Driver Distraction Detection using Convolutional Neural Networks

Thesis submitted to the SASTRA Deemed to be University
in partial fulfillment of the requirements
For the award of the degree of

B. Tech. Computer Science Engineering

BCSCCS801: MAIN PROJECT

Submitted by

Shaik Kanubai Mohammed Maheer (120003282)
Narla Sai Teja (120003203)
Pala Nitheesh Reddy (120003215)

JULY 2020

SASTRA
ENGINEERING · MANAGEMENT · LAW · SCIENCES · HUMANITIES · EDUCATION
DEEMED TO BE UNIVERSITY
(U/S 3 OF THE UGC ACT, 1956)
THINK MERIT | THINK TRANSPARENCY | THINK SASTRA

SCHOOL OF COMPUTING

THANJAVUR, TAMILNADU, INDIA – 613 401

SCHOOL OF

COMPUTING

THANJAVUR – 613 401

BONAFIDE CERTIFICATE

This is to certify that the thesis titled "**Driver Distraction Detection using Convolutional Neural Networks"** submitted in partial fulfillment of the requirements for the award of the degree of B. Tech. Computer Science to the SASTRA Deemed to be University, is a bona-fide record of the work done by **Shaik Kanubai Mohammed Maheer (120003282), Narla Sai Teja (120003203), Pala Nitheesh Reddy (120003215)** during the final semester of the academic year 2019-20, in the **School of Computing**, under my supervision. This thesis has not formed the basis for the award of any degree, diploma, associate-ship, fellowship or other similar title to any candidate of any University.

Signature of Project Supervisor :

Name with Affiliation :

Date :

Project *Viva voc*e held on

Examiner 1 **Examiner 2**

SCHOOL OF

COMPUTING

THANJAVUR – 613 401

Declaration

We declare that the project titled "**Driver Distraction Detection using Convolutional Neural Networks**" submitted by us is an original work done by us under the guidance of **Dr. Karthikeyan B, SAP, School of Computing, SASTRA Deemed to be University** during the final semester of the academic year 2019-20, in the **School of Computing**. The work is original and wherever we have used materials from other sources, we have given due credit and cited them in the text of the thesis. This project has not formed the basis for the award of any degree, diploma, associate-ship, fellowship or other similar title to any candidate of any University.

Signature of the candidate(s) :
Name of the candidate(s) :

Date :

ACKNOWLEDGMENTS

We would like to acknowledge everyone who had a role to play in making our humble efforts a success. First and foremost, we take pride in thanking the **Almighty** who gave us strength for the successful completion of this project. We would forever remain grateful to honorable Vice-Chancellor **Dr. S VAIDHYASUBRAMANIAM** and Registrar **Dr. R CHANDRAMOULI,** for their encouragement in our academic life at SASTRA Deemed to be university.

We are extremely thankful to our dignified Dean, School of Computing **Dr. UMAMAKESWARI A**, and Associate Dean of Computer Science and Engineering **Dr. SHANKAR SRIRAM V S,** for inspiring and motivating us to bring out a perfect and successful project work. We are grateful to our Guide, **Dr. Karthikeyan B ,** School of Computing**,** who was the driving force behind this whole project from the beginning. Her deep insight into the field and invaluable suggestions helped us in making progress through our project work.

We also thank all the **Teaching and Non-teaching staff**, and those who have directly or indirectly helped us by extending their moral support.

Table of Contents

List of Figures

List of Tables

ABBREVIATIONS

CNN	Convolutional Neural Networks
NN	Neural Networks
DNN	Deep Neural Networks
ANN	Artificial Neural Networks
ReLu	Rectified Linear Unit
NHTSA	National Highway Traffic Safety Administration
VGG	Visual Geometry group
LR	Learning Rate

Abstract

Distracted Driving is the major cause for several road accidents. According to the survey of the NHTSA, one among five motor vehicle crashes is due to driver distraction. To overcome this, our aim is to develop an accurate and robust system for detecting a distracted driver. In previous works, Transfer learning models like AlexNet, GoogleNet and ResNet have been used. To train the model, State Farm Driver Distraction Dataset has been used in which the actions of drivers are classified into ten categories. In the proposed work, a model from scratch is developed and trained on our dataset. Then the same AlexNet, GoogleNet, ResNet models have been used by doing Fine-Tuning and adding some extra layers to these models. Next other state-of-the-art models like Xception, VGG16, MOBILENET and ResNet50 are trained with our dataset by adding extra layers to these models and then Fine-Tuned the model. Finally, ensembled all the models to attain better accuracy and tested it on our test data. Our experimental results indicate the attainment of better accuracy.

Mohammed Maheer(120003282)

Specific Contribution

- Implemented pre-trained transfer learning models by adding some extra layers (Convolution, Pooling, Dense, Dropouts and Normalisation layers).
- Performed fine-tuning to the transfer learning models.
- Ensembled all the transfer learning models.

Specific Learning

- Detailed learning of transfer learning model.
- Understanding the action performed by each layer in Convolutional Neural Networks.
- Image Augmentation Techniques.
- Selecting the best Ensemble to bring better accuracy

Abstract

There is a continuous hike in the number of road accidents and fatalities due to distracted driving since past few years. According to the National Safety Council in a year around 1.6 million road accidents are occurring due to the usage of mobile phones by the driver while driving. To reduce this, our aim is to develop a better and an accurate system than the existing ones for detecting a distracted driver. In [1](Yang Xing et al. 2019), Pre-trained CNN models like AlexNet, GoogleNet and ResNet were used. In our proposed methodology initially a model from scratch is developed and trained on the dataset. Next the pre-trained models AlexNet, GoogleNet, ResNet are used. These are fine tuned and few extra layers like Convolution, Pooling, Dense, Dropouts and Normalization layers are added. In addition to the above models few other pretrained models like Xception, VGG16, MOBILENET and ResNet50 are also used to train on our data .In the end all these models are ensembled and evaluated on our test data. For training this proposed model, State Farm Driver Distraction Dataset has been used. In this data set the classification of the driver actions is done into ten categories. Our experimental results clearly depict a significant increase in accuracy of the ensembled model. Further there is scope to increase the accuracy by ensembling new pre-trained models.

Narla Sai Teja (120003203)

Specific Contribution:

- Implemented ResNet, VGG16 and ResNet50 models.
- Added extra layers(Convolution, Pooling, Dense, Dropouts and Normalisation layers) to above models and performed Fine-Tuning.

Specific Learning:

- Understanding the architecture of pre-trained CNN models.
- Detailed learning of transfer learning model.
- Understanding the ensembling of methods
- Understanding the action performed by each layer in Convolutional Neural Networks.

Abstract

Distracted Driving is the major cause for several road accidents. According to the survey of the National Highway Traffic Safety Administration (NHTSA), one among five motor vehicle crashes is due to driver distraction. To overcome this, our aim is to develop an accurate and robust system for detecting a distracted driver. In previous works, Transfer learning models like AlexNet, GoogleNet and ResNet have been used. To train the model, State Farm Driver Distraction Dataset has been used in which the actions of drivers are classified into nine categories.

In the proposed work, a model from scratch is developed and trained on our dataset. Then the same AlexNet, GoogleNet, ResNet models have been used by doing Fine-Tuning and adding some extra layers to these models. Next other state-of-the-art models like Xception, VGG16, MOBILENET and ResNet50 are trained with our dataset by adding extra layers to these models and then Fine-Tuned the model. Finally, ensembled all the models to attain better accuracy and tested it on our test data. Our experimental results indicate the attainment of better accuracy.

Pala Niteesh Reddy (120003215)

Specific Contribution:

- Implemented state-of-the-art models by adding some extra layers (Convolution, Pooling, Dense, Dropouts and Normalisation layers).
- Ensembled all the transfer learning models.

Specific Learning:

- Detailed learning of transfer learning model.
- Understanding the action performed by each layer in Convolutional Neural Networks.
- Understanding the ensembling of methods.
- Selecting the best Ensemble to bring better accuracy.

CHAPTER 1

INTRODUCTION

1.1 PROBLEM STATEMENT:

- Our intention is to achieve a decline in the number of vehicle crashes and aspire for an accident free driving environment.
- Many surveys have always been listing out that distracted driving is the prominent cause for the majority of automobile crashes.
- So our proposal of detection of driver distractions using CNN models can help in augmenting the government's endeavors to avert hazards caused by distracted driving.

1.2 MOTIVATION:

Distracted driving has always been so fatal because it doesn’t just affect the driver, but also the commuters and the uninvolved passersby as well. Activities like usage of mobile phones has been the top most factor for deaths due to crashes caused by distracted drivers and also is the huge concern of systems which are against this distracted driving.

1.3 SCOPE:

Our main focus is on how to implement transfer learning models and then classify whether the driver is distracted or not. Each and every individual image present inside the dataset contains only a single person. This study does not involve any of the hardware implementations. It is limited to just software implementation. Next, the outlook of this project is also bound to a pre-recorded video. This research does not cover the bounding box training as well as the segmentation.

1.4 CONTRIBUTIONS:

The bulk part of this study mainly depicts how to implement transfer learning models which are competent in working with large amount of labeled data and in extraction of feature representations from them. Then, this study depicts on how DNN for classification of objects into seperate classes uses the binary representation.To this an Additional attempt is made to fine-tune the transfer learning models and to optimize the hyperparameters of the DNN to enhance the performance of the model.

CHAPTER 2

OBJECTIVES

2.1 OBJECTIVES:

The main objective of this research is to identify whether the driver is distracted or not. One efficient way to minimize car accidents is averting distracted driving, which is the activity of driving while engaging in other actions such as messaging, speaking on the phone, etc. Our aim is to recognize unsafe behaviours and intimate the driver using short sound alerts, vibration in the seat or in the steering wheel, etc. Our aim is to develop a model using Convolutional Neural Networks with the help of various pre-trained models. Finally, the model is an ensemble of all the transfer learning models used. Our proposed model can classify the behaviour of a driver through a pre-recorded video.

2.2 RELATED WORKS:

This work is carried out with reference to the research carried out in [1](Yang Xing et al. 2019). This presents a Deep CNN model for driver pose recognition. Used a naturalistic data set and partitioned the data into seven classes with three being distracted and four classes as normal driving. In this Pre-trained models namely GoogLeNet, ResNet50 and AlexNet are adopted and implemented. [2] Presents a Single CNN model for driver distraction because of drowsiness and fatigue of a driver. Neural network models ResNet50, MobileNet and Inception are implemented(Whui Kim et al.2017). [3] Presents the ResNet50 model to identify whether the drivers behaviour looks to be distracted or not by interpreting the images. Used Tensorflow for the affirmations of the network and distraction recognition is shown to have an ideal accuracy(JianCheng et al. 2019). In [4] driver's drowsiness is detected by a system which is developed by using computer vision and deep learning. Deep CNN model helps in achievement of Image recognition(Kusuma et al. 2019).

CHAPTER 3

EXPERIMENTAL WORK / METHODOLOGY

3.1: ANALYSIS:

3.1.1 Data Exploration:

The dataset used in this project was provided by State Farm which contains snapshots from a video captured by a camera mounted in a car. This was the first dataset to consider a wide variety of distractions and was publicly available. The dataset for training has been divided into 10 classes, including 9 distracted and 1 safe driving class. These activities are classified as:

Table 3.1: Data Set Classes

c1	safe driving
c2	texting—right
c3	talking on the phone—right
c4	texting—left
c5	talking on the phone—left
c6	operating the radio
c7	drinking
c8	reaching behind
c9	hair and makeup
C10	talking to a passenger

The dataset had 22400 training images and 79727 testing images. Resolution was 640 x 480 pixels. The total size of the dataset is around 4GB.

Fig. 3.1 State Farm Dataset

3.1.2. Exploratory Analysis:

The class distribution of the dataset is uniform and can be seen in figure 3.1. . Due to the uniform distribution of the classes only one model is required for training the dataset. And if the dataset is not uniformly distributed then a separate binary classification model should be used for each class individually. Since the dataset is uniformly distributed performing validation split would be a simple task and yet yields a similar set of characteristics.

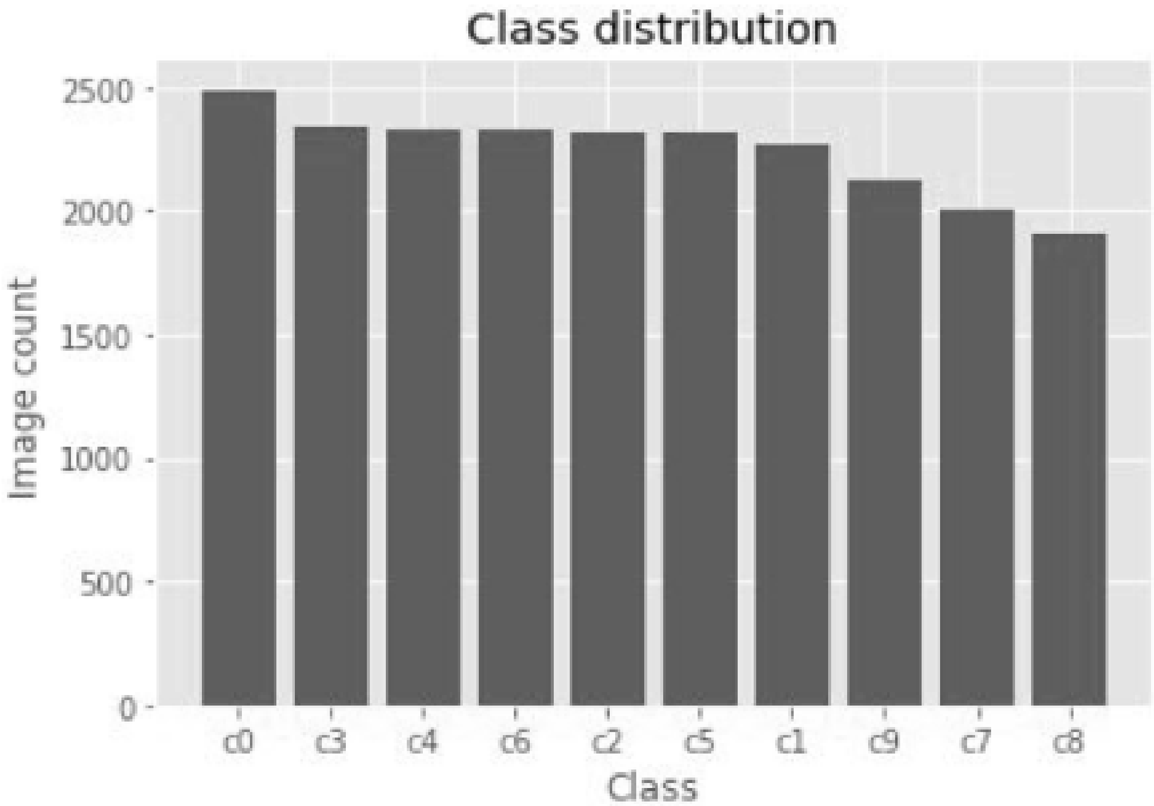


Fig. 3.2 Class Distribution

3.2 ALGORITHMS AND TECHNIQUES:

This project is implemented using Convolutional Neural Networks or CNNs which are made up of neurons which contain learnable weights and bias. The dataset was trained on a CNN model from scratch, VGG-16[5], AlexNet[6], ResNet50[7], Xception[8], MobileNet[9], and fine-tuned[12] all these pre-trained ImageNet models. We ensembled[11] all these transfer learning[10] models to attain better accuracy.

3.2.1 CNN Model:

Convolutional Neural Networks(ConvNet, or CNN) is a part of deep neural networks, mostly applied for image recognition, image clustering and classification, object detection etc. CNN is a methodology to study the images and it's pixel data. CNN's are made up of learnable weights and biases. When compared to other algorithms CNN's use less preprocessing. CNN consists of different layers. Input layer, output layer and some multiple hidden layers. The connectivity pattern of a CNN model is described in the figure 3.3.

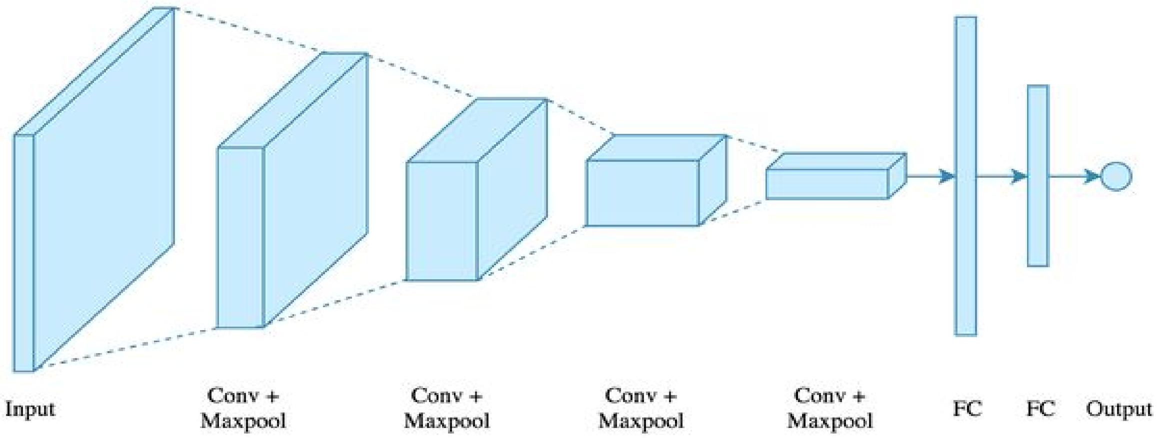


Fig. 3.3 Architecture of CNN

3.2.2 VGG-16:

VGG-16 is a CNN network which is 16 layers deep used in the ILSVRC-2014 competition. Thirteen convolution layers and three fully connected layers Input to the network is the image with the dimensions of 224 x 224 x 3. This network learnt rich feature representation for a wider range of images. This network is very slow to train and requires a GPU. It consists of 138M parameters and takes up about 528MB and hence it takes a lot of disk space and bandwidth which makes it inefficient. The architecture of the VGG-16 is described in the figure 3.4.

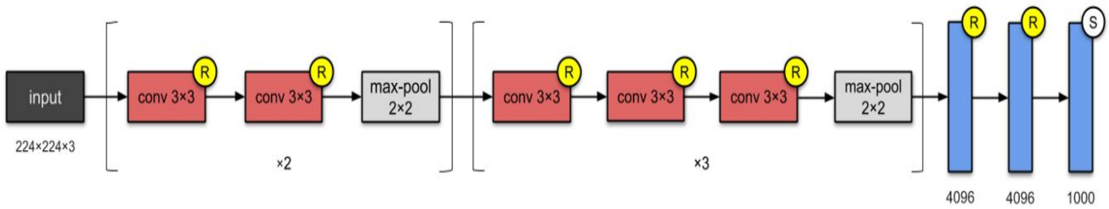


Fig. 3.4 VGG-16 Architecture

3.2.3 AlexNet:

AlexNet is a CNN network which is 8 layers deep. Five convolutional layers and three fully connected layers. It just stacked a few layers upon the LeNet-5. The Rectified Linear Unit (ReLu) activation function was first implemented in this network. It contains 60M parameters. The architecture of the AlexNet network is described in the figure 3.5.

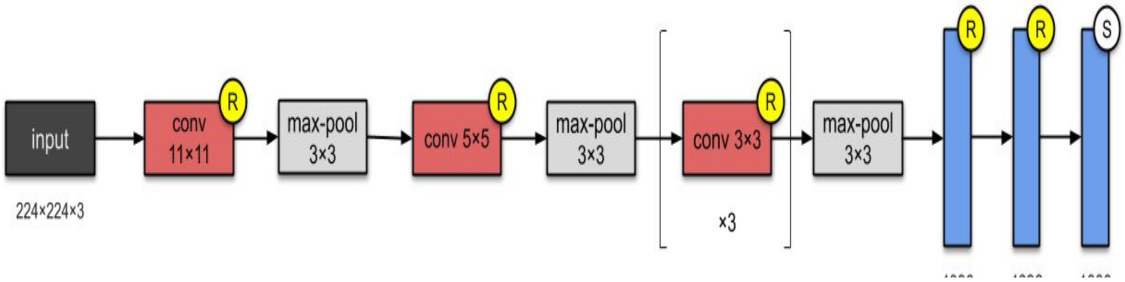


Fig. 3.5 AlexNet Architecture

3.2.4 ResNet50:

ResNet50 is a CNN network which is 50 layers deep. It is one of the early adopters of batch normalization. The basic building blocks for the resnets are the Conv and the identity blocks. It has 26M parameters. As we know that accuracy gets saturated as we go deeper. So, to overcome this problem Microsoft team implemented resnet using skip connections. The architecture of the ResNet50 network is described in the figure 3.6.

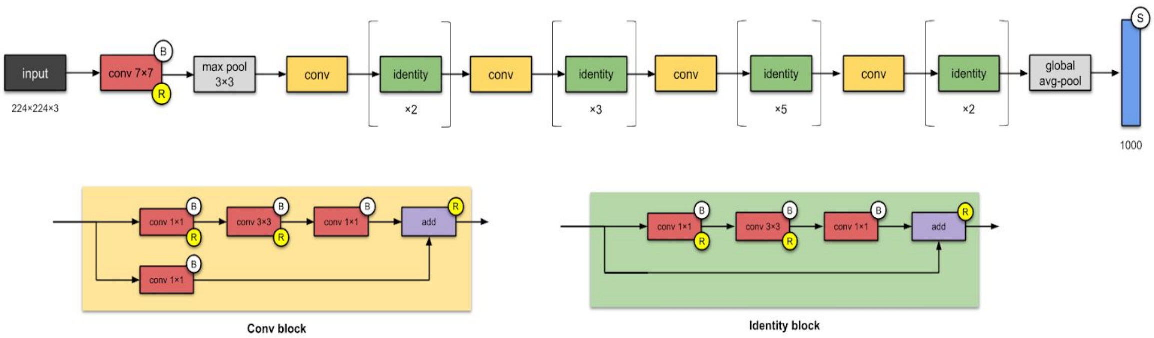


Fig. 3.6: ResNet50 Architecture

3.2.5 Xception:

Xception is a CNN network which is adapted from the Inception network. In this Xception Network depthwise separable convolutions replaced the the Inception modules . It has 23M parameters. The Xception Network introduced CNN entirely on depth wise separable convolutions. The architecture of the Xception Network is described in figure 3.7.

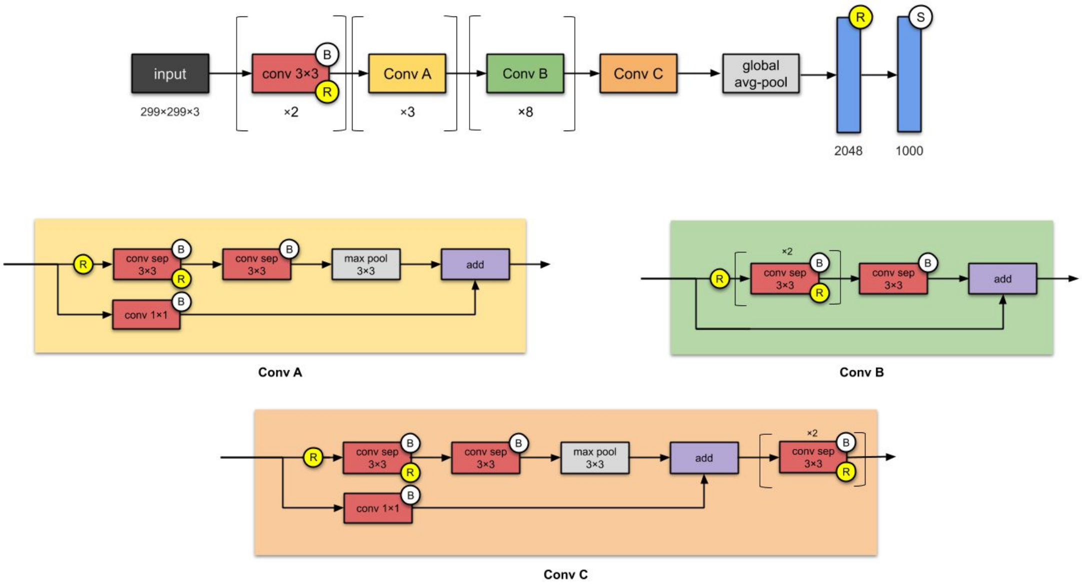


Fig. 3.7 Xception Architecture

3.2.6 MobileNet:

MobileNet is a CNN network which is lightweight in its architecture and is 30 layers deep. It uses depthwise separable convolution functions to build a light-weighted DNN that means a single convolution function on each colour channel(RGB) instead of combining all the three color channels and flattening them. The building block of MobileNet is described in figure 3.8.

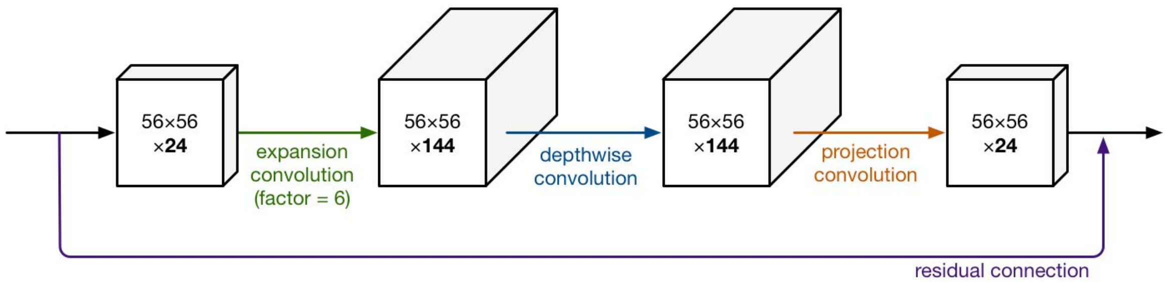


Fig. 3.8 Building block of MobileNet

3.3 METHODOLOGY:

3.3.1 Data Preprocessing:

The training set contains 22424 images of the ten categories. While preprocessing the data, the virtual instance ran out of memory due to the huge dataset. And the test set contains around 80000 images. This problem has been resolved when we splitted the training set into validation set and test set. In this project, keras with tensorflow as the backend is used so 4D tensors can be used as the input data. So, the images are then resized into 224x224x3 (height, width, channels).respectively.

Then this 3D tensor is converted into a 4D tensor (n,224,224,3) where n stands for number of images in the data. Then these 4D tensors have been rescaled.

3.3.2 Implementation:
Initially, a model was implemented from scratch which consists of Conv2D, BatchNormalization, MaxPooling2D, Dropout, Flatten and Dense layers. After implementing the model from scratch results were not satisfactory. Then the data was trained on the transfer learning models like VGG16, AlexNet, ResNet50, Xception, Mobilenet. Before training the models, Image augmentation is done by using keras inbuilt ImageDataGenerator function which generates lots of images without using any extra memory space. All these pre-trained models have been fine-tuned by unfreezing some top layers of these models. Finally, we ensemble all the modes to attain better accuracy.

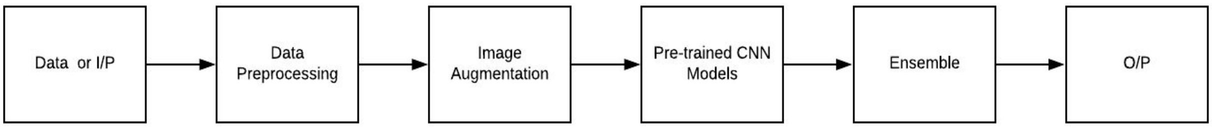


Fig. 3.9 Flow process of the model

3.4 HYPERPARAMETERS:

Hyperparameters are the parameters whose values are used to control the learning process. They can be thought of as tuning knobs for the model. A model even with a trillion parameters can't learn the features if the Hyperparameters are off.

3.4.1 Learning rate (LR):
LR controls how fast a neural net can learn. Our goal is to minimize the log loss function. So, if the LR is too high, the loss won't converge and it'll be fluctuating. And if the LR is too small, our model takes more time to converge and to obtain global minima. And it is illustrated in the figure 3.10.

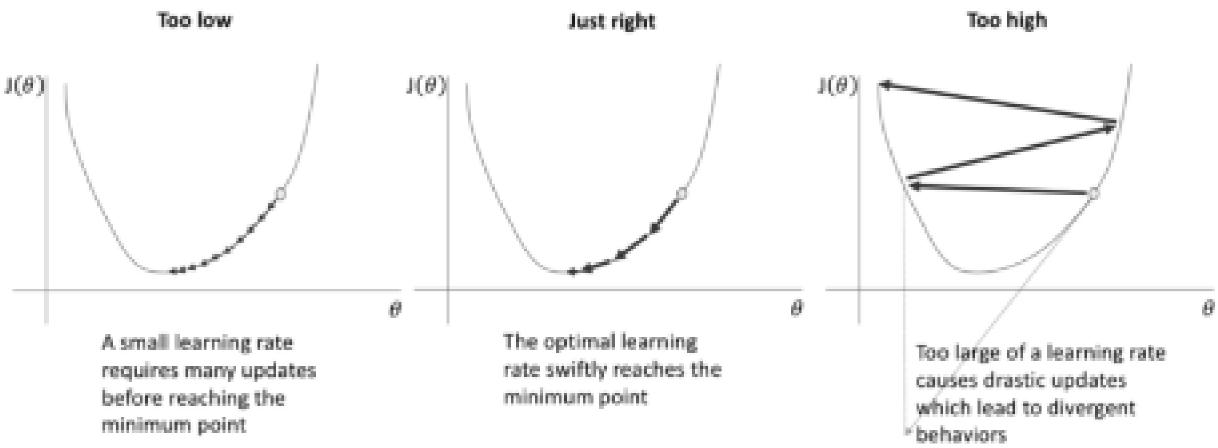


Fig. 3.10 Gradient Descent

3.4.2 Momentum:

Momentum is used to speed up the process. It accelerates the learning process and it can be thought of as the mass of the ball that's rolling down the surface of the loss function.

3.4.3 Dropout:

Dropout means deactivating a certain set of neurons chosen at random during the training phase. is used to prevent overfitting. It forces a NN to learn more robust features.

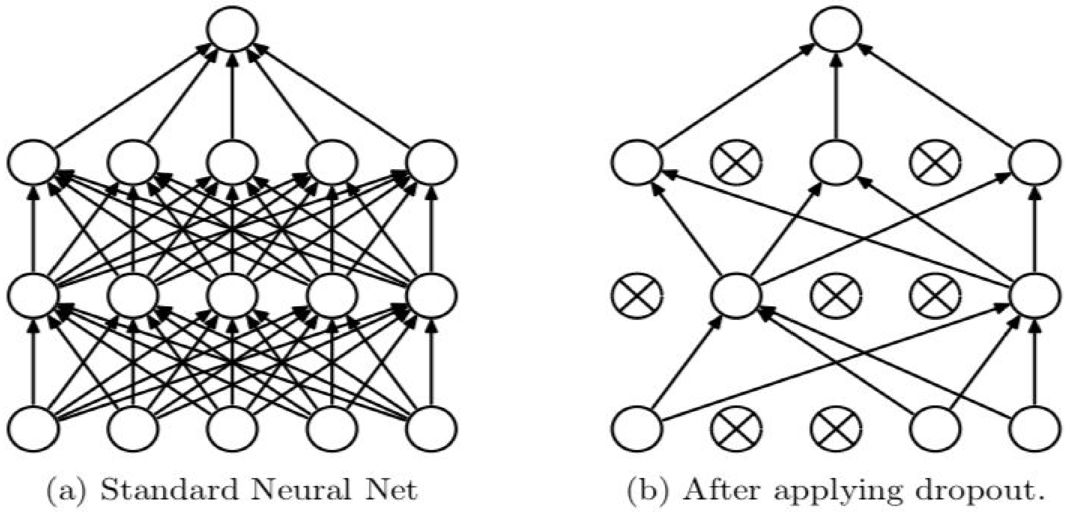


Fig. 3.11 Dropouts

In addition to these hyperparameters, Architecture of the NN itself is a hyperparameter. Number of layers, Neurons per layer, activation functions etc., all these are also considered as the hyperparameters.

3.5. PERFORMANCE EVALUATION METRICS:

Evaluation is the final stage of the process. This is where our proposed semantic keyword extraction model is being compared with other existing models.

3.5.1. Confusion Matrix:

It is a matrix that describes the model performance on a given data for which the true values are known.

Table 3.2: Confusion Matrix

	Positive	**Negative**
Positive	True Positive	False Negative
Negative	False Positive	True Negative

3.5.2. Accuracy:

It determines how good our model is performing..

Acc = (True Positive + True Negative) / (Positive + Negative)

3.5.3 Log Loss:

The log loss is a convex function and grows linearly for the negative values which make it less sensitive to the outliers.

$$f^*_{\mathrm{Logistic}} = \log\left(\frac{\eta}{1-\eta}\right) = \log\left(\frac{p(1 \mid x)}{1-p(1 \mid x)}\right).$$

CHAPTER 4

RESULTS & DISCUSSIONS

4.1 IMPLEMENTATION:

We have used python programming language for implementation of the proposed approach. We built an Image classifier using CNN, Keras with backend Tensorflow.

4.1.1 Software Requirements:

- Language: Python 3.7 & above
- Google Colab

4.1.2 Libraries Used :

- Keras, Tensorflow, pandas, numpy, matplotlib, opencv, sklearn, seaborn, dask.

4.1.3 Hardware Requirements:

- Processor: i7
- RAM: 16GB
- GPU: Tesla P100-PCIE-16GB
- Space on Hard Disk: 1TB

.

4.2 RESULTS:

This section includes our explanatory results of the proposed system using pre-trained CNN models. The effectiveness of the proposed plan has been evaluated using the State farm Driver Distraction data. The images are reshaped into (224,224,3) before training. The pre-trained CNN models are used to predict the class. There are ten classes out of which one class is safe driving while the remaining nine are distracted. The results of each model are described below.

4.2.1 Performance Evaluation:

Out[13]: <matplotlib.legend.Legend at 0x7f3bf2c484e0>

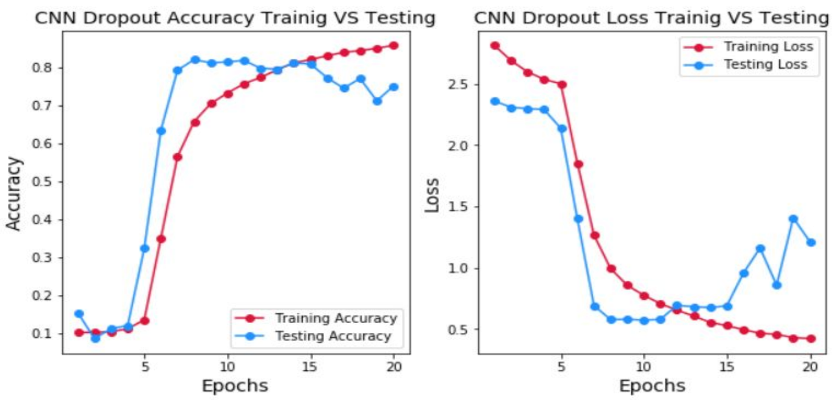


Fig. 4.1 VGG-16 accuracy and loss

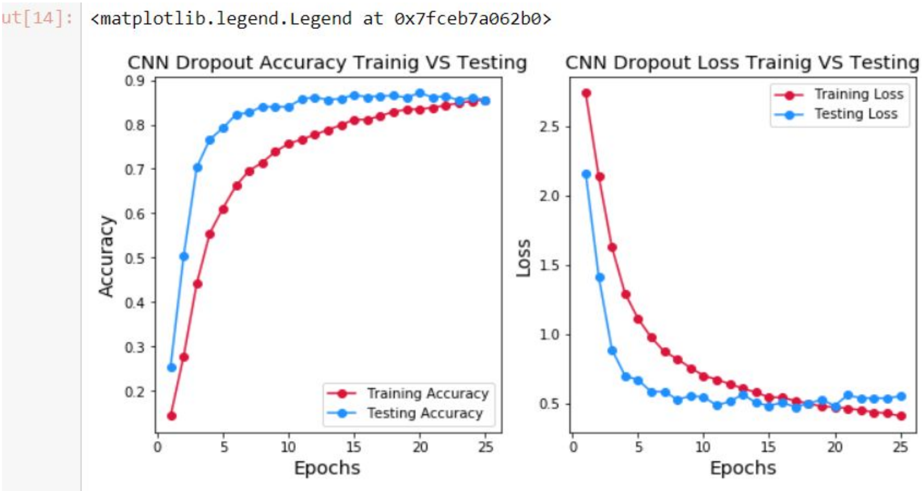


Fig. 4.2 ResNet50 accuracy and loss

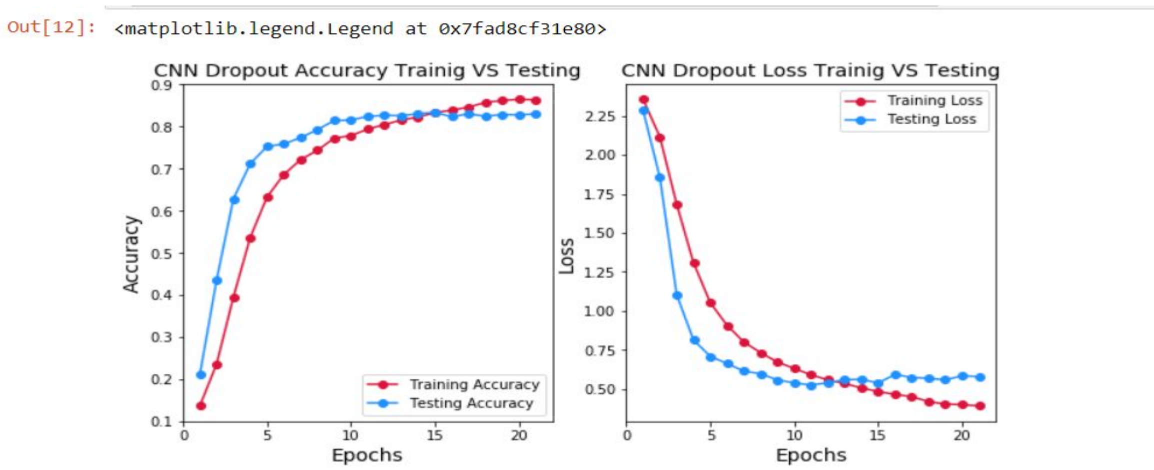


Fig. 4.3 Xception Accuracy and loss

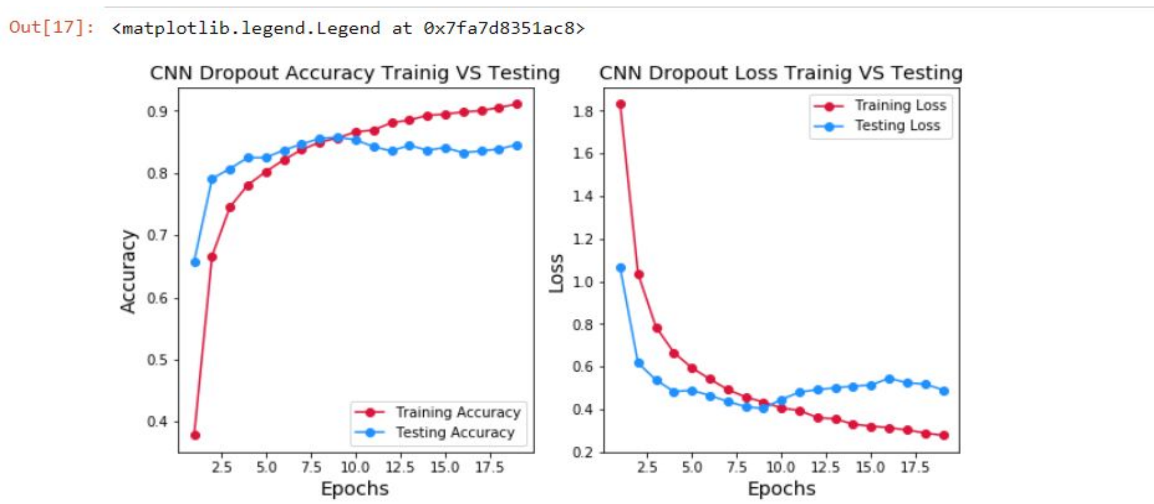
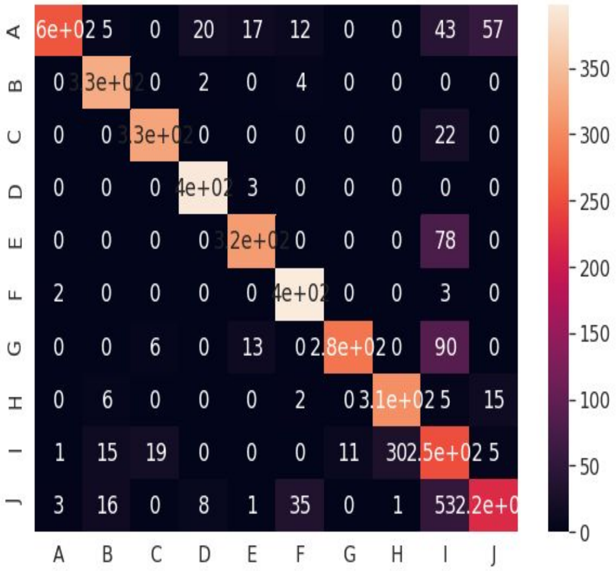
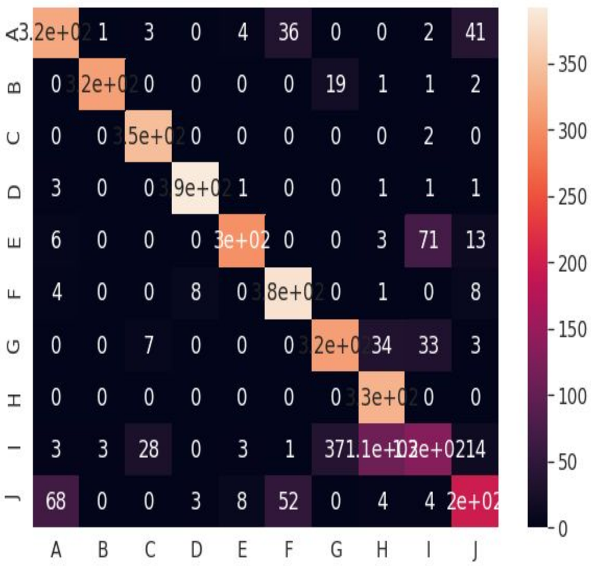


Fig. 4.4 MobileNetV2 accuracy and loss

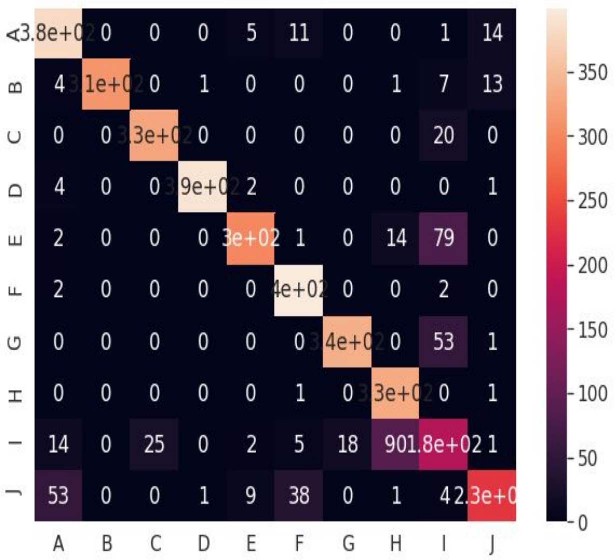
4.2.2 Confusion Matrix:



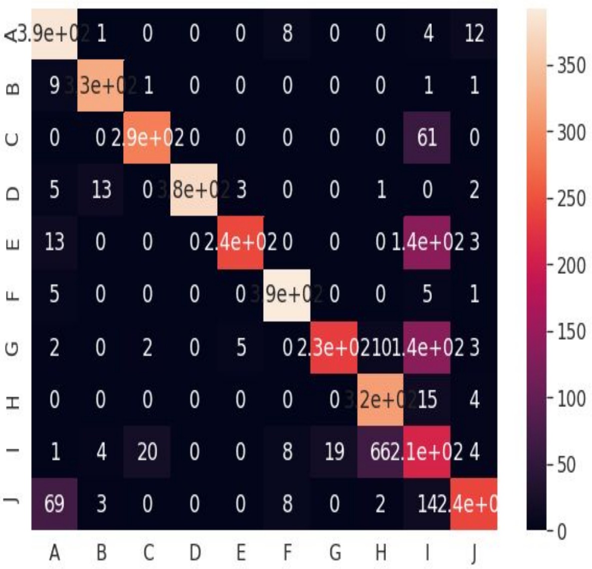
MobileNetV2



Xception



ResNet50



VGG-16

Fig. 4.5 Confusion Matrix

Table 4.1: Comparison of each Model

Model	Accuracy	Log Loss
VGG16(No extra layers)	84.7	0.5
VGG16(Extra layers)	81.5	0.57
Xception(no extra layers)	82.8	0.55
Xception(Extra layers)	82.5	0.52
RESNET50(No extra layers)	85.43	0.55
RESNET50(Extra layers)	86.43	0.47
MobileNet(No extra layers)	85.7	0.4
MobileNet(Extra layers)	83.6	0.63
Ensemble	92.1	0.24

For each model, all the individual performance evaluation metrics are compared. Among all these Models MobileNet gave the better results. But after ensembling all the models we attained better accuracy and Logarithmic loss .

We achieved an accuracy of over 92.1% and log_loss as 0.24 after the ensembling which is a better result compared to other models.

4.2.3Predictions:

Fig. 4.6 Predicting on Test Data

CHAPTER 5

CONCLUSION AND FURTHER WORK

In this paper, the dataset is trained with five models and with network architecture. And comparison is done among all these models out of which MobileNetV2 gives better accuracy. Then all these models are ensembled to attain the best accuracy. Using five different pre-trained models and ensembling them we have achieved the results.

We have demonstrated an ensemble of the pre-trained models to obtain better results by applying pre-trained weights. We have extended the base paper by using additional pre-trained models and adding extra layers like Convolution, Pooling, Dropouts, Dense Layers to these pre-trained models. We used Keras library with Tensorflow as the backend. For prediction, the input can be an image or even a pre-recorded video.

Further, there is a good scope to achieve even better accuracy by using other pre-trained models like EfficientNet, SpineNet etc. This research can be extended further such that a Live-Video Stream can be fed as an input to predict whether the driver is alert or not. Which can help us in achieving our aim to provide an accident free environment. Our experimental results clearly depict that this approach of Ensembling models provides very significant difference in the output when compared to individual trained Models.

CHAPTER 6

REFERENCES

1. Yang Xing , Chen Lv , Huaji Wang, Dongpu Cao , Efstathios Velenis, and Fei-Yue Wang (2019) Driver Activity Recognition for Intelligent Vehicles: A Deep Learning Approach. *IEEETransaction On Vehicular Technology*, **Vol. 68, No. 6**.

2. Whui Kim , Hyun-Kyun Choi , Byung-Tae Jang and Jinsu Lim (2017) "Driver distraction detection using single convolutional neural network" *International Conference on Information and Communication Technology Convergence (ICTC).*

3. JianCheng ZOU ,PeiZhou YAN and ZhengZheng LI (2019) "Detecting Distraction of Drivers Based on Residual Neural Network" *IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE).*

4. S. Kusuma, J. Divya Udayan and Aashay Sachdeva (2019) "Driver Distraction Detection using Deep Learning and Computer Vision" *2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT).*

5. Karen Simonyan, Andrew Zisserman and Visual Geometry Group, Department of Engineering Science, University of Oxford "Very Deep Convolutional Networks For Large-Scale Image Recognition"

6. Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton Presented by Tugce Tasci, Kyunghee Kim "ImageNet Classification with Deep Convolutional Neural Networks".

7. Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun Microsoft Research "Deep Residual Learning for Image Recognition".

8. Francois Chollet Google, Inc. "Xception: Deep Learning with Depth Wise Separable Convolutions".

9. Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto and Hartwig Adam "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications".

10. Chuanqi Tan, Fuchun Sun, Tao Kong, Wenchang Zhang, Chao Yang and Chunfang Liu "A Survey on Deep Transfer Learning".

11. John Elder (2018) “The Apparent Paradox of Complexity in Ensemble Modeling” *Handbook of Statistical Analysis and Data Mining Applications (Second Edition)*.

12. Alwaseela Abdalla, Haiyan Cen, Liang Wan, Reem Rashid and Yong He (2019) Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure *Computers and Electronics in Agriculture* ***Vol 167*** Article **105091**.

13. JianCheng ZOU, PeiZhou YAN and ZhengZheng LI (2019) Detecting Distraction of Drivers Based on Residual Neural Network *IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE)*.

14. Negar Moslemi, Reza Azmi and Mohsen Soryani (2019) Driver Distraction Recognition using 3D Convolutional Neural Networks *4th International Conference on Pattern Recognition and Image Analysis (IPRIA)*.

15. Jaeseok Kim, Heejin Choi, Jinuk Park, Juntae Kim and Minsoo Hahn (2018) Voice Activity Detection based on Multi-Dilated Convolutional Neural Network" *Proceedings of the 2018 2nd International Conference on Mechatronics Systems and Control Engineering - ICMSCE,2018.*

16. Mario Rivas-Sánchez, Maria De La Paz Guerrero-Lebrero, Elisa Guerrero and Guillermo Bárcena-Gonzalez (2017) Using Deep Learning for Image Similarity in Product Matching *Chapter 25 Springer Science and Business Media LLC.*