International Information and
Engineering Technology Association
Advancing the World of Information and Engineering

Utilizing a Hybrid Model for Human Injury Severity Analysis in Traffic Accidents

Uddagiri Sirisha*[ID], Bolem Sai Chandana[ID]

School of Computer Science and Engineering, VIT-AP University, Amaravati 522237, Andhra Pradesh, India

Corresponding Author Email: sirisha.u@vitap.ac.in

ABSTRACT

Road safety has been prioritized by governments globally, resulting in the implementation of numerous initiatives aimed at curtailing traffic accidents. Despite these efforts, the complete eradication of accidents remains unattainable. Therefore, swift and accurate responses to accident sites, accompanied by appropriate medical aid, are paramount in saving lives. Existing systems, primarily designed to alert medical personnel in the aftermath of an accident, rely solely on Vehicle Damage (V_d) to assess accident severity, neglecting Human Injury (H_i) considerations. This study proposes a hybrid model equipped with an attention mechanism, designed to classify accident severity based on both V_d and H_i. The proposed model accepts video or image inputs and classifies accident severity levels accordingly. Moreover, an extension of the model has been developed to obfuscate sensitive areas in accident imagery based on severity, particularly when such images are disseminated on public platforms without obtaining necessary consent. The proposed hybrid model, therefore, not only facilitates a more comprehensive severity assessment of traffic accidents but also ensures the protection of privacy and promotes ethical image sharing practices.

1. INTRODUCTION

The global increase in population, coupled with the proliferation of vehicles and the introduction of varied transportation methods, has led to a surge in road accidents. Several incidents occur daily, with causes ranging from driver distraction and fatigue to poor visibility and hazardous road design. Regardless of the diverse precipitating factors, the consequences frequently remain uniform, often culminating in substantial loss.

Over recent years, a marked escalation in accidents attributable to individual negligence or deficient maintenance has been observed. The National Crime Records Bureau (NCRB) reported a rise in accidental deaths in India to 1.55 lakhs in 2021. Although the total road fatalities showed a decline of over 14% from approximately 1.5 lakhs in 2019 to 1.3 lakhs in 2018, largely due to restrictions imposed by the Covid-19 pandemic, aspects such as the severity of accidents and the increasing ratio of two-wheeler occupant fatalities have raised significant concern. The percentage of fatalities involving two-wheeler occupants surged from 36% to 43.6% in 2020. Additionally, a 10% increase in fatal accidents was recorded in 2021, as documented by the Times of India. A comparative analysis of statistics from 2017 to 2021, as reported by the NCRB, is depicted in Figure 1.

Current accident detection systems effectively inform relevant personnel via mechanisms such as road buzzers and automated messages to nearby hospitals and police stations. However, a concerning trend has emerged wherein individuals capture and share accident scene imagery on social media platforms. Such practices, often involving the display of sensitive content, can have adverse effects on viewers, particularly children. Despite this, social media platforms continue to permit the sharing of diverse content types, including images, videos, text, and animations.

While users can adjust privacy and security preferences on social media platforms, the existing filters are only capable of restricting certain content types, neglecting other potentially unethical portions of the uploaded content. A notable example involves the dissemination of sensitive accident scene images/videos.

The proposed system in the present study, henceforth referred to as HUSA (Human Severity in Accidents), assesses accident severity and relays this information to medical personnel, enabling a more effective response. Furthermore, HUSA can be employed by content-sharing platforms to alert users to the sensitive nature of the content they are uploading, thereby preventing inadvertent sharing of unethical content.
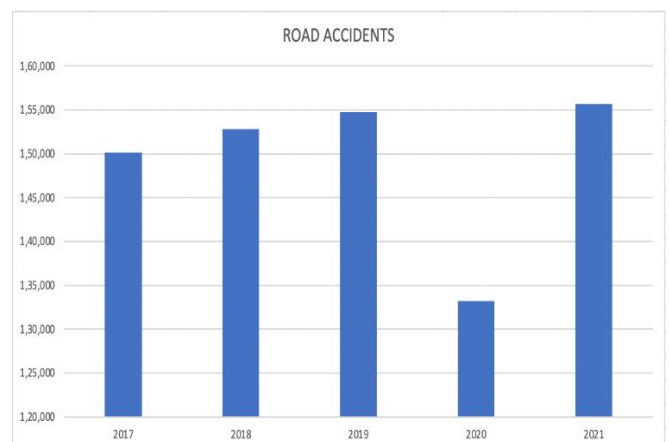

Figure 1. Statistics of road accidents in India

2. PRIOR RESEARCH AND NOVELTY OF HUSA

The novelty of HUSA is summarized in Table 1. It is evident that the proposed system performs all three activities which are not addressed by other works as a single entity i.e.

- The system inputs the video or images and continuously checks for abnormal activities (here, accidents).
- Severity analysis is not just assessed with vehicle damage but also with the human injuries at the crime scene.
- To reduce the chances of sharing unethical content containing sensitive objects, the proposed model can recommend blurring portions of the image/video being uploaded.

Accident detection Methods are divided into three categories: vehicle condition (hardware-based), accident video features-based, and accident image/text features-based.

2.1 Using the state of the vehicle to detect accidents

Approaches under this category monitor the change in vehicle parameters to detect accidents. In this method, various vehicle parameters, such as speed, acceleration, or angular velocity, are compared before and after an event to detect an accident. An accident is identified when the changes in the variables exceed certain threshold levels. For example, if the car's speed and the (phone) are more than the threshold value (24 km/h), both the user and the phone are likely to be inside the car. When a smartphone detects an acceleration event more significant than a particular threshold value (4G), it assumes an accident has occurred.

The authors [1] used the Deep Convolutional Neural Network model with trajectory-related variables like pitching angle, lateral position, vertical position, and longitudinal position. The authors [2] design smartphone sensors and UAV-based approaches to detect and alert medical personnel when they need help. Researchers [3] have used "Gaussian Mixture Model"; "Naive-Bayes Tree", "Decision Tree", and "Classification-Regression Trees" to detect the severity level of accidents, and then essential information about the accident is sent to the emergency services. The authors [4] modeled the traffic flow for accident detection using a "Support Vector Machine" and "Probabilistic Neural Network." The authors [5] used real-time entities with traffic network, demographic, and weather-related information. Finally, the authors created an IoT-based system called DeepCrash [6] to notify the appropriate authorities when a car accident happens in a place with few people or when there is simply a driver in the vehicle and experience a loss of consciousness.

The method's efficacy was verified by contrasting simulation data to real-world data on road traffic accidents, as used by the authors [7], who used a variety of smartphone sensors to detect accidents, including "accelerometers", "pressure sensors", and "microphones" to collect data on "vehicle speed", "gravity", "pressure", "sound", and "position". The authors [8] designed an accident detection, locating, and emergency reporting system using cell phones, accelerometers, and Google Maps. The authors [9] looked into harnessing cell phone sensors and processing power to identify road accidents without onboard sensors. The authors [10] proposed that the smartphone's accelerometer was used to detect an accident,

and the phone's GPS was used to monitor the location of the accident. The authors [11] introduced a smartphone app that used the accelerometer and gyroscope sensors to detect accidents and alert "emergency contacts", "local police", and "local emergency medical response teams".

In conclusion, this approach has practical real-time functionality and is affordable. Furthermore, this approach cannot accurately assess the accidents surrounding it, which impacts the multi-vehicle accident's severity.

2.2 Utilizing video characteristics to identify accidents

The extensive availability of high-end computational resources created an additional advantage of applying deep learning techniques to the live video stream data. "Motion detection", "extraction", and "feature analysis" are the three primary phases in video-based accident detection. The "velocity", "acceleration", "area", "location", and "direction" of a vehicle are among the retrieved features, and these features are compared to specified thresholds to identify accidents.

The authors [12] used support vector machines to identify accidents and alert emergency assistance. Accident detection models should be trained on large-sized real-time data collected by authors [13] and tested with customized fast-RCNN, resulting in average precision of more than 47%. The authors [14] proposed an unsupervised deep learning framework for detecting traffic accidents in first-person (egocentric) videos that capture the person's natural field of view. In videos, the backgrounds are dynamic and complex, and the accidents/abnormalities typically last a brief period and frequently follow a long-tailed distribution. As a result, it is challenging for these video-level techniques to deliver adequate performance for accident or abnormality detection. Therefore, the authors [15] classified the videos into three categories: "video level", "segment-level", and "frame-level" detection methodologies to solve the issue of long-tailed distribution in a traffic accident/abnormality detection.

The authors [16] used logistic regression over the Histogram of Flow Gradients to estimate the likelihood of an accident occurring in a video sequence. In addition, the authors [17] employed Farneback Optical Flow with a heuristic threshold-selection strategy for accident identification in a video sequence. The authors [18, 19] used a well-known object detection approach called YOLOv3 to detect traffic incidents from motorway cameras. For accident identification in movies, the authors combined the convolutional Long Short-Term Memory network-based temporal sequencer and spatial feature extractor [20, 21].

Table 1. Prior analysis of accident detection

References	Accident Detection	Severity Analysis w.r.t to Humans	Automatic Blurring of Images (Social Media)
[1]	Yes	Yes	No
[2]	Yes	Yes	No
[3]	Yes	Yes	No
[8]	Yes	Yes	No
[9]	Yes	Yes	No
[12]	Yes	Yes	No
[13]	Yes	Yes	No
[22]	Yes	Yes	No
[23]	Yes	Yes	No
[Present Work]	Yes	Yes	Yes

In conclusion, while using accident detection methods based on video data has grown rapidly as the number of cameras on the road has increased, data quality and accident detection accuracy remain poor. In addition, environmental factors and the data strongly influence the model's effectiveness is inconsistent because the cameras on the road cannot capture everything, severely limiting the method's application.

2.3 Utilizing image features to identify accidents

In this category, images act as input to various deep-learning models for detecting accidents. The authors [22] proposed an ensemble mixture of four learning algorithms containing "ELM", "KELM", "CELM", and "OSELM" to estimate the severity level of accidents based on vehicle damage. A better model was built using YOLO-CA and tested on the CAD-CVIS dataset, which resulted in an average precision of more than 90%. To automatically detect car accidents based on the CAD-CVIS dataset, the authors [23] developed the YOLO-CA deep neural network model. The results demonstrate that this model can detect car accidents in 0.0461 seconds (21.6FPS) with an average 90.02 per cent (AP) precision.

In conclusion, accident detection techniques based on images are used, and the accuracy of accident detection is high when compared to others.

2.4 Utilizing text features to identify accidents

Recently, unstructured social media data has been used to extract meaningful information using "machine learning" and "natural language processing" models to identify accidents.

Therefore, social networking platforms are essential for condition analysis and accident detection. However, taking the necessary information from social networking data and then studying them in road accident investigations can be challenging.

In this category, the techniques continuously scrape the data from content-sharing platforms and detect abnormal activities like accidents. For example, the authors [24-26] used sentiment analysis on the live feed to detect accidents without manual intervention.

In conclusion, incorporating social network data into the study of accidents opens up various research possibilities. The results suggest that social network data may be inconsistent, if not suspicious. Social networks might therefore complement current accident detection techniques rather than replace them. The identified limitations on accident detection and analysis are presented in Table 2.

3. A SYSTEM-LEVEL OVERVIEW OF HUSA FOR ACCIDENT SEVERITY ANALYSIS

The proposed system overview is shown in Figure 2. It involves multiple stages, and the same is discussed below.

3.1 Data set collection and labelling

There are datasets for accident detection [2, 4, 22], but none can be used to assess the state of injured humans' severity after the accident. We have collected the accident images/videos uploaded to social media to address this issue. Each image/video contains both vehicles and injured humans. The dataset collection process is shown in Figure 3.

Table 2. Prior analysis of accident detection w.r.t to various parameters

| Title Name	Data Sources	Type of Data	Summary
[1]	vertical position, lateral position, roll angle, pitching angle, yawing angle, longitudinal position, longitudinal velocity etc.	Based on hardware data	Based on vehicle trajectory data, this study utilizes a Deep Convolution Neural Network (DCNN) model with 95 per cent accuracy to identify and categorize traffic accidents.
[2]	Speed, Linear acceleration, Altitude, Roll and pitch.	Based on hardware data.	This study uses drones and smartphones to determine whether an accident occurred and notify emergency personnel.
[3]	microcontroller, GPS, and a group of sensors	Based on hardware data	This paper used the "Gaussian Mixture Model", "Decision Tree", "Naive-Bayes Tree", "Classification", and "Regression Trees" to detect the severity level of accidents, and the accident is intimated to emergency services.
[8]	Facebook data, Twitter data	Based on text data	Using an OLDA/Bi-LSTM model, it detects accident traffic events and achieves an accuracy of 97%.
[9]	Twitter data	Based on text data	This paper uses a Deep Belief Network (DBN) and Long Short-Term Memory (LSTM) to detect a traffic accident from Twitter data.
[12]	CCTV surveillance video dataset of Hyderabad City in India.	Based on video data	Spatial and temporal representation can locate the locations of accidents
[13]	CADP dataset	Based on video data	Focused on faster R-CNN model to predict accidents.
[22]	Images from various sources such as "surveillance cameras", "UAVs"	Based on image data.	Used an ensemble mixture of four learning algorithms containing "ELM", "KELM", "CELM", and "OSELM" to classify the accident severity.
[23]	CAD-CVIS dataset	Based on image data	Focused on car accident detection using "YOLO-CA" based on the "CAD-CVIS" dataset.
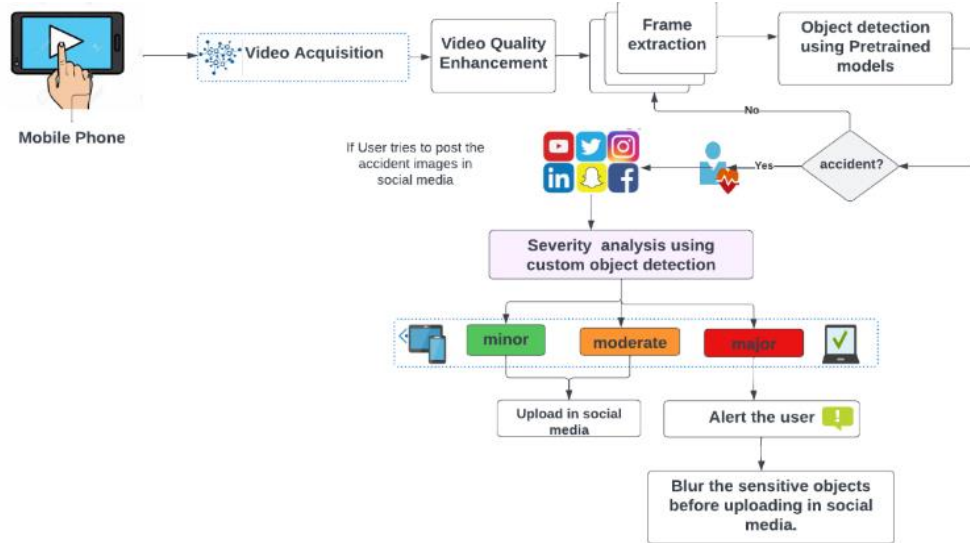
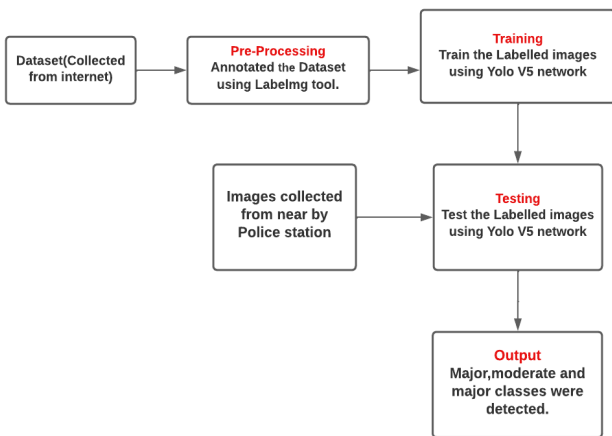**Figure 2.** System level overview of HUSA

Figure 3. Dataset collection process

Figure 4. Sample images from the dataset

We know real-world computer vision applications are increasing, such as disease diagnosis and categorization, traffic monitoring using CCTV cameras, human recognition for security purposes, and accident detection. Datasets are essential to the growth of many computational domains since they provide findings for any project with scope, robustness,

and confidence. However, datasets for accident detection emphasize the severity of the vehicle more than the human. Therefore, existing datasets are unsuitable for classifying severity levels in humans. To overcome this issue, real-time data collecting will be a wise strategy; we will create a dataset from videos. So, we developed a new dataset, namely humans in accidents (HIA), by collecting the sample videos from YouTube. We will use an OpenCV video capture module for frame extraction, which will give extracted frames (images) from videos, and some sample images are illustrated in Figure 4.

Pre-processing data is a crucial step that raises the calibre of the data. Data pre-processing is the process of arranging and cleaning raw data to produce the required outputs in an accessible and understood format. Image data can have issues with complexity, correctness, and sufficiency, to name a few. One of the most unexplored areas of data science is image data processing. Grayscale conversion, normalization, data augmentation, and image standardization are examples of image pre-processing techniques. We applied data augmentation to increase the size of the dataset and also rescaled the images to 514×514.

Labeling the data set was the subsequent step. We used the labeling tool to label the image. We started labelling images (drawing the bounding box and label the class). We receive a text file as output after correctly labelling the image. Each text file has five decimal values: the class of the bounding box is represented by the first value, followed by the centers of the x and y axes, and finally, the width and height. After that, the dataset is divided into train and test for further processing.

3.2 Content enhancement approaches

We have collected accident images/videos captured during the night to ensure model robustness in low-light conditions. Furthermore, content enhancement improves the visual appearance by performing analysis at every instance and correlating the time frames by layering the sparse entries. The most commonly used content enhancement algorithms are EnlightenGAN [27], KinD [28], and Retinex [29]. In addition, the proposed system uses Generative Facial Prior Generative Adversarial Network, which helps restore and enhance the images/videos.

3.3 Key frame extraction

A collection of frames that most accurately depicts the scene's visual content is identified using key frame extraction approaches like "clustering-based methods", "shot-based methods", "visual content-based methods", and "graph modelling-based methods". Frame extraction mainly depends on the frame rate. This is because frame rate impacts a video's style and viewing experience. The proposed system extracts the frames from the videos with a frame rate of 24 fps.

3.4 Pre-trained object detection models

Automatic accident detection is achieved through software and hardware-based approaches. A wide variety of deep learning models, such as "SSD", "YOLO", "Faster R-CNN", and "Mask R-CNN" are developed for the detection of accidents.

This paper [30] uses a pre-trained yolov7 Model for object detection that detects the dead animals. The authors in study [31] propose a Mini-YOLO architecture trained using knowledge distillation. The authors in study [32] stated an overview of YOLO based object detection. The authors in study [33] used a YOLO-CA model for accident detection. An ensemble architecture was used for accident detection by the authors in study [34] used the "YOLOv3" object detection algorithm and the Canny edge detection algorithm to classify the accidents.

3.5 Severity analysis using custom object detection

Prior research calculates the accident's severity and alerts the necessary emergency services about the accident. The current study categorises the severity based on the vehicle's damage. Instead of emphasising vehicle damage in accident detection, the proposed technique focuses on the human severity level. Three categories include minor, moderate, and major. They are used to categorize the severity level of an accident.

4. PROPOSED METHODOLOGY FOR HUSA

Figure 5 depicts the layout of the learning model employed by HUSA. The pre-processed images can be used as inputs for the object detection model. First, using the input images, the vehicle severity and human severity are identified. The model divides uploaded images into three categories: minor, major, and moderate if the user attempts to upload them to social media. The system generates a warning to blur the sensitive portion if the image belongs to a major class.

Algorithm 1. Severity Analysis in HUSA

Input: A set of images.
Output: Classify the images into minor, moderate and major classes, Blur the Sensitive portions.
Algorithm:
1. for I_i in images do
    $O_m \leftarrow (I_i, H_i)$
  endfor
    Using Object detection model (O_m), H_i (Human injuries) are detected through bounding boxes around the images.
2. Calculate the human severity in accident images by applying object detection model (O_m)
  $O_m \leftarrow (H_i, M_i, M_o, M_a)$
  M_i represents minor accident
  M_o represents moderate accident
  M_a represents major accident
3. if ($H_i == M_i \parallel H_i == M_o$) then
          Upload (H_i) in social media
  else
          Gaussianblur (H_i)
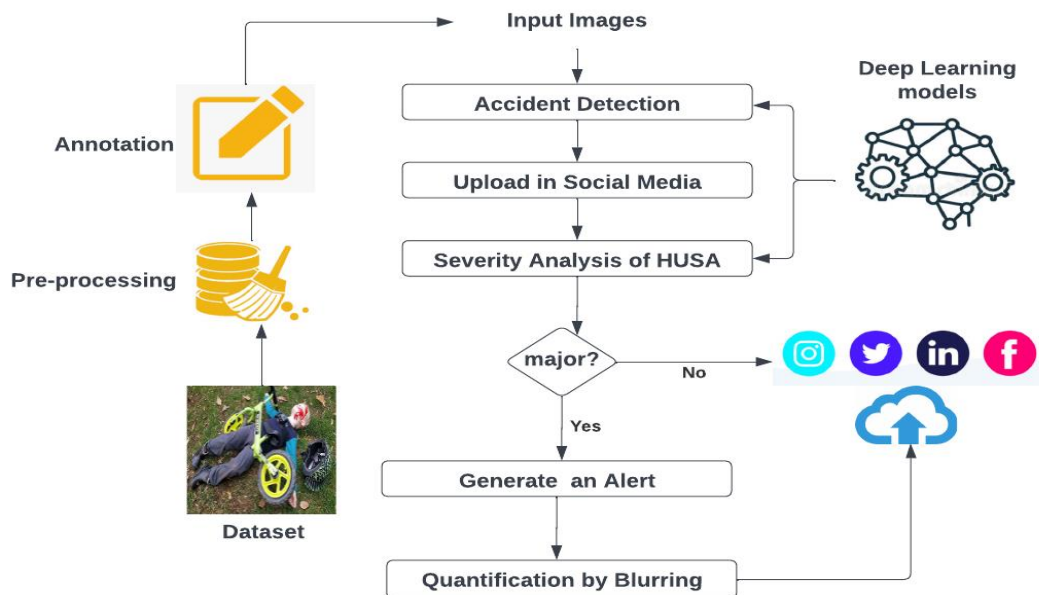   and then upload in social media
  endif

Figure 5. Workflow of HUSA

4.1 Deep learning models for accident detection and classification of HUSA

Finding and identifying various items in an image is known as object detection. Deep convolutional neural networks are used to identify objects in images by deep learning algorithms like YOLO, SSD, and R-CNN. Among the state-of-the-art object detection methods, we trained our model using YOLOv5, Faster R-CNN, SSD using Resnet, Mobilenet and also Efficientnet, by utilising those that consume less memory and have low CPU demands while still being highly accurate.

4.2 Applying YOLOV5 architecture with squeeze and excitation attention

The most recent iteration of YOLO i.e., YOLOv5, boosts great detection and inference speed accuracy. The weight file for YOLOv5 is 90% smaller than for YOLOv4. This makes it suitable for real-time detection in embedded devices. The YOLOv5 architecture aids in improving accident detection. Four various models, including YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, make up the architecture. The difference in the four architectures is in the feature extraction module, the convolutional kernel of the network, the size of the model, and model parameters. In our proposed method, we used YOLOv5s and the architecture is represented in Figure 6.

The SE channel attention method uses the CNN channel to extract features. For the model to recognize and retain significant details from all of the feature channels, a recalibration of features is required. First, the feature picture is spatially compressed using the squeeze approach to develop a model based on the channel correlation. Next, the excitation technique allows the feature channels' relative importance to be identified. The result is a matching set of channels stimulated from the original feature photos. The SE mechanism requires a few extra parameters and is easy to compute.

The following steps make up the training of a unique YOLOv5 model:

YOLOv5 Environment setup: Clone the YOLOv5 repository on GitHub. Torch, the platform used by YOLOv5, may be simply run using Kaggle or Colab. This will make a folder on the computer called "YOLOv5". The pre-trained weights for the model and the unique YOLO directory structure are also supplied.

Establishing the directory structure and the data: The YOLO folder and data folder are created on the same level as the directory of colab. We can directly upload images and labels in our google drive or make a new folder for images and labels inside the data folder. Create folders for "train" and "test" inside Images and Labels. The labels must be uploaded inside the "data/labels/train" folders and "data/labels/test". The labels file must have the same name as the image file but end in ".txt."

The bounding box includes the following items.

- The class number is represented by the first value. If there is only one class, the value is always zero.
- The bounding box's standard center pixel, measured in terms of width, is in second position.
- The third position denotes the height-standard center pixel of the bounding box.

The number of lines in a label file corresponds to the number of bounding boxes in a single image, and the number of label files corresponds to the number of images.

Set up the YAML file: We need a YAML file to train a YOLOv5 model. The YAML file provides the locations of the test data, train data, the number of classes.

Train the YOLOv5 model: The train.py command from the notebook is used to train the model. You can specify hyperparameters such picture size, epochs, and batch size. The weights for this model will be stored in a separate subdirectory in YOLOv5.

The detect.py script can be used to find the severity of the body that are present in the image.
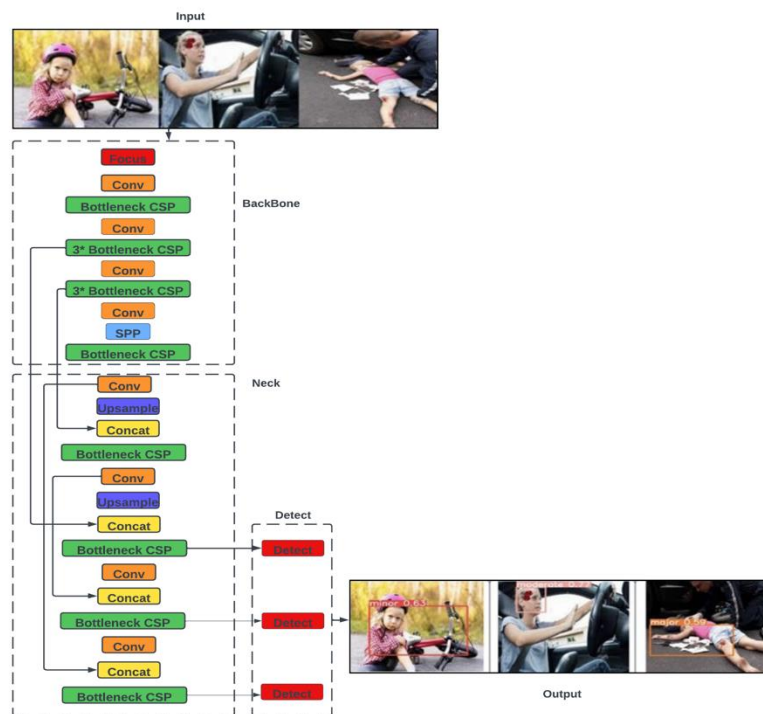

Figure 6. Accident detection using YOLOV5 architecture

4.3 Applying faster R-CNN architecture

Faster R--CNN is one of the Two-stage object identification models. First, it creates a sparse set of candidate object positions using a region pooling network built on shared feature maps. Then it uses the information from the RPN to classify each candidate proposition as either foreground or background. In the Initial stage, shared feature maps are accepted as input by RPN, extracting shared feature maps with a CNN and producing a set of bounding box candidate object locations, each with an objectness score. The size of each anchor is set using hyperparameters. Then, the suggestions are implemented in the region. Using the extracted shared feature maps as input, the first stage RPN creates a list of potential object locations with bounding boxes equipped with an objectness score of interest pooling layer (RoI pooling) to create subfeature maps. This is done after extracting shared feature maps using a CNN. The subfeature maps are converted into 4,096-dimensional vectors to feed forward into fully linked layers. A regression network built from these layers forecasts the bounding box offsets, and a classification network predicts the class for each bounding box proposal.

4.4 Applying SSD architecture

Single Shot Detection is an object detection technique that analyzes numerous things from a single image. Several things in an image are recognized and analyzed from a single frame. Comparing this too convoluted neural networks, the analysis is completed significantly more quickly. When using a multi-box, a one-shot detector may recognize several items in an image with just one shot. SSD has a higher object detection algorithm speed and accuracy.

4.5 Alert generation and blurring

Contrary to earlier publications, which mainly concentrated on personal privacy policies, privacy and security issues occur when photographs are posted on social media. Users that publish accident images to social media offend the relatives of the victims. A new programme named iUpload (imageUpload) is created to achieve automatic suggestions of privacy and security settings for image sharing. Therefore, those who post their images may quickly check the privacy settings. As a result, we divide the accident images into three classifications based on their severity: minor, moderate, and major. When the severity level is in major class, it is advised that the user uploading the image receive an alert. The sensitive area of the image is afterwards obscured before being uploaded to social media. The sample output is shown in Figure 7.
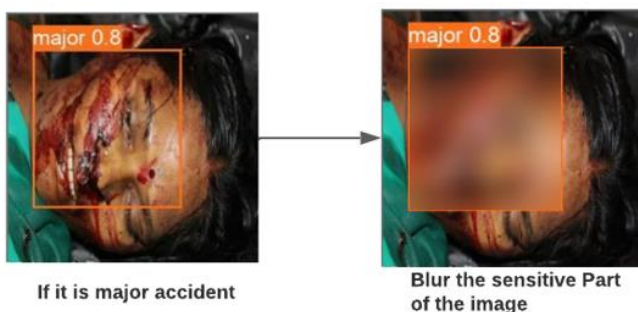

Figure 7. Quantification of major accident images by blurring

5. RESULTS AND ANALYSIS

The TensorFlow Object Detection API is used to train and test the model. Each model was evaluated after training on our custom data set using 25 accident images. The effectiveness of each model was assessed by looking at how precisely it identifies and places the suitable class of objects in a given test image. For example, in Figure 8, we plotted three photographs from several classes to demonstrate how each model predicted the outcomes of these images.
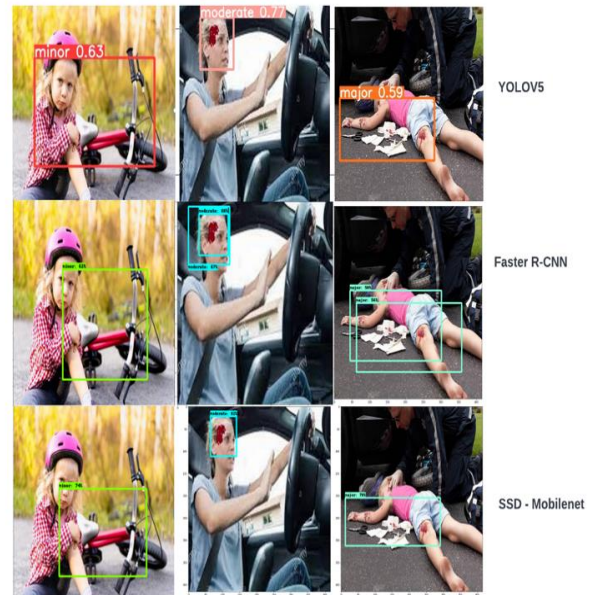

Figure 8. Results of various classes (severity level of humans) detected from test images

Figure 9. Vehicle severity assessment

The Metrics used for severity analysis in HUSA are Precision, Recall, Accurate Precision (AP), and Confidence. It is necessary to define fundamental terms like True Positive (Tp), False Positive (Fp), False Negative (Fn), and True Negative (Tn). The Intersection Over Union (IOU) measure assesses the overlap between two bounding boxes and is used in the following discussion. It is determined by the expected bounding box's overlapping area divided by the ground truth box's area of union.

Tp: A successful identification is True Positive. This happens when the IOU detection is larger than or equal to the predetermined threshold value.

Fp: A misclassified detection is False Positive. This happens when the IOU falls below the predetermined threshold amount.

Fn: When a ground truth cannot be found then it is False Negative.

Tn: True Negative is an outcome that could occur if the model accurately predicted the ground false. Due to the numerous bounding boxes that could be present in an image and should not be identified, object detection classifiers do not take this into account.
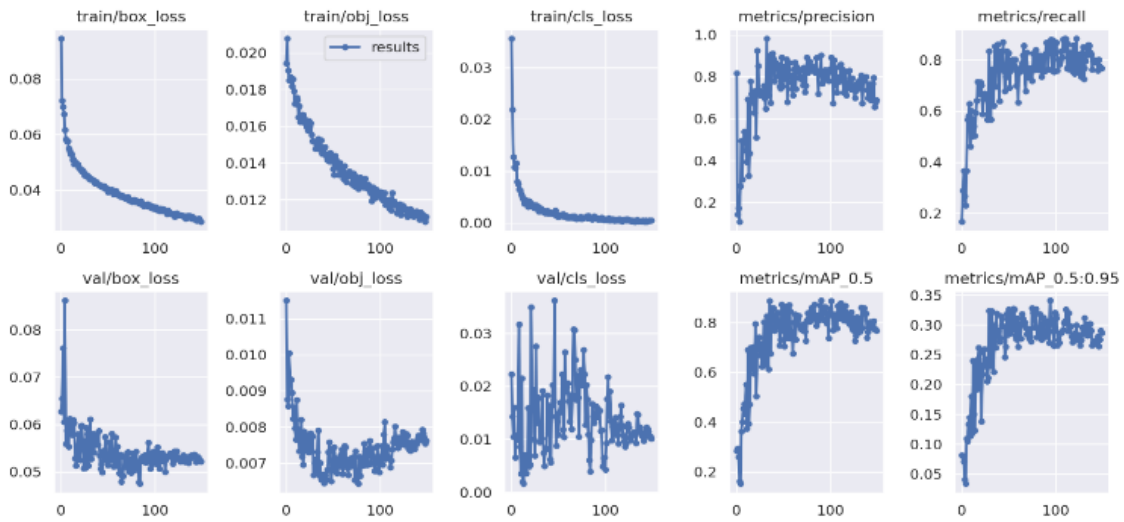

Figure 10. Different evaluation parameters in proposed model
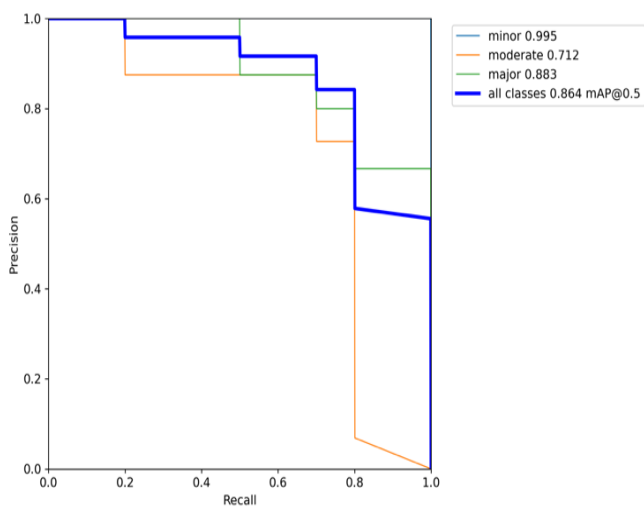

Figure 11. Estimation of MAP

Figure 9 shows the model-specific severity of the vehicles.

Existing literature largely focuses on vehicle severity, accident factors, accident location, emergency alarm systems, etc. Our approach considers both the human and vehicular injuries of an accident.

Precision: Precision is the capacity of a model to extract only the important objects from an image and is denoted by Pm.

$$P_m = \left(\frac{T_p}{T_p + F_p} X\ 100\right)$$

Confidence: It is used to assess the relationship between prediction and recall and rate the predictions. A confidence interval (C_m) is indicated as:

$$C_m = z \sqrt{\left(\frac{\beta.(1-\beta)}{N_s}\right)}$$

where, N_s denotes Sample size, z denotes critical value, β denotes accuracy.

Recall: Recall is the capacity of a model to identify all the relevant cases from the detected relevant objects and is denoted by R_m.

$$R_m = \left(\frac{T_p}{T_p + F_N} X\ 100\right)$$

Accuracy: The ratio of a model's correct detections to all of its detections can be used to define a model's accuracy and is denoted by A_m.

$$A_m = \left(\frac{T_p + T_n}{T_p + T_n + F_p + F_n}\right) X\ 100\%$$

Precision-Recall: Plotting a curve for each object class will allow you to assess the performance of the object detector when the confidence level is modified. A good object detection model has high precision and recall values that do not change when the confidence threshold is changed.

Average-Precision: It is the precision that has been calculated by averaging all recall values between 0 and 1. By calculating the precision-recall curve's area under the curve, AP aids in assessing a model's performance.

Figure 10 displays 100 epochs of training and validation losses for the Humans in accidents dataset using the proposed yolov5 algorithm for object detector and categorization. Maximum convergence is reached for precision and recall metrics during training and validation, whereas the mAP converges to 86% using a 0.5 threshold.

The mean average precision is shown in Figure 11. A model's accuracy in its detection is determined when its bounding box is compared with the ground-truth bounding box. By taking the mean of all class averages, we can calculate the mAP. So in Figure 10 the mAP is 0.864, which is the average of minor, major and moderate classes.

After initial evaluation, three superior models for detecting accidents were selected and their results compared. All three models-"Faster(R-CNN)," "SSD," and "YOLOv5"-combined squeeze and excitation attention mechanism, as shown in the preceding analysis. In terms of total detection accuracy, operation speed, and the number of parameters, the fastest R-CNN models performed the worst. The "SSD" models were fantastic for virtually instantaneous accident severity identification on the roads due to their quick operation speed and high accuracy. The "YOLOv5" models showed impressive resilience by maximizing detection precision with the fewest parameters and faring well while identifying both tiny objects and those obscured by background noise.

Certain variables, such as the dataset or the intended platform, will affect how well each model performs. And then, after conducting an objective comparison of the available models, to decide on one that best suits our requirements. The numbers in Table 3 and Figure 12 were acquired by applying several algorithms to the same dataset while running them on the same GPU.

Table 3. Comparison analysis of models w.r.t to different evaluation metrics

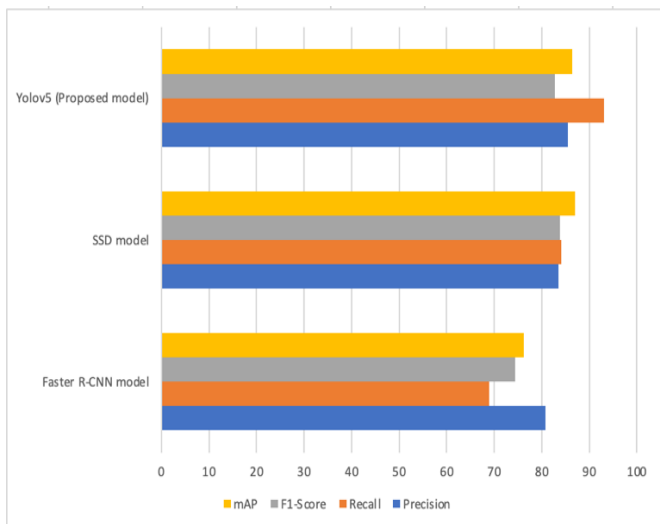| Model	Precision	Recall	F1-Score	mAP
Faster R-CNN model	80.74	68.87	74.33	76.23
SSD model	83.46	84.10	83.78	86.96
Yolov5 (Proposed model)	85.46	93.12	82.78	86.4

Figure 12. Comparision of different algorithms

6. CONCLUSION

Owning to the significant improvements on accident detection, researchers have focused on accident severity rather than human severity. To address this challenge, we developed a new dataset and classified the accidents in terms of severity level by using various proposed yolov5 model along with object detection techniques. From the comparison analysis the proposed yolov5 model with squeeze and excitation method performs better than other object detection models. At the end, we have given our inferences from the various models which would help the researchers to further streamline in accident detection.

REFERENCES

[1] Shah, A.P., Lamare, J.B., Nguyen-Anh, T., Hauptmann, A. (2018). CADP: A novel dataset for CCTV traffic camera based accident analysis. In 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), IEEE, pp. 1-9. https://doi.org/10.1109/AVSS.2018.8639160

[2] Yang, D., Wu, Y.Z., Sun, F., Chen, J., Zhai, D.H., Fu, C.Y. (2021). Freeway accident detection and classification based on the multi-vehicle trajectory data and deep learning model. Transportation Research Part C: Emerging Technologies, 130: 103303. https://doi.org/10.1016/j.trc.2021.103303

[3] Kumar, N., Acharya, D., Lohani, D. (2020). An IoT-based vehicle accident detection and classification system using sensor fusion. In IEEE Internet of Things Journal, IEEE, 8(2): 869-880. https://doi.org/10.1109/JIOT.2020.3008896

[4] Parsa, A.B., Taghipour, H., Derrible, S., Mohammadian, A.K. (2019). Real-time accident detection: Coping with imbalanced data. Accident Analysis & Prevention, 129: 202-210. https://doi.org/10.1016/j.aap.2019.05.014

[5] Parsa, A.B., Movahedi, A., Taghipour, H., Derrible, S., Mohammadian, A.K. (2020). Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. Accident Analysis & Prevention, 136: 105405. https://doi.org/10.1016/j.aap.2019.105405

[6] 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. In IEEE Access, IEEE, 7: 148163-148175. https://doi.org/10.1109/ACCESS.2019.2946468

[7] Bhatti, F., Shah, M.A., Maple, C., Islam, S.U. (2019). A novel internet of things-enabled accident detection and reporting system for smart city environments. Sensors, 19(9): 2071. https://doi.org/10.3390/s19092071

[8] Khan, A., Bibi, F., Dilshad, M., Ahmed, S., Ullah, Z., Ali, H. (2018). Accident detection and smart rescue system using android smartphone with real-time location tracking. International Journal of Advanced Computer Science and Applications, 9(6): 341-355. https://doi.org/10.14569/IJACSA.2018.09064

[9] White, J., Thompson, C., Turner, H., Dougherty, B., Schmidt, D.C. (2011). Wreckwatch: Automatic traffic accident detection and notification with smartphones. Mobile Networks and Applications, 16: 285-303. https://doi.org/10.1007/s11036-011-0304-8

[10] Shukran, M.A.M., Abdullah, M.N., Isa, M.R.M., Ismail, M.N., Khairuddin, M.A., Maskat, K. (2018). Developing a framework for accident detecting and sending alert message using android application. International Journal of Engineering & Technology, 7(4.29): 66-68.

[11] Zualkernan, I.A., Aloul, F., Basheer, F., Khera, G., Srinivasan, S. (2018). Intelligent accident detection classification using mobile phones. In 2018 International

Conference on Information Networking (ICOIN), IEEE, pp. 504-509. https://doi.org/10.1109/ICOIN.2018.8343170

[12] Tian, D.X., Zhang, C., Duan, X.T., Wang, X.X. (2019). An automatic car accident detection method based on cooperative vehicle infrastructure systems. In IEEE Access, IEEE, 7: 127453-127463. https://doi.org/10.1109/ACCESS.2019.2939532

[13] Singh, D., Mohan, C.K. (2018). Deep spatio-temporal representation for detection of road accidents using stacked autoencoder. In IEEE Transactions on Intelligent Transportation Systems, IEEE, 20(3): 879-887. https://doi.org/10.1109/TITS.2018.2835308

[14] Yao, Y., Xu, M.Z., Wang, Y.C., Crandall, D.J., Atkins, E.M. (2019). Unsupervised traffic accident detection in first-person videos. In 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, pp. 273-280. https://doi.org/10.1109/IROS40897.2019.8967556

[15] Pillai, M.S., Chaudhary, G., Khari, M., Crespo, R.G. (2021). Real-time image enhancement for an automatic automobile accident detection through CCTV using deep learning. Soft Computing, 25: 11929-11940. https://doi.org/10.1007/s00500-021-05576-w

[16] Sadeky, S., Al-Hamadiy, A., Michaelisy, B., Sayed, U. (2010). Real-time automatic traffic accident recognition using HFG. In 2010 20th International Conference on Pattern Recognition, IEEE, pp. 3348-3351. https://doi.org/10.1109/ICPR.2010.817

[17] Adeli, H., Samant, A. (2000). An adaptive conjugate gradient neural network-wavelet model for traffic incident detection. Computer-Aided Civil and Infrastructure Engineering, 15(4): 251-260. https://doi.org/10.1111/0885-9507.00189

[18] Chakraborty, P., Sharma, A., Hegde, C. (2018). Freeway traffic incident detection from cameras: A semi-supervised learning approach. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE, pp. 1840-1845. https://doi.org/10.1109/ITSC.2018.8569426

[19] Chong, Y.S., Tay, Y.H. (2017). Abnormal event detection in videos using spatiotemporal autoencoder. In Advances in Neural Networks-ISNN 2017: 14th International Symposium, Springer International Publishing, pp. 189-196. https://doi.org/10.1007/978-3-319-59081-3_23

[20] Sirisha, U., Chandana, B.S. (2023). GITAAR-GIT based abnormal activity recognition on UCF crime dataset. In 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), IEEE, pp. 1585-1590. https://doi.org/10.1109/ICSSIT55814.2023.10061116

[21] Yang, G.H., Feng, W., Jin, J.T., Lei, Q.J., Li, X.H., Gui, G.C., Wang, W.J. (2020). Face mask recognition system with YOLOV5 based on image recognition. In 2020 IEEE 6th International Conference on Computer and Communications (ICCC), IEEE, pp. 1398-1404. https://doi.org/10.1109/ICCC51575.2020.9345042

[22] Pashaei, A., Ghatee, M., Sajedi, H. (2020). Convolution neural network joint with mixture of extreme learning machines for feature extraction and classification of accident images. Journal of Real-Time Image Processing, 17: 1051-1066. https://doi.org/10.1007/s11554-019-00852-3

[23] Bejan, A. (2016). Constructal thermodynamics. International Journal of Heat and Technology, 34: S1-S8. http://doi.org/10.18280/ijht.34S101

[24] Ali, F., Ali, A., Imran, M., Naqvi, R.A., Siddiqi, M.H., Kwak, K.S. (2021). Traffic accident detection and condition analysis based on social networking data. Accident Analysis & Prevention, 151: 105973. https://doi.org/10.1016/j.aap.2021.105973

[25] Zhang, Z.H., He, Q., Gao, J., Ni, M. (2018). A deep learning approach for detecting traffic accidents from social media data. Transportation Research Part C: Emerging Technologies, 86: 580-596. https://doi.org/10.1016/j.trc.2017.11.027

[26] Jiang, Y.F., Gong, X.Y., Liu, D., Cheng, Y., Fang, C., Shen, X.H., Yang, J.C., Zhou, P., Wang, Z.Y. (2021). Enlightengan: Deep light enhancement without paired supervision. In IEEE Transactions on Image Processing, IEEE, 30: 2340-2349. https://doi.org/10.1109/TIP.2021.3051462

[27] Zhang, Y.H., Zhang, J.W., Guo, X.J. (2019). Kindling the darkness: A practical low-light image enhancer. In Proceedings of the 27th ACM International Conference on Multimedia, pp. 1632-1640. https://doi.org/10.1145/3343031.3350926

[28] Guo, C., Li, C.Y., Guo, J.C., Loy, C.C., Hou, J.H., Kwong, S., Cong, R. (2020). Zero-reference deep curve estimation for low-light image enhancement. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, pp. 1780-1789. https://doi.org/10.1109/CVPR42600.2020.00185

[29] Sundari, T.J.V., Aswathy, J.G., Jayakamali, S. (2021). Accident detection and severity analysis using deep learning. In 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), IEEE, pp. 1-5. https://doi.org/10.1109/ICAECA52838.2021.9675

[30] Sirisha, U., Chandana, B.S., Harikiran, J. (2023). NAM-YOLOV7: An improved YOLOv7 based on attention model for animal death detection. Traitement du Signal, 40(2): 783-789. https://doi.org/10.18280/ts.400239

[31] Lee, K.B., Shin, H.S. (2019). An application of a deep learning algorithm for automatic detection of unexpected accidents under bad CCTV monitoring conditions in tunnels. In 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), IEEE, pp. 7-11. https://doi.org/10.1109/Deep-ML.2019.00010

[32] Sirisha, U., Praveen, S.P., Srinivasu, P.N., Barsocchi, P., Bhoi, A.K. (2023). Statistical Analysis of design aspects of various yolo-based deep learning models for object detection. International Journal of Computational Intelligence Systems, 16(1): 126. https://doi.org/10.1007/s44196-023-00302-w

[33] Chung, Y.L., Lin, C.K. (2020). Application of a model that combines the YOLOv3 object detection algorithm and canny edge detection algorithm to detect highway accidents. Symmetry, 12(11): 1875. https://doi.org/10.3390/sym12111875

[34] Sirisha, U., Chandana, B.S. (2023). Privacy preserving image encryption with optimal deep transfer learning-based accident severity classification model. Sensors, 23(1): 519. https://doi.org/10.3390/s23010519