International Journal of Research Publication and Reviews

Road Accident Detection Using CNN

¹Dr. Harihara Santhosh Dadi, ²P. Manasa, ³S. Chiranjeevi, ⁴K. Yuvaraj, ⁵K. Somasekhar

[1]Associate Professor,[2,3,4,5] Student,

Dept of Electronics and Communication Engineering (ECE), AITAM (Aditya Institute of Technology and Management), Tekkali.

ABSTRACT:

We are employing deep learning techniques to carry out our project. In order to recognize the signs of collisions or accidents, this technology uses convolutional neural networks to analyze image or video data from cars. With this method, tagged datasets of photos with instances of accidents and non-accident events are used to train deep learning models. After being trained, these models can identify and react to collisions instantly, enhancing both emergency response times and general traffic safety.

Key Words: Artificial Intelligence, Deep learning techniques, Convolutional neural networks(CNNs), Image or video data analysis, Tagged datasets, Accidents and non-accident events, Training deep learning models, Instant identification and reaction to collisions

1. Introduction

Vehicle accidents ,also known as traffic collisions or crashes, occur when a vehicle collides with another vehicle collides with another vehicle, pedestrian, animal, road debris, or other stationary obstruction. These incidents result in varying degrees of damage. These accidents can occur due to various factors such as human errors, mechanical failure, adverse weather conditions or road infrastructure issues. we can identify these accidents through sensors, cameras and data analysis etc…

1.1 Convolutional Neural Network

A specific kind of deep learning model called a Convolutional Neural Network (CNN) is used to process and analyze visual input, especially photos and videos. It takes its cues from how the visual cortex of an animal is structured, which enables it to automatically learn from and extract information from unprocessed input data. For several computer vision applications, including object identification, picture segmentation, and image classification, CNNs have emerged as the cutting-edge method. Convolutional layers, pooling layers, and fully linked layers are the main parts of a CNN. In order to extract features and produce feature maps, convolutional layers apply filters to the input data. Pooling layers effectively reduce the computational effort and control overfitting by lowering the spatial dimensions of the data. Fully linked layers handle the final predictions or classifications based on the learned attributes. CNNs have the ability to learn local patterns and spatial hierarchies, which allows them to detect objects and features in an image regardless of their position. They excel in tasks requiring the understanding of complex visual patterns as a result. CNNs can consistently and effectively classify objects and forms, from low-level properties like edges and textures to high-level features like shapes, because to its hierarchical architecture.

2. Proposed System

The 16-layer Visual Geometry Group network is known as VGG16. The Oxford University's Visual Geometry Group proposed the convolutional neural network (CNN) architecture. Karen Simonyan and Andrew Zisserman's 2014 publication "Very Deep Convolutional Networks for Large-Scale Image Recognition" introduced VGG16. Convolutional Layers: VGG16 consists of thirteen convolutional layers, five max-pooling layers, and a ReLU activation function after each layer. Completely Networked Layers:the final layer classifies data using a softmax activation function after each of the three fully connected layers, which are each followed by a ReLU activation function. Depth and Filter Size: The term VGG16 comes from its fixed architecture, which consists of 16 layers. Max-pooling layers have 2x2 filters with stride 2 and all convolutional layers utilize 3x3 filters. Number of Filters: The number of filters increases by two after each max- pooling layer until it reaches 512 filters, with 64 filters in the first layer. Easy to Understand design: VGG16's simple design makes it a popular choice for instructional reasons and as a starting point architecture for a variety of computer vision jobs. Efficient Feature Extraction: VGG16 is well-known for its capacity to extract rich hierarchical features from images. This capability has been shown to be useful in a variety of applications, including object identification, image classification, and feature extraction for further tasks. Transfer Learning: VGG16 is frequently used for transfer learning because of its efficiency and the availability of pre-trained weights on large-scale datasets like ImageNet.

Training time and resources can be saved by fine-tuning the pre-trained weights on smaller datasets for certain tasks. Stability and Robustness: The stability and robustness of VGG16 have been demonstrated in a variety of tasks and datasets through considerable research and testing.
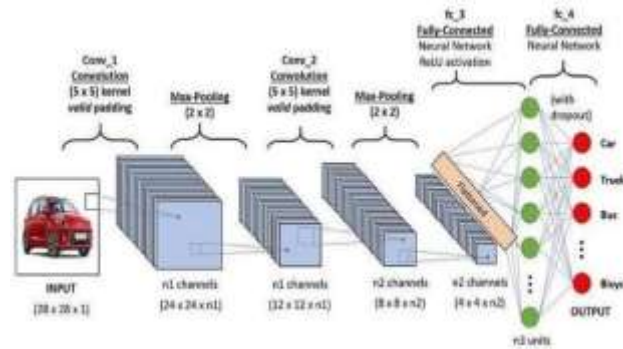
1: CNN

3. Literature Survey

[1] Detection And Prediction Of Traffic Accidents Using Deep Learning

This paper proposes a deep learning accident prediction model that uses Twitter data mining to detect and forecast transportation incidents. By using data such as sentiment analysis, emotions, weather, geocoded locations, and time information from twitter messages, the accuracy of accident detection is increased from 8% to 94%. The accuracy of the recommended strategy is 2% and 3% higher than the present methods, reaching 97.5% and 90%, respectively. Emojis, previously disregarded, are included to the study to provide context, and modern deep learning techniques like CNN and LSTM are employed to lower noise and increase accuracy.

[2] Early Warning Of Car Accidents Through Deep Learning

This project aims to address the challenge of assisting drivers in judging road conditions in order to lower the frequency of traffic accidents, with a specific focus on the issue of failing to pay attention to conditions ahead of the vehicle, which was identified as one of the leading causes of traffic accidents in Taiwan in 2021. By combining the YOLO and Deep Sort algorithms, this research seeks to provide drivers with real-time assistance in assessing road conditions and maintaining a safe distance from other objects on the road. This will promote safer driving practices and reduce the risk of collisions.

[3] Deep Learning Approaches for Vehicle and pedestrain detection in adverse weather

The difficulties of reliable vehicle and pedestrian recognition for road safety are discussed in this study, with particular attention paid to problems with different vehicle forms, congested areas, changing weather patterns, and driving habits. Using real-time datasets in bad weather, it assesses the efficacy of deep learning object-detection algorithms like faster F-RCN, SSD, HOG, and YOLO v7. The purpose of this research is to improve existing vehicle techniques in difficult situations by offering analyses.

[4] Real-Time Traffic Analysis Using Deep Learning Techniques And UAV Based Video

This paper uses deep learning techniques and UAV- captured footage to analyze traffic flow in real time and provide information about urban traffic congestion. These films contain moving objects, which are identified using sophisticated deep learning techniques. In order to assess traffic, pertinent mobility measures are calculated.

[5] Vehicle Detection And Counting

Sophisticated systems that employ real-time data from the global network of traffic cameras are needed for smart traffic control in cities. While collecting this type of data is easy, effectively utilizing it for traffic control is more challenging. As traditional vehicle detection systems like inductive loop detectors, infrared, and laser sensors become outdated, researchers are looking at deep learning and computer vision

technologies. The objective of this research is to review and compare various techniques for vehicle detection and counting in video-based traffic monitoring applications. To provide a comprehensive picture of the challenges and advancements in this industry, this paper delivers research findings and examines various methodologies.

4. METHODOLOGY

One type of artificial neural network that is particularly good at understanding visual data is called a convolutional neural network (CNN). CNNs have shown to be incredibly efficient in many different applications, such as picture classification, object identification, and image recognition. Convolution,

pooling, and fully linked layers are some of the layers that make up a convolutional neural network. Deep learning architectures called convolutional neural networks (CNNs) are mostly used to understand structured grid input, which includes photos and videos. In computer vision applications like object identification, facial recognition, photo classification, and so forth, CNNs have proven to be remarkably efficient An array of moveable filters that travel over the input data to carry out convolution operations makes up convolutional layers. These filters compute dot products between the filter and small portions of the input data in order to capture a range of patterns, such as edges, textures, and more intricate characteristics. Pooling Layers: After splitting the input feature map into rectangular, non-overlapping sections (usually 2 by 2 or 3 by 3), it keeps just the highest value in each sector. Fully Connected Layers: These layers enable high-level feature abstraction by connecting each neuron in the layers above and below.When generating final judgments or predictions, like categorizing an image, fully linked layers are frequently utilized. Flattening: Feature maps are usually flattened into a 1D vector before being transmitted through fully linked layers. With this enhancement, the data for the layers of conventional neural networks is made simpler. Batch normalization is a technique that normalizes each layer's input to speed up and stabilize training. By lessening overfitting, it can aid the network in learning more effectively. Output Layer: Depending on the goals and architecture of the network, the CNN's output—which is frequently a fully connected layer—is utilized for tasks like regression and classification.
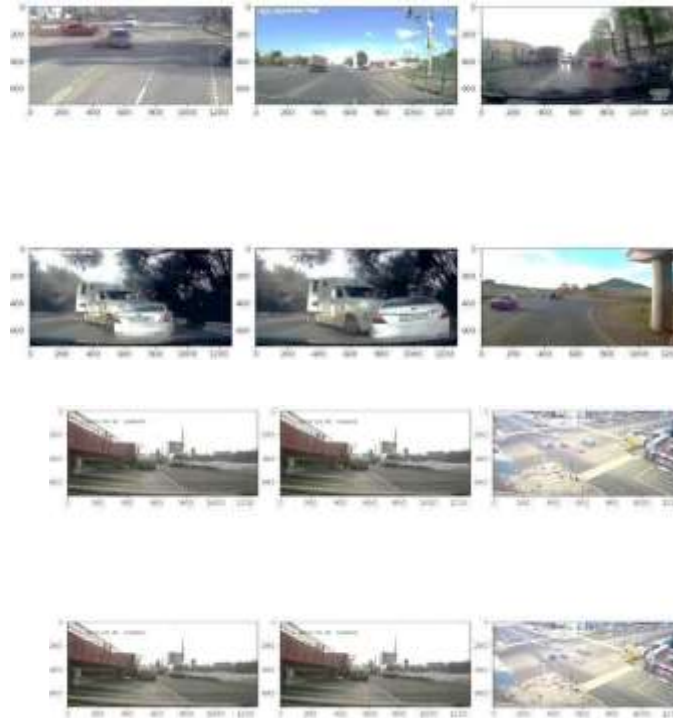
4.2 Architecture of CNN
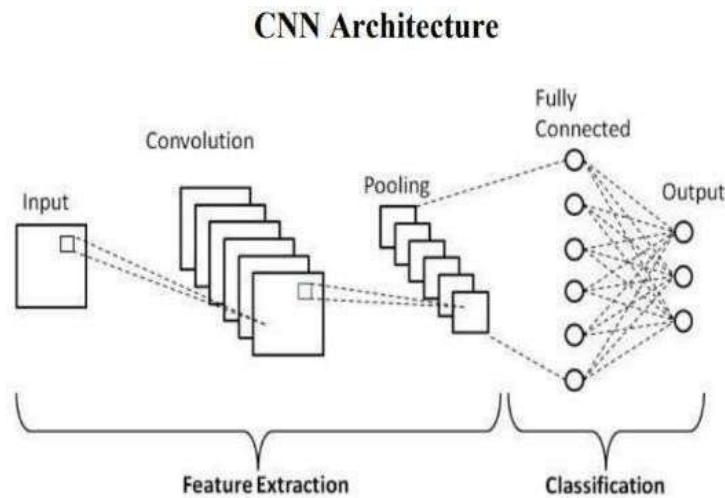
Block Diagram Of The Project
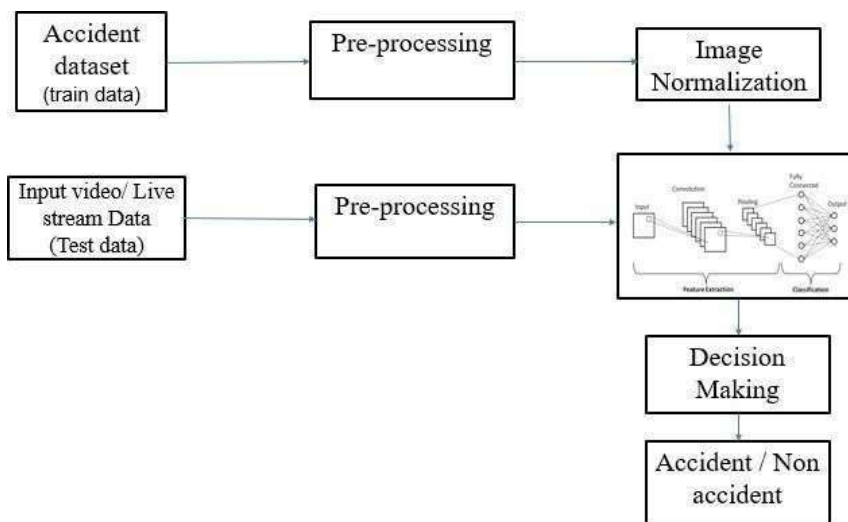
5. Results

Confusion Matrix

Epoch 1/50
8/8 [==============================] - 17s 428ms/step - loss: 0.9103 - accuracy: 0.5373 - val_loss: 0.6682 - val_accuracy: 0.5714
Epoch 2/50
8/8 [==============================] - 1s 185ms/step - loss: 0.5831 - accuracy: 0.6890 - val_loss: 0.5267 - val_accuracy: 0.7755
Epoch 3/50
8/8 [==============================] - 1s 182ms/step - loss: 0.4270 - accuracy: 0.8167 - val_loss: 0.4313 - val_accuracy: 0.8265
Epoch 4/50
8/8 [==============================] - 1s 184ms/step - loss: 0.2621 - accuracy: 0.9001 - val_loss: 0.2649 - val_accuracy: 0.8776
Epoch 5/50
8/8 [==============================] - 1s 196ms/step - loss: 0.1392 - accuracy: 0.9532 - val_loss: 0.2410 - val_accuracy: 0.8980
Epoch 6/50
8/8 [==============================] - 1s 188ms/step - loss: 0.0882 - accuracy: 0.9722 - val_loss: 0.2515 - val_accuracy: 0.8673

Epoch 44/50
8/8 [==============================] - 1s 184ms/step - loss: 0.0070 - accuracy: 0.9949 - val_loss: 0.1957 - val_accuracy: 0.9286
Epoch 45/50
8/8 [==============================] - 1s 179ms/step - loss: 0.0077 - accuracy: 0.9949 - val_loss: 0.1939 - val_accuracy: 0.9286
Epoch 46/50
8/8 [==============================] - 1s 184ms/step - loss: 0.0070 - accuracy: 0.9949 - val_loss: 0.1930 - val_accuracy: 0.9286
Epoch 47/50
8/8 [==============================] - 1s 181ms/step - loss: 0.0076 - accuracy: 0.9949 - val_loss: 0.1951 - val_accuracy: 0.9286
Epoch 48/50
8/8 [==============================] - 1s 178ms/step - loss: 0.0077 - accuracy: 0.9949 - val_loss: 0.1928 - val_accuracy: 0.9286
Epoch 49/50
8/8 [==============================] - 1s 183ms/step - loss: 0.0070 - accuracy: 0.9949 - val_loss: 0.1941 - val_accuracy: 0.9286
Epoch 50/50
8/8 [==============================] - 1s 188ms/step - loss: 0.0076 - accuracy: 0.9949 - val_loss: 0.1943 - val_accuracy: 0.9286

6. Conclusion

The design of a system that recognizes auto accidents automatically is the aim of our project. Enhancing emergency response to traffic accidents may be greatly enhanced by integrating cutting-edge deep learning networks—like convolutional and recurrent neural networks—with real-time visual identification and communication systems.

6.1 Future Scope

Multimodal Understanding: Include multimodal data in the model's extension. For example, to provide captions with more context, combine textual and graphic aspects. This can involve incorporating pre-trained language models like BERT or transformers to enhance the model's understanding of textual context. 2. minute-Grained Captioning: Acquire the skill of crafting captions that aptly communicate the interrelationships and intricate features present in images, including specific objects, actions, and spatial arrangements. This may mean incorporating modules into the captioning process for object recognition or scene graph generation. 3. Interactive Captioning Systems: Provide user-editable and augmentable captions for automatically generated content by developing interactive captioning systems. In this sense, human participation will enable the model to continuously improve and adapt over time**.**
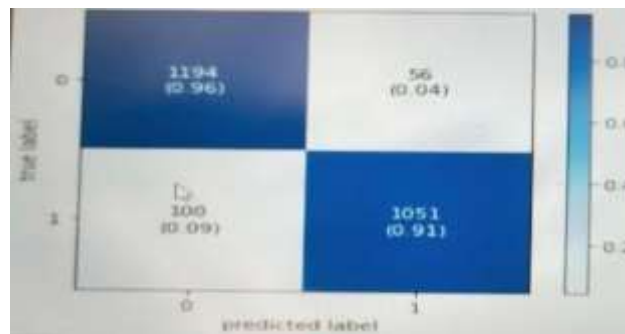
7. References

Abdel-Aty. M. Keller. J. Brady. P: Analysis of types of crashes at signalized intersections by using complete crash data and tree-based regression. Transp. Res. Rec. J. Transp. Res. Board 1905(1). 37-45 (2005).

Alhumoud. S: Twitter analysis for intelligent transportation. Comput. J. 62111. 1547-1556 (2015)

M. Hassaballah. M. A. Kenk. K. Muhammad. and S. Minace. "Vehicle detection and tracking in adverse weather using a deep learning frame work." IEEE Transactions on Intelligent Transportation Systems. 2020.

C. Badue. R. Guidolini. R. V. Carneiro. P. Azevedo. V. B. Cardoso. A. Forechi. L. Jesus. R. Berriel. T. M. Paixao. F. Mutz et al. "Self driving cars: A survey. Expert Systems with Applications. vol. 165, p. 113516. 2021.

G. Sullivan. K. Baker. A. Worrall. C. Attwood. and P. Remagnino. "Model-based Vehicle Detection and Classification Using Orthographic Approximations". J. Image and Vision Computing. Elsevier. pp. 649-54. 1997.

D. LI. B. Liang, and W. Zhang. "Real-time Moving Vehicle Detection. Tracking, and Counting System Implemented with OpenCV." Proceedings of IEEE ICIST. pp: 631-634. 2014.

Dedeoğlu. Yiğithan. Moving object detection. tracking and classification for smart video surveillance. Diss. bilkent university. 2004.

S. Sivaraman and M. M. Trivedi. "Looking at Vehicles on the Road: A Survey of Vision- Based Vehicle Detection. Tracking, and Behavior Analysis." in IEEE Transactions on Intelligent Transportation Systems. vol. 14. no. 4. pp. 1773-1795. Dec. 2013. doi: 10.1