JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Detection of Safety Helmets on Motorcycle Riders Using HIM: Helmet Identification Model

Laksh Goyal

Prabhavati Padamshi Junior International College, Maharashtra, India

Acknowledgment: Adinawa Adjgbodjou, Carnegie Mellon University.

Abstract

Safety helmets are essential to be worn while riding bikes, scooters, etc. for the rider's safety. In case of accidents, helmets reduce the chances of deaths and fatal injuries drastically. Despite this obvious fact there are still civilians who do not wear helmets while riding motorcycles and continue to risk their lives every time they drive. To make sure that people follow the rules of wearing helmets, manual checking takes place by the traffic police which is not at all effective, and maximum riders get away with it. Therefore, this paper proposes a classification model with custom vision that identifies whether a rider is wearing a helmet or not. Our model has a mAP of 80.2% and it has a fair precision for real-time object detection. This is much more cost-effective and impactful and if used correctly; it can act as a solution to tackle this issue of safety.

Keywords: *Safety helmet detection; supervised learning; Object detection; Microsoft Azure; Custom Vision*

1. Introduction

Technology is constantly being outdated; programs are becoming more efficient, and older systems are being made redundant**.** With the current global population of approximately 8 billion, it is obvious that more rudimentary facilities and infrastructure are required to cope with the increasing needs of individuals. While the needs are increasing, the issue of safety is rising as well. The reasons could be a shortage of monetary funds; lack of natural resources like water or even something like transport. Transport is quintessential in times like these where there are networks connecting different parts of the world. There are various different types of transport and rules and regulations to make sure that mobility on vehicles is safe. Despite those rules and regulations, there are various accidents, and with accidents come casualties. There are approximately 2 billion private vehicles, out of which 600 million are 2-wheelers. Out of this drastic number of individuals who ride motorcycles, 130,000 (approx. estimate) are injured with a toll of 1000 deaths annually. One of the primary reasons is the fact that these riders do not wear motorcycle helmets while riding their vehicle, even though there is a mandate for them to do so in a majority of the developed and developing countries of the world. For example, in countries like India where the frequency of bike accidents is the highest in the entire world, the rule for wearing a helmet while operating the vehicle; according to Section 129(b) in Motor Vehicles Amendment Act 2019 [1], isn't followed sincerely despite the quite obvious dangers. Hence measures to impose this rule upon individuals forcefully is something that needs to be done for the safety of the civilians.

According to the Bureau of Police Research and Development in India, there are about 72,000 police officers to manage 200 million vehicles, and 74.4% of these vehicles are two-wheelers. Additionally, traffic instructors are not present on highways hence there would be no supervision on expressways. India has 599 national highways at the moment; meaning 1, 32,500 km of unsupervised distance where the traffic police cannot stop reckless bikers who do not wear helmets. According to data by SaveLife Foundation, India's road crash severity in 2021 was 38.6, increasing from 37.5 in 2020. To add on, there are about 122,000 crashes on highways and 53000 deaths [12]. A US study states that helmets cut the risks of severe traumatic brain injury by half when riders suffer a brain injury. The report, in the American Journal of Surgery, also concluded that riders with helmets were 44% less likely to die from their injury, and 31% less likely to break facial bones. Hence, the absence of supervision results in more deaths.

Furthermore, the traffic police would have to check if the bikers are wearing helmets and the kind of helmets manually, which is more time-consuming. Also, at times the police may not notice the bikers on different sides of the road due to obstructions. Rush hours add to the problem. An automated system would definitely mean a more reliable and effective system that would capture images of traffic and identify bikers who are not wearing helmets, this can be further used to fine these bikers by identifying their vehicle to make sure the rule is followed diligently. Also, an automated system also prevents potential bribery and has access to highways where the traffic police are absent just by installing cameras.

Machine Learning and Artificial intelligence would offer an automated system to overcome the limitations of the current manual system of inspection and an efficient and effective system to make sure individuals follow the rule of wearing helmets. Machine learning here helps to create an algorithm that detects objects better than a human being would by using supervised learning. It is a branch of Artificial intelligence and eliminates almost all human intervention. While it is quite efficient and reduces the chances of error, it provides unbiased decisions and is available throughout the day. It also becomes easier to automatically store all the information about vehicles instead of doing it manually. Therefore, it will be more convenient for the traffic police to supervise the bikers with the help of an automated system, reducing their workload and making sure that everyone is following the traffic rules.

Object detection is the Artificial Intelligence technique being used to identify if the riders are wearing helmets in this paper. The idea is to use object detection to identify riders who do not wear a helmet, further identify the license plate, and fine the individual to encourage them to follow the rules – similar to the concept of how speed cameras helped in making individuals follow the speed limit. There have been various studies previously to find a solution to this problem. Past research papers have used deep learning mechanisms and Neural Networks as well to come up with a classification model to detect bikers wearing helmets. Shanshan Huang, Jianhui Huang and Yongqian Kong have designed an attention mechanism and paired it up with the YOLOv3 using an already existing dataset. This classification system does not detect whether the individuals wearing a helmet are riding a bike or not, which is the problem statement of this paper. The Image classification model being used in this paper is supposed to be implemented throughout the country and widely. Now, numerous ML platforms allow organizations to build secure ML apps with capabilities such as custom machine learning roles, role-based access, virtual networks, and private links. Organizations can also manage governance efficiently with policies, quotas, audit trails, and cost management. This paper will make use of one of these platforms in developing an image classification model of high accuracy for safety helmet detection to be used country-wide in hopes of alleviating this active issue.

This research paper will present the way the classification model works and how it identifies helmets with great accuracy, hence posing a potential solution to the current problem of individuals not realizing the importance of wearing a helmet while riding two-wheelers.

2. Related works

In this section, we will review the related work on object detection and the domain of safety helmet detection to understand what has been done in this field before. There have multiple approaches to making an object detection model to detect helmets and other entities in the safety sector. Presently, models using deep learning are on the up and up, for example, CNN [2] (Convolutional Neural Network) and YOLO (You Only Look Once) are a few of the most frequently used models which have a lot of variations. For CNN, there is R-CNN, Fast R-CNN, and Faster R-CNN. This is because the initial model took a lot of time to process each image which was brought down to a mere 47 seconds in the R-CNN. Although that still was not enough, hence Faster R-CNN tried to fuse CNN with RPN (Region Proposal Network) to solve the issue in the previous versions. Although there are papers that use machine learning classification models as well [3-4].

As witnessed in other papers, deep learning [1, 5-7] is used in real-time detection like crowd count [6-7], vehicle count [8], and other object detection applications. Moreover, there are papers that have used open-source libraries such as TensorFLow and Scikit-Learn and used Keras Functional API to develop multi-input-output networks to detect different characteristics of cows [9]. Also, this paper has underlined the importance of a clean and elaborate dataset and careful selection for the optimal training of the model.

There are also papers that specifically focus on detecting helmets and other safety equipment using alternate approaches. One of these papers divides YOLOv5 into 3 parts [11] which means that this paper uses deep learning and has an innovative approach to make a model which detects construction safety equipment. It has an overall precision of 92%, recall of 91.7%, AP of 92.6%, and F-1 score of 91.9% and it detects multiple entities. It also involves the real-time shooting of constructing sites which further enhances the dataset and the model has a greater variety of data to train from.

A YOLOv3 model is used along with the corporation of Arduino UNO as the main model to detect masks and helmets, as well as recognize voices [7]. Focusing on the detection of masks and helmets, the algorithm first divides the image into a 5x5 grid, and then a particular confidence threshold is set to eliminate the grid with low confidence results to get more precision.

YOLOv3 has been equipped with an attention mechanism in another case of safety helmet detection [10]. The paper does not include a variety of metrics to make a more detailed comparison, instead, it compares the performance capabilities of different object detection methods.

3. Dataset

The dataset for this paper to make the safety helmet detection model has been derived from various search engines like yahoo, Microsoft Bing, and Google. The search was really convenient as the keywords to collect images for this dataset were quite straightforward. Some of the keywords include: ‘bike+rider+helmet’ and ‘bike+traffic’. There were no real compatible datasets available on sources such as GitHub or Kaggle, hence this paper works with a completely new dataset that uses data from across different search engines.

Moreover, using search engines makes the data more reliable as the algorithm ranks the images according to their relevance to the keywords entered in the search bar. Hence, the most relevant data has been selected. The dataset is relatively small for such a complicated detection which also requires high precision, making the speculation and analysis much easier, and the model takes lesser time to be deployed. More importantly, as this is just a smaller scale of the actual model, this dataset is adequate and precise enough to showcase how the actual model should work and the output that it displays. This model consists of a total of 275 images, with two tags: bike and helmet. The helmets and bikes have been accurately selected and tagged in each and every image, and the images are all from a variety of angles with varying quality of the image. There are control images as well, pictures of just bikes and helmets separately, from different angles for a better analysis. The model was initially trained using about 65 images, and the accuracy and mean average precision (mAP) were quite low, hence an increased number of images are used. Existing research papers with object detection models have trained the model using over 1000 images but it would not be feasible in this paper. However, the current dataset has a variety of images hence it covers all the kinds of situations the model could potentially come across.

As the model is eventually going to work with plenty of images which may include images of moving vehicles – the quality of the images taken are going to be really low. There is a greater probability that the images would be blurred. Hence about 30% of this dataset contains images of low quality and images which are blurred to help this model cope up with its future limitations. Also, this dataset has been collated keeping in mind the fact that the images which will be used in the real world will be from a camera attached to the pole of a traffic signal, like that of a speed camera; therefore the selection includes a majority of images from the top view.

Various criteria have been kept in mind while curating the dataset, like the angle from which the traffic is being viewed; the clarity of images; the distance from which the image has been captured, and the number of entities in an image. The following images display the variety of the images in the dataset.

Figure 1: This image is an example of an image captured using a camera places like a speed camera.

Figure 2: This image is an example of a blurred and unclear image to show the possible images captured by the camera of moving traffic.

The dataset was checked multiple times for repeated images and images which were unclear to such an extent that it would act as an anomaly in the dataset. To elaborate further, all the images were of similar dimensions and the same file types to make it earlier for the dataset to be analyzed. The tags did not mark any extra or unnecessary areas other than the required entities. Most importantly, while tagging the helmet, the head of the individual is included hence the model will only identify the helmet if it is being worn by an individual on a bike, to avoid misunderstandings between civilians and bikers.

4. Methodology

The aim of this research paper is to identify whether civilians are wearing helmets while operating their 2-wheeler vehicles. Hence a classification system is used, more specifically an object detection model, an algorithm that leverages deep learning and machine learning (which is being used in this case) to replicate the way humans locate and identify objects. Object detection uses both classification and localization to detect images. It cannot solely depend on either of them because a localization model can detect only one entity per image (multiple entities are possible but it slows down the network), and a classification model does not mark entities and only states whether the object is present in the image without specifying its location.

Figure 3: How the various models differ- Classification sorts out images, while localisation shows where they are located using their coordinates on the image, and how object detection uses both of these techniques

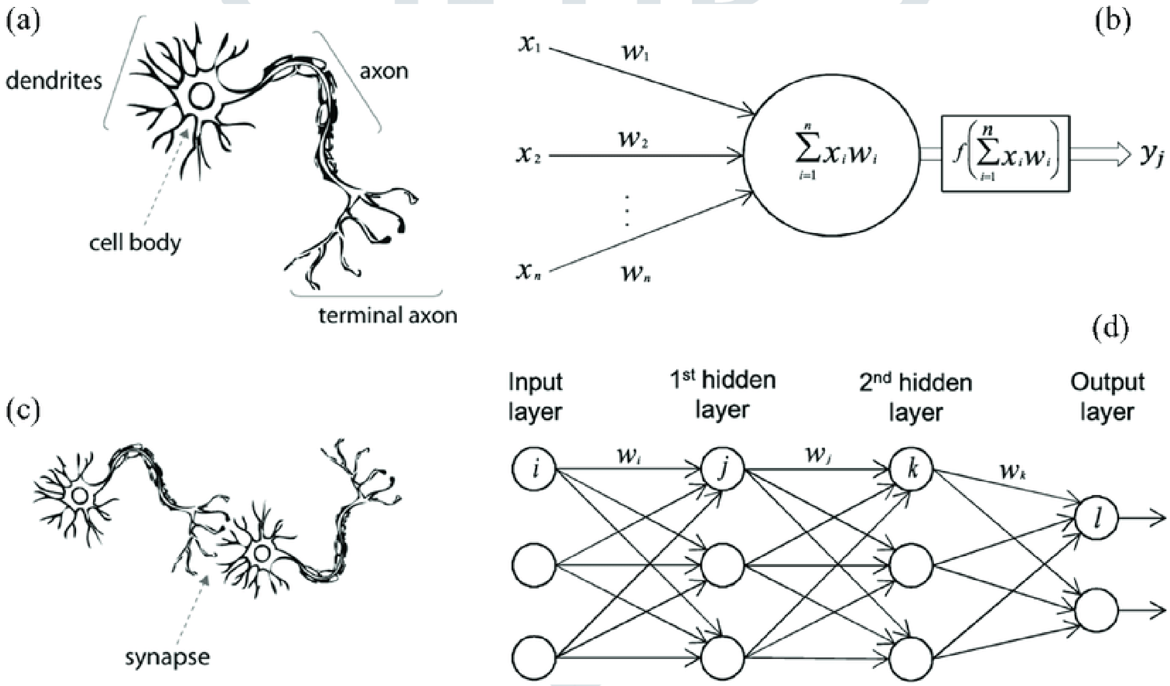

Figure 4: This is how the Artificial Neural Network acts like a biological neuron to imitate human intelligence.

Machine Learning is defined as a computational method that uses experience or past data to improve performance or make predictions. Experience refers to the dataset which is being used and its quantity and quality. On the other hand, Deep Learning is a part of machine le g but it relies on the concept of neural networks (which use neurons in the human brain to send s s in the form of electrical impulses). Therefore, Deep Learning uses Artificial Neural Networks (ANN) to mimic the working of t man brain. An example is the Convolutional Neural Network which is mainly used for object detection techniques.

The axon of the biological neuron represents the output layer. The weights or w_i, w_j, and w_k represent the synapse of the biological neuron. The nodes act as the cell nucleus and the input layer acts as the dendrites.

Machine learning, more specifically supervised learning is used instead of deep learning because deep learning requires a powerful GPU (graphics processing unit) and a lot more labelled training images, which was not in the scope of this paper. Although, due to the absence of GPU and an extremely large dataset, supervised learning is a better method to move forward with. Mo er, it is easy to use although it requires a lot more human intervention than deep learning does. However, in this case, that does not ke a difference at all as the model in this paper is made on a small scale with a smaller dataset than the model which will actually be used. Additionally, this model is vely cheaper to run than a deep learning model hence it will be accepted and implemented faster. More importantly, deep learning works well with only large amounts of data as it had to replicate the way the human brain thinks hence machine learning is again a better choice.

Machine Learning depends on three types of learning which are: supervised, unsupervised, and reinforcement. Object detection uses supervised learning where the dataset is decided and labelled so it can identify the labels in other images it sees. Unsupervised learning searches for trends that are not outlined in random data and can use it to predict future readings. Reinforcement uses trial and error where the machine is told when it has made the right decision.

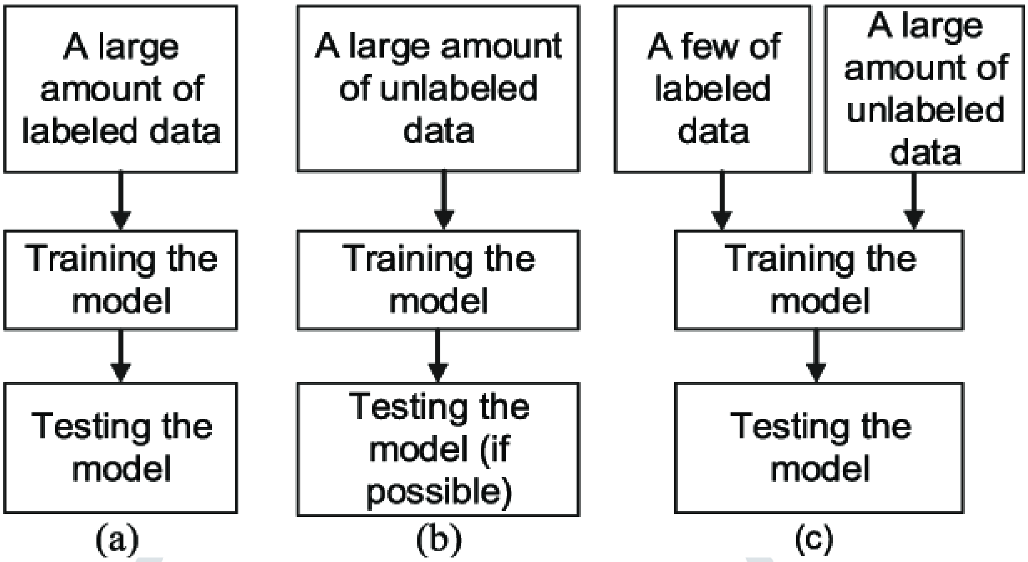

Figure 5: (a) Supervised learning, (b) Unsupervised learning, (c) Reinforcement learning

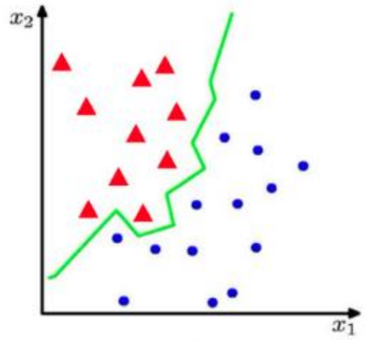

Figure 6: An example of how a graph produced by a classification model looks like. A training set of N observations {(x(i), y(i)) : i = 1,...,N}, where x ∈Rd and y ∈{−1,1}.

As mentioned before, this model uses supervised learning, more specifically a classification model using supervised learning. A classification model basically uses data which is labelled to mark data points, then it takes data and assigns it to a specific class according to the previous data used to train the model. To do so, the model will find the characteristics that correlate to a class and use a mapping function to identify whether the object is present or not.

This research paper uses Microsoft Azure's Custom Vision tool to make an object detection model where the platform has its pre-built machine learning algorithm which is used to train the model. This is a cognitive service on the platform Microsoft Azure where one can apply labels to the visual characteristics of images to train the model to detect that characteristic. After the images are labelled, the algorithm analyses the labels and the images to train the model. The practical benefit of the model is the fact that it can optimally work with small datasets as well, which allow the model to fit the scope of this paper.

While creating the model, one has to fill these criteria out. The project type determines the kind of model the algorithm works on; object detection is chosen because the model needs to be able to detect which individual is wearing the helmet. A pedestrian does not need to wear a helmet hence the localisation ability plays a great role here. Here, the 'General (compact) domain' is selected because helmets do not fit any other criteria like 'Logo'; other iterations have used 'General domain' which involve slightly different processing. Next, I made two tags: bike and helmet and began uploading all the images from the self-curated dataset. After uploading all the images included in the dataset, it is uploaded to the algorithm and certain images which cannot provide any value to the training are discarded. I started selecting the bikes and helmets in each picture and assigned them their respective labels; I made sure there were certain images which would act as control data (Data which did not include either tags). The purpose of the 'bike' tag was to identify whether a bike was in the image to decide whether a 'bike rider' is wearing a helmet or not, and that a pedestrian is not wrongfully identified.

Furthermore, this model works by identifying the entities by marking red boxes around them and displaying the confidence rate or probability that the model has whether the marked regions are correct. The greater the confidence, the more certain the model is. Moreover, it is virtually impossible to get a 100.0% confidence rate unless it is a case of overfitting (when one tests the model using training data); an accuracy of 100.0% is not possible in the real world.

This shows the results when the model goes through the quick train procedure. The image selected is clicked by the author in the real world, in India. The probability threshold is set as 85.0% for valid reasons as a model should have an accuracy of at least 70.0%. Over the course of making the model, it has been noticed that a confidence rate or probability of less than 85.0% would be a false positive hence with a lot of testing a probability of threshold of 89.0% is selected. As we can see, the model has predicted the entities by marking 2 red boxes which correctly identifies the bikes with a confidence rate of 99.6% for one of them - which is commendable as the testing image is one of mediocre quality and is taken in the real world, as mentioned earlier.

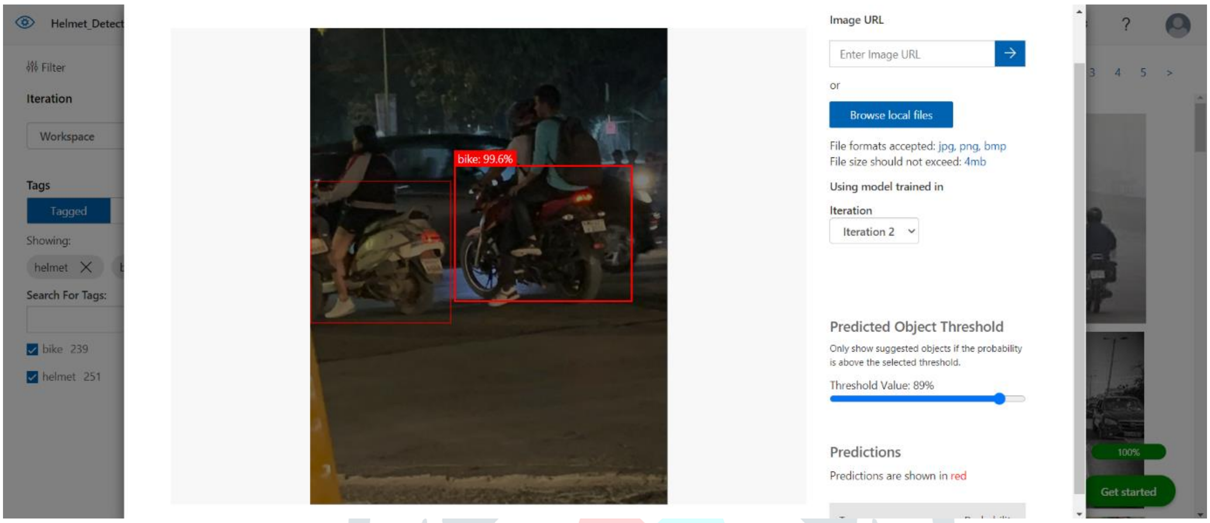

Figure 7: The quick train procedure

This model can be deployed and exported as an application or even published to be used by an organisation. This platform can train up to 10,000 images and detect a maximum of 50 labels. Also, the fact that a model of such accuracy can be created using a platform with its own machine learning algorithm is quite cost-effective and saves a lot of time. It can also be adapted to work with thousands of images at once to satisfy the requirements of the safety helmet detection project.

The idea is that cameras installed at busy junctions will click pictures of the vehicles at a certain time interval and these pictures will go through Image Pre-processing (resizing, colour corrections, etc.) and other correction processes to make the entities in the pictures clearer to the model. Also these pictures will go through image segmentation to separate vehicles so there is no overlapping and it becomes easier to detect small entities like helmets. Then the object detection model will run these images through the algorithm and identify any entities it recognises. This will determine whether the biker is wearing a helmet or not, which is the aim of making this model.

5. Results

It is important to understand the way Microsoft Azure analyses a model and rates it using a variety of different metrics, some of which include:

- Precision: percentage of identified classifications that were correct. Formula to calculate precision: Precision = TruePositives / (TruePositives + FalsePositives). For example, if the model identified 100 images as dogs, and 99 of them were actually of dogs, then the precision is 99 percent.

$$Precision = \frac{TP}{TP + FP}$$

- Recall: the percentage of actual classifications that were correctly identified. Formula to calculate recall: Recall = TruePositives / (TruePositives + FalseNegatives). For example, if there were actually 100 images of apples, and the model identified 80 as apples, the recall is 80 percent.

$$Recall = \frac{TP}{TP + FN}$$

- Mean average precision (mAP): the average value of the average precision (AP). AP is the area under the precision/recall curve (precision plotted against recall for each prediction made). The formula is: $mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i$

Here is a matrix in which the predictions made by the model can fit into. There are 4 classifications: True Positive, False Positive, True Negative, and False Negative. A True Positive is when the model correctly predicts the positive class (in this case it is the helmet). A False Positive is when the model predicts the positive class but the prediction is incorrect. A True Negative is the correct prediction of the negative class (No helmet) and False Negative is an incorrect prediction of the negative class which means that in reality there is a helmet but the model failed to detect it.

Figure 8: This table shows the matrix of the predictions in the case of helmets and the 4 types of cases that would occur when the model makes a prediction.

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Prediction</td></tr>
<tr><td>Positive (Helmet)</td><td>Negative (No helmet)</td></tr>
<tr><td rowspan="2">Actual</td><td>Positive (Helmet)</td><td>True Positive (Helmet is present in both the prediction and reality)</td><td>False Negative (No Helmet in prediction but the helmet is present in reality)</td></tr>
<tr><td>Negative (No Helmet)</td><td>False Positive (Helmet is present in prediction but helmet is not present in reality)</td><td>True Negative(Helmet is not present in both the prediction and reality)</td></tr>
</table>

I have used the same dataset to make 3 iterations of the model. Iteration1 has the project domain set as ‘General [A1]’ and the training time the model took was 18 hours. Iteration2 has the project domain set as ‘General [A1]’ as well but the training budget is 7 hours. Iteration3 has the project domain set as ‘General (compact)’ with a training budget of 24 hours.

Moreover, there are 2 other metrics that determine the precision, recall and mAP of the model. Probability threshold and Overlap threshold are 2 metrics that can be controlled manually to alter your model’s predictions. Probability threshold refers to the minimum confidence rate or probability score to make the prediction valid; this heavily determines the recall and precision. It is set as 90% because recall is as important as precision in this domain. It is better that the helmet is detected than for it to be incorrectly not detected as it is only a problem if the individual is not wearing a helmet, and a higher recall will decrease the chances of that happening . Overlap threshold refers to the localisation aspect of the model; it refers to the minimum percentage of overlap between the region marked in the prediction and the ‘ground truth’ boxes (actual regions marked according to the training and testing dataset). It is 50% in this case as it is very important that the prediction box matches the ground truth box as closely as possible. The localisation factor plays a great role here.

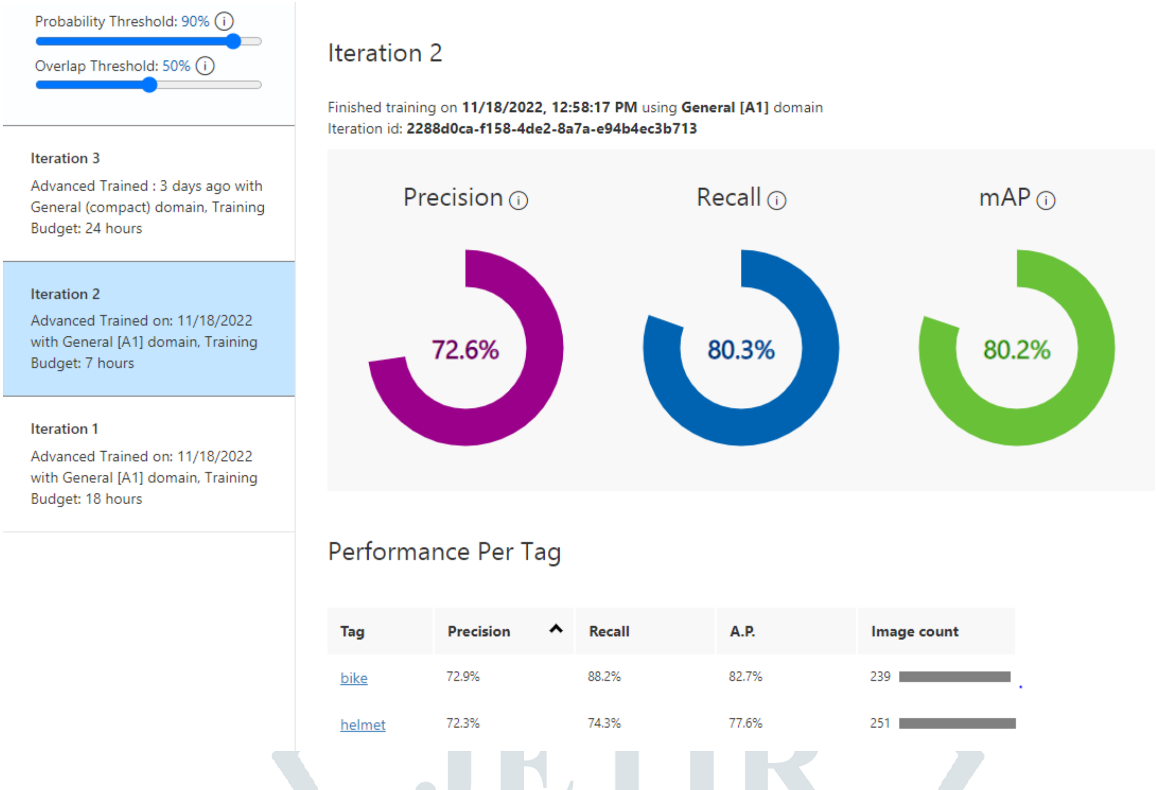

Figure 9: Iteration1 metrics

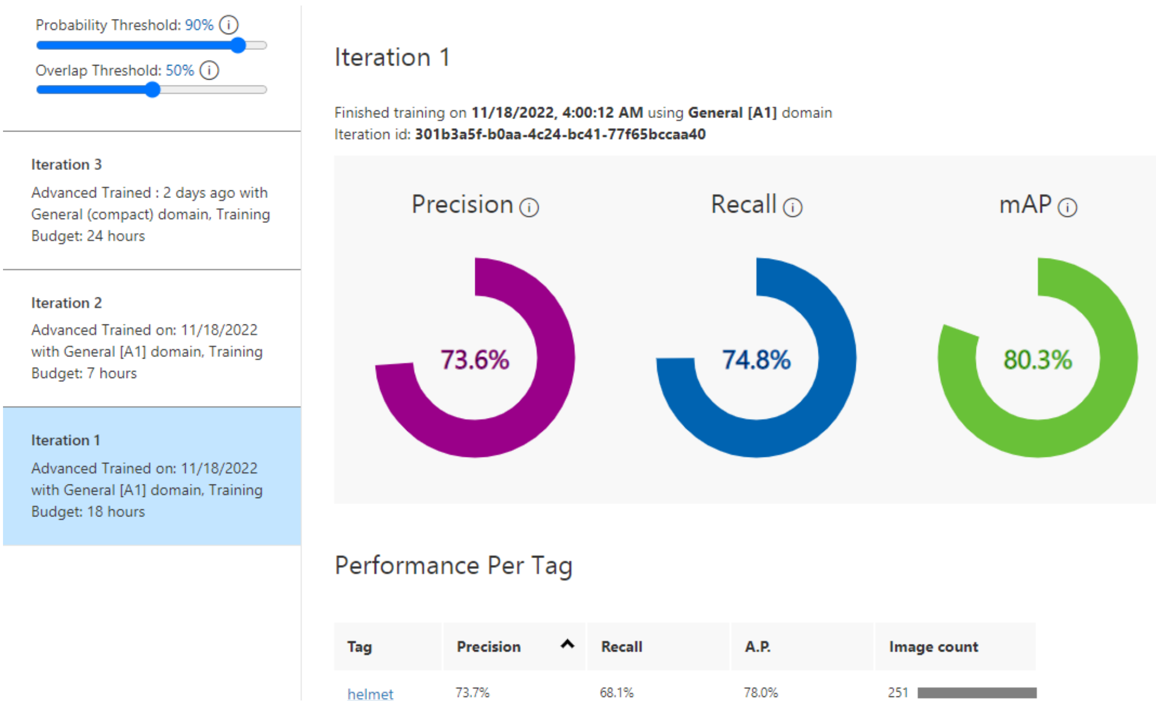

Figure 10: Iteration2 metrics

Iteration1 has a precision of 73.6%, recall of 74.8% and mAP of 80.3% with a probability threshold of 90% and overlap threshold of 50%. The mAP depends on the precision and recall. Moreover, the mAP is used as a standard metric to analyse the accuracy of an object detection model [14]. The higher the mAP is, the more accurate the model is. 80.3% is a fair mAP for a dataset of this size and for the training budget.

This iteration has the same thresholds with a precision of 72.6%, recall of 80.3% and mAP of 80.2%. It is different from Iteration1 by having a better recall rate while having a lesser but still similar precision value without sacrificing the mAP.

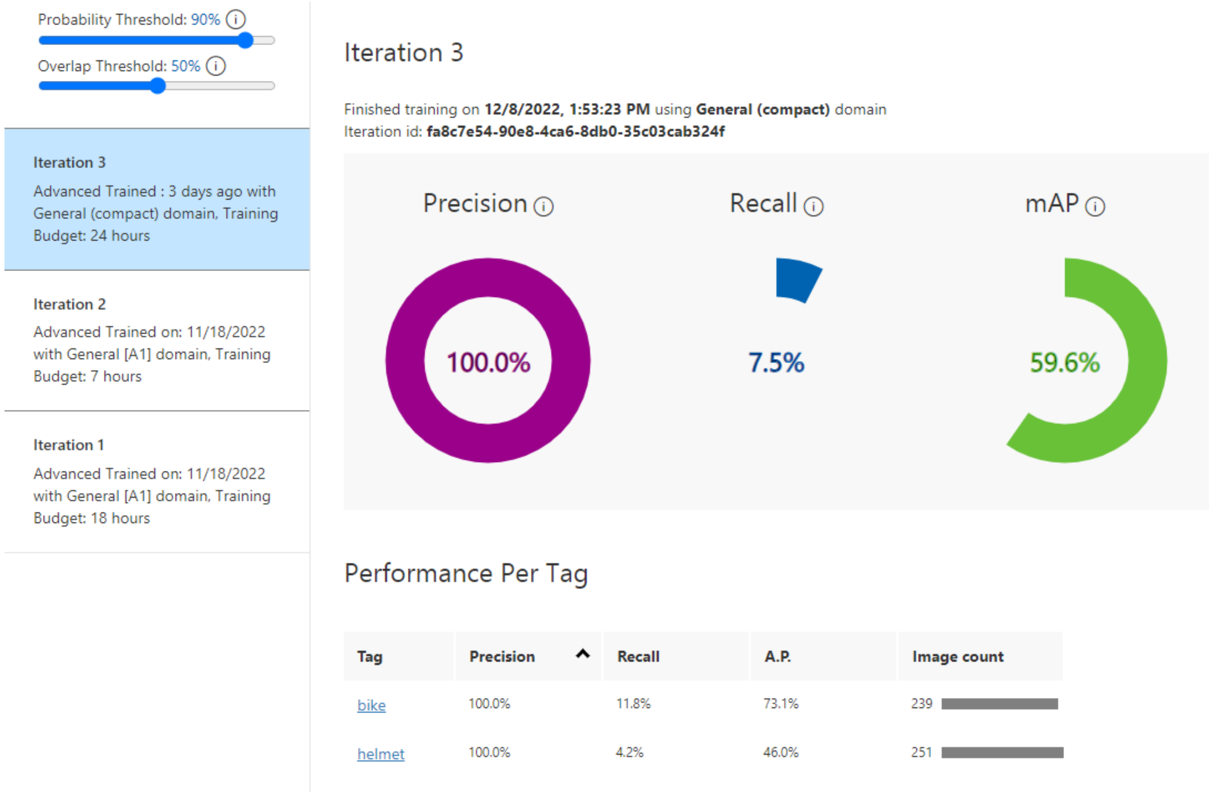

Figure 11: Iteration3 metrics

This iteration stands out because according to this threshold, the precision is 100% which is excellent for an ML model but it has a risk of overfitting. Although, the recall is terrible which makes it an invalid iteration to go through with. Also, the mAP is 59.6% which is inferior to the previous iterations as well.

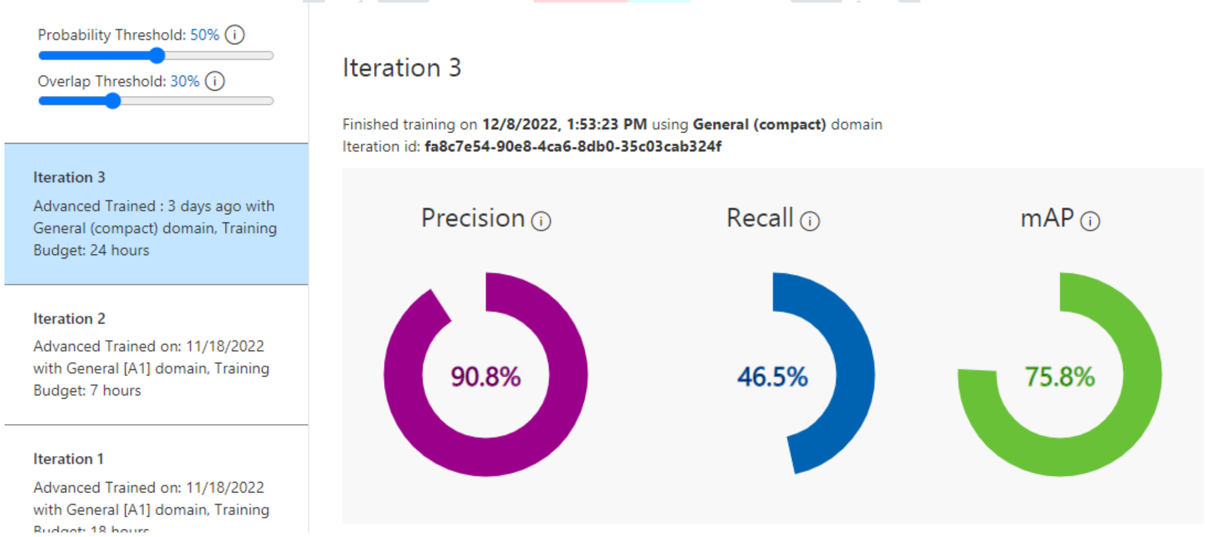

Figure 12: Iteration3 with a probability threshold of 50% and overlap threshold of 30%.

This displays that as the probability threshold and overlap threshold is decreased, the precision decreases slightly to give a drastic increase in the recall. The reduction in the overlap threshold gives a drastic increase in mAP as well. Although, reducing the threshold accepts a lot of false positives where the confidence rate or probability is relatively low than that of a true positive prediction.

Figure 13: Quick testing using an image with Iteration2. Despite the quality and the distance from which the image has been taken, the model has still managed to detect majority of the bikes and helmets with a confidence rate greater than that of 90.0%.

The best iteration to go forward with is definitely Iteration2. Iteration2 has the best combination of precision and it prioritizes recall over precision without affecting the mean average precision. Although, more iterations can be explored by making variations in the domains, training time budget and other settings which determine the metrics of the model. We will refer to Iteration2 as 'Helmet Identification Model' or 'HIM'.

6. Discussion

With the aim of this model being detecting helmets, HIM has good precision and recall to be used as a prototype for real-time detection to at least test how accurate the model will be and the impact it will have on making the public follow the Motor Vehicles Act. Even though it does not have 100.0% precision (which is virtually impossible despite of a case of overfitting), the frequency of bikes passing junctions will compensate for the bikers who don't wear helmets and don't get captured as well. There is a high probability of a biker passing busy junctions multiple times in a day so it increases the probability of the biker being identified automatically.

As the model also detects bikes, this will be able to eliminate pedestrians and other entities from consideration. Image segmentation and pre-processing also help to separate the different entities and then the model can detect bikes and filter the dataset to only identify helmets in those images. This model has used images where the label boxes only the head of the individual who is wearing a helmet and not empty helmets, hence this prevents HIM to detect helmets that may be just kept on the bike or held in the biker's hand.

This model specifically targets motorcycle riders and assists the Motor Vehicles Act 1988 [1]; this will lead to an increase in awareness and motivation for the riders to wear helmets and decrease the chances of deaths due to bike accidents.

7. Limitations and Future Work

As discussed previously, HIM's sole purpose is to detect whether a biker is wearing a helmet or not and this can be adapted into a non-committed fine system where the bikers not wearing a helmet can be identified using their vehicle license plate number and fined at the end of the month. This works similarly to the speed fine system but it is solely for bikers to encourage them to follow the rules. The Azure model has certain limitations which will need to be taken care of if a fining system is going to be implemented. HIM may predict a turban as a helmet due to the shape of turbans and because it is considered as a type of headgear.

Figure 14: Quick testing of an image which includes an individual with a turban

Even though the confidence rate is 79.3% of the helmet with this image the probability threshold can be adjusted but some images may have lower or higher confidence rate according to the angle and clarity of the image. Moreover, according to the Motor Vehicles Act, Sikhs (people who wear turbans by religion) are exempted from this law hence the model has to be trained to identify turbans as well which could help exempt them from the fining procedure.

Another limitation is that a bicycle may be identified as a bike as well. Hence, if a cyclist is not wearing a helmet, he or she may be fined as well (even though there is no way of identification). Even though it is encouraged for cyclists to wear helmets, it is not a legal requirement. So HIM needs to be trained to identify cycles as well which will require zoomed in images which show distinct comparison between bikes and cycles from a variety of angles with varying quality.

Figure 15: Quick testing of an image which includes an individual on a cycle.

The idea this research paper proposes is that the pictures clicked at busy junctions will be sent to be processed and then analysed by the model to detect whether bikers specifically have worn helmets or not, and if they haven't, these images will further be analysed to detect their license plate numbers and identify the vehicle owners. At the end of each month, they will receive a bill in the name of their vehicle fining them a certain amount every time they have been caught without wearing a helmet. This seems like a fair penalty for those who sacrifice their safety and others' safety on roads due to their own carelessness.

For HIM to be implemented, it needs to be trained using a more elaborate dataset where 2 more tags- bicycle and turbans- need to be added. To identify the vehicle and fine the rider, the model needs to first classify images where there are bikes detected but the rider is not wearing a helmet. Then these images need to go through a model is able to closely identify license plates and use OCR to read the number plate. This data can be saved in a database and these vehicle IDs can then be fined.

Furthermore, there can be a variety of iterations made for different regions according to the frequency of bikes that are present there. For regions with a higher number of bikes, a greater expense can be allotted to make an object detection model which works with images of higher clarity. A problem that can be encountered is that bikers not wearing helmets may cross multiple junctions and get captured multiple times, ending up getting fined multiple times as well. Therefore, the database needs to be configured in such a way that a particular vehicle license plate number is only recorded once a day.

There can be certain reservations the public may have that there are other reasons for the great number of accidents and that they may not agree with the penalty and argue against the sudden change even though the law existed. This is something that needs to be dealt with using awareness campaigns to encourage the public further to follow this rule or bear the consequences.

8. Conclusion

In this paper, a safety helmet detection model is proposed to give a potential solution to a safety concern that will just rise as the population's transportation need grows. There have been various research papers that use a different model and different branches of AI to tackle the same problem or a similar issue using object detection, so this paper intents to provide a model with a real-life implementation so that the number of deaths due to accidents reduce and the citizens understand the importance of following the rules for their own good.

Instead of using deep learning, this paper uses supervised learning and Microsoft Azure as the platform to build a custom vision classification model. As outlined previously, this model is more cost-effective and convenient to deploy plus the results that the model gave are fairly decent for real-time detection.

In the future, a more precise model can be deployed with a larger dataset by exploring the different domains and other criteria. The aim is for it to be put to use in real-time detection where installed cameras at busy junctions capture images to check whether the riders are wearing helmets and fining them at the end of the month for how many ever days they have been caught without wearing a helmet. While this system automates the process, police intervention can also help in encouraging wearing helmets.

References

[1] Motor Vehicles Act 1988 - https://www.bankbazaar.com/driving-licence/motor-vehicles-act.html

[2] Usha Mittal, Sonal Srivastava, and Priyanka Chawla. 2019. Review of different techniques for object detection using deep learning. In Proceedings of the Third International Conference on Advanced Informatics for Computing Research (ICAICR '19). Association for Computing Machinery, New York, NY, USA, Article 46, 1–8. https://doi.org/10.1145/3339311.3339357

[3] Chisulo Mukabe, Nalina Suresh, Valerians Hashiyana, Titus Haiduwa, and William Sverdlik. 2022. Object Detection and Classification Using Machine Learning Techniques: A Comparison of Haar Cascades and Neural Networks. In Proceedings of

the International Conference on Data Science, Machine Learning and Artificial Intelligence (DSMLAI '21'). Association for Computing Machinery, New York, NY, USA, 86–97. https://doi.org/10.1145/3484824.3484895

[4] Willy K. Portier, Yujian Li, and Bonzou A. Kouassi. 2020. Feature Selection using Machine Learning Techniques Based on Search Engine Parameters. In Proceedings of the 2020 3rd International Conference on Signal Processing and Machine Learning (SPML 2020). Association for Computing Machinery, New York, NY, USA, 28–34. https://doi.org/10.1145/3432291.3432308

[5] Chenqiang Gao, Pei Li, Yajun Zhang, Jiang Liu, and Lan Wang. 2016. People counting based on head detection combining Adaboost and CNN in crowded surveillance environment. Neurocomput. 208, C (October 2016), 108–116. https://doi.org/10.1016/j.neucom.2016.01.097

[6] Yi-Le Yang and Wei-Wei Gao (2019). A Method Of Pedestrian Counting Based On Deep Learning.

[7] Huaming Zhou, Aili Wang, Meixin Li, Yinghui Zhao, and Yuji Iwahori. 2022. Epidemic Prevention System Based on Voice Recognition Combined with Intelligent Recognition of Mask and Helmet. In 2021 3rd International Conference on Video, Signal and Image Processing (VSIP 2021). Association for Computing Machinery, New York, NY, USA, 8–15. https://doi.org/10.1145/3503961.3503963

[8] Allach Samir et al (2018). -Faster RCNN was used for the detection of vehicles followed by a CNN based on the atmospheric diffusion model.

[9] Mobasshir Bhuiya Shagor, Md. Zahidul Haque Alvi, and Khandaker Tabin Hasan. 2022. CID: Cow Images Dataset for Regression and Classification. In Proceedings of the 2nd International Conference on Computing Advancements (ICCA '22). Association for Computing Machinery, New York, NY, USA, 450–455. https://doi.org/10.1145/3542954.3543018

[10] Shanshan Huang, Jianhui Huang, and Yongqiang Kong. 2021. Attention Guided YOLOv3 for Wearing Safety Helmet Detection. In 2020 the 6th International Conference on Communication and Information Processing (ICCIP 2020). Association for Computing Machinery, New York, NY, USA, 65–69. https://doi.org/10.1145/3442555.3442566

[11] Weihong Jiang, Changzhen Qiu, Chunze Li, Dengxiang Li, Weibiao Chen, Zhiyong Zhang, Luping Wang, and Liang Wang. 2022. Construction site safety detection based on object detection with channel-wise attention. In 2021 The 5th International Conference on Video and Image Processing (ICVIP 2021). Association for Computing Machinery, New York, NY, USA, 85–91. https://doi.org/10.1145/3511176.3511190

[12] Road crash severity statistics: https://timesofindia.indiatimes.com/india/why-indian-roads-are-deadly/articleshow/94022723.cms#:~:text=According%20to%20data%20by%20SaveLife,death%20in%20a%20road%20crash.

[13] Accidents and deaths due to bike accidents: https://www.cdc.gov/transportationsafety/bicycle/index.html

[14] mAP is the standardised metric : https://towardsdatascience.com/on-object-detection-metrics-with-worked-example-216f173ed31e