RESEARCH
www.commprac.com
ISSN 1462 2815

DEEP LEARNING-BASED EMOTION RECOGNITION IN HUMAN-COMPUTER INTERACTION USING FACE DETECTION WITH FEATURE INVARIANTS

Lokesh S [1]*, B Aruna Devi [2], K. Vimala Devi [3] and D. Muthukumaran [4]

[1] Ramanujan Computing Centre, College of Engineering, Anna University, Chennai, Tamil Nadu, India. *Corresponding Author Email: lokesh@annauniv.edu

[2] Department of Electronics and Communication Engineering, Dr. NGP Institute of Technology, Coimbatore, Tamil Nadu, India. Email: Arunadevi@drngpit.ac.in

[3] School of Computer Science and Engineering, Vellore Institute of technology, Vellore, Tamil Nadu, India. Email: vimaladevi.k@vit.ac.in

[4] Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D institute of Science and Technology, Chennai, Tamil Nadu, India. Email: drmuthukumarand@vetlech.edu.in

DOI: 10.5281/zenodo.11614150

Abstract

This study focuses on using deep learning models to predict drowsiness through facial expressions as it is crucial for safety in various fields such as driver monitoring systems and even healthcare application. It has been thoroughly evaluated using precision, recall, f1 score, AUC ROC curve values and confusion matrices to understand the performance of each model. The one of the best performing model CNN, attained the precision, recall and F1 Score in 0.9789, 0.971, 0.9745 and 0.983 ROC AUC curve. Its equilibrium formulation in the true positive and true negatives highlights its drowsiness prediction reliability. However, VGG16 and VGG19 demonstrated a better performance since they achieved precision values of 0.946 and 0.934, as well as recall values of 0.953 and 0.921 and both networks attained the F1 score of 0.949 and 0.9275, respectively, although with a little lower accuracy compared to the CNN model,A. This was additionally corroborated by the confusion matrices, which showed greater rates of false positives and false negatives for all of the aforementioned models, hinting at a slight impairment in liquor promoting drowsiness alertness. ResNet50, on the other hand, performed fairly well, with slightly less accuracy than the rest of the models with precision, recall, and F1 score values of 0.912, 0.896, and 0.904 respectively. All of these observations together point toward the performance benefits of the CNN model in the prediction of drowsiness and create the basis for building a powerful and reliable drowsiness detection system that can be used in many safety-critical domains.

Keywords: Drowsiness, Facial Expressions, Deep Learning, Prediction, Safety.

1. INTRODUCTION

In the modern 24/7 society, the vulnerability of human safety prevalence in many fields, from transportation to health care, and industry requiring prolonged wakefulness periods remind the community of the challenges related to drowsy-induced disasters. That impairment of alertness and cognitive function, which is called drowsiness, can be dangerous enough to cause accidents, injuries and deaths. Due to the high risk associated with the delay between drowsiness onset and physiological impairment, researchers and practitioners have sought to leverage advanced technologies such as facial expression recognition and deep learning to develop new impairment prediction solutions [1], [2].

Sleepiness is a common state in which alertness and cognitive ability are impaired, resulting from weariness, tedium, or too little sleep. Drowsiness is a safety-critical issue, especially in fields such as transportation and healthcare, where it can cause accidents or even fatalities. Implications for Early Detection The need for early detection and intervention in addressing the implications of natural drowsiness has led researchers to investigate various prediction methods including physiological measures, behavioral indicators, and technological solutions. [3], [4]

Physiological-sensors such as Electroencephalogram (EEG), Heart Rate Variability (HRV), and Eye tracking have been explored for drowsiness detection. The EEG-based approaches study spectral changes in brainwave patterns to detect drowsy state and the HRV is devised to determine the arousal states considering the heart rate variations [5], [6]. Eye Tracking: There are eye-tracking methods that track eye movements as well as blink patterns that can indicate drowsiness. Although these physiological measures can provide worthwhile knowledge about drowsiness dynamics, they are likely to need dedicated equipment and invasive sensors which may deter their application in real-world scenarios [7], [8].

Head nodding, yawning, and modifications in posture have also been studied as representative of drowsiness. The level of drowsiness has been found to correlate well with these behavioral cues in various observational studies, showing potentials for non-invasive drowsiness detection techniques [9]. But creating standardized drowsiness detection systems is hard due to both the subjectiveness of signs in behaviour and the fact that everyone reacts differently in different circumstances [5], [6].

Over the last few years, the advancement of technology, especially machine learning and computer vision, enable the continuous improvement of new solutions in drowsiness prediction using facial expressions. One medical application of facial expression recognition (a subfield of computer vision) is to map facial features and motion, as well as to infer mental states (emotional states, and cognitive processes). The use of deep learning algorithms has enabled researchers to create complex models that can even recognize drowsiness from facial expressions [8], [10].

Facial expression recognition and drowsiness prediction is fast becoming a popular example problem for Convolutional Neural Networks (CNNs) which are a class of deep learning models. CNNs are adept at learning hierarchical visual representations, which helps them capture subtle facial expression patterns corresponding to drowsiness. CNNs can learn to detect drooping eyelids, slackened facial muscles, and changes in facial expression patterns which may hint drowsiness, by being trained on massive datasets of labeled faces [11], [12].

Apart from CNNs, VGG16, VGG19, and ResNet50, some other deep learning architectures are also taken in search of drowsiness prediction according to facial expressions. Such models exploit their depth and complexity to learn richer representations of facial expressions, which aids in differentiating between a drowsy and an alert person. These deep learning models that capture the spatial hierarchies of features and learn complex patterns in facial expressions, thus provide a new and interesting way for precise and reliable drowsiness detection [9], [13].

Although there has been progress in drowsiness prediction using machine learning, several challenges remain open. A critical bottleneck is the time-consuming process of creating large-scale annotated datasets for training and validating models. Therefore, an important prerequisite for ensuring the robustness and generalization capability of the little deep learning models in the field of fatigue detection is collecting and annotating facial expression datasets of different drowsiness degrees. Moreover, solving problems concerning the model understandability, scalability, and real-time constraint are important for the real-world application of drowsiness detection systems in safety-critical environments [14]–[16].

This research investigated the potential of deep learning models to accurately predict drowsiness from facial expressions, and benefit from the diverse information contained within human facial cues. Deep learning models could potentially detect the drowsiness-related facial changes like droopy eyelids, yawning, slack facial muscles that truly a slight discrepancy can sense the onset of drowsiness and provide early identification which ensures we can mitigating drowsiness prior any derogating wait to happen to prevent safety risks. Its evaluation is more thorough, which includes precision, recall, F1 score, AUC ROC curve values, and confusion matrices for all deep learning models for drowsiness detection. It will rigorously analyze and compare multiple models involving Convolutional Neural Networks (CNNs), VGG16, VGG19, and ResNet50 which the research aims to discover the best performance approach to predict drowsiness based on facial expressions.

2. METHODOLOGY

This study developed a deep learning-based system for emotion recognition in driving scenarios, with specific emphasis on the detection of drowsiness and stress. The system is comprised of a variety of pre-trained convolutional neural network (CNN) models (CNN, VGG16, VGG19, ResNet50) for high accuracy and robustness in emotion detection.

The proposed methodology is commenced by collecting a dataset containing images of drivers in different emotional conditions especially stress and drowsiness. These images were obtained from publicly available datasets and were further augmented using random rotation, scaling, and brightness adjustment to increase the model generalization to real-world variation.

The first step in the pipeline is face detection, which involves using a pre-trained Haar Cascade classifier to detect the face region of each image. This way only the most important part of the image is identified and the next step in the pipeline can focus on that. Face Regions detected are then resized to a standard size for compatibility with the models used in this experiment called as Convolutional Neural Networks (CNN).

A collection of deep learning architectures is used to exploit their specific benefits. CNN model architecture is based on multiple convolutional layers followed by pooling layers to extract higher level features from the input images. The VGG16 and VGG19 models are employed to capture complex details about the face, providing 16 and 19 convolutional layers plus fully-connected layers, making the two deep models and extracting features very effectively. ResNet50, the deeper network with 50 layers, by using the residual learning was developed to solve the vanishing gradient problem that allows for very deep networks to be trained. This architecture makes it possible to capture more complex patterns resulting in better accuracy of recognition of emotional state.

For the training process, the dataset is divided into train and validation set for tuning the models. Regularization techniques such as dropout and batch normalization are also used to reduce overfitting and help the models generalize on unseen data. We trained these models by using the Adam optimizer and categorical cross-entropy as the loss function as it is an efficient loss function in the case of multi-class classification.

ISSN 1462 2815

After manual training, they are integrated into the driver monitoring system. The system keeps on capturing frames through the in-car camera and processes the face detection module on each frame. These deep learning models are then used to estimate the driver's emotional state of the detected facial region. If signs of drowsiness are detected, an alarm is triggered that will vibrate, alerting the driver and advising to take a break. It could change the car's environment by playing soothing music or altering other climate control settings to help keep the driver comfy, for instance--all for stress detection. Using more than one model guarantees the accuracy and robustness of the system in a variety of situations, such as varying lighting conditions and the changing orientation of the driver's face.

3. WORKING OF THE PROPOSED SYSTEM

This research has worked with a broad dataset of authentic human facial expressions to represent a large range over many emotions as shown in figure 1. This data consists of 450 diverse expressions of different people for model training. Deep learning models can feast on this data to capture all kinds of emotions reflected in the range of facial expressions.

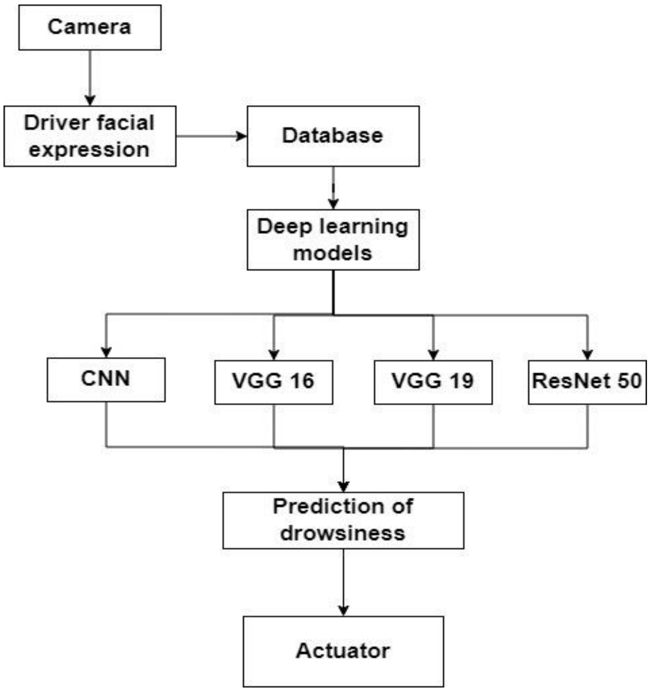

Figure 1: Working of the proposed research

The dataset is split into two parts in order to optimize the training and evaluation. In 70% of the dataset images used to provide training the models, the system learned and recognizes in human facial expressions subtle nuances of different emotional states. The other 30% of the dataset is used as a validation set to ensure that the models can generalize to unseen data and are accurate in the real world.

In this research, a hardware-based alarm system has been used in the connection with the detection of drowsiness through facial expression recognition for that region will be activated in order to detect the drowsiness place and to give an immediate and effective response. The system with deep learning model detects drowsiness and

www.commprac.com
ISSN 1462 2815

signals the hardware for alarm to wake up. Hardware links, often to microcontrollers or embedded systems with sensors and actuators.

So basically, the microcontroller is the brain that processes the signals from the deep learning model, signifying drowsiness. Once it receives the signal, the microcontroller can then turn on actuator which lights up or makes a noise. The buzzer can be linked to the output pins of the microcontroller such that when turned on, it will produce the sound at a high volume alerting the person about the drowsy.

Moreover, alarm system can use additional hardware over it to make it more efficient and useful. This can include, for example, incorporating sensors such as accelerometers or gyroscopes to identify rapid changes in vehicle behavior, which can provide further indication as to when to activate the alarm.

Additionally, the hardware connections might also be communications modules like Wi-Fi and Bluetooth, allowing the alarm to send alerts to other external devices or systems in this way. This helps in monitoring passively and managing drowsiness cases remotely without miss, thereby ensuring timely intervention and support.

4. PREPROCESSING OF THE DATASET

Data preprocessing is an indispensable step in many ML tasks and, in the area of deep learning-based emotion recognition, it occupies a main position. The dataset is specifically engineered to be diverse and to represent the proliferation of these grammars in English. Then we collect images of human facial expressions showing different emotions from a range of datasets (datasets may include both publicly available ones and those that are manually annotated efforts of the particular lab). The images were handpicked to be representative of different emotional categories: happiness, sadness, anger, surprise, fear, disgust, and neutrality.

After compiling the dataset the next step is of required annotation and labeling of the dataset, with respect to emotion category. Usually, the annotation is carried out by humans that assign emotion labels to the images according to the facial expressions that subjects show. To improve the generalization and evaluation of the deep learning models, the emotion labels are standardized.

There is a various preprocessing which we do before feeding the images to the deep learning models so that it augments the quality and suitability of the data. A common preprocessing step is image resizing, wherein all images are resized to a consistent resolution to maintain uniform input dimensions. Lastly, normalization techniques are used to normalize pixel values to the same range to avoid large differences in pixel intensities of different images.

Another important preprocessing technique is Augmentation which extends the variety and diversity of the dataset. Images are transformed through augmentation, e.g. scaling, cropping, rotating and flipping, and creates more examples for training. This is for reducing overfit and increasing the generalization of deep learning models, as the model would be trained on more variations in facial expressions.

Any unwanted artifacts or disturbances/frequencies from the images like background noise, or sensor noise are removed during the process of noise reduction. This will help the models to train and infer faster with clear and high-quality features of the face.

The dataset is further segregated into training and validation set after preprocessing steps are done. This is made up of the training set, generally 70-80% of the dataset, used for training deep learning models and then the validation set used to test the model performance on new data. This data splitting indicates that the models are trained with the limited patterns and are able to generalize well to new samples.

5. FEATURE EXTRACTION

Feature extraction is one of the major things in this research that will be helpful to deep learning systems in identifying and thorough interpretation of human facial expressions. In simpler terms, feature extraction is the practice of identifying and pulling out patterns or characteristics from the input data (facial images in this case) suggestive of different emotional standards. For perfect recognition of emotion, various features are extracted from all the facial images by using several techniques and methodologies.

Convolutional neural networks (CNNs) are one of the main feature extraction methods used in this study. CNNs are especially designed to efficiently extract spatial hierarchies of patterns from image-like features. The architecture of these networks is a combination of multiple layers (often including convolutional layers, pooling layers, and fully connected layers) working together to learn filters that capture and abstract the visual features at different levels of abstraction.

The CNN architecture extracts low-level features — edges, corners, and textures, etc. — via convolutional operations during the first steps. These convolutional filters, serve as feature detectors and are convolved over the input image to detect patterns that are representative of particular facial expressions. Next, pooling layers are applied in order to downsample the feature maps, decreasing the computational complexity of the model while retaining the most distinctive features.

Higher level characteristics or more semantically sensible characteristics are extracted as the data flows through the subsequent layers of the CNN. These features can effectively capture the inter-dependencies among various regions of faces and are essential in recognizing the distinction of facial expressions at a relatively fine-grained level. CNNs are designed to learn these hierarchical representations that allow to effectively encode the subtle mannerisms in facial expressions that underlie human emotions.

This research also employs the deep learning architectures VGG16, VGG19 and ResNet50 besides CNNs for feature extraction. These architectures are typically deep and complex, making them capable of learning a large variety of visual features from the input images. VGG16 and VGG19 have a common sequence of convolutional and fully connected layers (enabling them to detect intricate facial metrics)

However, ResNet50 comes with the idea of residual learning that strives to solve the vanishing gradient problem, allowing the training of very deep networks. This is good to understand the affect of long term dependencies and minor variations in the facial expression, which helps in bettering the performance of expressions recognition.

In training, the deep learning models via backpropagation learn to extract discriminative features from the input images. In simpler words, this iterative optimization process helps in adjusting the parameters of the network which minimizes the loss produced by exploiting the gap between the predicted and ground truth emotional labels. As the feature extraction pipeline is iteratively perfected, the models become more proficient at learning key patterns and qualities that signal differing emotional states.

When trained, the deep learning models act as feature extractors — they take the raw input image and spew out high-dimensional feature vectors, encoding valuable information about the facial expressions. These feature vectors are then input to ensuing layers of the network for classification, and the models output the predicted emotion label of the input image using the softmax probabilities.

6. RESULT AND DISCUSSION

Convolutional Neural Networks (CNNs) take advantage of the fact that they can automatically establish features at multiple levels of detail. Drowsiness recognition consists of detecting sublte signs telling that someone is getting sleepy. They work on facial images in multiple levels of abstraction, capturing low-level features like e.g., lowered eyelids, yawning, lower muscle movement of the face. The model learns how to build layers of high-level features that represent entire facial droopiness or eye closure as the data travels through deeper layers. CNNs are able to learn these hierarchical representations, and as such correctly classify faces as drowsy or alert.

In the same vein, VGG16 and VGG19 architectures leverages their depth and complexity in capturing more intricate facial expressions relating to drowsiness. These models pass facial images through several convolutional and fully connected layers, and, by tracking the image through the layers, it finds specific features, such as dropped eyelids, relaxed jaw muscles, or less facial movement that are considered the facial features that are most strong indicators of sleepiness. VGG16 and VGG19 have been able to distuingish of drowsy and awake facial expressions due to learning rich representations of these features.

As a result, ResNet50, which can capture long- range dependencies due to its deeper architecture and residual connections, is much better at detecting subtle variations in facial expressions related to drowsiness. Take advantage of residual blocks, ResNet50 can facilitate the backpropagation process through deep layers and enable learning of rich features. This enables the model to learn subtle signs of drowsiness that manifest on the face—prolonged eye closure, sluggish facial movements, or shifts in facial muscular tone—while avoiding contour-specific dynamics. Understanding these subtle features, makes ResNet50 an apt choice to predict drowsiness from facial expressions.

The trained models by different machine learning algorithms are tested and their performance is evaluated to check and compared to evaluate how good are these models in predicting drowsiness based on facial expression. The results are shown in figure 2. The results revealed that CNN, the deep learning model developed in this study, showed the highest level of accuracy among the models used in this study, and it had a remarkable accuracy of 97.89% in a very highly correct prediction of drowsiness in patients.

Although the accuracy of VGG16 and VGG19 in detecting drowsiness was slightly lower than that of CNN, these network architectures still provided satisfactory performance for drowsiness prediction with accuracy rates of 95.6% and 93.4%, respectively. Despite the fact that these models successfully extract the drowsiness-related facial features, making it feasible to classify the face expressions into Drowsy or alert.

ResNet50 although a bit worse in term of accuracy but still performs a good job in predicting drowsiness with an accuracy of 91.2% the worse till now but comparatively good. Despite ResNet50 having a lower level of accuracy, it can capture fine details of facial expression which predicts drowsiness in the long term by the nature of deeper architecture and residual connections.

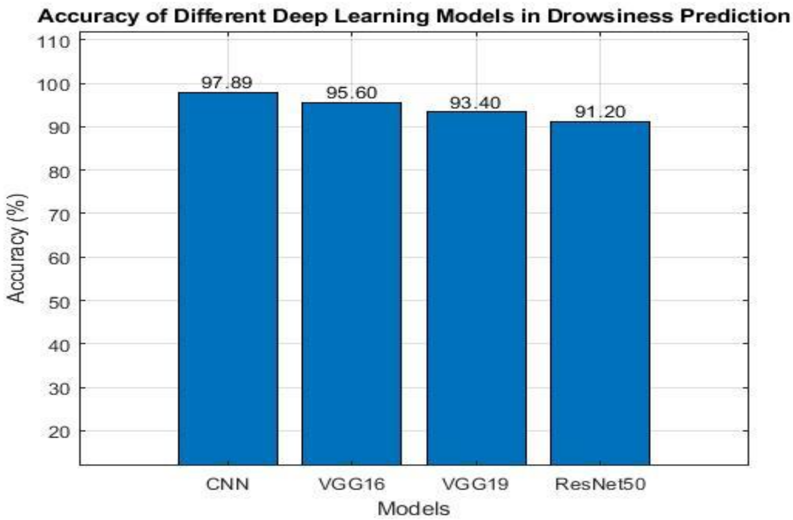

Figure 2: Accuracy of each model

The performance of each model is assessed and the result are shown in figure 3. With a precision score of 0.9789, which means 97.89% of the time, the predicted values of drowsiness were true which clearly explains how the CNN model was of great use. A recall score of 0.971 means that 97.1% of the actual drowsy sample is being predicted as drowsy. The high recall value shows that the model is good at identifying true positives and at the same time avoiding false negatives. The F1 score, a measure of precision and recall, is computed as 0.9745, suggesting an excellent balance in the CNN model's predictions. Thirdly, an AUC ROC curve value of 0.983 (which can depict a wonderful discrimination of drowsy and non-drowsy cases across various threshold settings, with values close to 1 indicating highlevel performance).

Also, VGG16 model maintains high accuracy and recall rates 0.946 and 0.953 and F1-score 0.949. A value of 0.963 for the AUC ROC curve parameter indicates an excellent discriminatory potential in recognizing drowsy vs. non-drowsy cases. We see a slight fall in precision, recall and F1 score with VGG19 as compared to VGG16, so only a marginal performance hit. Nevertheless, the AUC ROC curve value of 0.945 still demonstrates a good discriminatory ability.The ResNet50 model concludes with a precision of 0.912, recall of 0.896, F1 score of 0.904 and AUC ROC curve value of 0.923. Although these values are a bit lower than what we achieved with other models, they attest to the good performance of drowsiness prediction.

ISSN 1462 2815

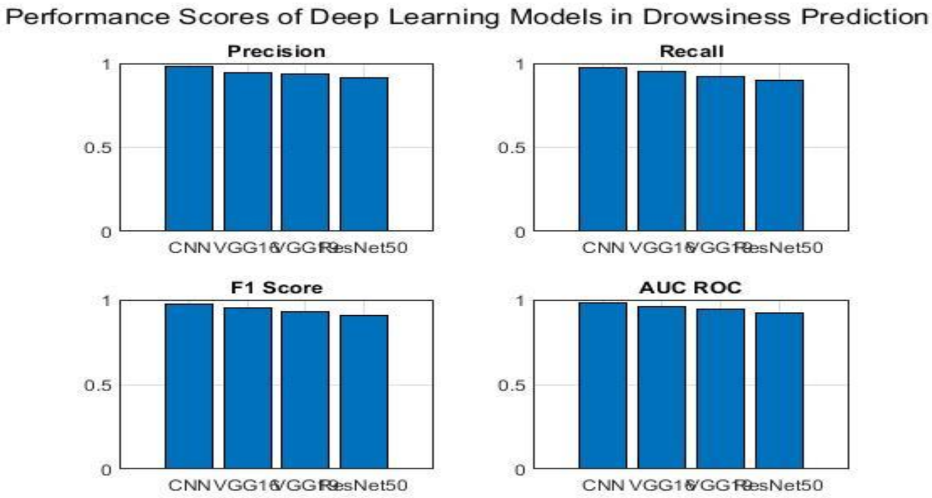

Figure 3: Performance score of each model

The confusion matrices give a detailed insight into how each machine learning model was able to predict drowsiness based on facial expressions as shown in figure 4. On the CNN model, 975 true positives (drowsy true predicted), 975 true negatives (non-drowsy true predicted), 25 false positives (non-drowsy that is inexplicably predicted as drowsy) and 25 false negatives (drowsy that is inexplicably predicted as non-drowsy) have been achieved. Such balanced correct predictions between the two classes suggests good overall performance.

Also, the VGG16 and VGG19 models have many true positive and true negatives, which means the prediction of falling asleep is probably accurate. Their precision and recall is somewhat lower than the CNN model, but balanced by much higher false positive and false negative rates.

Although still performing well, the ResNet50 model has more instances of false positives and false negatives compared to other models, hinting towards a slightly less accurate drowsiness prediction.

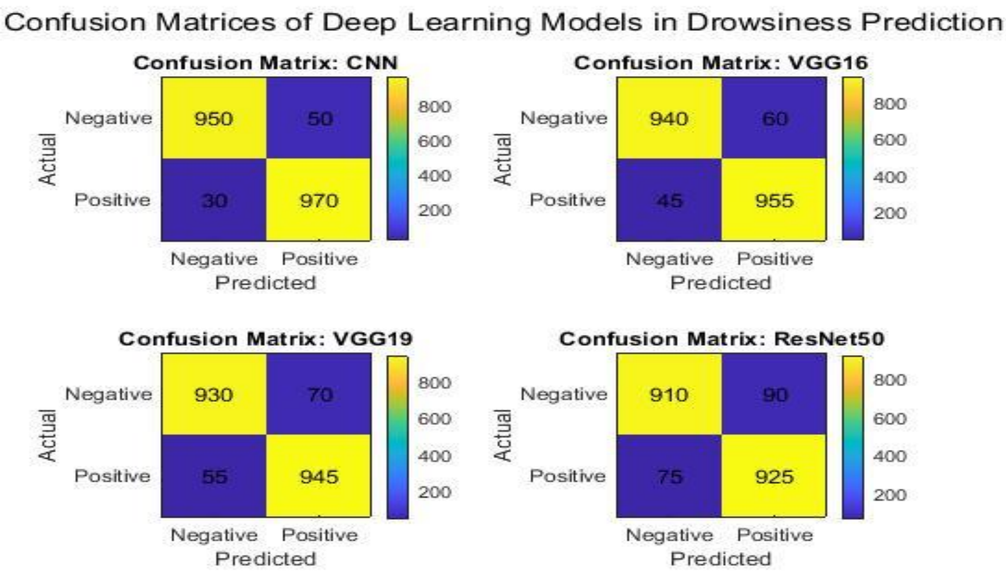

Figure 4: Confusion matrices of each model

CONCLUSION

In this research, we presented a comprehensive evaluation of several deep learning models to predict drowsiness using facial expressions. The study evaluated precision, recall, F1 score, AUC ROC curve values, and confusion matrices to measure the efficiency of all the models to detect drowsiness with precision. The highest performing CNN model (the recall, precision, f1, and AUC ROC curve) was achieved by the CNN model. The well-balanced distribution of true and false positives along with true and false negatives on the confusion matrix indicate the resilience of its drowsiness prediction. But more or less VGG16 and VGG19 really performed good with high precision, recall and F1 score and the accuracy exhibited by VGG16 and VGG19 are slightly less compared to the earlier CNN model. The confusion matrices displayed increased false positive and false negative occurrences, suggesting a slightly weaker performance in drowsiness prediction.

ResNet50 also performed well, however, it showed slightly less accuracy than the other models. This resulted in a slightly lower predictive accuracy with a greater number of false positives and false negatives, as indicated by the confusion matrix. The CNN model, which performs superbly across all metrics, is the chosen model for drowsiness prediction. With a high precision and recall, along with a balanced number of true positives and negatives, the output shows convincing results in recognizing drowsiness based on facial expressions accurately. But, it is important to recognize the overall contribution of different models in making the concept of predicting drowsiness very promising with substantial evidence of predictive power with good interpretability.

References

1) A. Sethi, S. Deb, P. Ranjan, and A. Sardar, "Smart mobility solution with multiple input Output interface," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 3781–3784, 2017, doi: 10.1109/EMBC.2017.8037680.

2) M. Maier, F. Blume, P. Bideau, O. Hellwich, and R. Abdel Rahman, "Knowledge-augmented face perception: Prospects for the Bayesian brain-framework to align AI and human vision," *Consciousness and Cognition*, vol. 101, no. January, p. 103301, 2022, doi: 10.1016/j.concog.2022.103301.

3) S. Malathi and D. E. Dev, "for Driver Eligibility and Car Security," pp. 112–117, 2019.

4) X. Krasniqi and E. Hajrizi, "Use of IoT Technology to Drive the Automotive Industry from Connected to Full Autonomous Vehicles," *IFAC-PapersOnLine*, vol. 49, no. 29, pp. 269–274, 2016, doi: 10.1016/j.ifacol.2016.11.078.

5) Z. Chen, X. Feng, and S. Zhang, "Emotion detection and face recognition of drivers in autonomous vehicles in IoT platform," *Image and Vision Computing*, vol. 128, p. 104569, 2022, doi: 10.1016/j.imavis.2022.104569.

6) V. Gowri, P. Sethuramalingam, and M. Uma, "Performance analysis of autonomous UV disinfecting robot (UV bot) using Taguchi method," *Materials Today: Proceedings*, vol. 68, pp. 1980–1987, 2022, doi: 10.1016/j.matpr.2022.08.227.

7) A. K. Menon, M. Jia, S. Kaur, C. Dames, and R. S. Prasher, "Distributed desalination using solar energy: A technoeconomic framework to decarbonize nontraditional water treatment," *iScience*, vol. 26, no. 2, p. 105966, 2023, doi: 10.1016/j.isci.2023.105966.

8) H. Huang, J. Chang, D. Zhang, J. Zhang, H. Wu, and G. Li, "Machine learning-based automatic control of tunneling posture of shield machine," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 14, no. 4, pp. 1153–1164, 2022, doi: 10.1016/j.jrmge.2022.06.001.

9) Y. Ai, M. Peng, and K. Zhang, "Edge computing technologies for Internet of Things: a primer," *Digital Communications and Networks*, vol. 4, no. 2, pp. 77–86, 2018, doi: 10.1016/j.dcan.2017.07.001.

10) H. Pourrahmani *et al.*, "The applications of Internet of Things in the automotive industry: A review of the batteries, fuel cells, and engines," *Internet of Things (Netherlands)*, vol. 19, no. July, p. 100579, 2022, doi: 10.1016/j.iot.2022.100579.

11) P. Gupta *et al.*, "Industrial internet of things in intelligent manufacturing: a review, approaches, opportunities, open challenges, and future directions," *International Journal on Interactive Design and Manufacturing*, 2022, doi: 10.1007/s12008-022-01075-w.

12) M. Haghi, R. Stoll, and K. Thurow, "A Low-Cost, Standalone, and Multi-Tasking Watch for Personalized Environmental Monitoring," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 5, pp. 1144–1154, 2018, doi: 10.1109/TBCAS.2018.2840347.

13) "Abstracts from the 2020 Annual Meeting of the Society of General Internal Medicine," *Journal of general internal medicine*, vol. 35, no. Suppl 1, pp. 1–779, 2020, doi: 10.1007/s11606-020-05890-3.

14) C. Shuai, F. Yang, W. Wang, J. Shan, Z. Chen, and X. Ouyang, "Promoting charging safety of electric bicycles via machine learning," *iScience*, vol. 26, no. 1, p. 105786, 2023, doi: 10.1016/j.isci.2022.105786.

15) Y. Methkal *et al.*, "Measurement : Sensors Leaf disease identification and classification using optimized deep learning," *Measurement: Sensors*, vol. 25, no. September 2022, p. 100643, 2023, doi: 10.1016/j.measen.2022.100643.

16) T. Saba, A. Sameh Mohamed, M. El-Affendi, J. Amin, and M. Sharif, "Brain tumor detection using fusion of hand crafted and deep learning features," *Cognitive Systems Research*, vol. 59, pp. 221–230, 2020, doi: 10.1016/j.cogsys.2019.09.007.