practical utility. Security personnel need to understand *why* the system flagged an event to make informed decisions.

Paper result vs Implementation result

The paper proposes an LSTM–CNN model for anomaly detection in surveillance videos, focusing on selecting only the most relevant features—mainly those related to human motion—while removing redundant background information. CNN layers extract spatial features from each video frame, and the LSTM captures temporal relationships across frames.

In the implemented code, video frames are resized to $64 \times 64 \times 3$ (RGB) and passed through TimeDistributed CNN layers with 64 filters and 3×3 kernels, followed by an LSTM layer with 64 units for temporal learning. A final softmax layer classifies activities. This setup aligns with the paper’s approach of combining spatial and temporal feature extraction for effective anomaly detection.

Implementation Flow

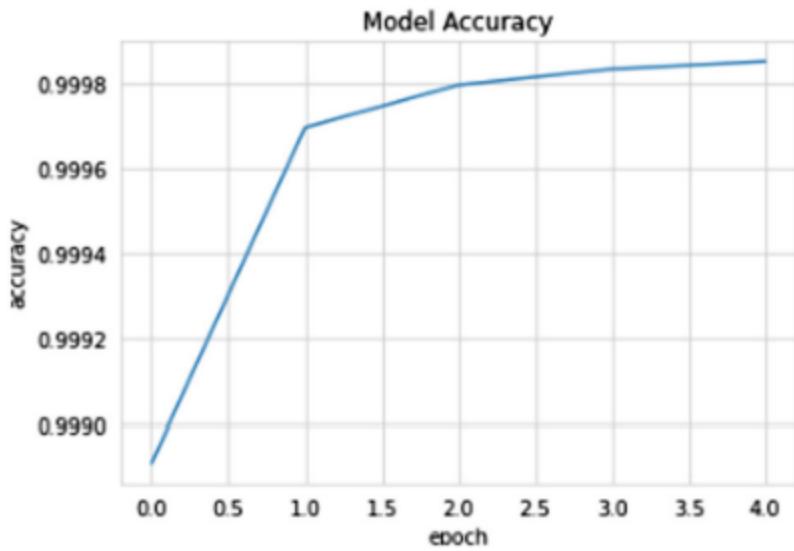


Parameter ,Dataset and Class:

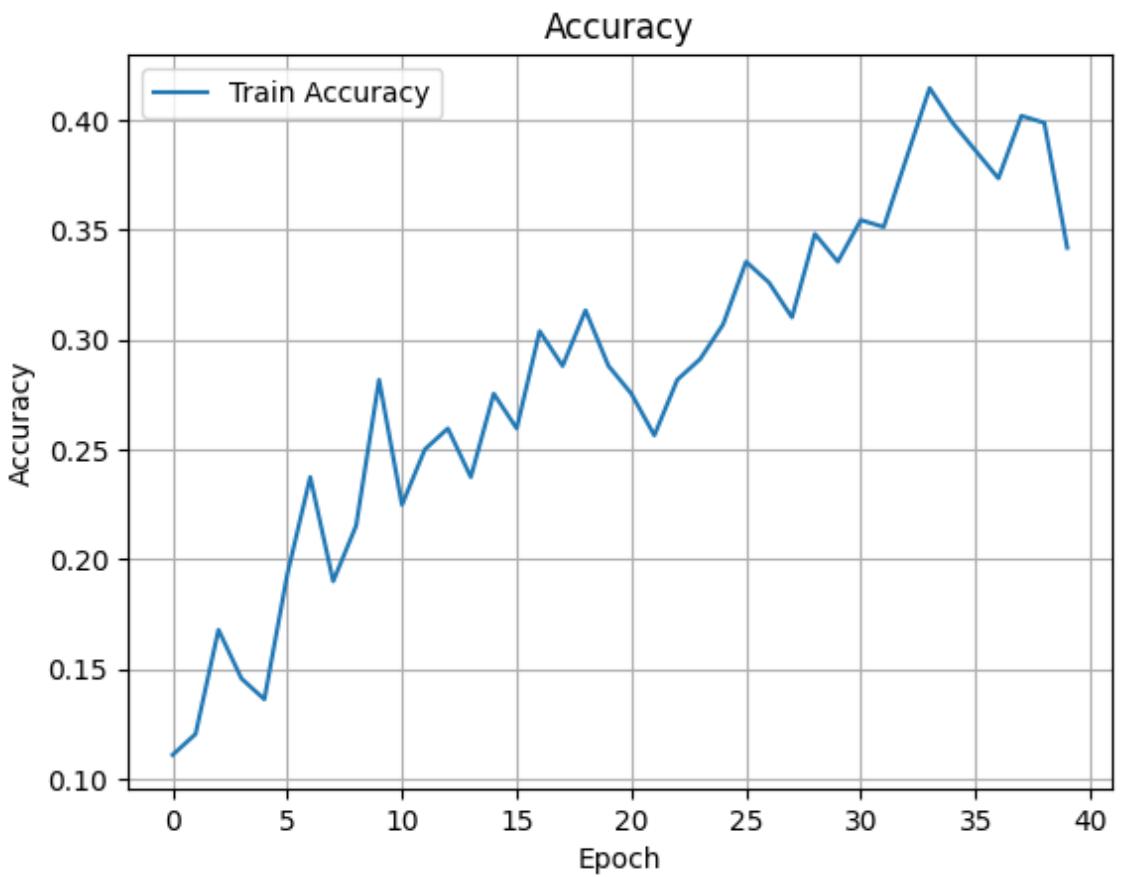
TYPE	Dataset	Classes	Total parameters	Trainable Parameter	Non-trainable parameters
Paper LSTM	UFC-Crime	14	7,170,510	5,586,254	1,584,256
Implementaion	Anomolous Activity	10	4,250,314	4,211,594	38,720

Accuracy Per Epoch:

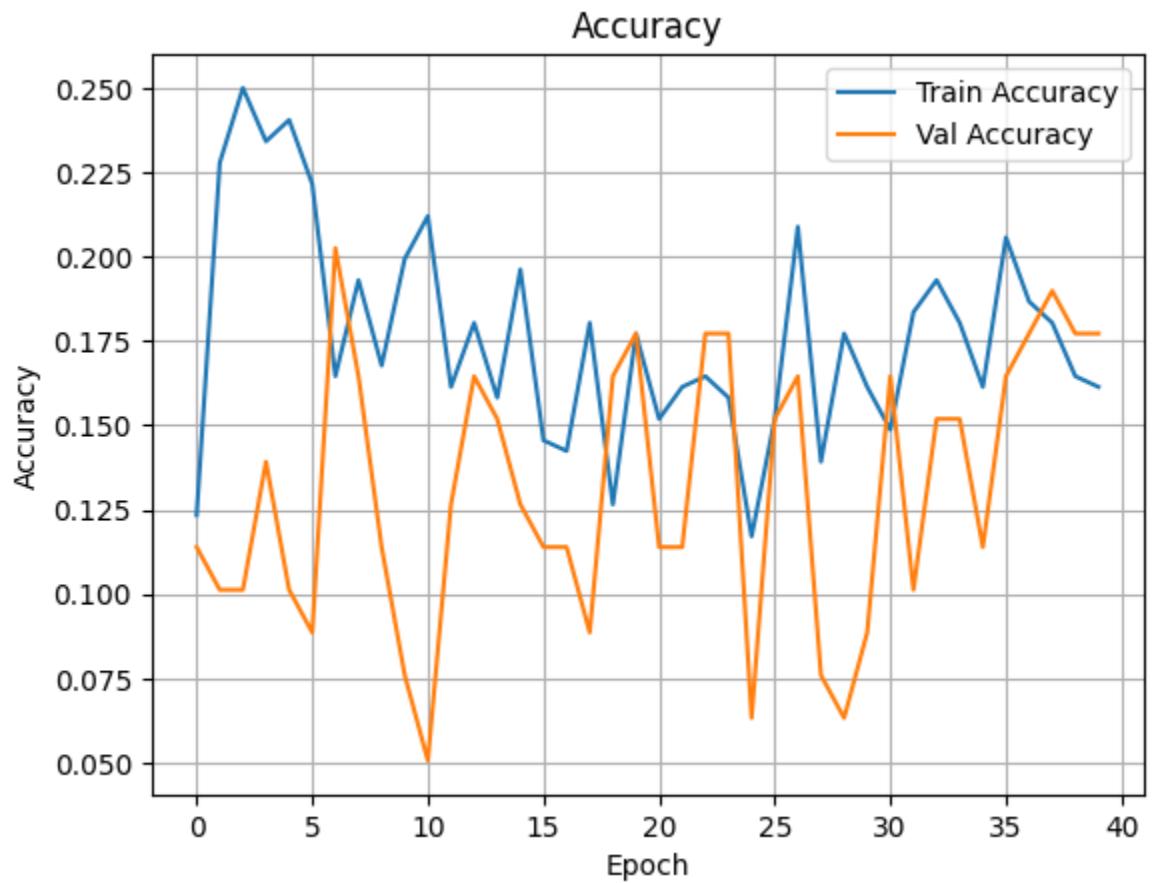
Paper:



Implementation:



Innovation:



Performance Metrics

Paper

	precision	recall	f1-score	support	
0.0	0.00000000000000	0.00000000000000	0.00000000000000	0.00000000000000	297
1.0	0.02922422954304	0.01634472511144	0.02096436058700	3365	
2.0	0.14473456683382	0.29967776584318	0.19519589552239	2793	
3.0	0.00000000000000	0.00000000000000	0.00000000000000	0.00000000000000	2657
4.0	0.03177570093458	0.01332114405119	0.01877243029355	7657	
5.0	0.15202444614209	0.03056835637481	0.05090164982734	6510	
6.0	0.00289136720363	0.00568643379366	0.00383351588171	1231	
7.0	0.71411345807385	0.77503387116640	0.74332754992802	64952	
8.0	0.11628348850466	0.22981599699587	0.15442846328539	2663	
9.0	0.00013559322034	0.00119760479042	0.00024360535932	835	
10.0	0.03797468354430	0.00157273918742	0.00302038761641	7630	
11.0	0.36850649350649	0.05955660501115	0.10254093732355	7623	
12.0	0.07916241062308	0.23437500000000	0.11835072537541	1984	
13.0	0.01370757180157	0.01890189018902	0.01589103291714	1111	
accuracy			0.47709957954505	111308	
macro avg	0.12075242928082	0.12043229517961	0.10196218242266	111308	
weighted avg	0.46450534360943	0.47709957954505	0.45679368812060	111308	

Implementation:

Classification Report:

	precision	recall	f1-score	support
0	0.21	0.20	0.21	15
1	0.00	0.00	0.00	8
2	0.35	0.50	0.41	12
3	0.00	0.00	0.00	6
4	0.00	0.00	0.00	4
5	0.00	0.00	0.00	4
6	0.27	0.75	0.40	4
7	0.50	0.14	0.22	7
8	0.23	0.71	0.34	7
9	0.64	0.58	0.61	12
accuracy			0.32	79
macro avg	0.22	0.29	0.22	79
weighted avg	0.27	0.32	0.27	79

Innovation:

Classification Report:

	precision	recall	f1-score	support
0	0.10	0.07	0.08	15
1	0.17	0.12	0.14	8
2	0.44	0.33	0.38	12
3	0.00	0.00	0.00	6
4	1.00	0.50	0.67	4
5	0.00	0.00	0.00	4
6	0.08	0.25	0.12	4
7	0.14	0.29	0.19	7
8	0.45	0.71	0.56	7
9	0.12	0.08	0.10	12
accuracy			0.22	79
macro avg	0.25	0.24	0.22	79
weighted avg	0.23	0.22	0.21	79

AUC Curve:

Paper:

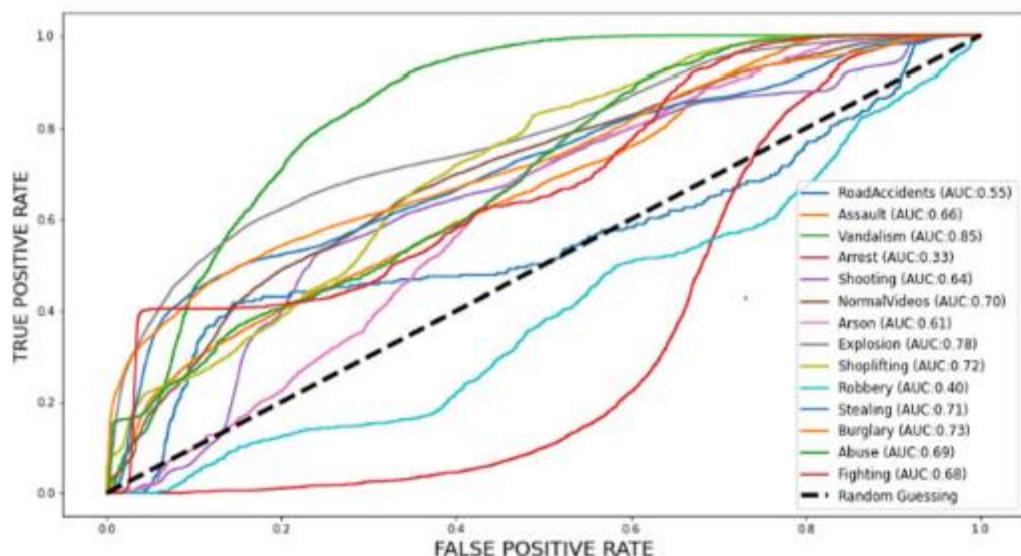
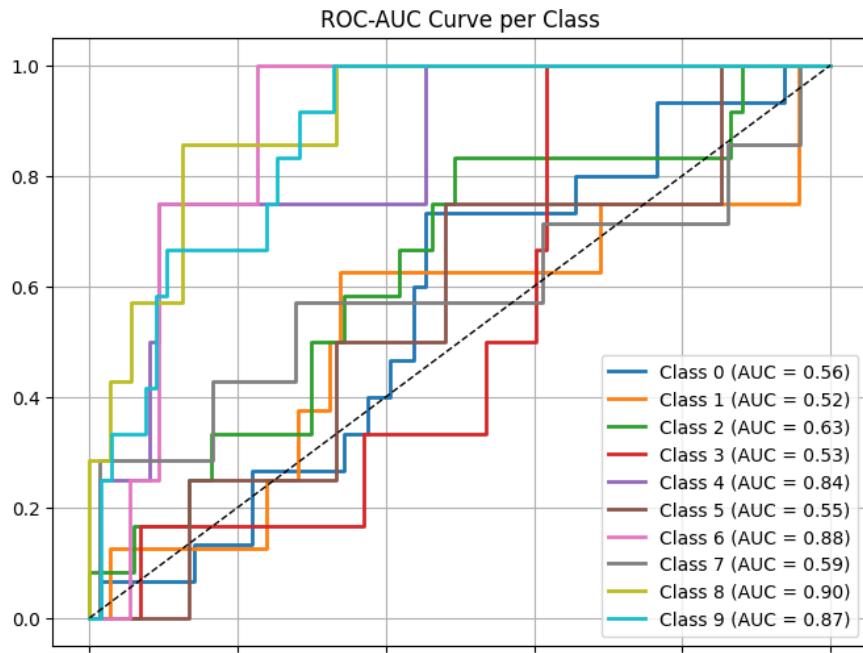
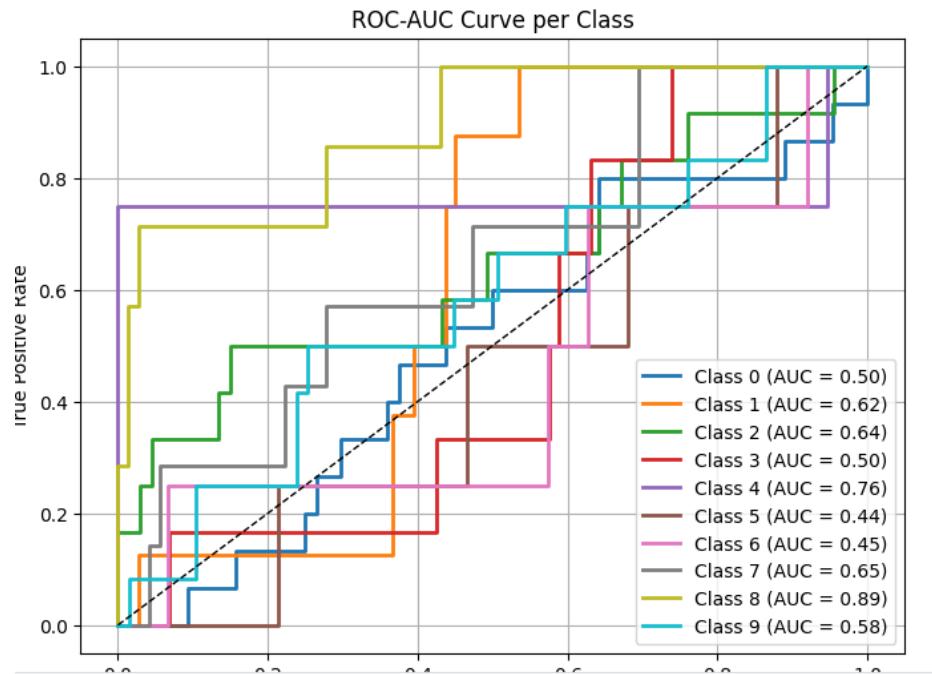


Fig. 8 AUC curve with learning rate 0.003

Implementation:



Innovation:



Conclusion:

Compared to the results reported in the referenced paper, the implementation produced relatively low accuracy and precision. The performance gap is primarily due to the limitation of small dataset size, insufficient training epochs, and suboptimal settings for the learning rate. These factors seriously limit the capabilities of this model to fully capture the spatial-temporal patterns within the video data due to underfitting.

Some improvements may be made to achieve better results: increasing the size of the dataset or using data augmentation to make it more diverse, training the model for more epochs so it can have a deeper learning process, and tuning the learning rate for smoother convergence.

Paper:

[1] K. Ganagavalli and V. Santhi (2024), “YOLO-based anomaly activity detection system for human behavior analysis and crime mitigation,2024

Code:

<https://github.com/mohib989/AI-Assignment>