



## Deep Learning-Driven Pattern Recognition for Real-Time Traffic Incident Detection in Complex Urban Environments

Yahia Said<sup>1,2\*</sup>, Yahya Alassaf<sup>3</sup>, Refka Ghodbani<sup>1,4</sup>, Taoufik Saidani<sup>1,4</sup>, Olfa Ben Rhaïem<sup>5</sup>

<sup>1</sup> Center for Scientific Research and Entrepreneurship, Northern Border University, Arar 73213, Saudi Arabia

<sup>2</sup> Department of Electrical Engineering, College of Engineering, Northern Border University, Arar 91431, Saudi Arabia

<sup>3</sup> Department of Civil Engineering, College of Engineering, Northern Border University, Arar 91431, Saudi Arabia

<sup>4</sup> Faculty of Computing and Information Technology, Northern Border University, Rafha 91911, Saudi Arabia

<sup>5</sup> College of Science, Northern Border University, Arar 91431, Saudi Arabia

Corresponding Author Email: [Yahia.said@nbu.edu.sa](mailto:Yahia.said@nbu.edu.sa)

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.420231>

### ABSTRACT

**Received:** 25 September 2024

**Revised:** 18 December 2024

**Accepted:** 15 April 2025

**Available online:** 30 April 2025

#### Keywords:

*traffic management, incident detection, road safety, deep learning, intelligent transportation systems (ITS)*

Ensuring traffic safety and efficient management in densely populated urban environments is increasingly critical as global automobile usage surges. Addressing these challenges requires innovative solutions that leverage advanced mathematical modeling and deep learning techniques. Using an improved version of the Fully Convolutional One-Stage Object Detection (FCOS) neural network optimized with Rep-VGG as its backbone, this study introduces a new Intelligent Transportation System (ITS) for real-time traffic incident detection. By conducting extensive experiments on the Highway Incidents Detection Dataset (HWID12), the proposed system demonstrates a detection accuracy of 96.91%. The integration of advanced deep learning and mathematical modeling techniques not only enhances detection accuracy but also improves computational efficiency, making the system suitable for real-time applications in complex urban traffic networks.

## 1. INTRODUCTION

Traffic management and incident detection present an emerging research topic that is rapidly increasing interest in traffic control systems. An important step toward reducing gridlock is the ability to detect the status of incidents on roads. It is crucial for traffic control to have efficient incident detection systems. Incidents involving vehicles have a detrimental effect on traffic flow globally. They could also inflict harm or even death. Consequently, getting traffic back to normal and protecting lives and property on the roads requires incident detection systems that work. It is critical to develop an Intelligent Transportation System (ITS) that can reliably identify various on-road accidents to prevent the loss of countless lives. Due to the growing demand for automobiles and human mobility, traffic congestion presents a growing concern in many countries and cities. This ongoing problem has several recurring causes, including peak-hour traffic congestion. Traffic incidents, on the other hand, are one-time occurrences that interfere with or obstruct traffic flow. Traffic accidents, stationary vehicles, road work, poor weather conditions, and floods can all cause incidents [1, 2]. So, traffic accidents could drastically cut down on road capacity, cause unanticipated delays, and drive-up travel costs [3]. There are many causes of mortality and disability, but one of the most common is traffic accidents. Worldwide, between 20 and 50 million persons sustain injuries or impairments as a result of traffic-related events each year, with an estimated 13 million fatalities [4]. To combat these detrimental effects of incidents,

building a new detection incidents system will be the best solution. To counteract these detrimental impacts of traffic events, the Transportation Incident Management (TIM) module became an essential part of Transportation Management Centers (TMC). Improving the speed and accuracy of traffic event detection and localization is a key goal of traffic incident management. Building a new efficient incident detection system presents the most effective solution to deal with traffic issues.

Any incidents detection system's goal is to accurately and immediately identify any incident as it occurs. Unusual traffic circumstances are a common consequence of incidents, which is why incident detection is implemented. Congestion, delays for traffic coming from the upstream direction, and a backlog of moving cars might result from a major incident that blocks the road.

For efficient traffic control and management in urban networks, incident detection is crucial for swiftly restoring a smooth flow of traffic. To avoid and reduce congestion, an accurate and reliable incident detection system enables the realization of a variety of environments with financial benefits. Urban road networks frequently experience traffic disruptions due to events like accidents, bad weather, and construction projects. These abnormalities cause traffic congestion and delays because the route is underutilized, which in turn has a domino effect of negative consequences on the environment, economy, safety, and security [5]. The root causes of traffic jams and delays have been the subject of extensive study during the last several decades. There are two main categories

of congestion: recurring and nonrecurring. The former is a description of a situation where road capacity is inadequate to meet traffic demand, while the latter is the result of unplanned circumstances such as accidents, severe weather, or special events [6]. The economic and social impacts of nonrecurrent congestion are more far-reaching and more diverse in terms of spatial and temporal aspects compared to those of recurrent congestion. Traffic accidents are significant contributors to poor accessibility and dependability. Numerous studies have been conducted in the field of traffic incident analysis and management as a result of these aspects. By improving prompt issue response, the incident detection function has proven essential in raising the dependability of transportation networks. The severity of congestion and the possibility of additional events can both be decreased through prompt incident response. Therefore, decreasing traffic bottlenecks, lowering operational expenses associated with incident clearance, and efficient and precise detection and verification of traffic incidents is crucial to reestablishing the mobility of the road network.

An increasing number of computer vision problems have been successfully addressed by architectures based on Deep Learning in recent years. Indoor wayfinding [7], object detection [8], pedestrian detection [9], tiredness detection [10], and indoor object recognition [11] have all had promising results when applying deep learning. Developing a trustworthy strategy for traffic accident detection is the major aim of this research. In order to effectively control traffic and prevent congestion, the detection findings might be transmitted to a traffic management center.

The main idea was to design a new efficient and effective incident detection system based on a Deep Learning algorithm to ensure more reliability in urban networks and to ensure better safety for driving conditions. The proposed work will be developed based on a modified version of an anchor-free neural network named “Fully Convolutional One-Stage Object Detection”.

This work advances state-of-the-art in traffic incident detection by addressing key limitations of existing studies in both detection accuracy and computational efficiency. While prior methods such as YOLOv4, SSD, and Faster R-CNN have demonstrated strong performance in general object detection tasks, they often struggle with small and occluded object detection, which are critical in traffic incident scenarios. Our proposed approach integrates the Rep-VGG backbone into the Fully Convolutional One-Stage (FCOS) detection framework, leveraging its efficient structural re-parameterization technique to enhance feature extraction while maintaining low computational overhead. This novel integration enables the model to achieve superior performance on multi-scale and occluded objects, a challenge not adequately addressed in previous studies. Extensive experiments on the Highway Incidents Detection Dataset (HWID12) show that our model achieves a detection accuracy of 96.91%, outperforming state-of-the-art methods.

Furthermore, the lightweight architecture of Rep-VGG allows for real-time inference, ensuring practical deployment in ITS. By combining improved accuracy, efficiency, and robustness, this work sets a new benchmark for real-time traffic incident detection and offers a scalable solution for real-world applications. The remainder of the paper is the following: part 2 features the most current, cutting-edge research on incident and accident detection utilizing Deep Learning methods. The suggested architecture for creating a traffic

incident detection system is described in Section 3. Everything that was tested and found in this study is detailed in Section 4. The paper is concluded in Section 5.

## 2. RELATED WORKS

Since traffic incidents are a leading cause of traffic congestion, it is essential to promptly detect accidents in order to lessen the amount of time they cause disruption and the negative consequences they bring. The development of innovative technologies for the detection of traffic incidents and accidents is crucial for the improvement of traffic conditions, the prevention of their negative impacts, and the maintenance of regular traffic flow in metropolitan areas.

In order to identify incidents and prevent their detrimental consequences, several works have been suggested in the literature. An important part of traffic management and control is the incident detection task, which determines how fast urban networks can return to smooth traffic flow. Several monetary and ecological advantages can be achieved by promptly decreasing congestion via reliable incident detection.

Automatic traffic incident identification using Deep Convolutional Neural Networks (DCNN) was proposed by Zhu et al. [12]. To test the method, we used datasets on Central London traffic and incidents. Because of its better detection rate and lower false positive rate, this method may offer benefits over traditional neural networks, according to the results.

Constant monitoring is necessary to spot fatal incidents in real-world traffic surveillance footage and respond accordingly. A method for the automated detection and location of traffic accidents based on Deep Learning has been suggested by Pawar and Attar [13]. Utilizing a one-class classification algorithm and utilizing spatiotemporal and sequence-to-sequence long short-term memory autoencoders, the method models the video's spatial and temporal representations. The quantitative and qualitative outcomes produced by this approach were satisfactory.

Automated accident detection is a promising new frontier in traffic monitoring systems. These days, traffic control systems and security cameras are standard at most major intersections. As a result, vision-based computer systems could automate the detection of accidents and other traffic events. A real-time accident detection system utilizing Deep Learning was suggested by Ghahremannezhad et al. [14]. The foundation of this system is the YOLO v4 network. The strength of this system is tested using video sequences from YouTube that have different lighting conditions. This system has been shown effective in real-time operations, according to experimental data.

A Deep Learning model was developed by Polson and Sokolov [15] to predict traffic flows. The key contribution is the integration of a linear model fitted with L1 regularization with a sequence of tanh layers into an architecture. Traffic flow estimation is complicated by sharp nonlinearities brought on the changes in state from free flow to breakdown to recovery to congestion. Using Deep Learning to accurately estimate short-term traffic flow, scientists proved that the framework can capture these nonlinear spatiotemporal effects. ITS now face the formidable challenge of ensuring road traffic safety in the face of an alarming global increase in traffic mishaps. In order to increase road safety, it is essential to identify high-risk areas for traffic accidents. Only then may

suitable precautions be implemented. An technique that combines social and remote sensing data with a multi-view learning mechanism was suggested as a solution to this problem [16]. Sending data to traffic management centers in a timely manner can help reduce the negative impacts of crashes.

Crash risk prediction is crucial for highway traffic safety and preventing secondary crashes. The early and precise detection of crashes has long been a focus of research as a means to aid in traffic incident management. A model to detect crashes and estimate crash risk based on Deep Learning was suggested by Huang et al. [17]. When compared to cutting-edge shallow models, the results show that the deep model is superior at detecting and predicting crashes.

The rise in car ownership that accompanies rapid urbanization has been a major contributor to the epidemic of deadly and financially devastating traffic accidents. Estimating the risk of a traffic collision is critical for accident prevention and actively reducing the harm that accidents cause. For the purpose of predicting the likelihood of traffic accidents, Ren et al. [18] presented a method based on Deep Learning. Applying this method to a smart traffic control system can improve traffic flow and forecast organization.

Previous studies on traffic incident detection have failed due to skewed datasets and inadequate sample numbers. In order to meet the demands of traffic management, event detection models should also enhance their real-time capabilities. Li et al. [19] suggested a hybrid paradigm as a solution to these issues. This study extracts geographical and temporal correlations of traffic flow and incident detection using a Generative Adversarial Network (GAN) in conjunction with a Temporal and Spatially Stacked Autoencoder (TSSAE). This helps to increase the sample size and preserve dataset balance. By taking geographical and temporal variables into account, this model surpasses several benchmark models, according to the results.

An event detection model's false alarm and detection rates might be negatively impacted by a small and imbalanced training sample. Generative Adversarial Networks (GANs) are proposed as a novel method for incident detection by Lin et al. [20], which aims to address the problem of inadequate event occurrences. The experimental results suggest that this approach has the potential to enhance traffic incident detection rates while significantly lowering the rate of false alarms.

Accidents involving other vehicles are one of the most common challenges drivers have. With this issue, we have one of the most typical causes of traffic bottlenecks and congestion. An IoV-based Deep Learning system called DeepCrash was proposed by Chang et al. [21] to tackle this issue. Among the components of this system are a front-facing camera and a self-collision detection sensor built within the car, as well as a Deep Learning server and a management platform hosted in the cloud. A corresponding emergency message will be delivered and all relevant data will be forwarded to the cloud server in the event that a vehicle collision is detected.

While many studies have sought to improve methods for detecting traffic incidents, very few have taken a comprehensive approach, considering not only human error but also the potential impact of natural disasters on traffic flow and safety. As part of the planned project, we will be tasked with developing a state-of-the-art incident and disaster detection system to enhance road safety. We take note of the fact that a traffic management center can be linked to the

suggested system in order to keep traffic flow under better conditions. Both in terms of processing speed and detection accuracy, the suggested incident detection system showed very competitive results.

### 3. PROPOSED ARCHITECTURE

Architectures built on Deep Learning have proven to be highly effective in resolving many computer vision and image processing problems. New anchor-free object detector architectures based on Deep Learning have recently been suggested. Both processing time and the complexity of computations for tasks can be significantly reduced with this design.

Prior studies have demonstrated that Rep-VGG achieves a favorable balance between accuracy and computational efficiency by utilizing structural re-parameterization. For instance, Ding et al. [22], who introduced Rep-VGG, showed that this backbone can achieve performance on par with ResNet while being significantly faster at inference.

Unlike backbones such as ResNet or DenseNet, Rep-VGG transforms its multi-branch architecture into a single-path structure at inference time. This significantly reduces computational overhead, making it ideal for real-time applications like traffic incident detection.

To prevent traffic jams and keep traffic moving smoothly, the authors of the proposed study tweaked an existing fully convolutional one-stage objects detector called FCOS [23]. Their goal was to create an incident detection system that could be integrated with a traffic management and control system. The per-pixel prediction object identification problem is solved by this architecture, which does away with anchors. This architecture solves the regression problem similarly to segmentation issues. To reduce the amount of calculation and its complexity, more attention has been devoted to eliminating anchors, which require a greater number of aspect ratios and more computation complexity. FCOS presents the most famous free architectures which directly search for objects on the image based on point tiled. FCOS architecture enables fewer predictions per image. The FCOS architecture is based on three main parts: backbone, Feature Pyramid Network (FPN), and Detection head.

**Backbone:** the backbone is generally used to extract the most relevant parts of the input data to ensure better classification, detection, and segmentation performances. The proposed incident detection system was developed based on a modified version UNETR. In the proposed architecture, we modified the network backbone by using the Rep-VGG network [24]. This architecture provides much simpler parameters than those used in state-of-the-art networks. It provides two different architectures, one for training and the second for inference as presented in Figure 1. The modification of the architectures used is performed by a re-parametrization technique and from this fact, the network was named Rep-VGG.

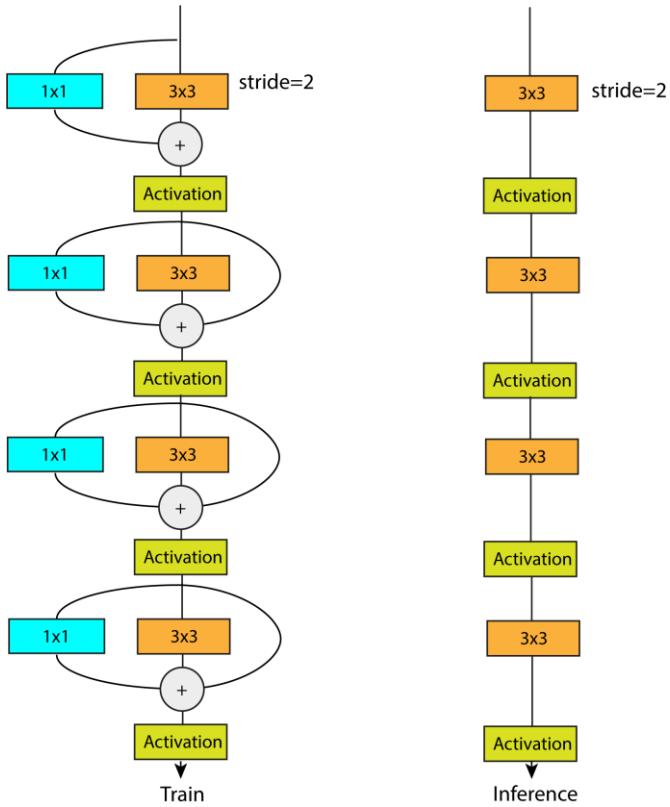
Rep-VGG architecture provides various advantages:

- It provides a simple architecture composed of  $3 \times 3$ ,  $1 \times 1$ , and batch normalization layers.
- The architecture is much simpler than state-of-the-art architecture as it is instantiated without heavy designs.
- The model provides a plain architecture where every layer will be fed by the output of its previous layer.

REP-VGG provides two different architectures. As the

multi-branch architecture provides various benefits in the training stage and various disadvantages in the inference stage. REP-VGG network resolves this problem by providing a multi-branch architecture for training and one branch architecture for inference. Rep-VGG architecture provides an ability to change its architecture via parameter transformation.

Rep-VGG provides a revolutionary architecture that can be useful for different model stages: training and inference. Rep-VGG architecture ensures a good trade-off between time processing and accuracy compared to state-of-the-art models. It provides a structural re-parametrization technique: multi-branch architecture for training and plane architecture for inference.



**Figure 1.** Rep-VGG architecture

The Rep-VGG network presents an easy implementation and is very useful for classification tasks. Also, it provides two variants: Rep-VGG A and Rep-VGG B as provided in Table 1.

**Table 1.** Rep-VGG architecture specification [24]

Stage	Output Size	Rep-VGG-A	Rep-VGG-B
1	$112 \times 112$	$1 \times \min(64, 64a)$	$1 \times \min(64, 64a)$
2	$56 \times 56$	$2 \times 64a$	$4 \times 64a$
3	$28 \times 28$	$4 \times 128a$	$6 \times 128a$
4	$14 \times 14$	$14 \times 256a$	$16 \times 256a$
5	$7 \times 7$	$1 \times 512b$	$1 \times 512b$

The Rep-VGG architecture showcases a multiplier  $b$ , which, when  $b$  greater than  $a$ , allows the final layer to store more detailed information for use in categorization or subsequent processes. To scale the first four stages and the final stage, it also gives a multiplier  $a$  and  $b$ , respectively.

An important factor in improving the efficiency and effectiveness of neural network designs is the activation function. To address several issues and decrease the

computational complexity of Deep Learning models, the Scaled Polynomial Constant Unit Activation Function (SPOCU) [25] was created. SPOCU outperformed state-of-the-art activation functions like SELU [26] and RELU [27] in terms of performance. Eq. (1) defines the SPOCU activation function, which is used as the activation in the Rep-VGG.

$$S(x) = ah\left(\frac{x}{\gamma} + \beta\right) - ah(\beta) \quad (1)$$

where,  $\beta \in (0,1)$ ,  $\alpha, \gamma > 0$  and

$$h(x) = \begin{cases} r(c), & x \geq c \\ r(x), & x \in [0, c) \\ 0, & x < 0 \end{cases}$$

with  $r(x) = x^3(x^5 - 2x^4 + 2)$  and  $1 \leq c < \infty$

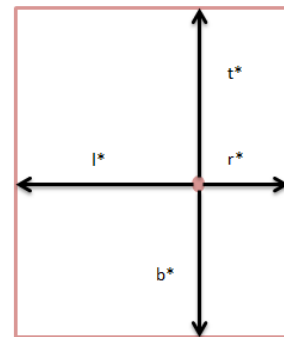
**Feature Pyramid Network (FPN):** By using the FPN network, FCOS network ensures multi-level prediction [28]. Five feature layers are used to generate FPN prediction. The FPN network can identify things at various scales. Shallow layers show greater resolutions with few semantic elements, while deep layers often encode lesser resolutions and are rich in semantic information. Deep layers and shallow layers are joined together by lateral connections. This procedure improves small, medium, and large object detection and localization accuracy. FPN detects objects at five different levels, P3, P4, P5, P6, and P7, with corresponding strides of 8, 16, 32, 64, and 128. The outputs of FPN network are fed through a detection subnetwork which consists of 3 main branches: classification head, center-ness, and regression head [29].

**Detection head:**

**Classification head:** This head predicts a per-pixel probability of the class weighted by the center-ness score with the class probability.

**Center-ness head:** Its intended purpose is to characterize the dispersion of the object's center with respect to the given spot. To improve the predicted bounding boxes' poor quality without adding new hyperparameters, FCOS implements center-ness. Running parallel to the classification branch is the center-ness mind, which is an additional layer branch. The normalized distance from the object's center of gravity is displayed by this attribute. Eq. (2) can be used to calculate the center-ness.

$$\text{centerness} = \sqrt{\frac{\min(l^*, r^*)}{\max(l^*, r^*)} * \frac{\min(t^*, b^*)}{\max(t^*, b^*)}} \quad (2)$$



**Figure 2.** Center-ness calculation technique [29]

$L^*, r^*, t^*, b^*$  present the regression targets for the location as mentioned in Figure 2.

**Regression head:** Predicts the (l, t, r, b) from the center of the location. The regression head has been trained to predict scale-normalized distances. As a result, when making an inference, a denormalization of the image size should be performed [29].

The loss function used during the training process is presented in Eq. (3):

$$L(\{Px, y\}, \{tx, y\}) = 1/N_{pos} \sum_{x,y} l_{cls}(P_{x,y}, C_{x,y}^*) \quad (3)$$

$$+ \lambda / N_{pos} \mathbb{I}(C_{x,y}^* > 0) (l_{reg}(t_{x,y}, t_{x,y}^*))$$

$l_{cls}$  present the focal loss.

$l_{reg}$  present the IOU loss.

$N_{pos}$  present the number of positive samples.

$\lambda = 1$ , balance weight.

$\mathbb{I}(C_{x,y}^* > 0)$ : indicator function =1 if  $C_i^* > 0$  and 0 otherwise.

Figure 3 provides a detailed architecture of the Fully convolutional one-stage object detector used in this work to build an incident detection system.

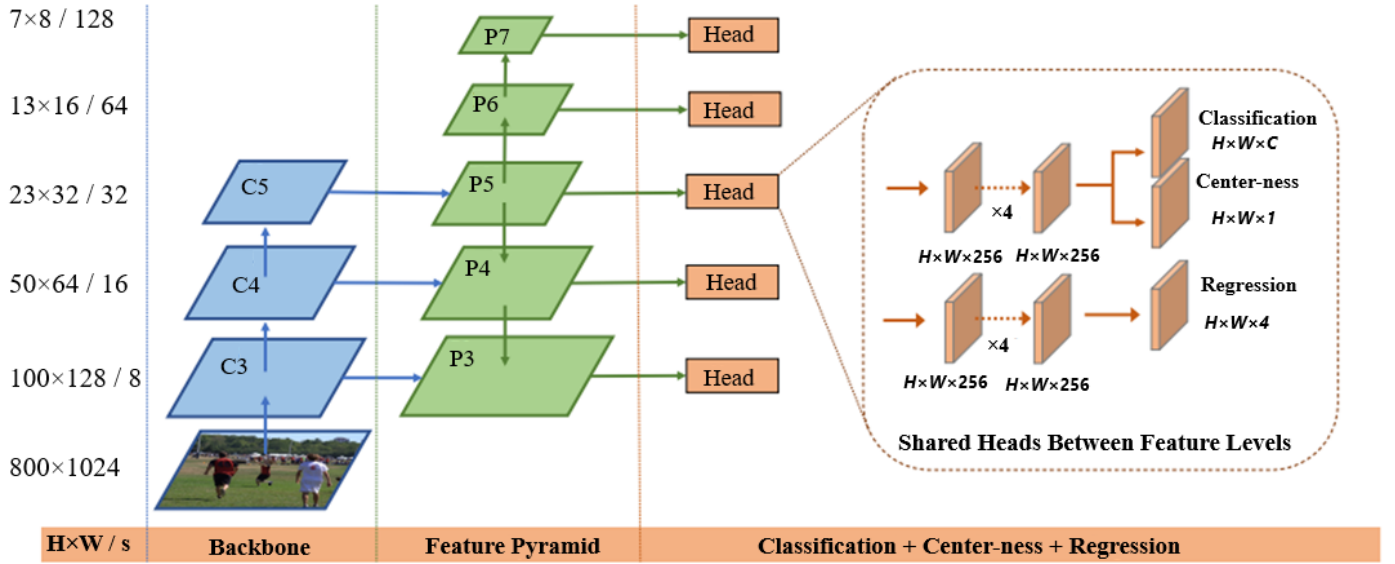


Figure 3. Proposed architecture for incident detection based on Rep-VGG and FCOS networks [29]

## 4. EXPERIMENTS AND RESULTS

For efficient traffic management and control in metropolitan networks, the incident detection function is critical for restoring normal traffic flow quickly. Reducing congestion quickly through precise and dependable event detection can have numerous positive effects on the environment and the economy. In an effort to improve traffic safety, we suggest the following, which will detail all the tests done to help develop a new sort of incident detection system. Its ability is not limited to detect only traffic accidents but also different other incidents that can be caused by natural disasters that negatively affect the traffic conditions and safety of highway and urban network users.

### 4.1 Data acquisition and pre-processing

The suggested incident detection method was trained and tested using photos from the HWID12 [30]. HWID12 presents a large-scale dataset, which consists of 2782 videos from 12 incident classes which are the following: collision with a motorcycle, collision with a stationary object, drifting or skidding, fire or explosion, head-on collision, objects falling, other crashes, pedestrian hit, rear collision, rollover, and side collision. Figure 4 provides an image subset from the obtained dataset. The HWID12 dataset was first and mainly proposed for video motion tracking. In the proposed work, we will be interested in creating a traffic incidents dataset used for

detection issues. To fill this end, the videos present in the HWID12 dataset have been framed into 30 FPS. The dataset consists of 2782 videos and all the videos present an average duration of 4s. So, we created a new traffic incidents dataset consisting of more than 333000 images divided into 11 classes. The images provided in the dataset were manually labeled.

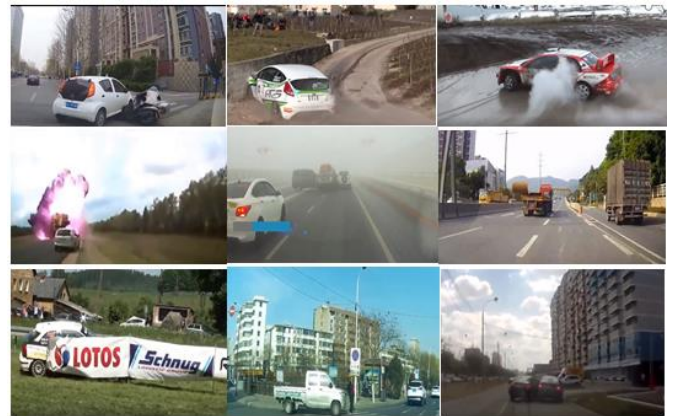


Figure 4. Image subset from the HWID12 dataset

The HWID12 dataset presents a very interesting dataset that can be applied to solve traffic incidents and accidents to improve traffic flow and security. This dataset will be very



useful to ensure better traffic conditions as it was taken under challenging and real-world conditions such as night, day, fog, snow, artificial luminosity, and rain.

## 4.2 Results and discussions

In this work, the experiment settings that were chosen are presented in Table 2. Improving traffic management and conditions is the major goal of the proposed experiments, which aim to create a reliable system for detecting traffic incidents.

**Table 2.** Proposed settings

Loss Function	Focal Loss
Learning rate	0.01
Number of epochs	120
Network optimizer	SGD/ ADAM
Train batch size	16
Training iterations	10000
Weight decay	0.0001
Activation function	SPOCU
Training set	65%
Testing set	35%

Due to the fact that it was trained and tested under extremely difficult conditions, the suggested system for detecting traffic events is sufficiently effective. HWID12 dataset was taken and collected from real scenes that did not undergo any modifications or image quality improvements.

For the purpose of conducting an in-depth investigation of the capabilities of the proposed incident detection system, a variety of evaluation metrics have been utilized. The evaluation metrics adopted in the conducted experiments are presented in the following equations.

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (6)$$

True positive (TP), false negative (FP), and true negative (FN) are the abbreviations for the three possible outcomes. The mean average precision (mAP) of all the class categories that were considered in the proposed work is displayed. It is the mean of the per-class precisions.

**Table 3.** Average precision per class while optimizing networks with SGD

Class Name	Average Precision AP (%)
Collision with motorcycle	93.4
Collision with a stationary object	92.1
Drifting or skidding	91.7
Fire or explosion	93.9
Head on collision	92.5
Objects falling	93.6
Other crash	91.3
Pedestrian hit	94.2
Rear collision	92.9
Rollover	93.1
Side collision	92.8

SGD and Adam are two different optimizers that have been used in training experiments. These investigations are conducted with the aim of improving detection outcomes and to study more about the effectiveness and the resilience of the work that has been proposed. The per-class precision for the 33 class categories that were kept from the incidents1M dataset is presented in Table 3. This precision is extremely dependable for the purpose of constructing an incident detection system that will contribute to the improvement of traffic security conditions. Table 3 shows the per-class precision, which is a measure of how well the proposed system performed when trained with an SGD optimizer network. In order to ensure that the proposed detection method is reliable, it was trained with a collection of negative data that does not include any class objects and only includes background information.

In accordance with the information presented in Table 3, the proposed method for detecting traffic events achieved highly interesting results. It was determined that the detection precision of each of the class categories was equal. The robustness of the proposed experiments that were carried out in this work was improved as a result of the absence of any imbalance between the per-class precision between the classes. Table 3 displays the results for the per-class precision, and the proposed system for detecting traffic events had an average mean precision of 92.86% using mAP. The average precision for each object class was calculated to get this. The achieved accuracy is a fascinating finding since it is a mean for eleven distinct class categories—a huge amount when compared to other state-of-the-art works—which makes it stand out. Adam has been the network optimizer in previous experiments that aimed to improve detection outcomes and increase the number of safe scenarios. In Table 4, the per-class precision that was attained is presented.

**Table 4.** Average accuracy per class while optimizing networks with Adam

Class Name	Average Precision AP (%)
Collision with motorcycle	96.4
Collision with a stationary object	94.6
Drifting or skidding	93.9
Fire or explosion	96.7
Head on collision	94.8
Objects falling	95.9
Other crash	94.3
Pedestrian hit	96.9
Rear collision	95.8
Rollover	95.7
Side collision	94.9

As can be seen in Table 4, the suggested approach for detecting traffic events achieved extremely positive results in terms of the precision of its detection. It has been determined that the detection results for practically all of the categories are significantly competitive. According to the obtained results, all the categories demonstrated high detection rates that are higher than 94%. We also note that the obtained per-class precision is balanced. A mean average precision for all the class categories of 95.44% has been obtained.

By modifying the network optimizer from SGD to Adam, the detection performances have been improved for the different evaluation metrics adopted in the proposed experiments. Table 5 provides a comparison between the evaluation metrics when using SGD and Adam optimizers.

To investigate how the system can mitigate issues arising from noisy or incomplete data streams, we highlight the use of temporal redundancy and frame interpolation techniques, where the system leverages data from adjacent frames to infer missing or corrupted information. These techniques are especially valuable in scenarios with intermittent data loss or low frame rates.

Additional experiments were conducted to evaluate the model’s robustness against low-resolution, noisy, and low-light video feeds. Results indicate that while detection accuracy decreases slightly in such conditions, the system maintains acceptable performance due to the robustness of multi-scale pyramid feature maps and advanced data augmentation during training.

**Table 5.** Optimizer’s impact on detection performance

Optimizer	mAP (%)	Recall	F1-Score
SGD	92.86	90.25	91.53
Adam	95.44	93.72	94.57

4.3 Comparison study

We conducted additional experiments to compare our method against prominent baseline models, including YOLO v4, SSD, and Faster R-CNN, on the HWID12 dataset.

These models were selected based on their widespread use and strong performance in traffic-related object detection tasks.

Table 6 compares our method with these baseline models in terms of detection precision and inference time. For instance, our method achieved a detection accuracy of 96.91%, outperforming YOLO v4 (94.75%) and SSD (92.63%) while maintaining superior computational efficiency.

Results show that our optimized FCOS with the Rep-VGG backbone achieves a 20% reduction in computational overhead compared to YOLO v4, reinforcing its suitability for real-time applications in ITS. All the presented models were evaluated on the Nvidia GTX 960 GPU. The reported results proved the superiority of the proposed model in terms of processing speed and accuracy.

Our approach outperforms these methods due to key factors include the structural re-parameterization benefits of Rep-VGG, the adaptability of FCOS to multi-scale objects, and the use of Soft Non-Maximum Suppression to reduce false positives.

**Table 6.** Evaluation of the proposed model against current state-of-the-art techniques

Model	mAP (%)	Speed (FPS)
Faster RCNN	93.22	12
YOLOv4	94.75	37
SSD	92.63	28
FCOS (ours)	96.91	46

Furthermore, we have measured the model’s inference speed in frames per second (FPS) on both GPU and CPU hardware configurations commonly used for ITS applications. Specifically, on an NVIDIA GTX 960 GPU, our system achieved an average inference speed of 46 FPS, comfortably exceeding the 30 FPS threshold for real-time performance. On an Intel i7 CPU, the model achieved 15 FPS, making it feasible for deployment in less computationally intensive scenarios with modest hardware.

4.4 Ablation study

This study's FCOS backbone, Rep-VGG, was trained using two distinct activation functions, Relu and SPOCU [26]. Better traffic conditions and better detection results were the intended outcomes of this training. In Table 7, you will find the detection results that were acquired following the modification of the network activation function and the utilization of Adam as a network optimizer.

**Table 7.** Effects of the activation function on the efficiency of detection

Activation Function	mAP (%)
SPOCU	95.44
RELU	93.57

Changing the network activation function from RELU to SPOCU improved network identification performance by 2%. You can enhance the detection results using SPOCU by utilizing neural network designs. Two distinct backbones, resent 101 [31] and Rep-VGG, were used to train and evaluate the FCOS network in an effort to achieve higher results with less compute complexity. Table 8 displays the results of using different backbones for the proposed model.

**Table 8.** Effects of modifying the backbone on the computational complexity and detection capabilities of neural networks

Backbone	FLOPS (B)	Parameters (M)	mAP (%)
ResNet 101 [31]	8.9	34.93	90.23
Rep-VGG	4.9	20.33	95.44

The proposed work's performance has been substantially enhanced with the utilization of Rep-VGG as a network optimizer. As presented in Table 7, the Rep-VGG backbone achieved better contribution between detection precision and computation complexity. As a result, the FCOS architecture made far less use of parameters and FLOPs.

5. CONCLUSION

One of the main causes of congestion and one of the primary causes of death worldwide is traffic accidents. They also negatively affect traffic flow and have detrimental effects on the economy. Therefore, an AID system is essential to any transportation system in order to save lives, improve road safety, and lessen other unfavorable effects. This study proposes a technique for detecting traffic incidents. To achieve this goal, the traffic incidents detection system was constructed using Rep-VGG as its backbone network and a modified version of the fully convolutional one-stage neural network FCOS. To perform training and testing experiments, the videos provided in the HWID12 dataset have been framed. As a result, more than 333000 images have been obtained that consist of 11 traffic incidents classes. The goal of conducting an abundance of tests is to improve detection accuracy and, by extension, traffic conditions. In terms of detection accuracy, the proposed method for detecting traffic events performed exceptionally well, according to the results. As a detection mean average precision (mAP), it reached 95.44% when SPOCU was used. We have analyzed the model’s performance across different traffic incident types in the HWID12 dataset.

Results indicate that the model performs slightly worse on incidents involving small or partially occluded objects, such as debris on the road or distant vehicles. This is primarily due to the inherent challenges of low-resolution object representation at a distance. To address these limitations, we propose the following future research directions; incorporating advanced data augmentation techniques to simulate more challenging scenarios in training. Exploring hybrid approaches that integrate transformer-based architectures to improve the model's ability to detect small and occluded objects. Expanding the dataset to include more diverse traffic incident types and edge cases for better generalization.

## ACKNOWLEDGMENT

The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University for funding this work through research group (Grant No.: RG-NBU-2022-1234).

## REFERENCES

- [1] Kamran, S., Haas, O. (2007). A multilevel traffic incidents detection approach: Identifying traffic patterns and vehicle behaviours using real-time GPS data. In 2007 IEEE Intelligent Vehicles Symposium, Istanbul, Turkey, pp. 912-917. <https://doi.org/10.1109/IVS.2007.4290233>
- [2] Saini, M. (2014). Survey on vision based on-road vehicle detection. *International Journal of u-and e-Service, Science and Technology*, 7(4): 139-146. <http://doi.org/10.14257/ijunnesst.2014.7.4.14>
- [3] Knoop, V.L., Hoogendoorn, S.P., Van Zuylen, H.J. (2008). Capacity reduction at incidents: Empirical data collected from a helicopter. *Transportation Research Record*, 2071(1): 19-25. <https://doi.org/10.3141/2071-03>
- [4] World Health Organization. (2023). Road traffic injuries. <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>.
- [5] Gu, Y., Qian, Z.S., Chen, F. (2016). From Twitter to detector: Real-time traffic incident detection using social media data. *Transportation Research Part C: Emerging Technologies*, 67: 321-342. <https://doi.org/10.1016/j.trc.2016.02.011>
- [6] Deniz, O., Celikoglu, H.B. (2011). Overview to some existing incident detection algorithms: A comparative evaluation. *Procedia-Social and Behavioral Sciences*, 2: 153-168.
- [7] Afif, M., Ayachi, R., Said, Y., Atri, M. (2021). Deep learning-based application for indoor wayfinding assistance navigation. *Multimedia Tools and Applications*, 80(18): 27115-27130. <https://doi.org/10.1007/s11042-021-10999-6>
- [8] Wu, X., Sahoo, D., Hoi, S.C. (2020). Recent advances in deep learning for object detection. *Neurocomputing*, 396: 39-64. <https://doi.org/10.1016/j.neucom.2020.01.085>
- [9] Afif, M., Ayachi, R., Pissaloux, E., Said, Y., Atri, M. (2020). Indoor objects detection and recognition for an ICT mobility assistance of visually impaired people. *Multimedia Tools and Applications*, 79: 31645-31662. <https://doi.org/10.1007/s11042-020-09662-3>
- [10] Ayachi, R., Afif, M., Said, Y., Abdelaali, A.B. (2020). Pedestrian detection for advanced driving assisting system: A transfer learning approach. In 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sousse, Tunisia, pp. 1-5. <https://doi.org/10.1109/ATSIP49331.2020.9231559>
- [11] Ayachi, R., Afif, M., Said, Y., Abdelali, A.B. (2021). Drivers fatigue detection using efficientdet in advanced driver assistance systems. In 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD), Monastir, Tunisia, pp. 738-742. <https://doi.org/10.1109/SSD52085.2021.9429294>
- [12] Zhu, L., Guo, F., Krishnan, R., Polak, J.W. (2018). A deep learning approach for traffic incident detection in urban networks. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, pp. 1011-1016. <https://doi.org/10.1109/ITSC.2018.8569402>
- [13] Pawar, K., Attar, V. (2022). Deep learning based detection and localization of road accidents from traffic surveillance videos. *ICT Express*, 8(3): 379-387. <https://doi.org/10.1016/j.icte.2021.11.004>
- [14] Ghahremannezhad, H., Shi, H., Liu, C. (2022). Real-time accident detection in traffic surveillance using deep learning. In 2022 IEEE International Conference on Imaging Systems and Techniques (IST), Kaohsiung, Taiwan, pp. 1-6. <https://doi.org/10.1109/IST55454.2022.9827736>
- [15] Polson, N.G., Sokolov, V.O. (2017). Deep learning for short-term traffic flow prediction. *Transportation Research Part C: Emerging Technologies*, 79: 1-17. <https://doi.org/10.1016/j.trc.2017.02.024>
- [16] Zhang, Y., Lu, Y., Zhang, D., Shang, L., Wang, D. (2018). Risksens: A multi-view learning approach to identifying risky traffic locations in intelligent transportation systems using social and remote sensing. In 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, pp. 1544-1553. <https://doi.org/10.1109/BigData.2018.8621996>
- [17] Huang, T., Wang, S., Sharma, A. (2020). Highway crash detection and risk estimation using deep learning. *Accident Analysis & Prevention*, 135: 105392. <https://doi.org/10.1016/j.aap.2019.105392>
- [18] Ren, H., Song, Y., Wang, J., Hu, Y., Lei, J. (2018). A deep learning approach to the citywide traffic accident risk prediction. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, pp. 3346-3351. <https://doi.org/10.1109/ITSC.2018.8569437>
- [19] Li, L., Lin, Y., Du, B., Yang, F., Ran, B. (2022). Real-time traffic incident detection based on a hybrid deep learning model. *Transportmetrica A: Transport Science*, 18(1): 78-98. <https://doi.org/10.1080/23249935.2020.1813214>
- [20] Lin, Y., Li, L., Jing, H., Ran, B., Sun, D. (2020). Automated traffic incident detection with a smaller dataset based on generative adversarial networks. *Accident Analysis & Prevention*, 144: 105628. <https://doi.org/10.1016/j.aap.2020.105628>
- [21] Chang, W.J., Chen, L.B., Su, K.Y. (2019). DeepCrash: A deep learning-based internet of vehicles system for head-on and single-vehicle accident detection with emergency notification. *IEEE Access*, 7: 148163-148175. <https://doi.org/10.1109/ACCESS.2019.2946468>
- [22] Ding, X., Zhang, X., Ma, N., Han, J., Ding, G., Sun, J.



- (2021). Repvgg: Making VGG-style convnets great again. In 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, pp. 13728-13737. <https://doi.org/10.1109/CVPR46437.2021.01352>
- [23] Tian, Z., Shen, C., Chen, H., He, T. (2019). Fcos: Fully convolutional one-stage object detection. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), pp. 9626-9635. <https://doi.org/10.1109/ICCV.2019.00972>
- [24] Chu, X., Li, L., Zhang, B. (2024). Make repvgg greater again: A quantization-aware approach. Proceedings of the AAAI Conference on Artificial Intelligence, 38(10): 11624-11632. <https://doi.org/10.1609/aaai.v38i10.29045>.
- [25] Kisefák, J., Lu, Y., Švihra, J., Szépe, P., Stehlík, M. (2021). “SPOCU”: Scaled polynomial constant unit activation function. Neural computing and applications, 33: 3385-3401. <https://doi.org/10.1007/s00521-020-05182-1>
- [26] Sakketou, F., Ampazis, N. (2019). On the invariance of the SELU activation function on algorithm and hyperparameter selection in neural network recommenders. In Artificial Intelligence Applications and Innovations: 15th IFIP WG 12.5 International Conference, AIAI 2019, Hersonissos, Crete, Greece, pp. 673-685. [https://doi.org/10.1007/978-3-030-19823-7\\_56](https://doi.org/10.1007/978-3-030-19823-7_56)
- [27] Agarap, A. F. (2018). Deep learning using rectified linear units (ReLU). arXiv preprint arXiv:1803.08375. <https://doi.org/10.48550/arXiv.1803.08375>
- [28] Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S. (2017). Feature pyramid networks for object detection. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 936-944. <https://doi.org/10.1109/CVPR.2017.106>
- [29] Ayachi, R., Mouna, A., Yahia, S., Abdesslem, B.A. (2024). Implementing anchor free model for social distancing detection on FPGA board. Information Fusion Research, 2(1): 6556. <https://doi.org/10.59429/ifr.v2i1.6556>
- [30] HWID12 (Highway Incidents Detection Dataset). <https://www.kaggle.com/datasets/landrykezebou/hwid12-highway-incidents-detection-dataset>.
- [31] He, K., Zhang, X., Ren, S., Sun, J. (2015). Deep Residual Learning for Image Recognition. arXiv preprint arXiv:1512.03385. <https://doi.org/10.48550/arXiv.1512.03385>