International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Advanced Driver Drowsiness Detection: Integrating CNN and ANN Technologies for Proactive Road Safety

Dr. Aruna Varanasi [1] & Varun Puram [2]

Submitted: 28/01/2024 **Revised:** 06/03/2024 **Accepted:** 14/03/2024

Abstract – This paper delves into the critical realm of automated driver drowsiness identification, presenting a pivotal stride in advancing road safety through preemptive driver alerts. Employing an auto camera system, real-time images of the driver are captured, and a neural network, encompassing both Artificial Neural Network (ANN) and Convolutional Neural Network (CNN) technologies, analyzes each frame independently. The temporal dimension is introduced by averaging characteristics from the last 20 frames, aligning with approximately one second in both training and testing datasets. The research critically examines image segmentation methods, anchoring a robust model in CNN technology. With a meticulously curated dataset of over 2000 annotated image slices, the study pioneers an innovative approach to pre-emptive drowsy driving interventions, seamlessly integrating ANN and CNN analyses, thereby promising tangible contributions to road safety efforts.

Keywords *–Convolutional Neural network, Artificial Neural Network,* user behavior

1. INTRODUCTION

The last 10 years have seen an unparalleled advancement in digital information, communication, and cognitive technologies that are having a profoundly transformative impact on the quality of life on Earth. Research, development, and the introduction of novel intelligent techniques for identifying crucial components in modern automobiles, as well as new system modifications and ways of accomplishing things are now necessary due to global Innovation trends in sophisticated embedded digitalization and computerization systems in intelligent automobiles.Because soft computing is a key component of the building of contemporary smart cars, it is used to a very high degree in modern automotive architecture. Large-scale changes in the automobile industry are realized and named through smart car architecture and intelligent platforms. These changes include ICT integration, automation, and digitization at every stage of the vehicle control and diagnostic process.

1Professor and Head of Department
Computer Science and engineering
Sreenidhi Institute of Science and Technology
Hyderabad, India
arunavaranasi@sreenidhi.edu.in
2Student of Mtech, Sreenidhi Institute of Science and Technology
Computer Science Department
Hyderabad, India
varunraopuram@gmail.com

The proposed investigation focuses on the creation of efficient soft computing strategies that monitor, identify, and model stressful situations and emotional states experienced by a driver of a smart automobile through the use of convolution neural networks and deep methods of learning.For both safety and comfort when driving, rapid identification and monitoring of the chosen automobile characteristics in addition to the driver's emotional and attentive state are essential. In many contexts, face recognition—a common biometric identity technology—is acknowledged as being crucial to establishing safe management.Over the past few decades, face recognition has gained significant interest from researchers and engineers because of its numerous applications in various industries, like Access control systems, smart cards, law enforcement, identity authentication, and information security are just a few examples.The two primary phases in the face recognition procedure are feature extraction and classifier design. These two procedures significantly influence the dependability and efficacy of different recognition techniques. This introduction aims to give a foundation and context for the major developments in digital information, communication, and intelligent technologies that have taken place during the past decadeand illustrates how these advancements have revolutionized lifestyles and quality of life all around the world. In particular, the implications for smart car technology are discussed, emphasizing the necessity of doing research, developing, and implementing novel system solutions as well as clever techniques for identifying crucial components of contemporary cars.

The present introduction delineates the dominant worldwide patterns in intricately embedded computerization and digitization procedures in intelligent automobiles, culminating in the creation of novel system designs and sophisticated platforms. It emphasizes how crucial soft computing approaches—in particular, convolutional neural networks and deep learning techniques—are for tackling problems with tracking, identifying, and simulating stressful scenarios and emotional states in pilots of smart cars.

The introduction also emphasizes how important it is to quickly identify and monitor certain car parameters as well as the driver's emotional and attentive state to guarantee both Safety as well as comfort when operating a vehicle. It argues that one essential biometric authentication technique is face recognition.with numerous uses in law enforcement, access control systems, identity authentication, and information security, among other domains.To put it briefly, the introduction prepares the reader for the main objective of the proposed work, which is to develop efficient soft computing techniques to improve smart car technology, particularly for resolving driver-related issues. It highlights how critical it is to adjust to the continuous transformations in the automotive sector, such as automation, digitization, and the integration of information and communication technology (ICT) at every stage of vehicle diagnostics and control.

2. LITERATURESURVEY

This work develops an innovative video-based technique for identifying sleepy drivers[1] independent of lighting conditions. The device is capable of accurately identifying sleepy situations even when a driver wears glasses. The suggestedsystem is divided into two sequential computational processes: the detection of sleepy drivers and the detection of driver eyes.which are both detected by a near-infrared-ray (NIR) camera. The accuracy of the drowsy status detection is up to 91%, and the average open/closed eyes detection rates with/without glasses are 94% and 78%, respectively. Following software optimizations, the 640x480 video format can be processed at up to 16 frames per second (fps) on an FPGA-based integrated platform.

A non-intrusive vision-based system for the identification of driver weariness is described in this study [2]. The device employs a color video camera to track the driver's eyes and identify micro-sleeps, or brief intervals of sleep, by directly rewarding the driver's face.

Working with skin-color data, the algorithm finds the face in the input space[3]. The precise location of the face is ascertained using blob processing, which comes after the pixels with skin-like color have been segmented. By examining the face's horizontal gradient map and accounting for the fact that eye regions exhibit a significant shift in the horizontal intensity gradient, we can narrow down the search space.

We present a non-intrusive vision-based approach for driver fatigue detection in this study [4]. Taking advantage of a color video camera, the system watches the driver's eyes and rewards their face to identify micro-sleeps or brief intervals of sleep. To identify the face in the input space, the system processes skin tone data. Blob processing is used to pinpoint the precise location of the face after skin-like color segmentation of the pixels. Detecting drowsiness has numerous benefits, one of which is a decrease in the frequency of traffic accidents. One of the new and reliable strategies in Sleepy Face is the use of image processing techniques. Using a virtual reality driving simulator, the current pilot study[5] was conducted to look at tiredness and offer photographs of drivers' expressions. One benefit of this study over others is its ability to recognize and use multiple criteria intelligently over an extended period. Raising the alert, helps in the early detection of tiredness and stops irreversible losses.

DATASET

Figure1

The data set consists of 7800 images out of which 5000 images are Eyes open and the remaining 3800 images are Eyes closed

Figure 2

The second data set consists of 4300 images out of which 2500 images are mouth opened and the remaining 1800 images are mouth closed.

Data preprocessing is an extremely crucial step for both ML and DL methods. It improves the accuracyofthedataset.Theoriginalimagesizeis50*50*3 ,weconvertallimagesinto100*100*3. After resizing the image, we got better results when compared with the original pixel values.

3. METHODOLOGY

Convolutional neural Networks (CNN) needa large number of data to train the model. Alternatively, if yourdatasetis small, choose a pre-trained model. The convolutional layer extracts specific features from the images. Convolution contains different filters to perform operations. Pooling layers reduce the height and width of feature maps. The fully connected layer takes input vectors from the previous layers and flattens them into a single vector that can be fed into the next layers. An example diagram of CNN is shown in Figure 6.

In this research work, we used pre-trained models for image classification because they were developed on large datasets for similar problems. The benefit of employing a pre-trained model is that these models have been built on huge data sets, and they have learned very accurate discriminative features, allowing us to apply them to transfer learning tasks.

Transfer learning is the process of applying information from one task to another to enhance learning in a new one. It improves productivity while cutting down on training time.

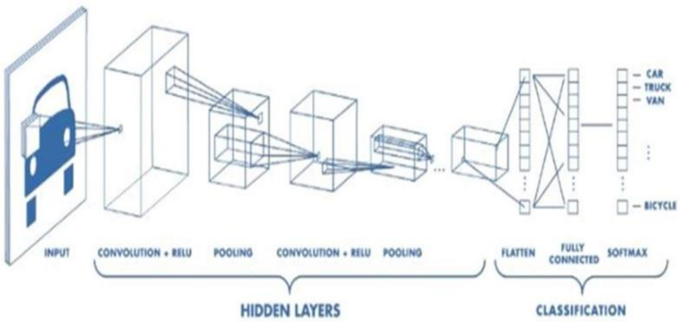
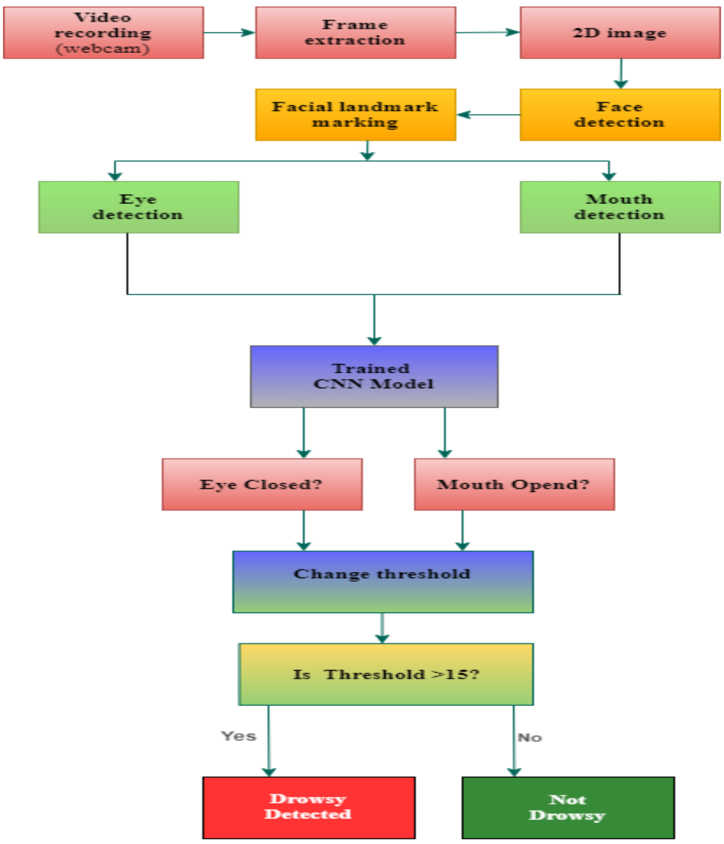


Figure 3Convolutional Neural Networks

The seven layers that make up the CNN training process are input, convolution, pooling, and activation functions. First, input is given in the form of feature vector size. Second, dimensions are determined for matrix convolution. Third, pooling and max pooling are carried out. Fourth, stack-to-line conversion is completed. Finally, in the remaining layers, activation functions such as the sigmoid function are employed. The output error is reduced by using the activation function.Each layer updates its weight and bias, and the new parameter is fed into the subsequent layer. CNN's implementation is significantly impacted by the 1*1 feature mapping, which enables scalar operation of weight and bias updates. CNN is a fully connected layer. A number of factors need to be taken into account, including the size of the pool and the number and dimensions of the filters.In order for the activation function to receive the maximum pooled output and the subsequent layers to minimize error, the filter dimensions must be specified to achieve maximum pooling.

SYSTEM ARCHITECTURE



Here, first, this system will run the webcam and read the frames continuously. Thereafter, the Haar cascade classifier will detect facial features, from these features, the eyes, and mouth landmarks will be detected based on their cascade classifiers. The detected eyes and mouth features will be forwarded to the trained CNN model then this model will retune the status of the eye opened or closed and mouth opened and closed as well. Based on this status the threshold value will be increased or decreased. Finally, when the threshold value exceeded greater than the score of 15 value then this system will be detected drowsiness and generate an alarm.

Face Detection In this module, after taking input

as an image from the camera then it can detect the face coordinates with the help of the Haar cascade classifier.

Facial Landmark:

The completion of face detection with the help of a shape predictor network model file can be loaded by the *dlib* library. This method can take the input image as a face detection image and invoke the predictor method for detecting facial landmarks like a mouth. As well as based on the left-eye and right-eye cascade classifiers, this system will also detect the left-eye and right-eye landmarks.

Drowsy Detection:

The trained models belong to eyes and mouth, they will perform the classification of eyes closed and opened from the image of the webcam and increase the threshold value when this system detects the eye closed status. As well as when this system detects the mouth opening status then also the threshold value will be increased. So, based on the threshold value, where if the threshold value is exceeded the defined threshold value of 15, then this system will detect the drowsiness and generate an alarm sound to alert the driver.

Artificial Neural Network:

It is a similar deep learning method that use brain neurons. In comparison to the other machine learning classifiers, it must improve the classifier. Three layers will be processed by the ANN classifier: an input layer, a middle layer that functions as a hidden layer, and an output layer that is the last layer. The input layers will receive these values once the attribute values from feature extraction have been gathered. In this case, the values from each input layer will be passed to each hidden layer for processing the network weight values. Afterwards, the hidden layers will compute all of the weight values and forward them to the output layer. By comparing the highest weight values with the weight values of every hidden layer, the output layer will ultimately determine the anticipated result value.

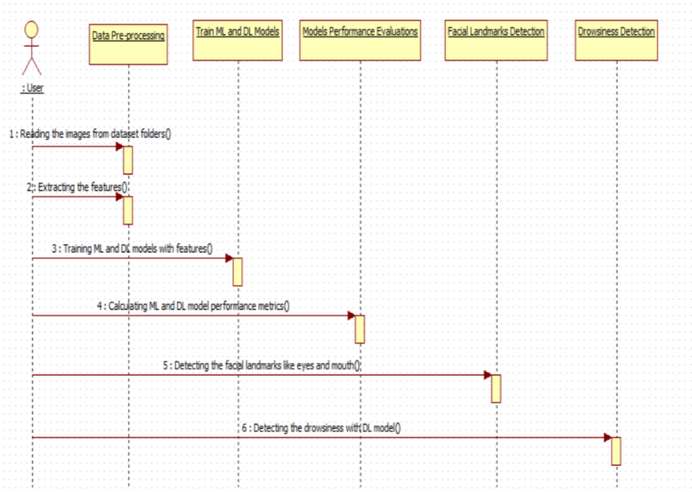


Figure 5:State Sequence diagram

In the initial phase, the user engages in data retrieval from the dataset, delving into the intricacies of the information. Subsequently, a meticulous application of data preprocessing techniques ensues, focusing on extracting pertinent features from images. This critical step involves refining and preparing the data to enhance its suitability for analysis. Moving forward, the extracted features become the foundation for training machine learning and deep learning models. This training process is instrumental in imparting intelligence to the models, enabling them to discern patterns and make informed predictions. Following the training phase, an assessment of model performance transpires, involving the calculation of performance metrics. This comprehensive workflow ensures a robust and well-informed approach to leveraging machine learning and deep learning for effective analysis and decision-making.

4. SYSTEM DESIGN

In this phase, the system design is prepared and the need specifications from the previous phase are studied. Determining the hardware and system requirements, in addition to the overall system architecture, is aided by system design. The software code that will be built in the next step is created in this way.

Execution:

The system is first developed in units, or tiny programs, using inputs from the system design. These units are then incorporated into the subsequent phase. Making and evaluating each unit to make sure it's functional is known as unit testing.

Integration and Testing:

After each unit is tested, all of the units created during the implementation phase are combined into a single system. To check for bugs or defects, the software design must undergo continuous software testing. Testing is carried out to ensure that the client encounters no issues when installing the software.

Deployment of System:

The product is either released into the market or deployed in the client environment once both functional and non-functional testing is completed.

Maintenance:

This step, which comes after installation, entails modifying the system or a particular component to alter features or increase performance. Either customer-initiated change requests or issues discovered during real-time system usage are the causes of these modifications. The client assists in providing ongoing maintenance for the developed program.Additionally, the system uses deep learning methods to enable adaptive learning from various driving scenarios and individual actions. This flexibility ensures that the system continues to be effective in handling the dynamic nature of real-world driving scenarios while also improving the accuracy of the system over time. Facial recognition is integrated for a couple of reasons: it offers safe biometric authentication and adds another level of customization to the driving experience.

5. OUTPUT AND RESULTS

```
Epoch 1/25
73/73 [==============================] - 47s 644ms/step - loss: 0.1989 - acc: 0.9367 - val_loss: 0.4419 - val_acc: 0.9122
Epoch 2/25
73/73 [==============================] - 39s 532ms/step - loss: 0.1316 - acc: 0.9538 - val_loss: 0.2397 - val_acc: 0.9382
Epoch 3/25
73/73 [==============================] - 41s 557ms/step - loss: 0.1191 - acc: 0.9570 - val_loss: 1.2767 - val_acc: 0.8832
Epoch 4/25
73/73 [==============================] - 39s 529ms/step - loss: 0.1093 - acc: 0.9605 - val_loss: 0.5535 - val_acc: 0.9276
Epoch 5/25
73/73 [==============================] - 39s 531ms/step - loss: 0.1226 - acc: 0.9575 - val_loss: 0.6102 - val_acc: 0.9175
Epoch 6/25
73/73 [==============================] - 39s 537ms/step - loss: 0.1196 - acc: 0.9577 - val_loss: 0.5770 - val_acc: 0.9063
Epoch 7/25
73/73 [==============================] - 39s 530ms/step - loss: 0.1023 - acc: 0.9632 - val_loss: 0.5210 - val_acc: 0.9248
Epoch 8/25
73/73 [==============================] - 40s 550ms/step - loss: 0.1026 - acc: 0.9633 - val_loss: 1.7190 - val_acc: 0.8799
Epoch 9/25
73/73 [==============================] - 40s 550ms/step - loss: 0.1238 - acc: 0.9570 - val_loss: 1.9751 - val_acc: 0.8761
Epoch 10/25
73/73 [==============================] - 41s 559ms/step - loss: 0.1035 - acc: 0.9632 - val_loss: 1.0540 - val_acc: 0.8950
```

figure 6: EPOCHS ruined

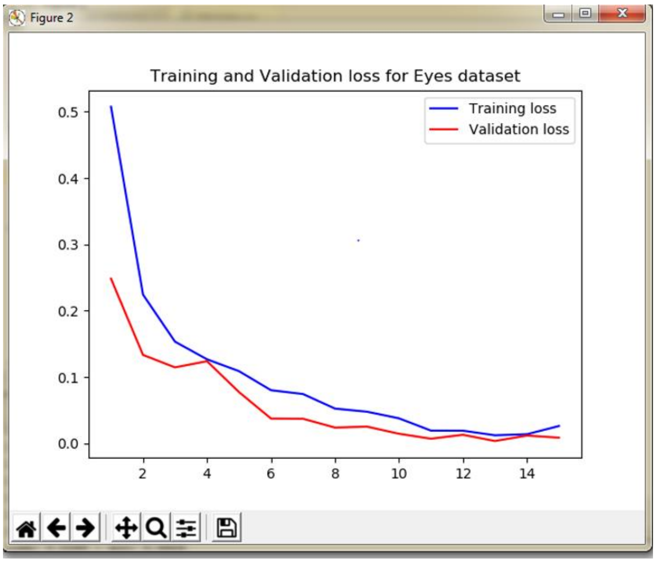


figure 7

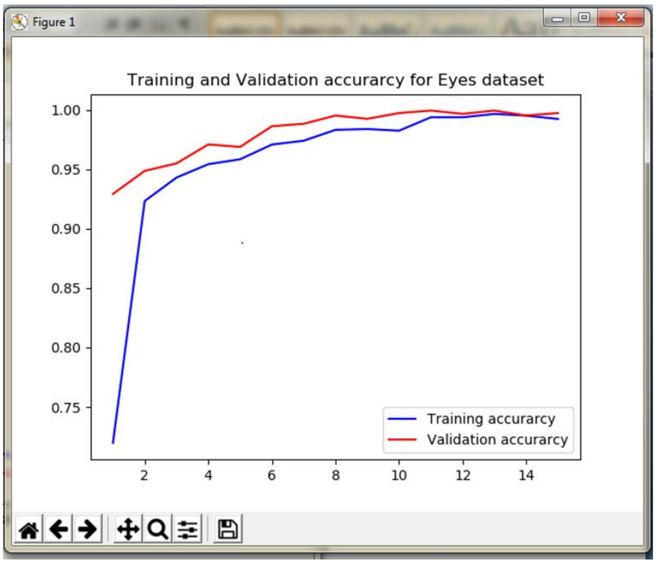


figure 8

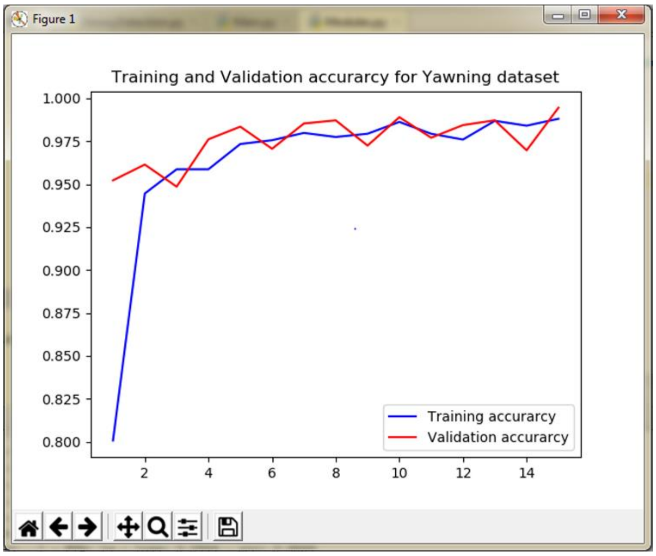


Figure 9

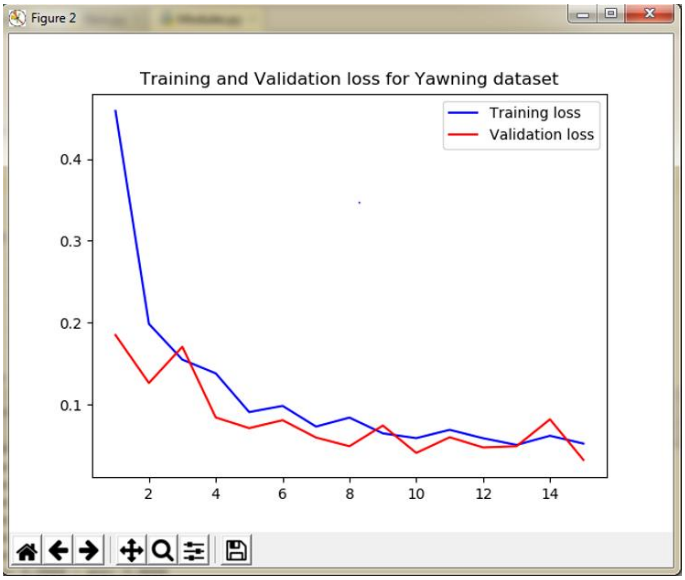


Figure10

6. CONCLUSION

In the realm of automotive safety, the Driver Drowsiness Detection project emerges as a pioneering initiative with the potential to reduce road accidents significantly. The project's foundation lies in the adoption of Convolutional Neural Networks (CNNs), a cutting-edge technology in image analysis. By employing an auto camera to capture real-time images of the driver, the neural network becomes adept at distinguishing between states of wakefulness and fatigue. The intricate process of image segmentation refines this distinction, ensuring a nuanced understanding of the driver's emotional status.

One unique aspect of the study is the use of a model that was trained on a large dataset with more than 2000 annotated image slices. The rigorous training procedure, which considers the average of the previous 20 frames and evaluates each frame separately, attests to the dedication to precision and dependability in sleepiness detection. This method, which requires an evaluation time of around a second, is consistent with the temporal requirements that are essential for real-time applications.

The implications of this project extend beyond the technical intricacies of neural networks and image segmentation. Integrating such a driver monitoring system into smart cars holds the promise of a proactive approach to road safety. The ability to swiftly detect and alert a drowsy driver not only mitigates the risk of accidents but also exemplifies the transformative potential of artificial intelligence in ensuring passenger well-being.

Looking forward, the project hints at a future where smart cars are equipped with advanced technologies that go beyond conventional safety features. As research and development in this domain progress, refinements and expansions of the drowsiness detection system are anticipated. This could lead to even more sophisticated applications, contributing to the broader landscape of intelligent transportation systems and redefining the standards of safety on our roads. In essence, the Driver Drowsiness Detection project not only exemplifies the power of neural networks in safeguarding lives but also underscores the transformative role of artificial intelligence in shaping the future of automotive technology.

7. REFERENCES

[1]. Kirti Dang, Shanu Sharma, “Review and comparison of face detection algorithms”, 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, 2017

[2]. Jongkil HyunCheol-Ho Choi, Byungin Moon, “Hardware Architecture of a Haar Classifier Based Face Detection System Using a Skip Scheme”, IEEE International Symposium on Circuits and Systems (ISCAS),2021

[3]. jain, “face detection in color images; r. L. Hsu, m. Abdel-mottaleb, and a. K. Jain.

[4]. Wang Yang;Zheng Jiachun, “Real-time face detection based on YOLO” 1st IEEE International Conference on Knowledge Innovation and Invention (ICKII), 2018

[5]. Kartika Candra Kirana,Slamet Wibawanto,Heru Wahyu Herwanto, “Redundancy Reduction in Face Detection of Viola-Jones using the Hill Climbing Algorithm”, 4th International Conference on Vocational Education and Training (ICOVET), 2020