

Optimizing Deep Learning for Video Violence Detection: A Comparative Analysis of Recent Techniques

Vaishali M Bagade

Research Scholar, Department of Electrical and Electronics Engineering, School of Engineering and Technology, Sandip University Nashik, India

Dr. Jagdish B Helonde

Department of Electrical and Electronics Engineering, School of Engineering and Technology, Sandip University Nashik, India

Abstract:

Automated violence detection in surveillance video is essential for timely intervention in public-safety, transportation, healthcare, and smart-city environments. However, video-based detection must simultaneously achieve reliable recognition, low inference latency, limited memory consumption, and robustness to illumination changes, occlusion, camera motion, and visually similar nonviolent actions. This paper reviews and compares five recent optimization-oriented approaches for deep-learning-based violence detection: Modified Shuffled Shepherd Optimizer (MSSO)-optimized CNN-LSTM, Golden Jackal Optimization (GJO)-optimized ShuffleNetv2-LSTM, attention-enhanced three-dimensional convolutional neural networks (3D-CNNs) with unsupervised keyframe selection, Tuna Swarm Optimization (TSO)-optimized Residual-DenseNet, and identity-aware graph convolutional networks (GCNs). Metaheuristic optimizers improve hyperparameter selection and convergence, whereas keyframe, attention, and graph-based approaches reduce input redundancy or model interactions among individuals. Comparative analysis indicates that lightweight backbone models optimized through GJO or MSSO are appropriate for latency-sensitive edge applications, while graph-based and 3D-CNN approaches offer stronger spatiotemporal interaction modeling but impose higher computational demands. The study identifies accuracy, speed, memory, interpretability, and cross-dataset robustness as the principal selection criteria. It concludes that hybrid lightweight architectures incorporating selective frame sampling, attention, and constrained hyperparameter optimization offer a practical direction for real-time deployment.

Keywords:

Violence detection, video surveillance, deep learning, metaheuristic optimization, CNN-LSTM, graph convolutional network, 3D convolutional neural network, key frame selection.