International Journal of Advanced Trends in Computer Applications
www.ijatca.com

Detection of Motorcyclists without Helmet using Convolutional Neural Networks

Taiwo Ajagunsegun[1], Parvinder Kaur[2]
[1,2]Department of Computer Science and Engineering,
Shaheed Udham Singh College of Engineering and Technology,
Tangori, Mohali, India.
[1]tajagunsegun@yahoo.com, [2]parvinderkaurcse@sus.edu.in

Abstract: *Road traffic injuries are one of the highest public health hazards and in order to bring down the mishaps, one should be well aware of the road safety rules. Most highway users ignore road safety rules probably because they believe it is perfectly okay to violate them, or they get a feeling of accomplishment from being able to violate them and not get caught. The helmet is the main safety equipment for motorcyclists, but many riders do not use it. In order to enhance the enforcement of obeying road safety rule of wearing a suitable helmet for every motorcycle user on the highway, an automated system that captures highway users' state and report defaulters to appropriate authorities is important. Deep learning method known as transfer learning with Convolutional Neural Networks can be highly efficient in expediting detection of helmet on the highway.*

Keywords: Helmet Detection, Image Classification, Convolutional Neural Network, Transfer Learning.

I. INTRODUCTION

Although road safety rules are universal the world over, some of the rules in India vary from state to state. The differences are mostly peculiar to signage, speed limits, and the punishments for breaking certain laws. It is therefore expedient to familiarize oneself with the traffic laws in one's society. Irrespective of one's intention toward driving, learning the rules of the road keeps highway users safe even in an unfamiliar terrain. While there is sometimes an aura of relaxation and leisure about the driving experience, traffic laws in India are very strict and straying from these can cause serious trouble. Asides getting arrested by the police, an accident is likely to happen which might lead to serious injury or loss of life.

Here are some of the reasons why most highway users ignore traffic laws:

- Some drivers believe that traffic laws and regulations apply only to other people.
- Some drivers believe it is perfectly okay to violate traffic laws and regulations.
- Some drivers get a feeling of accomplishment from being able to violate traffic laws and not get caught.
- Some drivers actually do not care whether their driving harms other people.

Road traffic injuries are one of the highest public health hazards and in order to bring down the mishaps, one should be well aware of the road safety rules. Here are two major road safety rules in India a motorcycle user should always remember:

- Wear a helmet:– While wearing a helmet for a long period can be uncomfortable for some, it keeps one protected not only from collisions at the time of a fall but even from wind blasts when riding at high speeds.
- Avoid swerving between lanes:– Many motorcycle users love to ride across lanes and squeeze in tiny gaps. Nevertheless, it is advisable to follow lane discipline and avoid sudden maneuvering to prevent accidents.

Every state in India has mandated the use of a helmet while riding a motorcycle. However, in December 2019, the Gujarat high court made helmets temporarily optional. Not wearing a suitable helmet while riding a motorcycle is something that is not just illegal but also puts one at risk of suffering from serious head injuries in an event of an accident. Hence, one should wear an ISI-approved helmet of recommended specifications while riding a motorcycle or a scooter.

In order to enhance the enforcement of obeying road safety rule of wearing a suitable helmet for every motorcycle user on the highway, an automated system that captures highway users' state and report defaulters to appropriate authorities is important.

II. RELATED WORK

Some researchers have proposed several methods to solve the problem of real-time helmet detection on the highway. Chiu et al., 2007 [6] proposed a system that detects motorcyclists in surveillance videos. This system segments the moving object and then tracks motorcycles and heads using a probability-based algorithm which handles the occlusion problem but it is unable to handle small variations due to noise and illumination effects. Also, it uses Canny edge detection with a search window of certain size in order to detect head.

Chiverton et al., 2012 [7] used edge histogram-based features to detect motorcyclists. This method performs well even if there is low illumination in videos due to the use of edge histograms near the head but since the edge histograms used circular hough transforms to compare and classify helmets, it leads to a lot of misclassifications among motorcyclists with helmet as helmet-like objects were also classified as helmet. To overcome this misclassification problem, Silva et al., 2013, 2014 [17], [18] proposed a system which tracks the vehicles using Kalman filter [12]. This Kalman tracking system tracks objects even if they are lightly occluded but when there are two or more motorcyclists in the same frame, Kalman filter fails because it mostly works well for linear state transitions (that is, tracking single object at a time). In order to track multiple objects, non-linear functions are needed.

Dahiya et al., 2016 [8] proposed a system which first uses Gaussian mixture model to detect moving objects. This model is robust to slight variations in the background. It uses two classifiers, one for the separating motorcyclist from moving objects, and another for separating without helmet from the upper one fourth part of the motorcyclists. However, it uses only hand engineered features such as SIFT [13], HOG [9], LBP [11] along with kernel SVM in both classifications. Their approach was promising as it accurately classifies motorcyclists and non-motorcyclists but was not able to accurately classify between helmet and non-helmet riders under difficult conditions. Singh et al., 2016 [19] proposed a visual big data framework which scales the method in [8] to a city scale surveillance network. Experimental results show that the framework is able to detect a violator in less than 10 milliseconds.

All these existing methods suffer from several challenges such as occlusion of objects, illumination effects, and poor localization on images with less pixels.

III.OVERVIEW OF TRANSFER LEARNING

Transfer learning is the reuse of a pretrained model on a new problem. In transfer learning, a machine exploits the knowledge gained from a previous task to improve generalization about another but related task as shown in figure 1. For instance, in training a classifier to predict whether an image contains food, one could use the knowledge it gained during training to recognize drinks. The architecture as well as the weights of the pretrained model are transferred. Either the complete model is transferred or only part of the early layers.

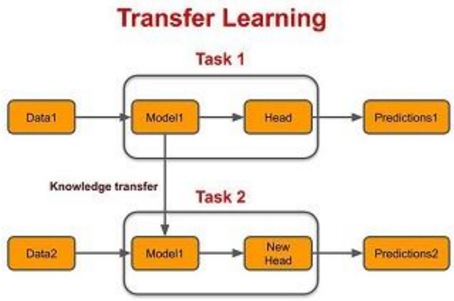
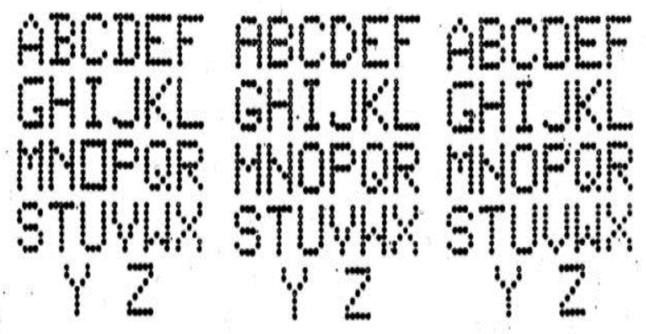


Figure 1: Pictorial illustration of Transfer learning

Early works after 1986 used other terms to describe transfer learning. One of such terms was “Sequential Learning”, where negative transfer learning was covered with the term “Interference” [14]. Other terms used were “Adaptive Generalization” [16], “Learning by Learning” [15], and “Lifelong Learning” [20].

(Stevo Bozinovki, 2020) gave a review of the initial work on transfer learning in neural networks which took place between 1972 [2] and 1985 [3],[4]. Initial experiments were with a dataset containing images of alphabets A, B, E, F, and T taken from the terminal IBM29 card puncher. Those experiments were carried out on the IBM 1130 computer. Later experiments were carried out with two datasets. One dataset contained 40 images, consisting of 26 alphabets, 10 numbers, and 4 special characters from the terminal IBM29. The other dataset can be described as Computer Terminals dataset, shown in figure 2, consisting of 78 images, taken from three computer terminals (IBM29 card puncher, VR14 video screen, and VT50 video screen). The experiments were carried out on a computer VAX/VMS. Figure 2 shows the Computer Terminals dataset which points out that the alphabets in the three terminals are mostly identical on an image with resolution 7x5, with differences in alphabets A, B, D, G, J, M, N, O, V, and W.



IBM29 **VR14** **VT50**

Figure 2: The Computer Terminals dataset used in experimental investigation [5].

These experiments showed that the training time of the model trained with transfer learning is 36% of the training time of the model trained from scratch and the speed of learning is 2.8 times faster.

Previous works performed on image recognition tasks such as the simple digit image recognition [14] showed that the architecture of a network strongly influences the ability of the network to generalize. Good generalization on complex tasks can be achieved by designing a network architecture that has a certain amount of a priori knowledge about the task. The basic design principle is to minimize the number of free parameters in the network as much as possible without overly reducing the computational power of the network. This principle increases the probability of correct generalization because it results in a specialized network architecture that has a reduced entropy.

There are two common approaches for developing models through transfer learning. In the first approach, the pretrained model is used as a feature extractor. Some specific layers in the pretrained model are used for the generation of features. These features are then used as input to another machine learning algorithm to train the model on a new task. The weights of the entire network are frozen except that of the last fully-connected layer. The last fully-connected layer is replaced with a new one with random weights and only this layer is trained.

The second approach is known as fine-tuning. The weights of the new network are initialized with the weights of a pretrained model which are then used to train the network allowing for slight adjustments to more abstract representations of the model. Fine-tuning adapts some of the representations previously learned by an existing model to a new problem which helps the model to perform better.

In most cases, the pretrained models were initially trained on big datasets with millions of images which would have learnt useful features in the lower layers. These features could be useful for similar tasks and help to improve the performance of new models because the weights are well initialized to detect common

IV. CONCLUSION

This study reviewed the deep learning method known as transfer learning using convolutional neural network and how it can be useful in classifying highway images of motorcyclists into either “With Helmet” or “Without Helmet”. The two approaches for developing models through transfer learning were identified as feature extraction and fine-tuning.

Convolutional neural networks are proving to be the best type of deep learning algorithm when faced with image recognition and classification tasks yielding higher and more accurate results than other previously used methods. They are particularly favorable when working with small datasets and are able to still yield decent results. However, one of the drawbacks with small datasets is overfitting, but this can be mitigated with data augmentation.

REFERENCES

[1] An introduction to convolutional neural networks [Online] available at: http://white.stanford.edu/teach/index.php/An_Introduction_to_Convolutional_Neural_Networks.

[2] Bozinovski, S. (1972). Perceptrons: Training in Pattern Recognition. (original in Croatian: Perceptroni i obucavanje u prepoznavanju oblika) unpublished student scientific competition paper, University of Zagreb.

[3] Bozinovski, S. (1985a). Adaptation and training: A viewpoint. Automatika 26 (3-4) 137-144.

[4] Bozinovski, S. (1985b). A representation theorem for linear pattern classifier training. IEEE transactions on systems, man, and cybernetics, (1), 159-161.

[5] Bozinovski, S. (2020). Reminder of the first paper on Transfer Learning in Neural Networks, 1976. Informatica, 44(3).

[6] Chiu, C. C., Ku, M. Y., & Chen, H. T. (2007, June). Motorcycle detection and tracking system with Occlusion Segmentation. In Eighth International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS'07) (pp. 32-32). IEEE.

[7] Chiverton, J. (2012). Helmet presence classification with motorcycle detection and tracking. IET Intelligent Transport Systems, 6(3), 259-269.

[8] Dahiya, K., Singh, D., & Mohan, C. K. (2016, July). Automatic detection of bike-riders without helmet using surveillance videos in real-time. In 2016 International Joint Conference on Neural Networks (IJCNN) (pp. 3046-3051).

[9] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (Vol. 1, pp. 886-893).

[10] Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., & Shet, V. (2013). Multi-digit Number Recognition from Street

View Imagery using Deep Convolutional Neural Networks. URL http://arxiv.org/abs/1312.6082.

[11] Guo, Z., Zhang, L., & Zhang, D. (2010). A completed modeling of Local Binary Pattern Operator for Texture Classification. IEEE transactions on image processing, 19(6), 1657-1663.

[12] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering, vol. 82, no. 1, pp. 35–45.

[13] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), 91-110.

[14] McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In Psychology of learning and motivation (Vol. 24, pp. 109-165). Academic Press.

[15] Naik, D., Mammone, R. (1993). Learning by learning in Neural Networks, In R. Mammone (ed.) Artificial Neural Networks for Speech and Vision, Chapman and Hall, London.

[16] Sharkey, N. E., & Sharkey, A. J. (1993). Adaptive generalisation. Artificial Intelligence Review, 7(5), 313-328.

[17] Silva, R., Aires, K., Santos, T., Abdala, K., Veras, R., & Soares, A. (2013, October). Automatic detection of motorcyclists without helmet. In 2013 XXXIX Latin American Computing Conference (CLEI) (pp. 1-7).

[18] Silva, R. R. V., Aires, K. R. T., & Veras, R. D. M. S. (2014, August). Helmet detection on motorcyclists using image descriptors and classifiers. In 2014 27th SIBGRAPI Conference on Graphics, Patterns and Images (pp. 141-148).

[19] Singh, D., Vishnu, C., & Mohan, C. K. (2016, December). Visual big data analytics for traffic monitoring in smart city. In 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 886-891). IEEE.

[20] Thrun, S., & Mitchell, T. M. (1995). Lifelong robot learning. Robotics and autonomous systems, 15(1-2), 25-46.

[21] Zakhareuski, A. (2021, May). U.S. Rules of the Road. Website. https://driving-tests.org/beginner-drivers/rules-of-the-road

Author Profile

Taiwo Ajagunsegun received his B.Eng. degree in Electrical and Electronics Engineering from the prestigious Federal University of Technology Akure, Nigeria in 2014. He is currently a masters' degree student at Shaheed Udham Singh College of Engineering and Technology, Tangori, Mohali, Punjab with major in Machine Learning and Artificial Intelligence.

Parvinder Kaur: Assistant Professor, Department of Computer Science and Engineering at Shaheed Udham Singh College of Engineering and Technology, Tangori, Mohali, Punjab.