Detection of Driver's Mobile Phone Usage

Hayrullah Yasar, *Phd Student, De La Salle,*

***Abstract*—While driving, mobile phone usage is dangerous that it may cause traffic accident. Detection and proof of usage should be done by a system. Anti-Distracted Driving Act that became a law last August 1, 2016 will now be enforced starting May 18, 2017 in the Philippines. So drivers may get penalized if they use mobile phone while driving. On this study it is intended to develop a neural network application that can detect mobile phone usage. Sample pictures used for the system training and testing. Positive pictures and negative pictures were used to train the Cascade Object Detector on MATLAB.**

***Index Terms*—Cascade Classifier, Object Detection, neural network, computer vision, cascade object detector, mobile phone usage detection of driver.**

I. Introduction

THIS research paper is intended to explore a new technique for detection of mobile phone while driving. Anti-Distracted Driving Act that became a law last August 1, 2016 will now be enforced starting May 18, 2017 in the Philippines. for this reason such an enforcement system is needed to apply on the live traffic. When it comes to penalization of the driver, it must be proved by an enforcement. It is of course better with a picture that driver's usage will be indicated. Using a mobile phone is a daily part of our lives any time we may get text messages, IM, direct call or social media using. So any time people get used to practice it. But when it comes to driving, especially in the crowded city center in the Philippines, it might be more dangerous and even people might die just because of using a mobile phone. Manila has one of the highest population density in the world so the law is prohibited drivers to use mobile phone and it should be practiced on the roads by detecting of the driver and penalize to give up this badly habit.

Day by day artificial intelligence application increase when we check the latest news on the media and research papers. Deep learning is the most emerged topic under the artificial intelligence study. Especially image processing trend is so fast to grow and there are many studies. Traffic management systems and traffic enforcements systems uses many cameras. All around the world city centers are monitored by many security cameras. Local governments and traffic enforcement groups are using cameras for security and monitoring purposes. Image processing applications are giving a chance to security systems to detect some objects. Object detection of image processing is an emerging topic that provide many application. So this idea come up when it comes to an enforcement system design in the Philippines to penalize the drivers according to anti-Distracted Driving Act if they are using mobile phone while driving. Such a system can be designed by deep learning and neural network application with the training of the system. There are many different types of vehicles and different types of gestures to detect for mobile phone usage. According to D. Wang and M. Pei and L. Zhu, Detecting Driver Use of Mobile Phone Based on In-car Camera, they made a study that requires and input camera inside the car[1]. On Philippines condition it can not be possible to apply because it is not indicated on the law and drivers may easily block it while using a mobile phone. So internal camera application is not practical in the Philippines condition. The other details on those study that, generally there are types of detection; detection of hand state, detection of phoning action and phoning activity detection in image sequences.

H.Yasar was with the Department of Elecronics and Communication Engineering, De La Salle University, Manila, Philippines

978-1-5386-0912-5/17/$31.00

The other study is Driver Cell Phone Usage Detection From HOV/HOT NIR Images, they intended to get a pictures on a view that front windshield of the vehicle when it approaches on the target of the camera view, getting pictures of the vehicle and firstly detecting the drivers face[2]. According to face of the driver, ROI is identified and mobile phone usage gestures detected. Region of Interest(ROI) approach is used also on this study to create a table on MATLAB to provide positive pictures to program. The detection approach is to train the system by positive and negative pictures. Positive pictures are with mobile phone usage and negative pictures are without mobile usage.

Prior to application details, neural network fundamentals details explanied on the introduction part.

Deep neural networks are a set of functions that every function is called a layer. A layer has a linear function and output function is non-linear[3]. Gradient descent is one of the popular algorithm that used in neural network. The algorithm provides optimization. So a gradient descent based optimizer can be used for training of function parameters in deep neural network. For the classification task two equations defined that objective function is cross-entropy(eq. 1) and probability distribution is softmax (eq. 2).

$$H(p,q) = \sum_x p(x) log q(x) \tag{1}$$

$$\sigma(z)_j = \frac{e^z j}{\sum_{k=1}^{K} e^z k} \quad for \quad j = 1, ..., K \tag{2}$$

According to the article[4], deep neural network application procedure is explained to solve classification problem type by firstly selecting a network architecture, secondly training the network and thirdly applying it to the problem at hand.

Network architecture might vary to application and there are many types so far mainly introduced four major type network architectures by Josh Patterson and Adam Gibson

- Unsupervised Pretrained Networks (UPNs)

- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks
- Recursive Neural Networks

A. Neural Network and Deep Learning

Deep machine learning is a new type of machine learning that the concept is based on representation of data. On that model it possible that higher level of data comes from low level with hierarchically and the process is like human brain function. Deep learning concept derived from artifical neural network research. Multilayer perceptron with many hidden layers is one of the good example for the deep architecture model. Deep belief networks(DBNs) was introduced on 2006 as an unsupervised learning algorithm. DBN provides better results than with the random weights. Additional advantages of DBNs[5]:

- Provides effective usage of unlabeled data
- Possible to interprete as Bayesian probabilistic generative model
- Provides efficient computation on the deepest layer variables
- Effective addressing of underfitting problem that happens in deep networks by the generative pretraining step.

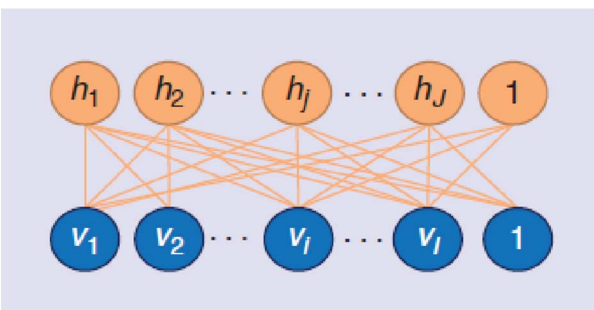

Fig. 1. An RBM with *I* visible units and *J* hidden units[5]

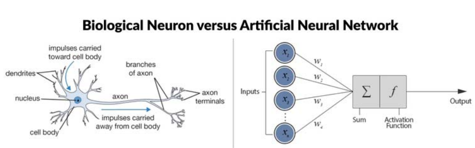

Fig. 2. Biological Neuron vs. Artificial Neural Network[6]

There are six components of to artifical neurons. From left to right, according to figure 2, these are :

1) **Input Nodes**; As it so happens, each input node is associated with a numerical value, which can be any real number.
2) **Connections**; each connection that departs from the input node has a weight associated with it and this can also be any real number.
3) **Weighted Sum**; All the values of the input nodes and weights of the connections are brought together: they are used as inputs for a weighted sum:

$$y = f\left(\sum_{i=1}^{D} w_i * x_i\right) \tag{3}$$

4) **Transfer Function**; This result will be the input for a transfer or activation function. In the simplest case, transfer function would be an identity function, $f(x) = x$ or $y = x$. In this case, x is the weighted sum of the input nodes and the connections. However, just like a biological neuron only fires when a certain treshold is exceeded, the artificial neuron will also only fire when the sum of the inputs exceeds a treshold.
5) **Output Node**; As a result we have the output node, which is associated with the function
6) **Bias**; The perceptron maybe an additional parameter, called a bias, which you can actually consider as the weight associated with an additional input node that is permanently set to 1. The bias value is important because it allows you to shift the activation function to the left or right, which can make a determine the success of your learning.

B. Detection of Mobile Phone Usage

The main objective is to detect drivers hand and fingers near his or her head with a mobile phone. Mobile phone can not be just detected so a specific region of interest should be identified around the face of the driver. MATLAB image processing and computer vision system toolbox will be used to develop such system. and MATLAB Trainin Image Labeler application will be used to specifize the region of interest(ROI) to detect mobile phone usage as shown in figure 3.

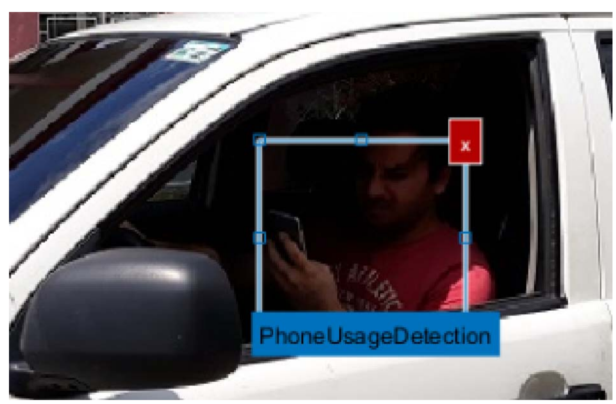

Fig. 3. ROI samples on a positive sample for mobile phone usage

II. Object Detection with Cascade Classifier

Object detection with cascade classifier is proposed by Paul Viola and Michael Jones in their paper ”Rapid Object Detection using a Boosted Cascade of Simple Features” in 2001. They proposed a new algorithm and new frameworks to implement object detection application with high accuracy. And they used Haar-like features for face detection. According to

Paul Viola and Michael Jones, their object detection classifies based on the value of simple features because features-based system has mainly two advantages on pixel-based system[7].

- Easy learning
- Faster operation

Features-based system approach illustrated on figure 4.

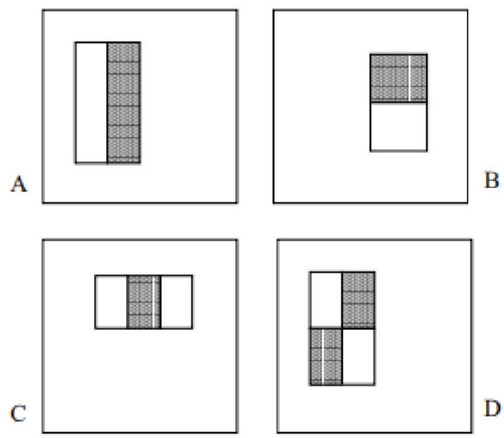

Fig. 4. (A) and (B) two-rectangle features, (C) three rectangle feature, (D) four-rectangle feature[7]

Cascade object detector can detect object categories whose aspect radio does not very significantly. Cascade classifier works with multiple stages.Every stage is an ensemble of weak learners. The weak learners are simple classifiers called decision stumps. Each stage is trained using a technique which is boosting. Boosting provides the ability to train a highly accurate classifier by taking a weighted average of the decisions made by the weak learners. The system stages are working on the assumption that positive images have the region of interest and negative images have no region of interest. so processing time is mainly spent on the positive images. positive indicates that an object found and negative indicates no object were found. In order to have higher positive rates, need to add more stages. As added more stages need to add more negative pictures as well[8].

III. Experiments

On this paper, training of the mobile phone usage detection done by 49 positive samples and 30 negative samples. Positive sample pictures are with a driver using a mobile phone and negative sample pictures are either without driver or driver without mobile phone. To have a good training of the system, negative samples should be more or less same with positive samples for this reason almost same pictures but without mobile phone usage dataset trained. Region of interest(ROI) selected and for all the 49 positive samples. According to experiment, it is understood that more ROI selection gives better result to detection. Resolution of samples are all same that 615x461 pixels. If the samples are more higher resolution, training of the system is getting very long process. Low resolution images make the training faster and more reliable. the reason of that ROI is also small rectangular of point(low resolution) that includes mobile phone and driver face on the positive samples. On the MATLAB program ”trainCascadeObjectDetector” function is used for training. The general parameters of the function is given that number of stage is 8, false alarm rate 0.1 and feature type is HAAR-Feature.

One of the output of the system showed on the figure 6. Mobile phone detection is correctly done the the label ”mobile phone usage” is displayed with yellow color.

Fig. 5. Mobile Phone Usage detection

Depends on the vehicle type, sometimes system might detect multiple usage of mobile phone like on figure 6. On that case, cascade detector detected three times mobile phone usage, the reason for that vehicle type and background of the vehicle.

Fig. 6. Multiple Detection of Mobile Phone Usage

For the testing of the system 8 pictures were used. Details of the test pictures shown on the table 1.

According to test pictures and system result, if mobile phone nar the driver’s head, detection can be done but if the mobile phone on the hand of driver, detection can not be done. Actually the result depends on the data set given to system. If possible to add more sample pictures, cascade classifier may learn also to detect the phone on hand. For the test picture 2, detection of usage is three times as shown on the figure 6. Vehicle type 2 causes the problem. Test picture 7 is not detected by the system that the reason might be the face of

TABLE I
INPUT PICTURES AND SYSTEM RESULT

Test Image	Phone	Vehicle	Usage	System Detection
Test Pic 1	Near Head	Vehicle 1	YES	Detected
Test Pic 2	Near Head	Vehicle 1	YES	Detected
Test Pic 3	Near Head	Vehicle 1	YES	Detected
Test Pic 4	On Hand	Vehicle 1	YES	Not Detected
Test Pic 5	Near Head	Vehicle 2	YES	Detected
Test Pic 6	-	Vehicle 3	NO	Not Detected
Test Pic 7	Near Head	Vehicle 1	YES	Not Detected
Test Pic 8	-	Vehicle 1	NO	Not Detected

the driver is not so visible. According to experiment, generally mobile phone detection dependent on those parameters;

- Number of Positive pictures with identified ROI
- Positive and negative samples
- Number of cascade stages
- Picture resolution
- False alarm rate

The initial training of the system it is taught that higher resolution may give good result. However the practical approach was not good. Considering the process time of training on the high resolution, the idea came up that low resolution training would be better. So after resizing of the pictures manually the expected result of detection goes higher. The other reason might be the small dataset. Because of some limitations such as limited vehicle, limited people for testing, etc. big data set could not been generated. Installing a camera to live traffic would give more dataset and to train the system. The accuracy would be definitely be different on real time application because of the parameters that windshield, vehicle type, direct sun on day time and darkness on night time, windshield tint etc.

IV. CONCLUSION

On that study, a novel method was applied for the detection of mobile phone usage. The study might me an enforcement system for Philippines LTO or MMDA to use. To implement such a system, cameras should be installed on a proper place to get good view. Real time application of the system is not considered so system analysis and testing procedures needed for real time application. The MATLAB code may modified to accomplish for real time application. On the other hand it is assumed that program input would be picture only. so on the real time application it should be considered to get snapshot of vehicle. So getting the good view of picture can be done only to detect vehicle first and than getting picture of the view of driver. So trigger algorithm might be needed to send command to camera to take snapshot. Detection of vehicle might be done by loop detector. Loop detector may easily detect vehicle when vehicle approach so on that point trigger signal will be sent to camera. And camera might take picture and the system may evaluate if the driver using mobile phone. A general trigger point of system shown on figure 7.

Program output accuracy is 6 of 8 so the percentage of the output accuracy is 75 achieved result.

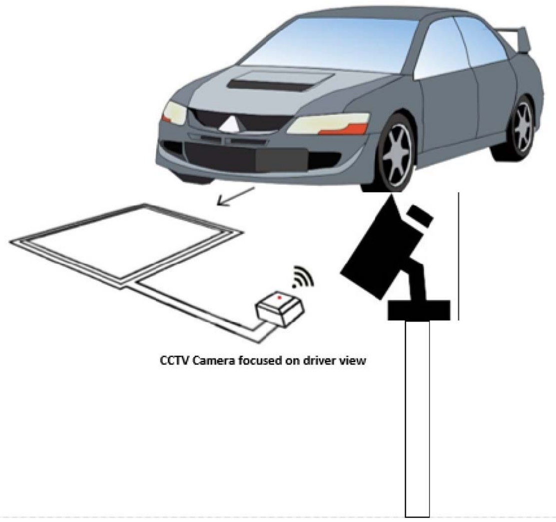

Fig. 7. Vehicle Detection and Camera Snapshot

ACKNOWLEDGMENT

The author would like to thank his superior Dr. Elmer P. Dadios and my friends Ms. Jessica Velasco and my classmates to all of their help and support.

REFERENCES

[1] D. Wang, M. Pei, and L. Zhu, "Detecting driver use of mobile phone based on in-car camera," in *2014 Tenth International Conference on Computational Intelligence and Security*, Nov 2014, pp. 148–151.

[2] Y. Artan, O. Bulan, R. P. Loce, and P. Paul, "Driver cell phone usage detection from hov/hot nir images," in *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, June 2014, pp. 225–230.

[3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[4] N. E. West and T. J. O'Shea, "Deep architectures for modulation recognition," *CoRR*, 2017.

[5] D. Yu and L. Deng, "Deep learning and its applications to signal and information processing [exploratory dsp]," *IEEE Signal Processing Magazine*, vol. 28, no. 1, pp. 145–154, Jan 2011.

[6] K. Willems. Keras tutorial: Deep learning in python. [Online]. Available: https://www.datacamp.com/community/tutorials/deep-learning-python

[7] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 1, 2001, pp. I–511–I–518 vol.1.

[8] traincascadeobjectdetector. [Online]. Available: https://www.mathworks.com/help/vision/ref/traincascadeobjectdetector.html