**Fight Detection**

[1] Nasaruddin, Nasaruddin, Kahlil Muchtar, Afdhal Afdhal, and Alvin Prayuda Juniarta Dwiyantoro. "Deep anomaly detection through visual attention in surveillance videos." Journal of Big Data 7, no. 1 (2020): 87.

The paper discusses the growing importance of automated video anomaly detection for public safety. It proposes a hybrid approach combining background subtraction and bilateral filtering to localize attention regions efficiently. These regions are then fed into a 3D CNN for action recognition and anomaly detection. The method aims to overcome challenges such as occlusions and motion blur in untrimmed public video footage. Experimental results demonstrate the effectiveness of the proposed approach on the UCF-Crime dataset.The proposed anomaly detection method achieved 99.25% accuracy on the UCF-Crime dataset using a standard PC setup. It combines background subtraction and bilateral filtering, outperforming previous methods in detecting various anomalies.

[2] Zhou, P., Ding, Q., Luo, H. and Hou, X., 2017, June. Violent interaction detection in video based on deep learning. In Journal of physics: conference series (Vol. 844, No. 1, p. 012044). IOP Publishing.

Existing literature lacks attention on detecting violent interactions despite their significant impact on security. This study introduces FightNet, a specialized convolutional network for automatically detecting violent behaviors in videos using computer vision techniques. By leveraging acceleration fields alongside RGB images and optical flow fields, FightNet achieves high accuracy on the proposed Violent Interaction Dataset (VID). Compared to traditional methods, FightNet demonstrates superior performance and efficiency, underscoring the need for further research in this area.FightNet achieves up to 97.05% accuracy on VID, outperforming traditional representations, with 100% accuracy on the "Movies" dataset. These results underscore the efficacy of deep ConvNets for robust violent interaction detection in diverse scenarios.

[3] Bermejo Nievas, Enrique, Oscar Deniz Suarez, Gloria Bueno García, and Rahul Sukthankar. "Violence detection in video using computer vision techniques." In Computer Analysis of Images and Patterns: 14th International Conference, CAIP 2011, Seville, Spain, August 29-31, 2011, Proceedings, Part II 14, pp. 332-339. Springer Berlin Heidelberg, 2011.

Recent advancements in violence detection in video have focused on leveraging local descriptors and spatiotemporal analysis for improved accuracy and applicability. This study introduces a hockey fight dataset and evaluates STIP and MoSIFT descriptors for violence detection, achieving around 90% accuracy with the bag-of-words approach. These findings highlight the potential of these methods for practical applications in surveillance and content moderation, underscoring the importance of robust video analysis techniques in ensuring online safety and security.

[4] Zhou P, Ding Q, Luo H, Hou X. Violence detection in surveillance video using low-level features. PLoS one. 2018 Oct 3;13(10):e0203668.

The literature survey addresses the challenge of detecting violent activities in surveillance videos. Traditional methods include violence-related characteristics and spatiotemporal descriptors like STIP and MoSIFT. Recent advancements involve deep learning approaches like multi-stream deep neural networks and convolutional LSTM models. However, challenges like overfitting and computational complexity persist. The proposed method extracts low-level visual features (LHOG and LHOF) from motion regions and processes them with the BoW framework, showing superior performance across datasets.The proposed method outperforms state-of-the-art techniques in violence detection, achieving 100% accuracy on the BEHAVE dataset and significantly improving accuracy on the Crowd Violence dataset, demonstrating effectiveness in detecting violent behaviors in diverse scenarios.

[5] Sudhakaran, Swathikiran, and Oswald Lanz. "Learning to detect violent videos using convolutional long short-term memory." 2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS). IEEE, 2017.

Existing techniques for violence detection in videos have relied on handcrafted features, lacking the ability to generalize effectively. Deep learning methods have shown promise but have been underexplored in this domain. We propose an end-to-end trainable deep neural network model leveraging convolutional long short-term memory (convLSTM) for encoding spatial and temporal changes. This approach outperforms traditional methods and is validated on benchmark datasets including the Hockey Fight Dataset, Movies Dataset, and Violent-Flows Crowd Violence Dataset.  
The deep neural network achieved 97.1% accuracy on Hockey Fight Dataset and 100% on Movies Dataset, surpassing prior methods. Leveraging convLSTM, it outperformed LSTM with fewer parameters (9.6M vs. 77.5M), ensuring better generalization.

[6]Abdali, Al-Maamoon R., and Rana F. Al-Tuma. "Robust real-time violence detection in video using cnn and lstm." In 2019 2nd Scientific Conference of Computer Sciences (SCCS), pp. 104-108. IEEE, 2019.

The proposed violence detection system integrates CNN + LSTM and Human Pose Estimation + LSTM approaches, evaluated on Hockey Fight, Movies, and Violent-Flow datasets. Hyperparameter tuning reveals ResNet50 as the optimal CNN architecture, with learning rate, augmentation, and sequence length significantly affecting model performance. Future work includes audio inference integration and priority-based scheduling for efficient computation allocation based on violence probability. ResNet50 performed best with 90% accuracy, followed closely by InceptionV3 at 89%. A learning rate of 0.0001 yielded optimal results, with 4.5% accuracy improvement from augmentation. However, dropout of 50% did not enhance performance, resulting in 86% accuracy. Unretrained CNN weights led to only 61% accuracy.

[7] Accattoli, Simone, Paolo Sernani, Nicola Falcionelli, Dagmawi Neway Mekuria, and Aldo Franco Dragoni. "Violence detection in videos by combining 3D convolutional neural networks and support vector machines." Applied Artificial Intelligence 34, no. 4 (2020): 329-344.

The study introduced a novel violence detection method using a pre-trained 3D Convolutional Neural Network (C3D) and SVM classifier, achieving high accuracy in both person-to-person and crowd violence scenarios, surpassing existing techniques. It showcased deep learning's potential in violence detection and emphasized feature extraction and classification. Future research aims to improve accuracy, optimize real-time implementation, and explore multi-class classification for different violent behaviors.The proposed approach achieved high accuracy rates of 98.51% on the Hockey Fight dataset and 97.3% on a combined dataset, outperforming existing methods. False negatives were primarily due to absence of aggression in some video segments, while false positives often stemmed from friendly behaviors resembling violence.

[8] Freire-Obregón D, Barra P, Castrillón-Santana M, Marsico MD. Inflated 3D ConvNet context analysis for violence detection. Machine Vision and Applications. 2022 Jan;33(1):15.

The literature survey presents an advanced approach for violence detection in videos, leveraging deep learning models like Deep SORT and SiamRPN+ for subject tracking and context analysis. Results show the importance of context in classification accuracy, with a trade-off between performance and computational efficiency. Additionally, the study highlights challenges in cross-dataset classification, emphasizing the need for further research. Overall, the survey offers valuable insights for applications like CCTV surveillance and content filtering.Simplifying context leads to significant accuracy reductions: about 5% on the AVD Dataset and 2% on the Hockey Fight Dataset. Despite this, computational advantages emerge, with a 36% reduction in frames for the Hockey Fight Dataset and 10% for the AVD Dataset.

[9] Li, Hongchang, Jing Wang, Jianjun Han, Jinmin Zhang, Yushan Yang, and Yue Zhao. "A novel multi-stream method for violent interaction detection using deep learning." Measurement and Control 53, no. 5-6 (2020): 796-806.

Recent advancements in automatic detection of violent interactions in videos have been driven by deep learning techniques, particularly through multi-stream fusion models. These models combine attention-based spatial RGB, temporal (using optical flow), and local spatial streams to enhance detection accuracy despite challenges like occlusion and motion scale changes. Exploring various fusion modes has provided insights into optimizing multi-modal strategies for violent interaction detection tasks.Our multi-stream fusion model outperforms existing methods in detecting violent interactions. For example, on the hockey fight dataset, we achieve 99.5% accuracy compared to FightNet's 97.0% and TSN's 98.0%. Similarly, on the VID dataset, our model achieves 98.58% accuracy, surpassing TSN's 98.32%. These results highlight the effectiveness of our approach in improving detection accuracy for violent interactions.

[10] Vijeikis, Romas, Vidas Raudonis, and Gintaras Dervinis. "Efficient violence detection in surveillance." Sensors 22, no. 6 (2022): 2216.

The paper proposes an efficient AI-based method for detecting violent activities in surveillance footage. It combines spatial feature extraction using MobileNet V2 with temporal analysis via LSTM, achieving high accuracy (82%) on real-life datasets like RWF-2000. Compared to existing methods, it offers a lightweight solution with promising results. The study highlights the importance of real-time violence detection for public safety and suggests future research directions for improving model performance and applicability in diverse scenarios.

[11] Naik AJ, Gopalakrishna MT. Deep-violence: individual person violent activity detection in video. Multimedia Tools and Applications. 2021 May;80(12):18365-80.

The literature survey explores methods for video surveillance and human activity recognition, focusing on violence detection. It discusses optical flow-based, histogram-based, space-time interest points-based, and convolutional neural network-based approaches. The proposed framework integrates Mask-RCNN for single-person detection, key-point detection for pose estimation, and LSTM for temporal modeling. The proposed model demonstrates superior performance compared to existing methods, achieving up to 93.4% accuracy on the KTH dataset. Confusion matrices show accurate classification, with some misclassifications attributed to detection issues. AUC-ROC analysis confirms the model's effectiveness in distinguishing between violent and non-violent activities.

[12] Deepak, K., L. K. P. Vignesh, and S. J. I. E. Chandrakala. "Autocorrelation of gradients based violence detection in surveillance videos." ICT Express 6, no. 3 (2020): 155-159.

The study explores a simplified approach using auto-correlation of gradients for violence detection in surveillance videos, eschewing deep learning models. By leveraging STACOG features and SVM classifiers, it achieves competitive accuracies of 91.38% and 90.40% on the Crowd Violence and Hockey Fight datasets, respectively. Despite its shallow architecture, the method offers efficient feature extraction comparable to more complex methods like Hough Forests, with higher true positive rates than KNN classifiers.

[13] Omarov, Batyrkhan, Sergazy Narynov, Zhandos Zhumanov, Aidana Gumar, and Mariyam Khassanova. "A Skeleton-based Approach for Campus Violence Detection." Computers, Materials & Continua 72, no. 1 (2022).

The study tackles the global issue of violence in schools, advocating for automated violence detection in surveillance videos using human skeletal data for swift processing. Recent advancements in violence detection systems, particularly those employing deep learning techniques, are highlighted. The proposed skeleton-based approach aims for real-time violence detection, promising practical application potential in schools and other settings without requiring high-performance hardware.The study proposes a skeleton-based approach for real-time violence detection, achieving an accuracy of 95%-99% in video-based violence detection.

[14] Ye, Liang, Tong Liu, Tian Han, Hany Ferdinando, Tapio Seppänen, and Esko Alasaarela. "Campus violence detection based on artificial intelligent interpretation of surveillance video sequences." Remote Sensing 13, no. 4 (2021): 628.

The survey explores AI and deep learning methods for violence detection and emotion recognition from video to audio data, including 3D CNNs, multi-task fusion learning, and cepstrum-based emotion recognition. Existing studies prioritize social violence, prompting the paper's proposal of a novel fusion algorithm to enhance campus violence detection by integrating video and audio evidence.The paper achieves 92.00% accuracy in video-based physical violence classification, while audio-based bullying emotion recognition yields accuracies exceeding 88% across three databases. An enhanced Dempster–Shafer fusion algorithm outperforms other methods, particularly when combined with the Finnish emotional database, achieving the highest average recognition accuracy.

**Road Accidents(New)**

[1] Desai, Rutik, Akash Jadhav, Suraj Sawant, and Neha Thakur. "Accident detection using ml and ai techniques." Engpaper Journal (2021).

The literature survey highlights various approaches to accident detection using computer vision technologies. Studies by Machaca Arceda et al., Gour et al., Gupta et al., Ravindran et al., and Ghosh et al. have explored techniques like YOLO algorithm, convolutional neural networks, and image processing for accident detection with notable accuracies ranging from 89% to 98%. These methods demonstrate the potential for vision-based systems to accurately identify accidents from CCTV footage. Future directions could involve extending these systems to detect other factors such as alcoholic driving or vehicle identification for enhanced safety measures and emergency response.

<https://www.engpaper.com/download/accident-detection-using-ml-and-ai-techniques.pdf>

[2] Ahmed, Shakil, Md Akbar Hossain, Md Mafijul Islam Bhuiyan, and Sayan Kumar Ray. "A comparative study of machine learning algorithms to predict road accident severity." In 2021 20th International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/SmartCNS), pp. 390-397. IEEE, 2021.

The study conducted a literature review on road accident prediction models, comparing the performance of various machine learning (ML) algorithms. It utilized road accident data from New Zealand to evaluate algorithms like logistic regression, k-nearest neighbor, naive Bayes, random forest, extreme gradient boosting, and adaptive boosting for binary and multiclass injury severity prediction. Results showed that ensemble mode algorithms, particularly random forest, outperformed single mode algorithms. This highlights the significance of driver attributes and the effectiveness of ensemble ML approaches in road accident prediction.

<https://ieeexplore.ieee.org/abstract/document/9719675>

[3] Bokaba, Tebogo, Wesley Doorsamy, and Babu Sena Paul. "Comparative study of machine learning classifiers for modelling road traffic accidents." Applied Sciences 12, no. 2 (2022): 828.

The study compared six machine learning classifiers using real-life road traffic accident (RTA) data from Gauteng province, South Africa. It aimed to identify the most effective predictive model for RTAs, employing evaluation metrics such as accuracy, RMSE, precision, recall, and ROC curve. Additionally, dimensionality reduction techniques like PCA and LDA were utilized to improve model performance, aligning with Sustainable Development Goals (SDGs) for road safety. The study analyzed six classifiers with real road traffic accident data from Gauteng, South Africa. Random Forest (RF) consistently outperformed others, achieving 97% accuracy and 93% precision, with the lowest RMSE and highest AUC (99%). Future research could explore hyperparameter tuning for SVM and test additional ML classifiers to enhance predictive capabilities for transportation safety stakeholders.

<https://www.mdpi.com/2076-3417/12/2/828>

[4] Megnidio-Tchoukouegno, M., & Adedeji, J. A. (2023). Machine learning for road traffic accident improvement and environmental resource management in the transportation sector. Sustainability, 15(3), 2014.

This study explores traffic accident severity and its causes, using machine learning to analyze UK traffic data from 2020. It highlights the potential of ML techniques like Decision Tree, LightGBM, and XGBoost in predicting accidents. Evaluation metrics assess model performance. Overall, the study contributes to understanding and mitigating traffic accidents.This study analyzed 135,453 UK traffic accidents in 2020, highlighting the role of vehicle characteristics in predicting severity. Decision Tree models performed well, emphasizing the need for vehicle inspection and policy enforcement. Machine learning accurately predicted fatal (AUC: 0.96), serious (AUC: 0.86), and slight (AUC: 0.96) injuries, suggesting their potential for accident prevention.

<https://www.mdpi.com/2071-1050/15/3/2014>