An Automatic Detection of Helmeted and Non-helmeted Motorcyclist with License Plate Extraction using Convolutional Neural Network

Jimit Mistry, Aashish K. Misraa, Meenu Agarwal, Ayushi Vyas, Vishal M. Chudasama, Kishor P. Upla
Electronics Engineering Department,
Sardar Vallabhbhai National Institute of Technology (SVNIT),
Surat, India.
{mistryjimit26, iaashish7, mee.agarwal97, vyas.ayushi6, vishalchudasama2188, kishorupla}@gmail.com

***Abstract*—Detection of helmeted and non-helmeted motorcyclist is mandatory now-a-days in order to ensure the safety of riders on the road. However, due to many constraints such as poor video quality, occlusion, illumination, and other varying factors it becomes very difficult to detect them accurately. In this paper, we introduce an approach for automatic detection of helmeted and non-helmeted motorcyclist using convolutional neural network (CNN). During the past several years, the advancements in deep learning models have drastically improved the performance of object detection. One such model is YOLOv2 [1] which combines both classification and object detection in a single architecture. Here, we use YOLOv2 at two different stages one after another in order to improve the helmet detection accuracy. At the first stage, YOLOv2 model is used to detect different objects in the test image. Since this model is trained on COCO dataset, it can detect all classes of the COCO dataset. In the proposed approach, we use detection of person class instead of motorcycle in order to increase the accuracy of helmet detection in the input image. The cropped images of detected persons are used as input to second YOLOv2 stage which was trained on our dataset of helmeted images. The non-helmeted images are processed further to extract license plate by using OpenALPR. In the proposed approach, we use two different datasets i.e., COCO and helmet datasets. We tested the potential of our approach on different helmeted and non-helmeted images. Experimental results show that the proposed method performs better when compared to other existing approaches with 94.70% helmet detection accuracy.**

***Index Terms*—helmet detection, YOLOv2, license plate extraction, COCO**

I. INTRODUCTION

Motorcycles being one of the most convenient modes of transport has led to increasing use and thus accounts for the highest share of total road accidents. As per one survey carried out in the year of 2014, 30% of all road accidents related to deaths were of riders on two-wheelers [2]. Also, as per the record of Chennai, India, between the period of January 1, 2013 to June 28, 2015, total 1,453 two-wheeler riders who died in road accidents were not wearing helmet [3]. Also, according to the World Health Organization (WHO), wearing a helmet can reduce the risk of severe injury by 72% and the risk of death by 39% [2]. In India, road safety has been a neglected area. Due to this wearing a helmet has been made mandatory and violation of that attracts a hefty fine. Currently, all major cities employ surveillance system on roads to identify violators which needs a system that can automatically detect whether a motorcyclist is wearing a helmet or not. This paper aims to an automatic helmet and non-helmet detection of motorcyclists.

II. RELATED WORK

Since past several years, many researchers have solved the problem of helmet detection. Wen *et al.* [4] proposed a circular arc detection method based on Hough transform for helmet detection. Chiu *et al.* [5] use a Canny edge detector [6] to detect helmets. Similarly, authors in [7] use a background subtraction to detect moving vehicles. Here, authors calculate a image based statistical information by isolating the approximate region of the head of the riders. In order to extract features, histogram of gradient (HOG) descriptors is used from the isolated regions and then results are passed to a support vector machine (SVM) classifier for classification purpose. Furthermore, authors in [8] proposed a model for vehicle detection, tracking and classification from roadside CCTV cameras. In order to make their model invariant against illumination changes and camera vibrations, they use a new background Gaussian mixture model. Silva *et al.* in [9] propose a hybrid descriptor model based on geometric shape and texture features for automatic detection of motorcyclists without helmet. They use Hough transform with SVM to detect the head of the motorcyclist. Additionally, they extend their work in [10] by multi-layer perception model for classification. In [11], Waranusat *et al.* [11] propose a system to detect moving objects with k-NN classifier over the head region of motorcyclist to classify helmet. Similarly, authors in [12] use background subtraction and object segmentation with k-NN classifier to detect bike-riders. All these methods which are employed with engineered features have limitations in accurate detection of helmets.

Alex *et al.* [13] introduced a convolutional neural network (CNN) based method for object classification and detection. Since then CNNs became the standard model for classification as it outperformed all the traditional models such as HOG [14], scale-invariant feature transform (SIFT) [15] and local binary patterns (LBP) [16]. Recently, A. Hirota *et al.* [17] use a CNN for classification of helmeted and non-helmeted riders. Although they use CNN, their helmet detection accuracy is poor with limitations of helmet color and multiple riders on a single motorcyclist. Literature shows that the CNN can learn engineered features from raw data effectively and it outperforms over it's hand-crafted counterparts. Therefore, we use CNN to solve helmet detection problem. The overall contributions of our paper are as follows:

- Use of person detection instead of motorcyclist in order to increase the helmet detection accuracy in the test image.
- Training a model using our database of helmeted images.
- Extraction of license number plate for non-helmeted motorcyclist.

III. Proposed Approach

The block schematic of the proposed approach is depicted in Fig. 1. Here, we use two YOLOv2 models [1] one after another in order to detect helmeted and non-helmeted motorcyclist. First YOLOv2 model is trained with COCO dataset [18] which detects number of classes in an image. The cropped images of person class from the first YOLOv2 are used as input to the second YOLOv2 model which was trained on helmet dataset. This helmet dataset is prepared by us and it consists of helmeted images only. Also, we modify the number of filters to 30 in the final convolutional layer in order to detect single helmet class. At last, we perform a license plate detection using OpenALPR [19].

Person Detection: The first step in the proposed approach is to detect persons from the input image. In order to perform this, input test image is passed through first YOLOv2 model which is trained on COCO dataset [18]. YOLOv2 [1] is the state of the art object detector and it has the capability to detect all classes of COCO dataset [18]. It can also detect classes such as person, car, motorbike, etc., in addition to other classes in the dataset. We use a person detection class out of all other classes in order to detect helmeted motorcyclists. All other detected classes from the first YOLOv2 model have been discarded.

Here, it is worth to note the reasons for selecting the bounding box of a person i.e. person detection instead of motorcycle in the first step which is based on our empirical tests. Since the input image consists of motorcycles, cars and persons we found out that when the camera is either front or back facing, then the detector does not always detect the motorcycle or it detects motorcycle with very less confidence score. In such cases, if motorcycle detection criteria is used then it leads to many helmets as undetected in the test image. However, in the criteria of person detection, if a person is sitting on a motorcycle, then the head region covers the helmet, and the foot region contains the license number plate area in motorcycle as shown in the Fig 1. For this reason, we use person detection criteria instead of detecting motorcycle or person and motorbike together. This also increases the detection score of helmet in the input test image.

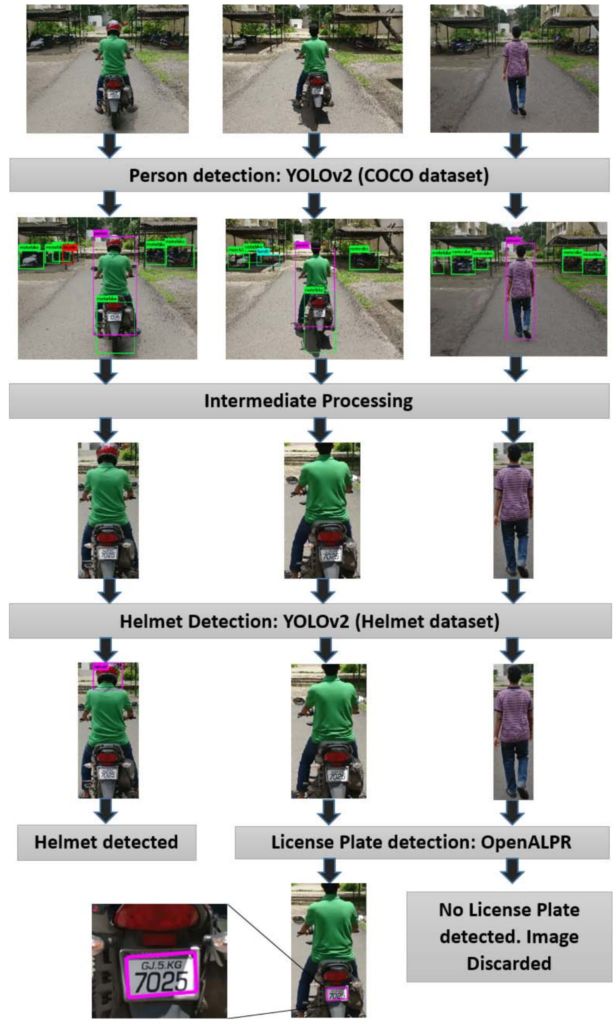


Fig. 1: Block schematic representation of the proposed approach.

Intermediate processing: As discussed earlier, YOLOv2 detects all the classes from the COCO dataset and in the proposed approach we use person detection only in order to detect helmeted and non-helmeted motorcyclist. All the other classes except person are discarded by intermediate processing. Also, the detected person's bounding box is cropped automatically and that image is used for further processing.

Helmet Detection: The main step of the proposed approach is to detect helmeted and non-helmeted motorcyclists in an image. For this step, we again use YOLOv2 model which is trained on dataset of helmeted image. The cropped images of detected person are used as input here to YOLOv2 model. Since YOLOv2 model is trained on helmeted image dataset, whenever a test image consists of helmets, same is classified and detected in that cropped image.

License Plate Detection: If helmet is detected from the output of the second YOLOv2 stage, then the process is stopped for that cropped image. However, if helmet is not detected from detected person's image, then the cropped image is required to pass through the license plate detector, OpenALPR

[19] which detects the license plate with coordinates. We use these coordinates further in order to extract the license plate. Note that if the number plate is not detected, then it means that instead of a person on the motorbike, the first YOLOv2 had detected a pedestrian.

A. You Only Look Once-YOLOv2 [1], [20]

During the past several years, the advancements in deep learning models have drastically improved performance of object detection. One such model is YOLO which combines both classification and object detection in a single architecture [20]. It's upgraded version i.e., YOLOv2 [1] with 23 convolutional layers and 5 max pooling layers is very fast which can detect the helmets in real time with high confidence score. Also, it has a capability to learn general representations of objects, which is very helpful in helmet detection as helmets can be of various shapes and sizes. This motivates us to use this model for our approach. Moreover, YOLOv2 predicts bounding boxes and class probabilities directly from full images in one evaluation. It also learns about the contextual information about classes in addition to their appearance. Due to this, during training, when we give images of person wearing a helmet, it learns that helmet is generally worn in the head region which enables it to detect helmet more accurately.

B. Open Automatic License Plate Recognition (OpenALPR) [19]

OpenALPR is an open-source library used for automatic license plate recognition [19]. It is the state-of-the art license plate library currently available for commercial or research purpose. Due to the robust documentation and vast features we use the same to detect license plate on the motorcyclist. The default detector has been trained for US and EU number plates but we can add a country by training the detector on large example dataset of that country's license plates. Instead of adding a new layout of license plate, which would have been cumbersome, we relied on the region extraction of license plate. The region extraction feature works better and it results the coordinates of the number plate in the image.

IV. Experimental Results

The experiments were conducted on a machine with Intel i7 7700k CPU, 16GB RAM and NVIDIA GeForce GTX1050 GPU. The programs for helmet detection are written in Python 3.5.2 with the help of the various libraries such as OpenCV 3.0 and Darknet [21] which is a neural network framework.

A. Database

In the proposed approach, we use two datasets in order to detect helmeted and non-helmeted motorcyclists. At the first stage of YOLOv2 model, we use COCO database which consists number of images of different classes. The second database which consists of helmeted images is used at the second YOLOv2 model.

Training: For person detection at first YOLOv2 stage, we use COCO dataset [18] in order to train that network. For detecting helmet at the second stage, we trained YOLOv2 model with helmeted image dataset. This dataset is prepared by us and it has total 3054 helmeted images. Here, we trained the YOLOv2 model for one class detection only i.e., helmet. Since our dataset contained around 4000 images only, we chose to train the model using weights pre-trained on ImageNet [13]. Network pre-trained on a huge and manifold dataset like the ImageNet, captures features like curves and edges in its early stages which are useful to most of the classification problems. Also, this speeds up the training process as model has already learned the elementary features. For both training and testing the batch size is set to 64 with the learning rate of 0.001. The momentum and decay are set to 0.9 and 0.0005, respectively.

Testing: For testing purpose, the helmeted and non-helmeted images have been downloaded from the ImageNet dataset [22]. It is worth to mention that these images were not used in the training dataset at second YOLOv2 model i.e. helmet dataset. This testing dataset consists of total 409 helmeted and 403 non-helmeted images (i.e., person, car, motorcycle, etc.,). According to the norms, one can use 700 helmeted and 100 non-helmeted images, i.e. approximately 20% of training data for testing. However, we use equal amount (almost 50%) of helmeted and non-helmeted data since our model was already aware of faces (transfer learning).

B. Result and Discussion

In this section, we present experimental results obtained using proposed approach. To check the robustness of the algorithm, we tested it on images having both helmeted and non-helmeted motorcyclists. Results obtained using proposed method for different scenarios are displayed in Fig. 2.

As one can see from figure that the helmets are detected accurately in crowded areas and also in the images with single motorcyclist. Additionally, our approach can distinguish between cap and helmet despite both of them having similar features. Also, it can differentiate scarves from helmets which may pass off as helmets in some cases (see first image in Fig. 2). It is important to note that the proposed approach also detects the helmets accurately in the motorcyclists captured from side-view camera. Results for the side-view motorcyclists are displayed in the third row in Fig. 2. The license plate extraction using OpenALPR for non-helmeted images for different environments are displayed in Fig. 3. Furthermore, the helmet detection approach proposed by A. Hirota *et al.* fails when color of helmet is black, same as hair color. However, this limitation is overcome in the proposed method and our method can detect helmets of all colors and shapes. From the results displayed in Fig. 2, one can note that the helmets are detected and the license plates are extracted for non-helmeted images by the proposed method. Thus, our approach is robust and reliable in different scenarios.

The quantitative measures such as accuracy, precision and recall during training phase are evaluated using a testing dataset of 409 helmeted and 403 non-helmeted images and

Fig. 2: Helmet detection in different scenarios.

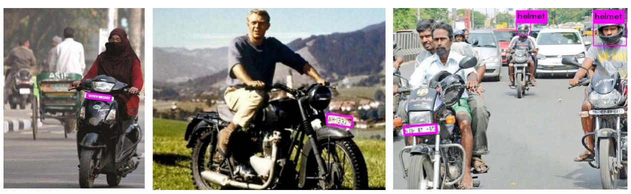


Fig. 3: Detecting licence plates in different scenarios.

TABLE I: Quantitative measures for helmet detection.

Measure	Value
Accuracy	0.9470
Precision	0.9463
Recall	0.9486

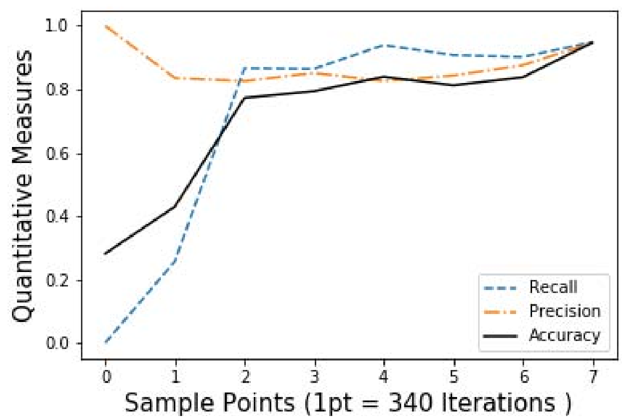


Fig. 4: Quantitative measures during training phase.

same is depicted in Fig. 4. We tested the model at various checkpoints to obtain the results. One can observe in Fig. 4 that starting from 500 iterations, recall and accuracy increase rapidly. However, these values remain steady after 800 iterations. The graph of precision begins with high value since most of the predictions are false in the initial phase of training then it decreases slightly and it remains constant for the most part of the graph. The final values of those quantitative measures are displayed in Table I. One can note that helmet detection accuracy using proposed method is 94.70% which is better when compared to other previous state of the art approaches [9]–[11]. Silva *et al.* [9] obtain helmet detection accuracy of 94.23% using SVM classifier. Authors in [10] use multi-layer perceptron model for classification. They obtain the helmet detection accuracy of 91.37% on their own database of 255 images. Furthermore, Waranusat *et al.* [11] obtain accuracy of 74% by using k-NN classifier over the head region of motorcyclist.

In Fig. 5, we display the graph of average error during training iterations i.e., number of batches completed. Here, one can see that at the beginning of the training, the average error started very high, i.e., approximately 160, and then it reduced exponentially and finally it becomes stagnant, with a very little but significant change, after 500 iterations. We stop our training around 2500 iterations where the error was not changing much. Here, we use a learning rate of 0.001.

Additionally, the graph of average of intersection-over-union (IOU) against number of iterations is depicted in Fig. 6. Here, we use average of IOU over every 150 iterations, and plot the same against the sampled units which is obtained by dividing the number of iterations by 150. Here, the average IOU increases rapidly in the initial stage of training. Then it starts to fluctuate by 2% to 4%, while still generally increasing, and finally reaches to around 80%, after which the error and IOU remains constant. Finally, Fig. 7 shows the results of helmet detection at various phases of training. Fig. 7(a) displays the result obtained when 150 iterations of training for helmet detection are completed. Here, no helmet is detected. When iteration reach to 300 iterations, some helmets are detected as displayed in Fig. 7(b). Interestingly, proposed approach detects all helmets properly when 950 iterations are over as displayed

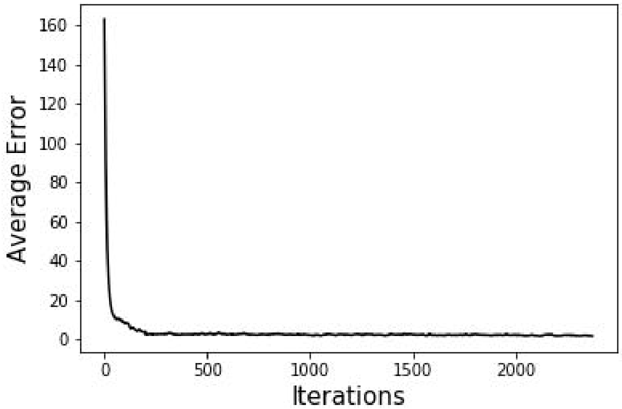


Fig. 5: Average error Vs number of iterations.

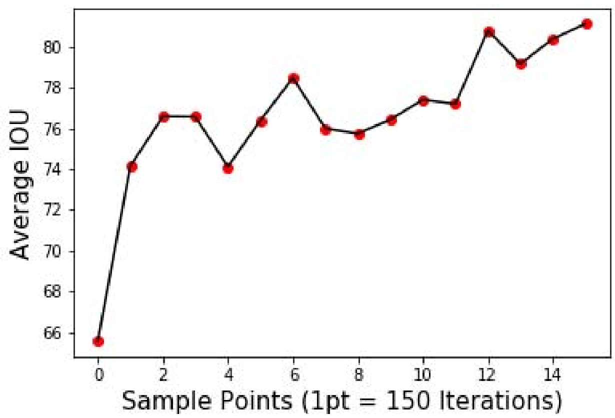
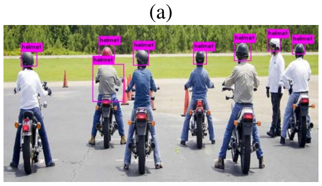


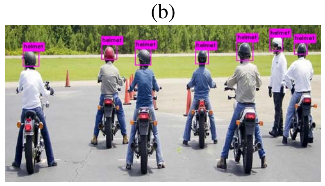
Fig. 6: Average IOU Vs number of iterations.

(a)

(b)



(c)



(d)

Fig. 7: Detection of multiple helmets for different iterations: (a) 150 iterations (b) 300 iterations (c) 950 iterations (d) 2500 iterations.

in Fig. 7(c). However, the objects which are not helmet are also detected as class 'helmet'. Finally, Fig. 7(d) depicts the helmet detection results obtained when 2500 iterations are over which shows that the final weights detect all the helmets correctly without any false positive.

V. Conclusion

We propose an approach for automatic helmet detection using CNN. In order to increase the helmet detection accuracy we use two stage of YOLOv2 models. First YOLOv2 model ensures the person detection which is trained on COCO dataset. This step decreases the number of helmets being undetected. The cropped images of detected person are used as input to the second YOLOv2 model which was trained on our helmeted image dataset. The proposed method has been tested on different helmeted image scenarios. Also, we evaluated the quantitative measures on the test images and same is compared with other state of the art methods. Experimental results and quantitative measures on different scenarios show that our approach is robust and reliable with a high helmet detection accuracy.

Acknowledgement

Authors are gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

References

[1] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," *CoRR*, vol. abs/1612.08242, 2016. [Online]. Available: http://arxiv.org/abs/1612.08242

[2] [Online]. Available: https://scroll.in/article/757365/road-accidents-kill-382-in-india-every-day-1682-times-more-than-terrorism

[3] [Online]. Available: http://timesofindia.indiatimes.com/city/chennai/98-6-of-bikers-who-died-didnt-wear-a-helmet/articleshow/47904790.cms

[4] C.-Y. Wen, S.-H. Chiu, J.-J. Liaw, and C.-P. Lu, "The safety helmet detection for atm's surveillance system via the modified hough transform," in *IEEE 37th Annual 2003 International Carnahan Conference onSecurity Technology, 2003. Proceedings.*, Oct 2003, pp. 364–369.

[5] C.-C. Chiu, C.-Y. Wang, M.-Y. Ku, and Y. B. Lu, "Real time recogniton and tracking system of multiple vehicles," in *2006 IEEE Intelligent Vehicles Symposium*, 2006, pp. 478–483.

[6] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, Nov 1986.

[7] J. Chiverton, "Helmet presence classification with motorcycle detection and tracking," *IET Intelligent Transport Systems*, vol. 6, no. 3, pp. 259–269, September 2012.

[8] Z. Chen, T. J. Ellis, and S. A. Velastin, "Vehicle detection, tracking and classification in urban traffic," *2012 15th International IEEE Conference on Intelligent Transportation Systems*, pp. 951–956, 2012.

[9] R. Silva, K. Aires, T. Santos, K. Abdala, R. Veras, and A. Soares, "Automatic detection of motorcyclists without helmet," in *2013 XXXIX Latin American Computing Conference (CLEI)*, Oct 2013, pp. 1–7.

[10] R. R. V. e. Silva, K. R. T. Aires, and R. d. M. S. Veras, "Helmet detection on motorcyclists using image descriptors and classifiers," in *2014 27th SIBGRAPI Conference on Graphics, Patterns and Images*, Aug 2014, pp. 141–148.

[11] R. Waranusast, N. Bundon, V. Timtong, C. Tangnoi, and P. Pattanathaburt, "Machine vision techniques for motorcycle safety helmet detection," in *2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013)*, Nov 2013, pp. 35–40.

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

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[14] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, vol. 1, June 2005, pp. 886–893 vol. 1.

[15] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004. [Online]. Available: http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94

[16] T. Ahonen, A. Hadid, and M. Pietikäinen, *Face Recognition with Local Binary Patterns*. Springer Berlin Heidelberg, 2004, pp. 469–481.

[17] A. Hirota, N. H. Tiep, L. Van Khanh, and N. Oka, *Classifying Helmeted and Non-helmeted Motorcyclists*. Cham: Springer International Publishing, 2017, pp. 81–86.

[18] T. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B. Girshick, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, “Microsoft COCO: common objects in context,” *CoRR*, vol. abs/1405.0312, 2014. [Online]. Available: http://arxiv.org/abs/1405.0312

[19] GitHub. (2015) Openalpr/openalpr”. [Online]. Available: https://github.com/openalpr/openalpr

[20] J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” *CoRR*, vol. abs/1506.02640, 2015. [Online]. Available: http://arxiv.org/abs/1506.02640

[21] J. Redmon, “Darknet: Open source neural networks in c,” http://pjreddie.com/darknet/, 2013–2016.

[22] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 248–255.