Accident Detection using Convolution Neural Network and MobileNetV2

Chowdhury Sabir Morshed
Roll no-232IT009
National Institute of Technology,Surathkal
Karnataka, India

Jai Narayan Singh
Roll no- 232IT016
National Institute of Technology,Surathkal
Karnataka, India

***Abstract*—Innovation has made it easier to reach people in trouble quickly. In countries like India, where transportation via vehicles is increasing, traffic accidents are also on the rise. When accidents occur, getting medical help to the victim fast can greatly improve their chances of survival. This paper proposes a system that can send important information from the accident scene to nearby hospitals promptly, using technologies like Convolutional Neural Networks (CNN) and TensorFlow. These tools help in detecting accidents and sending email alerts to hospitals, ensuring the victim gets immediate medical attention. The system has been tested and achieves an accuracy of 84 percent, making it reliable and efficient without needing constant human oversight.**

I. Introduction

The rapid advancement of technology has revolutionized the automobile industry, yet it has also led to a concerning increase in road accidents, particularly evident in countries like India where negligence towards traffic rules and poor road conditions contribute significantly to alarming rates of fatalities. According to the National Crime Records Bureau (NCRB), India recorded 3,54,796 instances of road accidents in 2020 alone, resulting in 1,33,201 deaths and 3,35,201 injuries, with over speeding accounting for more than 60 percent of these accidents. Moreover, the delayed arrival of emergency medical care exacerbates the situation, highlighting the urgent need for innovative solutions to enhance emergency response systems. Various existing systems for detecting accidents employ a range of technologies such as sensors, cameras, and machine learning algorithms, aiming to swiftly identify accidents and alert emergency responders. In our proposed system, we aim to harness cutting-edge technologies such as Convolutional Neural Network (CNN) and TensorFlow to improve the efficiency of accident detection and emergency response. The system will receive real-time inputs from various sources, including roadside cameras, vehicle-mounted sensors, and mobile devices, and upon detecting an anomaly suggestive of a road accident, such as sudden deceleration or collision, CNN-based image analysis will confirm the presence of a vehicle collision and assess the severity of the incident. Subsequently, an alert will be generated and automatically transmitted to nearby hospitals, providing essential details such as the accident location, the number of vehicles involved, and the estimated severity of injuries. This proposed system holds immense potential for significantly reducing the time gap between accident occurrence and emergency medical response, thereby mitigating the impact of road accidents on human lives. By leveraging state-of-the-art technologies like CNN and TensorFlow, we can achieve higher accuracy and efficiency in accident detection, enabling prompt dispatch of emergency medical services and seamless coordination between various stakeholders, ultimately saving lives and reducing the burden on healthcare infrastructure. However, further research and development are needed to optimize the system's performance, ensure its reliability in real-world scenarios, and explore potential integration with existing traffic management systems. Nevertheless, the proposed system holds immense promise in transforming the landscape of road safety and emergency medical services, ushering in a new era of efficiency and effectiveness in accident response.

II. DEEP LEARNING TECHNIQUES

Deep Learning techniques have captivated researchers due to their remarkable advancements in various computer vision applications. A subset of machine learning, Deep Learning utilizes neural networks to autonomously learn from unstructured or unlabeled data. These strategies offer versatile solutions to a myriad of problems across diverse domains.

A. Generative deep architectures This architecture captures intricate interconnections within data for pattern recognition, elucidating the joint probability distribution of perceived information and associated classes. By leveraging Bayes' rule, the architecture can be transformed into a preferential model, enhancing its predictive capabilities.

B. Convolutional Neural Networks (CNNs) Specifically designed for image clustering, CNNs feature an input layer typically composed of two-dimensional arrays of neurons and an output layer with a single layer of result neurons. They excel in extracting features from images through convolutional operations, enabling efficient pattern recognition.

C. Recurrent Neural Networks (RNNs) RNNs form a category of artificial neural networks structured as directed graphs, facilitating chronological sequence modeling with interconnected nodes. This architecture's unique ability to retain information over time enables it to exhibit temporal dynamic behavior, making it suitable for tasks involving sequential data.

D. Deep Boltzmann Machine (DBM) DBM is a framework characterized by undirected connections between nodes across multiple hidden layers. It learns hierarchical representations of features from raw data, with hidden variables serving as input for subsequent layers. This hierarchical feature extraction process enables DBMs to effectively model complex datasets and discover latent structures within them.

III. Methodology

A. *Image Pre-Processing*

Image Pre-Processing involves a series of transformations applied to raw data prior to its utilization in a deep learning algorithm. The primary objective of pre-processing is to enhance image quality, facilitating more effective analysis. By suppressing undesired distortions and amplifying pertinent features, pre-processing ensures that the images are optimized for the specific application at hand. The necessary enhancements may vary depending on the requirements of the application, ultimately leading to improved performance and accuracy of the deep learning model.

B. *CNN*

A Convolutional Neural Network (CNN), a form of deep learning architecture, is specifically designed to handle structured data such as images. CNNs exhibit remarkable proficiency in identifying patterns within input images, ranging from basic shapes like lines and circles to more complex features like eyes and faces, making them highly effective for tasks in computer vision. One of the key advantages of CNNs is their ability to operate directly on raw images without requiring preprocessing. Typically, a CNN comprises multiple layers, often including convolutional layers that play a critical role in feature extraction. These layers are stacked on top of each other, with each layer capable of recognizing increasingly complex shapes. Although CNNs can be deep architectures, they typically consist of fewer than 20 layers. The effectiveness of a CNN stems from its use of convolutional layers, which are adept at extracting features from images. These convolutional layers are typically followed by pooling layers and fully connected layers in the CNN architecture, with activation layers usually preceding the convolutional layers. The preprocessing required by a ConvNet bears similarities to the organization of neurons in the human brain's visual cortex, serving as inspiration for its design. In our study, we utilized CNNs for image classification and recognition tasks, leveraging their robust feature extraction capabilities. The specialty of CNNs lies in their convolutional ability, particularly in extracting relevant features from images. Once these features are extracted by the network, they can be utilized in a variety of ways to achieve our objectives. Overall, CNNs represent a powerful tool in the field of computer vision, offering a sophisticated approach to feature extraction and image analysis. Through their unique architecture and capabilities, CNNs enable effective processing and interpretation of visual data, paving the way for advancements in various domains such as image recognition, object detection, and image synthesis.

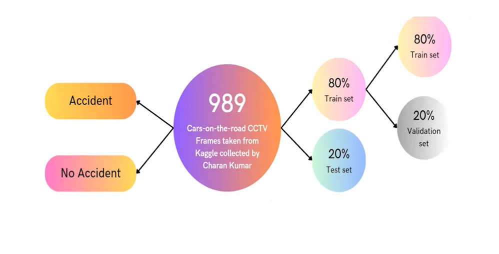


Fig. 1. Dataset Statistics

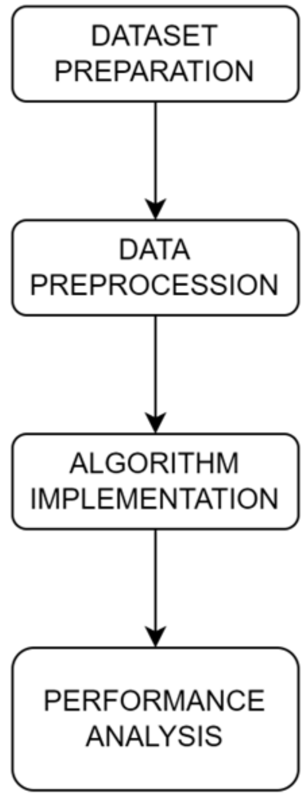


Fig. 2. Flow Diagram

- **Input Layer:**
  Artificial input neurons make up the information layer of a neural network, which gives the underlying information into the framework for handling by succeeding layers of counterfeit neurons.

- **Conv2D:**

The Conv2D layer, crucial in convolutional neural networks, learns from input channels and determines output channels in convolutions. Typically, the number of channels specified influences both learning and output. Post-convolution, Max Pooling reduces spatial dimensions of the output volume, aiding in feature extraction. This process, pivotal in deep learning, enhances the model's ability to recognize patterns and features within data. By iteratively adjusting convolutional parameters, such as channel count and kernel size, the network can effectively extract meaningful information from input data, facilitating tasks like image recognition and classification in computer vision applications.

- **Relu Activation:**
  ReLU (Rectified Linear Unit) is a widely used activation function in neural networks. It introduces non-linearity by outputting zero for negative inputs and maintaining positive inputs unchanged. This simplicity accelerates training and helps mitigate the vanishing gradient problem. ReLU's effectiveness lies in its ability to promote sparsity and improve network convergence. Despite its simplicity, ReLU has proven highly effective in deep learning tasks such as image classification and natural language processing.

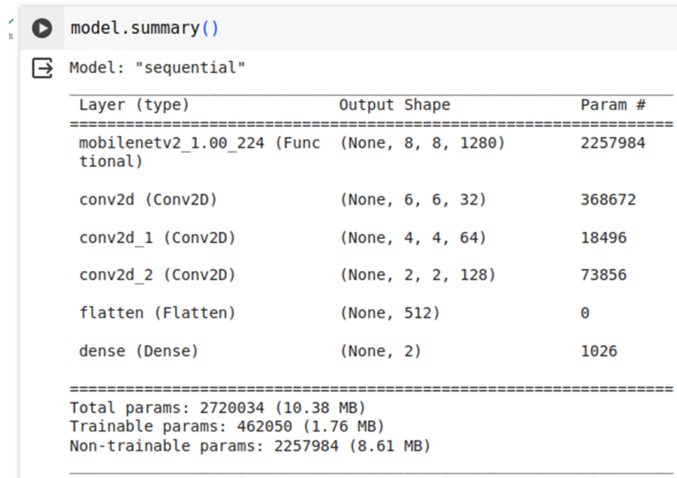
- **MaxPooling2D:**
  involves selecting the maximum value from each region of the feature map covered by the filter. This process effectively downsamples the input, capturing the most significant features. Consequently, the output of the MaxPooling layer comprises feature maps containing the most prominent features extracted from the preceding layer.

- **Flatten:**
  Flattening is the prevalent method used to convert information into a single-layered format for input into the subsequent layer. The output from convolutional layers is transformed into a single elongated feature vector. This vector is then linked to the final classification model, commonly referred to as a fully connected layer.

- **Dense Layer**
  The dense layer is a neural network layer characterized by significant interconnections, implying that each neuron within the dense layer receives input from every neuron in the preceding layer. Widely regarded as one of the most intricate layers in neural network models, the dense layer plays a crucial role in information processing and feature extraction.

C. *MobileNetV2*

MobileNetV2 is a variant of deep learning neural networks, specifically designed for handling large volumes of data,



```
model.summary()

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 mobilenetv2_1.00_224 (Func  (None, 8, 8, 1280)        2257984
 tional)

 conv2d (Conv2D)             (None, 6, 6, 32)          368672

 conv2d_1 (Conv2D)           (None, 4, 4, 64)          18496

 conv2d_2 (Conv2D)           (None, 2, 2, 128)         73856

 flatten (Flatten)           (None, 512)               0

 dense (Dense)               (None, 2)                 1026

=================================================================
Total params: 2720034 (10.38 MB)
Trainable params: 462050 (1.76 MB)
Non-trainable params: 2257984 (8.61 MB)
_________________________________________________________________
```

Fig. 3. CNN Model Summary

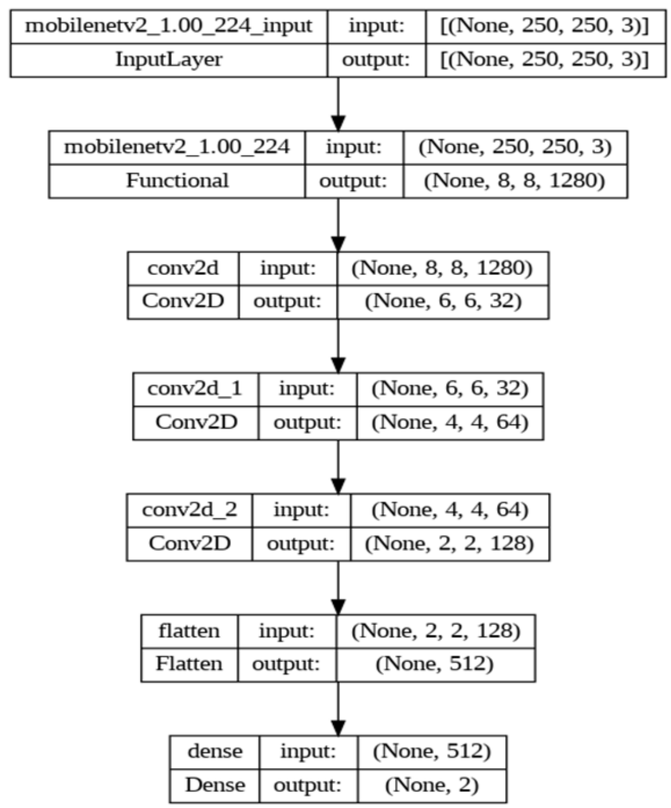


Fig. 4. CNN Model Architecture

particularly images. It excels at detecting patterns within input images, ranging from basic shapes like lines and circles to more complex features such as facial expressions and objects. Its architecture is particularly suited for tasks in computer vision, as it can operate directly on raw images without requiring preprocessing. MobileNetV2 typically consists of multiple layers, with the number of layers seldom exceeding 20, making it relatively lightweight compared to other architectures.

At the core of MobileNetV2's effectiveness lies its unique

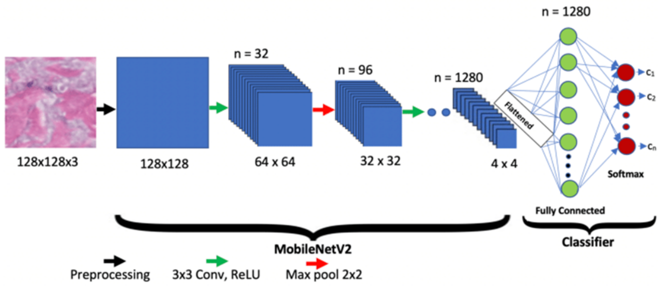


Fig. 5. MobileNetV2 architecture

architecture, which heavily utilizes a specialized layer known as the depthwise separable convolutional layer. These layers are stacked on top of each other, with each layer capable of capturing increasingly intricate patterns within the input data. MobileNetV2 is constructed as a multi-layered feed-forward neural network, with each layer building upon the hierarchical representations learned by the previous layers. Typically, convolutional layers are followed by depthwise separable convolutional layers, and activation layers follow convolutional layers.

Similar to the structure of the human brain's visual cortex, MobileNetV2 requires minimal preprocessing, as it can adaptively learn and extract features from raw input data. This architecture is inspired by the organization of neurons in the human brain's visual cortex, enhancing its efficiency in processing visual information.

In our study, we employed MobileNetV2 for image classification and recognition tasks. One of MobileNetV2's key strengths lies in its ability to efficiently extract features from images through its depthwise separable convolutional layers. Once these features are extracted, they can be utilized in various downstream tasks according to the specific requirements of the application.

IV. EXPERIMENTAL RESULTS

In this section, we'll look at the CNN algorithm's results, which is a classification technique. In this study, 989 accident and non-accident images were used. The system detects the accident and provides us with textual feedback such as "accident detected" or "accident not detected.". The classifier's accuracy and f1 score were used to assess its effectiveness. Fig 6 depicts the accident prediction wheather it an accident or non accident based upon its learning. Figure 7 shows a graphic portrayal of the current structure for training loss and training accuracy. Figure 8 shows a graphic portrayal of the current structure for test lost and test accuracy. To make an informed decision about classifier effectiveness, it's crucial to consider metrics beyond just classification accuracy. That's where the F1 score comes in. It evaluates the balance between precision and recall, giving insight into how well the classifier performs in handling positive instances in the data. This comprehensive view helps in assessing the classifier's performance more accurately. The

Fig. 6. Detection Result

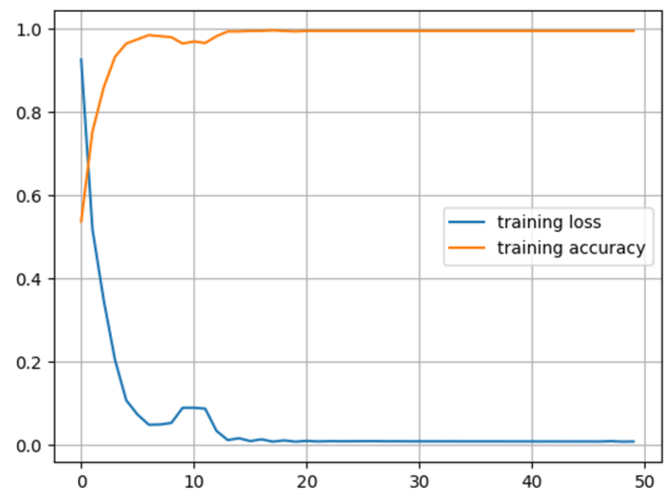


Fig. 7. Graphical portrayal of the created framework with train loss and training accuracy

f1 score is 91 percent, indicating that the system is highly efficient same is represented by fig 9.

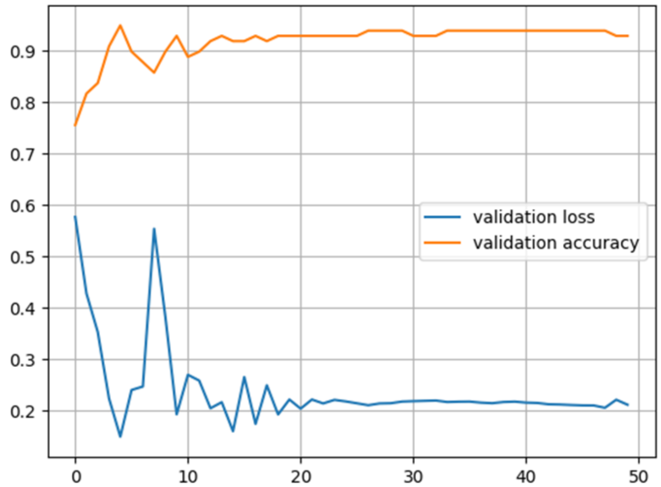


Fig. 8. Graphical portrayal of the created framework having test loss and test accuracy

```
              precision    recall  f1-score   support

     Class 0       0.90      0.91      0.91        47
     Class 1       0.92      0.91      0.91        53

    accuracy                           0.91       100
   macro avg       0.91      0.91      0.91       100
weighted avg       0.91      0.91      0.91       100

Micro-Accuracy: 0.91
Macro-Accuracy: 0.91
```

Fig. 9. Model precision, recall and f1-score

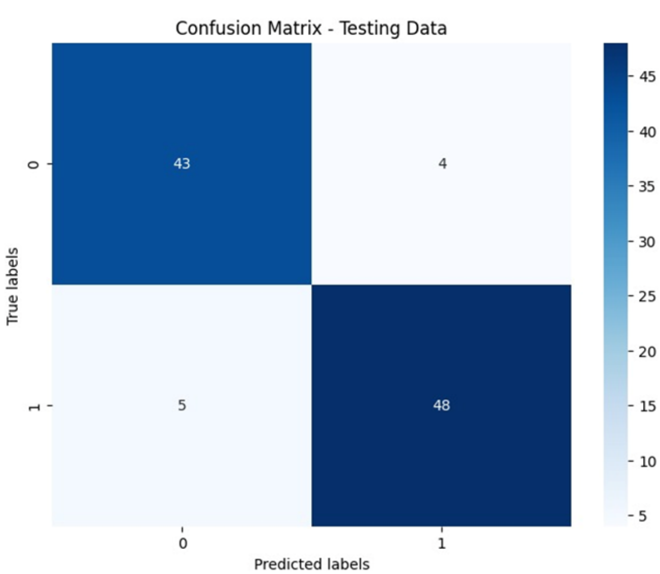


Fig. 10. Confusion Matrix on testing data

V. CONCLUSION AND FUTURE WORK

This paper outlines the development of an accident detection empowered by machine learning technology. It leverages Convolutional Neural Network (CNN) architecture implemented through TensorFlow and MobileNetV2to construct a model capable of detecting accidents.The system achieves an accuracy rate of 91 percent, indicative of its enhanced performance. For future scope we can incorporate a model which would help in detecting the accident and will send an email alert or notification to the nearby hospital as well as to nearest police station and helping the victim to get immediate medication and help. So that the human lives can be saved thus reducing the death rate caused due to accidents.

REFERENCES

[1] NCRB data on road accident . Available https://morth.nic.in/sites/default/files/RA$_2$022$_3$0$_O ct.pdf$.

[2] Unaiza Alvi, Muazzam A. Khan Khattak, Balawal Shabir, Asad Waqar Malik, Sher Ramzan Muhammad “A Comprehensive Study on IoT Based Accident Detection Systems for Smart Vehicles”, (IEEE 2020).

[3] Akanksha Jadhav, Shruti Jadhav, Archana Jalke, Kirti Suryavanshi “Road Accident Analysis and Prediction of Accident Severity using Machine Learning”, (IJERT 2020)

[4] Sathiya Devi S, Parthasarathy G, Mohamed Asif L, Ned Curtis Dylan P, Parasuraman B, Vinoth R “Prediction of Accident Prone-Zone and Alert Messages to the Hospitals Using Machine Learning and Internet of Things”, (IJERT 2020).

[5]] Shraddha S. More, Mansi Ashok Mange, Mitheel Sandip Sankhe, Shwethali Santosh Sahu. ”Convolutional Neural Network based Brain Tumor Detection”, 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021.

[6] Alexandra Fanca (POP), Adela Puscasiu, Honoriu Vălean, Silviu Folea “Trauma Accident Detecting and Reporting System”, (IEEE 2018).

[7] Sherwin Joseph Sunny; Rohan Roney “Accident Detection Using Convolutional Neural Networks”,(IEEE 2019).

[8] Vaishali Shrivastava, Manasi Giancana, “A Review Paper on Pre and Post Accident Detection and Alert System: An IoT Application for Complete Safety of the Vehicles” (2020).