Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Intelligent Road Safety and Accident Detection System

P. Sucharitha, A. Japnika, B. Prem Sai, K. Kavya, C. Vijaya Lakshmi

Department of CSE RGMCET, Nandyal, India
Email: sucharitha.puvvada@gmail.com

In order to improve road safety, this study presents a brand-new technique for identifying accidents in security footage. Our hybrid system, which combines convolutional neural networks and YOLOv10 object detection, produces a solution that not only quickly detects accidents in real time but also automatically notifies the nearby police station. The CNN enhances the classification of detected incidents, while the YOLOv10 framework is used for precise object identification while preserving a high degree of detection speed and accuracy. Furthermore, when recording an incident, we have created a vehicle license plate detection system that combines to recognize and extract the characters in the license plate number, use YOLOv10 with Easy OCR.

Keywords: Accident Detection, Surveillance Video Monitoring, YOLOv10, Real-Time Response, License Plate Recognition, Intelligent Transportation Systems.

1. Introduction

The world of security and surveillance systems has undergone significant change in recent years due to the astounding advancements in technology, particularly with the current AI and ML technologies. Anomaly detection in video surveillance systems is the most crucial component of these technologies and is essential for ensuring safety and security.

Anomaly detection is the term used to describe the detection of any abnormality or out-of-norm activities in video streams. This feature is an important one, which enables immediate response in cases that range from an automobile accident to violent acts or any other emergency.

Historically, there relied heavily on personnel to look at video feeds throughout the day due to existing surveillance practices. This routine of people constantly watching the TV screens was not only monotonous but also made way for human errors being committed, largely because of tiredness and boredom. Hence, several crucial events remained unattended or were noticed much later in time than when they occurred. The emergence of anomaly detection systems has effectively resolved this limitation by utilizing advanced technologies capable of processing videos and detecting events as they occur. These systems can precisely identify anomalies and events. This means that the amount of human supervision becomes less dependent, and the reaction time is improved towards such abnormalities.

Video analytics has changed and improved since the advent of deep learning methods, particularly Convolutional Neural Networks (CNN). CNNs can recognize particular patterns in video frames because they are known to be visualoriented. A large amount of training data, including both normal and abnormal actions, can be used to perform normal activities. It is possible to automate the remaining tasks that require attention. CNNs can be trained to identify traffic accidents based on abrupt vehicle stops or collisions between vehicles, for instance, in traffic monitoring applications.

It is beyond detection of an accident because it off ers an importance with multiple applicabilities in o their fields and use. For example, the alerts triggere d by occurrences in public safety due to violent acts or intrusions contain problems before they develop to severe crises. In the same fashion, a patient activity for example is video analytics exploitation that makes patients graze in care facilities, falls of patients, wandering avatars thus greedy for p atients' blood to enhance outpatient wellness. By this, AI, and ML based technology, especially when put through video surveillance systems have brought into us a whole different level into the detection of anomaly that increases the safety rate of public life and in how they run.Dedicated for desi gning a hybrid model from YOLOv10 and CNNs for detecting accidents in real time. Vehicle identification also takes place on license plates. Urban safety and law and order are being aided by the continuous growth of these technologies and the expansion of their application sectors. Improving these systems to lower false alarms and address other privacy and data protection concerns is the current challenge. We can only realize the full potential of anomaly detection technologies by using them responsibly and continuously improving them.

2. Literature Review

In the field of surveillance system anomaly detection, there is growing pressure to enhance the detection algorithms' time and accuracy. This review of the literature examines seven noteworthy works that advance the field of study and talks about their methods, conclusions, and inherent limitations.

1. Real-Time Object Detection Using SSD

The goal of this study is to develop a real-time object detection system using deep learning and the SSD algorithm. After using several pre-trained models, the authors have shown remarkable performance of 97% precision in identifying moving and stationary objects using datasets like PASCAL VOC and MS COCO. The system works well for object detection, but it has several drawbacks. For example, it cannot detect objects that are very small or closely packed because it lacks contextual information beyond the bounding boxes that defines area.

2. Improved SSD for Object Detection in Real Time

According to the authors of this work, depth-wise separable convolutions were added to the SSD algorithm to improve classification accuracy without compromising speed. Their model effectively classified objects in real time across a range of devices. Aside from these enhancements, the study found that the SSD algorithm had trouble detecting small objects, which are frequently absent because of insufficient contextual knowledge.

3. YOLO: A Comprehensive Guide

Due to its effectiveness and speed, the YOLO (You Only Look Once) framework is arguably the most widely used object detection algorithm. The most recent version of YOLO, YOLORYOLOv7, can achieve an astounding 3.5ms inference time per frame, according to this article's review of the latest iterations. Despite its remarkable speed, the YOLO algorithm struggles to identify cluttered objects or strange shapes that are closely spaced apart, which could lead to detection failures in complex environments.

4. Real-Time Detection in Stacked Environments Using Chamfer's Algorithm

5. Deep learning for real-time object detection in complex stacked scenarios is the main focus of this study. The authors examined current algorithms and identified persistent problems, such as dynamic target scale and target occlusion with overlapping occludes. To address these issues, they suggested methods like multi-resolution features and hyperpa.

6. Algorithms for Lightweight Object Detection

The goal of this research is to develop efficient, lightweight object detection algorithms for realtime applications in resource-constrained environments. By reducing model complexity without sacrificing performance, the study presents a novel approach that strikes a balance between accuracy and computational efficiency. This approach may perform poorly in scenarios requiring high precision or where larger models could be used, despite its benefits for use on edge devices.

7. Advanced Object Detection Techniques

The issue of detection in a range of complex plane objects is addressed in this work. The authors emphasize that there are engineering applications for real-time performance in this context, so it is crucial that the services offered be suitable for the limitations and lightweight. Despite this optimism, these tools still usually need a lot of fine-tuning and don't appear to generalize across even a small number of datasets.

YOLOv10: Unquestionably Not a Slim Choice Tag-based methods have been replaced by articulated V10 evolution of real-time object detection frameworks, which are aimed at speed and reduced power consumption. The authors showed that YOLOv10 could be more efficient without sacrificing competitive performance. It is evident from the reviewed literature that there have been improvements made to the architecture of anomaly detection systems within the purview of surveillance systems, particularly in relation to the application of deep learning anomalies such as YOLOs and SSDs. Although there is variation in the authors' hypotheses regarding how to increase detection accuracy and overall detection efficiency, some commonalities include difficulties in detecting small or overlapping objects and variations in efficacy among various environmental categories. Future studies should combine experimental and sophisticated model designs to improve situational awareness while maintaining real-time detection capabilities.

3. Methodology

We here suggest a hybrid model based on YOLOv10 and Convolutional Neural Networks since the stateof-the-art for real-time security systems accident detection necessitates a sophisticated level for vehicle recognition. The UCF dataset can be used to identify the anomaly, and a particular Kaggle dataset can be used to recognize the license plate.

A. DATASET

We use a Kaggle dataset and the UCF dataset to build our hybrid number plate recognition model. The UCF dataset is well-known because it contains a large number of short films that generate different mobile activities, such as interacting with vehicles. For the purpose of training models to identify and distinguish between various number plate formats and styles, this dataset provides a wide variety of vehicle types and license plates.

The Kaggle dataset, as an alternative is providing images that are captured at different angles and also using different lighting conditions and episodes that focus on number plates. These areas have license plates with widely disparate fonts, colors, and even designs. These datasets allow us to efficiently train our hybrid model, which combines CNN for character recognition and YOLOv10 for object detection.

In the training process first, all the images have to go through some post processing operations to delete unwanted information. This step of post processing helps in further segmenting the license plates in the remaining portion of the picture. In order to allow the model to learn from the unique characters that multiple license plates present, these images are then added to the model as training samples. Even in complex situations found in most real-world settings, our system is still able to comprehend the license plates thanks to the combination of all these factors.

B. FEATURE EXTRACTION

The feature extraction process forms a significant p art of the hybrid model, which can be seen as supporting the identification of vehicles and even in detecting accidents. We established a consistent and methodical way to preprocess and note data for both number plate recognition and auto accidents during this project. Our model will therefore undoubtedly learn efficiently and generalize the data in a suitable manner.

Data Preparation Process:

We looked through the UCF dataset, which included videos and frames from which information about auto accidents could be gleaned, in order to carry out the project's accident detection component. This involved going over each frame separately and emphasizing regions of interest, which are places that draw attention to significant occurrences like vehicle crashes, abrupt braking or acceleration, irregular driving, etc. These areas are identified in the video frames where the bounding boxes of the vehicles that took part in the odd events are visible, along with this explanation of anomalies. This kind of event description enables the model to pinpoint the salient characteristics associated with various accident types.

The other research parts required a different annotation procedure, and as such, the Kaggle dataset increased in the ULCV classification. Meanwhile, with t his, marking and recognizing the license plate regions in different images became the prime purpose of the annotation. Each of the number plates has been applied with boxes and texts showing the letters and figures on the plates.

This stage also plays a crucial role for the training of the YOLOv10 model.

Data cleaning Steps:

Following annotation, several preprocessing steps were undertaken to enhance the quality of the data before feeding it into our model. For both datasets, we applied image normalization techniques to ensure consistent brightness and contrast levels across different images. This helps mitigate issues caused by varying lighting conditions during data collection. Furthermore, we carried out Using data augmentation methods like flipping, scaling, and rotation, we were able to artificially expand the size of our training dataset.

This not only increases the diversity of training samples but also helps improve the robustness of our model against overfitting. For the accident detection dataset, we also extracted temporal features by analyzing sequences of frames to capture motion patterns associated with accidents. This temporal analysis is critical for understanding dynamic events in video footage.

In summary, our feature extraction process combines thorough annotation with effective preprocessing techniques to create high-quality training data for our hybrid model. By carefully preparing both datasets— one for detecting accidents and another for recognizing number plates— Our goal is to improve our system's overall accuracy and performance in practical applications.

C. DEEP LEARNING MODELS

In our project, we employ a hybrid approach that utilizes the YOLOv10 architecture for both accident detection and vehicle number plate recognition. YOLOv1, In real-time object detection tasks, the most recent version of the You Only Look Once (YOLO) series is well-known for its effectiveness and precision. It is appropriate for use in surveillance systems because of its architecture, which strikes a balance between speed and accuracy.

The YOLOv10 architecture consists of three main components : the head, neck, and backbone.

1. Backbone: The backbone of YOLOv10 is based on CSPDarknet53, a convolutional neural network that excels in feature extraction. Cross Stage Partial (CSP) connections are used in this backbone to improve information flow between layers, improving gradient propagation during training. By capturing hierarchical features from input images, It successfully conveys high-level semantic information and low-level textures that are essential for precise object detection.

2. Neck: A Path Aggregation Network (PANet) is used by the neck component of YOLOv10 to fuse and refine multi-scale features that were extracted by the backbone. Through the facilitation of information flow across various spatial resolutions, this structure improves the model's capacity to detect objects of different sizes. The PANet design allows for improved feature integration, which is essential for detecting both large and small objects in complex environments.

3. Head: Final predictions, such as bounding box coordinates, class probabilities, and objectness scores, are produced by the YOLOv10 head. The use of an anchor-free bounding box prediction method, which streamlines the prediction procedure and lowers the number of hyperparameters, is a significant innovation in this version. This change enhances the model's capacity to represent objects with varying aspect ratios and scales.

In addition to YOLOv10 for accident detection, we integrate Convolutional Neural Networks (CNN) to enhance our model's capability in recognizing characters on vehicle number plates. The CNN processes the segmented images of license plates extracted by YOLOv10, focusing on character recognition tasks. This combination allows for efficient detection of vehicles involved in accidents while simultaneously identifying their registration details.

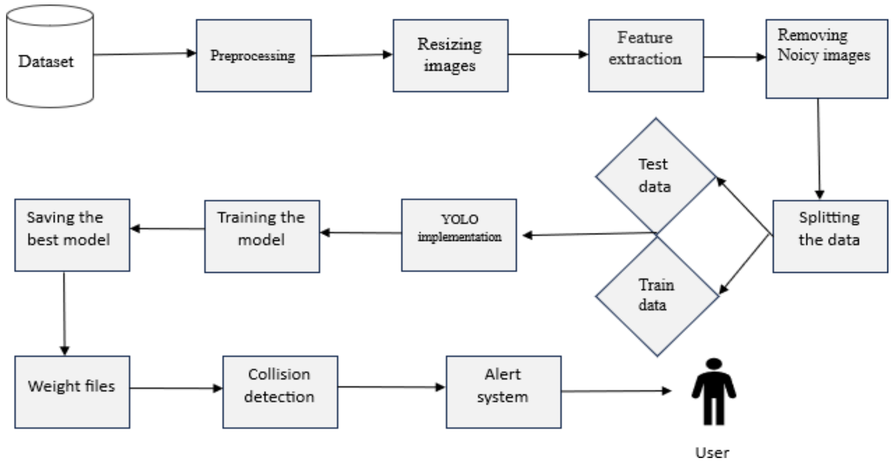

Fig 1 - Architecture Diagram

D. EVALUATION OF MODEL

In this step, we assess the models by feeding each model a dataset containing predictor variables. Based on the prediction outcomes, the models will then forecast the target variable, which we will then compare to actual values.

4. Experimental Results

To validate the effectiveness of the hybrid models developed for accident detection and vehicle number plate recognition in our project, We employed accuracy as the main evaluation metric. In tasks involving classification, accuracy is a commonly used metric that gives a clear picture of a model's performance by calculating the percentage of accurate predictions the model makes.

The following is an expression for the accuracy formula:

$$Accuracy = \frac{Numberofcorrectpredictions}{Totalnumberofpredictions} \qquad (1)$$

The following formula can also be used to determine accuracy in a more specific context:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (2)$$

Whereas

FPFP stands for False Positives, TNTN for True Negatives,

FNFN for False Negatives, and TPTP for True Positives.

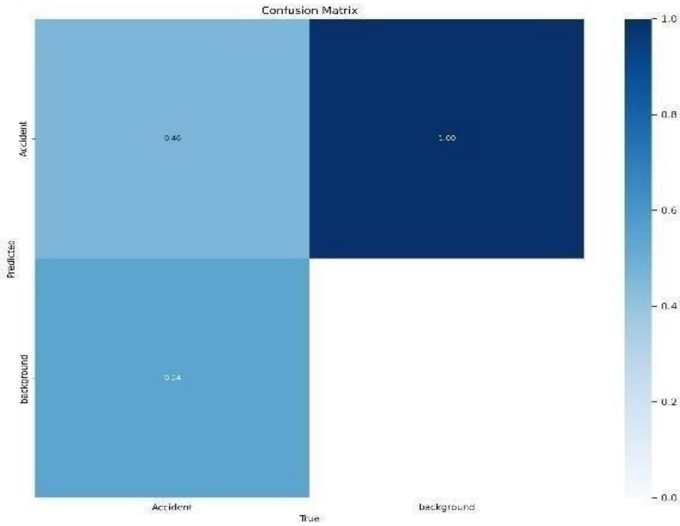

Fig 2 – Confusion Matix

In this project, we trained separate models for two distinct tasks: accident detection and vehicle number plate recognition. Each model was evaluated based on its ability to accurately classify instances as either accidents or non-accidents, and to recognize vehicle registration numbers from images of license plates. The experimental evaluation yielded the following results:

Accident Detection Model: By implementing the YOLOv10 architecture in conjunction with CNN techniques, this model achieved an impressive accuracy of 94% in identifying car accidents from video feeds. The model demonstrated strong performance in recognizing various types of accidents, including collisions and abrupt stops.

Number Plate Recognition Model: Utilizing YOLOv10 specifically for license plate detection, this model successfully recognized vehicle registration numbers with an accuracy of 98.94%. The system effectively identified plates under different lighting conditions and angles, showcasing its robustness in real-world scenarios.

The aforementioned findings demonstrate how well the chosen models identify vehicle number plates and detect accidents. The high accuracy rates demonstrate the robustness of our hybrid approach, which combines advanced deep learning techniques to enhance public safety through timely incident detection and vehicle identification.

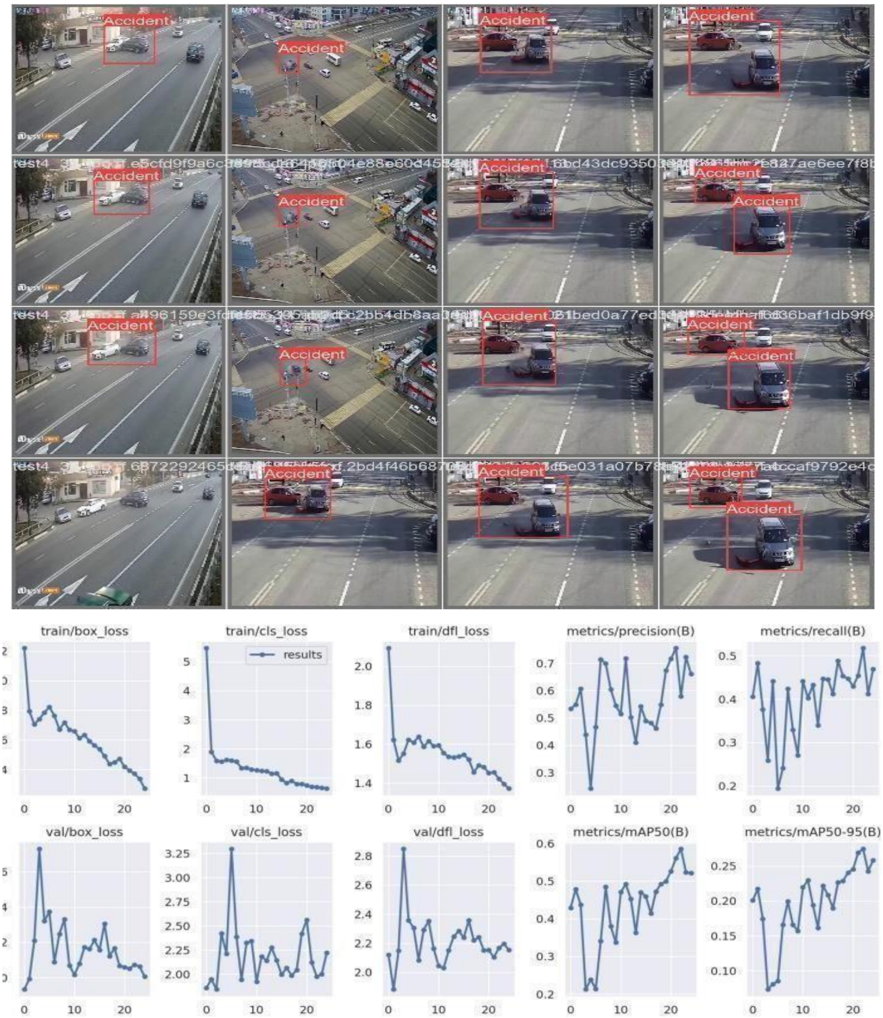

Fig 3 - Loss and Performance Metrics

5. Conclusions and Future Work

In this project, we created a reliable hybrid model for real-time accident detection and vehicle number plate recognition by combining YOLOv10 and CNNs, or convolutional neural networks. Our method successfully tackles important surveillance system issues like precise vehicle identification and prompt incident detection.

94% accuracy rates were shown in the experimental results for accident detection and 98% for number plate recognition, underscoring the effectiveness of our methodologies. The successful implementation of YOLOv10 for both tasks highlights its versatility and efficiency in processing visual data in real-time. By leveraging advanced deep learning techniques, our system not only enhances public safety but also streamlines response efforts by providing law enforcement with immediate alerts and vehicle information during incidents.

Although the results of our current model are encouraging, there are a number of directions that future research could go and improvement. First, we plan to enhance the dataset used for training by incorporating more diverse scenarios, including various weather conditions and lighting situations. This will enhance the model's generalizability and resilience in various contexts.

Second, Our goal is to investigate how extra features, like contextual data, can be integrated. from surrounding traffic conditions or integrating data from other sensors (e.g., radar or LIDAR) to further enhance detection accuracy.

Moreover, addressing the issue of false positives is crucial for improving the reliability of our system. Future work will focus on refining the model's algorithms to reduce misclassifications, particularly in crowded scenes where multiple vehicles are present.

Finally, we intend to investigate the deployment of our model on edge devices to facilitate real-time processing in practical applications. This would allow for quicker response times and broader accessibility in various surveillance settings.

By pursuing these enhancements, we aim to further advance our hybrid model's capabilities, ultimately contributing to safer urban environments and more efficient law enforcement operations.

References

1. Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.
2. Krizhevsky, A., Sutskever, I., & Hinton, G.
3. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 10971105).
4. , R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587)
5. Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., ... & Murphy, K. (2017). Speed/accuracy trade-offs for modern convolutional object detectors. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7310-7311).
6. Law, M. T., & Deng, J. (2018). Cornernet: Detecting objects as paired keypoints. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 734-750).
7. Li, Y., Huang, J., & Yang, W. (2021). Object detection in videos: A survey and a practical guide. arXiv preprint arXiv:2103.01656.
8. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988).
9. Ma, Y., & Zhang, Y. (2021). Deep learning for object detection: A comprehensive review. Neurocomputing, 441, 289-302.
10. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster RCNN: The paper focuses on the development of real-time object detection using region proposal networks. It was published in the conference proceedings of Advances in Neural Information Processing Systems, specifically covering pages 91-99.
11. Wang, X., & Gupta, A. (2018). Videos as space-time region graphs. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 399-417)