LITERATURE SURVEY

The frequency of road accidents is on the verge of increasing worldwide in recent years. As per the survey of National Highway Traffic Safety Administrator, nearly one in five motor vehicle crashes are caused by distracted driver. The research conducted by B. Baheti, S. Gajre and S. Talbar **[1]**uses VGG-16 architecture which is modified for this particular task and various regularization techniques are implied in order to improve the performance. Experimental results have shown that their system has outperformed previous methods in the literature achieving an accuracy of 96.31% and it processed 42 images per second on GPU.they also study the effect of dropout, L2 regularization and batch normalization on the performance of the system. Next, they presented a modified version of the architecture which achieved 95.54% classification accuracy with the number of parameters reduced to 15M which was originally 140M.

The research paper **[2]** focuses on developing a CNN-based method to detect distracted drivers and identify the cause of distractions such as talking, sleeping, or eating through face and hand localization. Four architectures—CNN, VGG-16, ResNet50, and MobileNetV2—were employed for transfer learning. To validate the effectiveness of the proposed model, it was trained using thousands of images from a publicly available dataset sourced from the State Farm Distracted Driver Detection competition hosted by State Farm (Kaggle, 2016). This dataset comprises a significant number of images depicting ten different postures of distracted drivers, which were analyzed using various performance metrics.The results of the performance evaluation indicated that ResNet50 and MobileNetV2 achieved higher accuracy levels of 94.50% and 98.12%, respectively.

S. Kusuma, J. Divya Udayan and A. Sachdeva**[3]** In their research, they developed a system which detects the driver's drowsiness by using deep learning and computer vision. The results of their experiments displayed outperformance of their system in literature and achieved 96.31% accuracy and 42 images were processed per second on GPU. They also studied batch normalization,effect of dropout and L2 regularization on the system's performance. In this particular research, they estimated the position of the eyes and the face using the state of the art model for detecting better and reducing positives and false negatives. ResNet-152 yielded the most accuracy of 84.2% utilizing the softmax layer as a classifier which is superior to VGG-16 precision and much superior to Alex Net.

In this study **[4]** Sign Language Detection using LSTM Deep Learning Model and Media Pipe Holistic Approach the aim is to develop a deep learning model to recognize alphabet signs depicted in images and videos. We collected hand movement data using media pipe, which extracts coordinates of 21 points on the palm.These coordinates are then converted to numpy array and then for the sign detection it is given input to a Long Short-Term Memory(LSTM) model. The output of the model is a representation of strings ranging from 'A' to 'Z' of the detected alphabet.

The research conducted by C. Chanachan, P. Thanesmaneerat, T. Mahasukon, J.

Polvichai and S. Dumkor**[5]** aims to develope a Car Driver's Behavior Detection System using deep learning. Video data is captured from the cameras which are placed inside the passenger's compartment and is given to the system for analysis. It uses LSTM,YOLOv5 and Mediapipe models for analysis. The system's performance is evaluated using F1 scores. The system alarms the driver and identifies inappropriate behaviour and thus aims to reduce driver distraction and fatigue and therefore reduce accidents. However variation in brightness level in different types of enviornments may affect system's performance.

The research paper by G. -W. Gao and M. -Y. Chen**[6]** [they use a few key models and using media pipe holistic to be able to extract key points. This is going to allow them to extract key points from our face. The system uses tensorflow and keras and builds up a long short-term memory(LSTM) model to be able to predict the action which is being shown on the screen. they need to do is collect a bunch of data on all of our different key points, so they collect data on their face and save those as numpy arrays. This face detection method was based on deep neural network using LSTM layers to go on ahead and predict that temporal component, which is able to predict action from several frames not just a single frame. Integrate using OpenCV and then proceed to make real-time predictions using the webcam.

The paper **[7]** by B Sundar, T Bagyammal on American Sign Language Recognition for Alphabets Using MediaPipe and LSTM aims to develop, a simple and efficient vision-based approach for American Sign Language (ASL) alphabets recognition utilized to recognize static and dynamic gestures. Mediapipe introduced by Google has been used to get hand landmarks. A custom dataset has been generated and used for the experimental study. Hand gesture recognition has been done by using LSTM. The proposed system has been tested with 26 alphabets and an accuracy of 99% has been achieved. This work can be used to convert hand gestures into text.The dataset required for the experimental study of the proposed algorithm is created using MediaPipe which is an open-source library. It captures the hand gestures for 30 frames and stores them as a NumPy array for training the model. This dataset has 30 different permutations of each alphabet. The same process is repeated with four different demographics of humans to get a diverse dataset. The four different humans are two male hands, one female hand, and one kid's hand. This makes a total of 93,600 NumPy arrays, these are indexed and stored. The dataset has different parameters which are x ,y, and z axes of the hand joints. The dataset is divided into 90% for training and 10% for testing.

The paper **[8]** by A. Shaout, B. Roytburd and L. A. Sanchez-Perez presented a method for detecting driver distraction using computer vision methods within an embedded environment. By using SqueezeNet architecture, which is optimized for embedded deployment, and benchmarking it on a Jetson Nano embedded computer, this paper demonstrates a viable method of detecting driver distraction in real time. The method demonstrated here involved modifications to SqueezeNet to be trained on an AUC Distracted Driver Dataset, yielding max accuracy as 93% when detecting distracted driving.

The paper proposed by Aljasim M, Kashef R.**[9]**proposes E2DR, a new scalable model that uses stacking ensemble methods to combine two or more deep learning models to improve accuracy, enhance generalization, and reduce overfitting, with real-time recommendations. The most successful E2DR variant, which integrates ResNet50 and VGG16 models, demonstrated a significant test accuracy of 92%. This accuracy was attained when evaluating the model on prominent datasets, such as the State Farm Distracted Drivers dataset, utilizing innovative data splitting techniques

The detection algorithm**[10]**in the paper consists of two modules. The first module predicts the boundaries of the driver's right hand and right ear from images. The second module takes the boundaries as input and predicts the type of distraction. 106,677 frames were collected from videos of twenty participants in a simulator. These frames were used for training (50%) and testing (50%). The distraction classification results indicated that the proposed framework detects normal driving, using the touchscreen, and talking with a phone with F1-scores of 0.84, 0.69, and 0.82 respectively. For overall distraction detection, it achieved F1-score of 0.74. The whole framework is executed at 28 frames per second.

The pape**r** **[11]** have attained an appreciable number of deep learning algorithms on the dataset like convolutional neural network (CNN) and VGG16 to detect what the driver is doing in the car as given in the driver images. This process can also be completed by predicting the likelihood of the driver's actions in each picture. Of all models, they distinguished that the VGG16 Algorithm has conquered CNN with a loss of 0.298 and an Accuracy of 91.7%.

The paper **[12]** presents a distracted driver detection system that utilizes ensemble techniques to classify various types of distracted activities. Different convolutional networks are trained on images with their final layers removed to obtain feature vectors. These feature vectors are then stacked using the stacking ensemble technique and trained on a convolutional network. This stacking approach achieves an impressive accuracy of 97% in detecting distracted driver postures. The study provides insights into how the models predict the desired classes, demonstrating the potential for real-time application in detecting driver activities. The proposed model, based on various CNN architectures including VGG19, InceptionV3, Xception, and ResNet50, outperforms pre-trained models and requires less computational time. The stacked ensemble approach demonstrates superior performance compared to earlier research models.

In another paper **[13]**, a detection method based on Histogram of Oriented Gradient (HOG) and Support Vector Machine (SVM) is proposed for identifying driver distraction behavior. The method involves detecting the driver's interest region from video images, enhancing, denoising, and normalizing the image, followed by extracting features using HOG. Cross-validation is utilized to optimize parameters in SVM. The effectiveness of the proposed method is validated by comparing it with classical SVM and Local Binary Pattern (LBP) based SVM algorithms. Experimental results show that the proposed HOG+SVM method achieves better classification accuracy, with an average recognition rate of 93.33% for driver distraction behavior.

In this paper**[14]**, they suggest an ensemble approach for detecting driver distraction using the State Farm Driver Distraction dataset. To accomplish reliable and accurate classification, they combine NasNet-A large, ResNext-101, and EfficientNet-B0 Convolutional Neural Network (CNN) models. Various driver distraction scenarios are shown in the State Farm Driver Distraction dataset, which is a sizable collection of labeled images. The ensemble model aims to achieve higher accuracy and robustness in detecting driver distraction. The ensemble model uses fusion techniques that enhance the overall performance and reliability of the system. they evaluate the performance of our ensemble model from accuracy and loss function. Findings show that the ensemble model performs better than individual CNN models.

The study investigates**[15]** the driver pose distraction. The distracted driver pose is detected through pose detection using landmark feature extraction. The dataset used is taken from State Farm Distracted Driving Detection which is a secondary dataset. The extracted driver poses are in numeric form and used as the input for the classification model to detect ten kinds of driving distractions ranging from safe driving to talking to passengers. The classifiers used to detect distracted drivers are CatBoost, Decision Tress, Random Forest, and SVM. The model performance was evaluated using Precision, Recall, Accuracy, and F1-score. Furthermore, prediction time was recorded to determine the best model to detect driver distraction. The results show that the CatBoost model outperformed all other models with an F1-score and prediction time of 0.899 and 0.014 seconds. Random Forest model performance is the second highest compared to four others with F1-score and prediction time of 0.881 and 0.028 seconds. The catBoost model is the best even though Cat Boost model execution time is not the highest and the model performance is not significantly different compared to Random Forest.

REFRENCES

[1] B. Baheti, S. Gajre and S. Talbar, "Detection of Distracted Driver Using Convolutional Neural Network," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 2018, pp. 1145-11456, doi: 10.1109/CVPRW.2018.00150. keywords: {Vehicles;Computer architecture;Vehicle crash testing;Convolutional neural networks;Task analysis;Wheels;Roads},

[2]Md. Uzzol Hossain, Md. Ataur Rahman, Md. Manowarul Islam, Arnisha Akhter, Md. Ashraf Uddin, Bikash Kumar Paul,Automatic driver distraction detection using deep convolutional neural networks,Intelligent Systems with Applications,14,2022,200075,ISSN 2667-3053,https://doi.org/10.1016/j.iswa.2022.200075.(https://www.sciencedirect.com/science/article/pii/S2667305322000163)

[3]S. Kusuma, J. Divya Udayan and A. Sachdeva, "Driver Distraction Detection using Deep Learning and Computer Vision," 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kannur, India, 2019, pp. 289-292, doi: 10.1109/ICICICT46008.2019.8993260. keywords: {Deep Learning;Convolutional Neural Network;Object detection;Computer Vision},

[4]M. Deshpande et al., "Sign Language Detection using LSTM Deep Learning Model and Media Pipe Holistic Approach," 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), Greater Noida, India, 2023, pp. 1072-1075, doi: 10.1109/AISC56616.2023.10085375. keywords: {Deep learning;Image recognition;Statistical analysis;Data preprocessing;Gesture recognition;Media;Data models;Deep Learning;Image Processing;LSTM;Media pipe;NumPy;OpenCV;Python;Sign Language;TensorFlow},

[5]C. Chanachan, P. Thanesmaneerat, T. Mahasukon, J. Polvichai and S. Dumkor, "Car Driver's Behaviors Detections Using Ensemble Model," 2023 International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC), Jeju, Korea, Republic of, 2023, pp. 1-6, doi: 10.1109/ITC-CSCC58803.2023.10212694. keywords: {Deep learning;System performance;Fatigue;Data models;Behavioral sciences;Automobiles;Integrated circuit modeling;Vehicles;Accidents;Logistics;Driver's Behaviors Detections;Mediapipe;YOLO;Image Processing;Deep Learning;LSTM},

[6]G. -W. Gao and M. -Y. Chen, "Base on Long Short-term Memory Network for Fatigue Detection," 2022 International Conference on System Science and Engineering (ICSSE), Taichung, Taiwan, 2022, pp. 055-058, doi: 10.1109/ICSSE55923.2022.9948238. keywords: {Deep learning;Visualization;Solid modeling;Webcams;Computational modeling;Glass;Predictive models;fatigue detection;mediapipe;tensorflow;long short-term memory network},

[7]B Sundar, T Bagyammal,American Sign Language Recognition for Alphabets Using MediaPipe and LSTM,Procedia Computer Science,Volume 215,2022,Pages 642-651,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2022.12.066.(https://www.sciencedirect.com/science/article/pii/S1877050922021378)

[8]A. Shaout, B. Roytburd and L. A. Sanchez-Perez, "An Embedded Deep Learning Computer Vision Method for Driver Distraction Detection," 2021 22nd International Arab Conference on Information Technology (ACIT), Muscat, Oman, 2021, pp. 1-7, doi: 10.1109/ACIT53391.2021.9677045. keywords: {Deep learning;Computers;Performance evaluation;Computer vision;Image coding;Neural networks;Real-time systems;Computer Vision;Embedded Software;Deep Learning;Distracted Driving;SqueezeNet},

[9]Aljasim M, Kashef R. E2DR: A Deep Learning Ensemble-Based Driver Distraction Detection with Recommendations Model. Sensors (Basel). 2022 Feb 26;22(5):1858. doi: 10.3390/s22051858. PMID: 35271004; PMCID: PMC8914716.

[10]Li, Li & Zhong, Boxuan & Hutmacher, Clayton & Liang, Yulan & Horrey, William & Xu, Xu. (2020). Detection of driver manual distraction via image-based hand and ear recognition. Accident; analysis and prevention. 137. 105432. 10.1016/j.aap.2020.105432.

[11]Reddy, Shiva & Neelima, Pilli & Priyadarshini, Vaishnavee & Chigurupati, Swaroop. (2022). An approach to classify distraction driver detection system by using mining techniques. Indonesian Journal of Electrical Engineering and Computer Science. 27. 1670. 10.11591/ijeecs.v27.i3.pp1670-1680.

[12]K. R. Dhakate and R. Dash, "Distracted Driver Detection using Stacking Ensemble," 2020 IEEE International Students' Conference on Electrical,Electronics and Computer Science (SCEECS), Bhopal, India, 2020, pp. 1-5, doi: 10.1109/SCEECS48394.2020.184. keywords: {Computational modeling;Stacking;Wheels;Sensors;Smart transportation;Automobiles;Vehicles;stacking ensemble;distracted driver;CNN},

[13]Q. Bu, J. Qiu, H. Wu and C. Hu, "Research on Driver's Distracted Behavior Detection Method Based on Multiclass Classification and SVM," 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO), Dali, China, 2019, pp. 444-448, doi: 10.1109/ROBIO49542.2019.8961551. keywords: {Driver distraction;histogram of oriented gradient;cross validation;support vector machine},

[14]S. Subbulakshmi, A. K. Arathy and P. Madhu, "Ensemble model for Driver Distraction Detection," 2023 4th IEEE Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2023, pp. 1-7, doi: 10.1109/GCAT59970.2023.10353386. keywords: {Real-time systems;Robustness;Road safety;Safety;Convolutional neural networks;Ensemble learning;Injuries;Driver distraction;Road accidents;NasNet-A large;ResNext-101;EfficientNet-B0;Convolutional Neural Network;Driver distraction scenarios},

[15]F. A. Bachtiar, I. Arwani, R. A. B. Prasetyo, G. Pangestu, R. S. Soetedjo and A. R. M. Mahesa, "Distraction Detection in Driving Using Pose Extraction and Machine Learning Techniques," 2023 IEEE 7th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Purwokerto, Indonesia, 2023, pp. 279-284, doi: 10.1109/ICITISEE58992.2023.10404467. keywords: {Human factors;Predictive models;Feature extraction;Safety;Decision trees;Random forests;Vehicles;distraction;detection;driving;pose;landmark},