Detection of Helmetless Bike Riders Using Deep Learning Techniques

Rathnayake, C.†, Kokul, T.§, Ramanan, A.†

†Department of Computer Science, University of Jaffna, Sri Lanka

§Department of Physical Science, University of Vavuniya, Sri Lanka

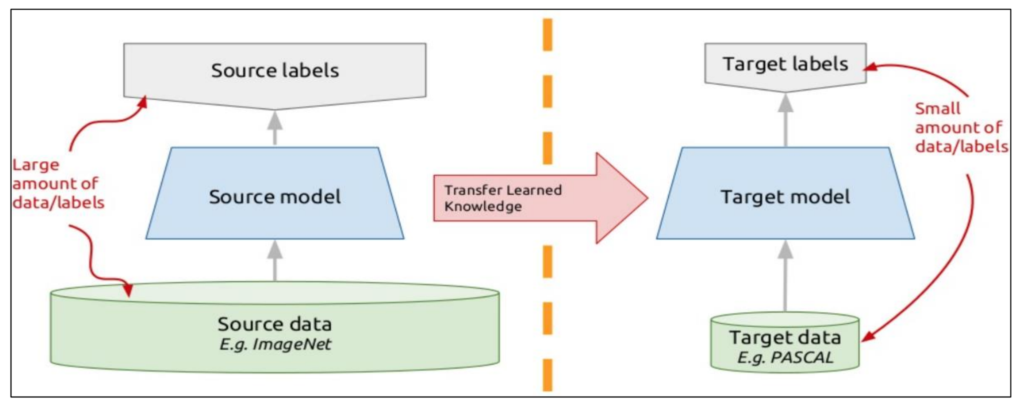
chamod94715@gmail.com, kokul@vau.ac.lk, a.ramanan@univ.jfn.ac.lk

INTRODUCTION

- Motorbike accidents are rapidly growing throughout the years in many countries. Wearing a helmet is the most effective way to reduce head injuries and fatalities from motorbike crashes, but many bikers don't use it.
- Monitoring video surveillance by humans proves ineffective as the duration of monitoring of videos increases, the errors made by humans also increases.
- In this study, we have proposed a deep learning based approach to identify whether a biker is wearing a helmet or not. The deep transfer learning technique is utilized in this approach to classify the helmet images from other images.

BACKGROUND

- Deep transfer learning techniques address the data deficiency limitation by transferring the acquired knowledge from one problem domain to another.
- Such pre-trained models are trained on large benchmark datasets.
- We used the pre-trained neural network as a feature extractor by discarding the fully connected layers and place a new fully connected layer architecture according to the application.
- Furthermore, we fine-tuned some last layers of the pre-trained model based on the availability of training samples since the early convolutional layers of a pre-trained model extracts more generic features and, deeper convolutional layers extract more application-oriented or class-specific features.



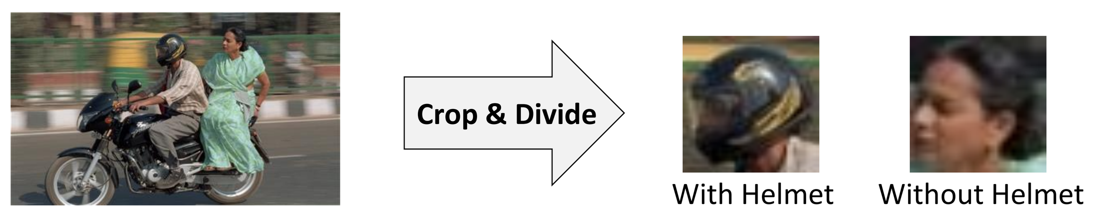
An illustration of Deep Transfer Learning

DATASET

- We used a helmet detection dataset [4] which consists of 764 images from kaggle.com.
- Resolution: 400x210 pixels to 400x400 pixels.

➢ Dataset Preparation

- The images were cropped using annotated bounding boxes and divided into two classes automatically using Python code.



- Total images after cropping: 1428 and images were resized into 224 x 224.
- Train/Validation/Test datasets were selected randomly.

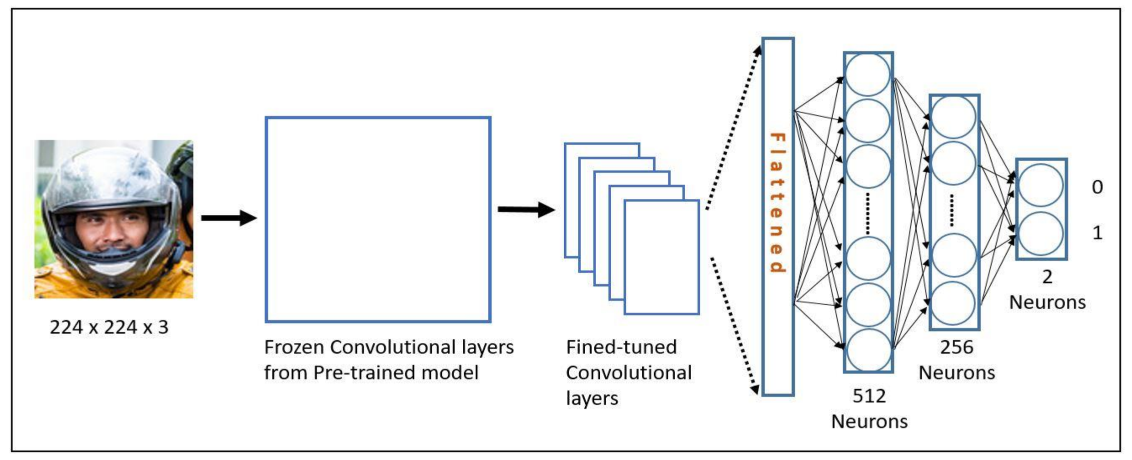
Summary of the Dataset

	With Helmet (Class-0)	Without Helmet (Class-1)	Total
Train	605	308	913
Validation	152	77	229
Test	190	96	286

Samples of the cropped images

METHODOLOGY

- The proposed architecture consists of two main components: The convolutional layers and three fully connected layers.
- The convolutional layers were obtained from the pre-trained CNN models, which were trained on the ImageNet dataset.
- The number of fully connected layers, number of neurons in an FC layer, and dropout values were set experimentally.
- Input size: 224 x 224 x3.



The proposed Network Architecture

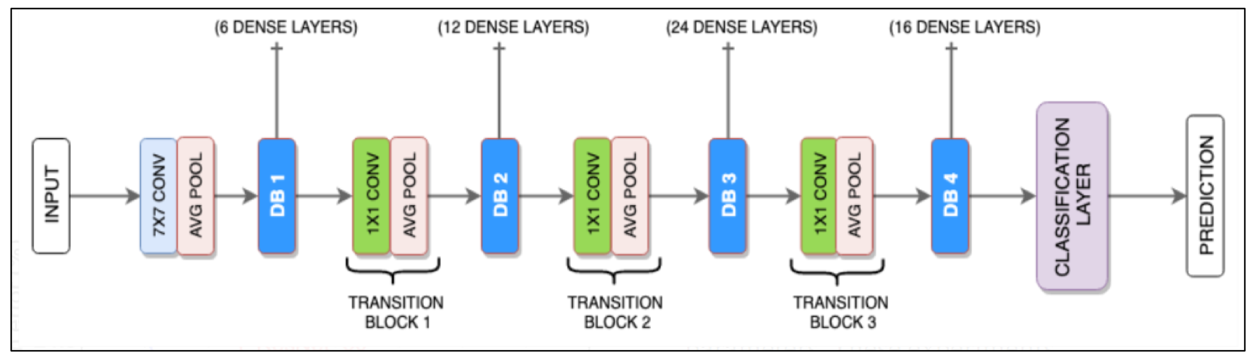
- The hyperparameters were fine-tuned, such as learning rate, batch size, dropout value, optimizer, number of epochs, etc.
- We have performed data augmentation to increase the number of training data. Transform operations: rotation, zoom, width shift, height shift, shear, and horizontal flip.
- We have set the class weights to handle the data imbalance.
- Several pre-trained models were used to evaluate the proposed approach: VGG16, VGG19, EfficientNet, MobileNetV2, and DenseNet121.
- The last few convolutional blocks were fine-tuned in these models.

➢ VGG16 & VGG19 are innovative object-recognition models which conclude with incorporate multiple non-linear rectification layers instead of a single rectification layer.

➢ EfficientNet is a convolutional neural network architecture which follows a uniform scaling technique to scale all dimensions of the network (depth/width/resolution) using a compound coefficient.

➢ MobileNetV2 is a lightweight convolutional neural network architecture that introduced the inverted residual structure where the residual connections are between the bottleneck layers.

➢ DenseNet121 is an architecture that focuses on making the deep learning networks go even deeper, but at the same time making them more efficient to train, by using shorter connections between the layers.



DenseNet121 Architecture

EXPERIMENTAL SETUP

- The experiments were done separately by several structures based on: VGG16 | VGG19 | EfficientNet | MobileNetV2 | DenseNet121
- Optimizer: Adam | Batch Size: 16 | Number of epochs: 100
- Learning Rate: 0.001 | Dropout: 0.7, 0.5, 0.2
- Class Weights: {0:0.001057,1:0.002074}
- Loss Function: Binary Cross-entropy Loss
- Activation Function: ReLU & Softmax
- Language: Python | Framework: Keras-Tensorflow | IDE: Google Colab

RESULTS

- Classification accuracy based on a single run is used to evaluate the performance of the proposed approach.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

TP - True Positive
TN - True Negative
FP – False Positive
FN – False Negative

- The last few convolutional blocks were fine-tuned according to the experimental results.

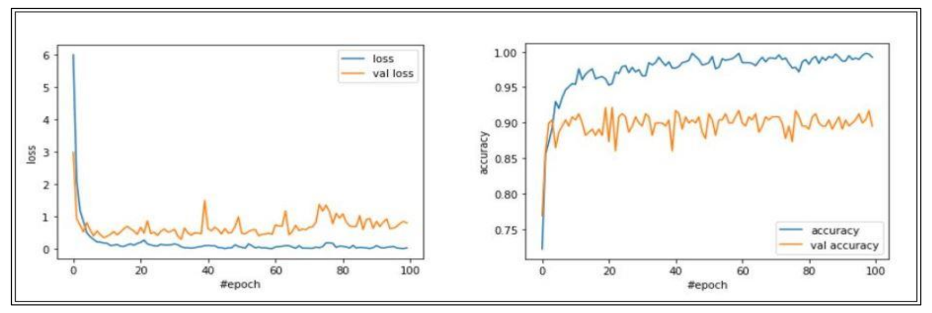
Comparison of fine-tuned convolutional layers for pre-trained models

No. of fine-tuned Convolutional Blocks	Classification Accuracy (%)				
	VGG16	VGG19	EfficientNet	MobileNetV2	DenseNet121
0	**89.81**	**83.46**	**68.13**	94.50	93.98
1	80.66	50.00	51.03	**94.77**	95.31
2	50.00	50.00	50.00	93.73	95.56
3	50.00	-	50.00	93.72	94.00
4	-	-	-	92.69	93.73
5	-	-	-	-	94.51
6	-	-	-	-	**96.34**
7	-	-	-	-	95.04
8	-	-	-	-	94.52
9	-	-	-	-	94.23

The best classification accuracies of the fine-tuned pre-trained models

Model	Accuracy (%)
EfficientNet	68.13
VGG19	83.46
VGG16	89.81
MobileNetV2	**94.77**
DenseNet121	**96.34**

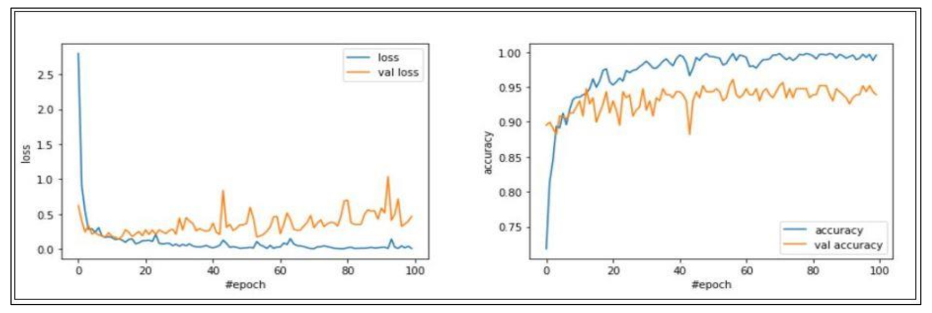
- MobileNetV2 & DenseNet121 show better performance



The loss and the accuracy of MobileNetV2 on the training and validation datasets over training epochs

The confusion Matrix for MobileNetV2

		Predicted Classes	
		With Helmet	Without Helmet
Actual Classes	With Helmet	180	10
	Without Helmet	05	91



The loss and the accuracy of DenseNet121 on the training and validation datasets over training epochs

The confusion Matrix for DenseNet121

		Predicted Classes	
		With Helmet	Without Helmet
Actual Classes	With Helmet	186	04
	Without Helmet	05	91

- We have compared the accuracy of the proposed approach with the reported results of some similar approaches. It has been noted that different authors used their own datasets.

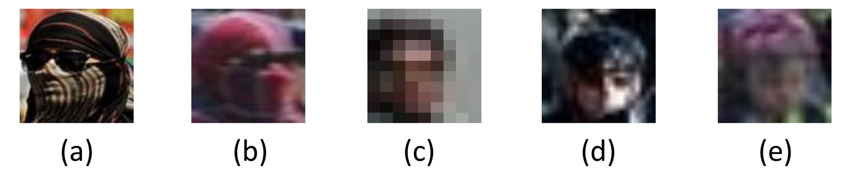
Comparison with similar approaches

Approach	No. of Images	Accuracy (%)
Vishnu *et al.*, (2017) [1]	3573	87.11
Boonsirisumpun *et al.*, (2018) [2]	493	85.40
Maliye *et al.*, (2021) [3]	4000	90.84
Ours	1428	96.34

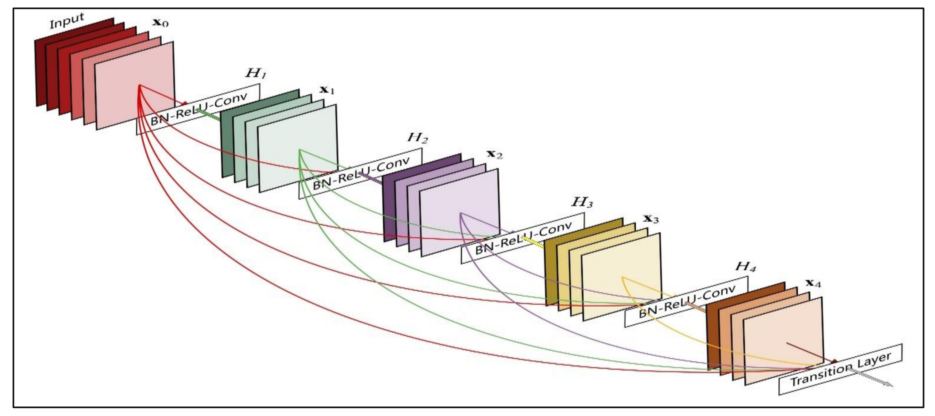
DISCUSSION

A few misclassifications occurs in the following cases:

- Some bikers wear a cap or cover their head with a scarf or other clothing
- Sometimes, the colour of the rider's head and the helmet are similar
- Some images lose their information due to the low resolution



Samples (a), (b), (c) predicted as "with helmet" and (d), (e) predicted as "without helmet".



A 5-layer dense block. Each layer takes all preceding feature-maps as input

➢ The DenseNet121 was able to achieve the highest classification accuracy. Each layer in a dense block receives feature maps from all the preceding layers and passes its output to all subsequent layers. These forms of connections allow better gradient flow.

➢ Each layer has direct access to the gradients of the loss function and the original input signal, thus helps to alleviate the vanishing gradient problem.

➢ Dense connections also encourage feature re-use and reduce the number of parameters since it's re-using previous feature-maps information instead of generating more parameters.

CONCLUSION

- The proposed deep learning-based approach showed better performance against recently reported works in the literature to identify whether a biker wearing a helmet or not.
- We conducted a comprehensive study to achieve better classification results using deep convolutional neural networks for tiny dataset with low resolution images.
- The performance of transfer learning was measured by verifying the number of fine-tuned convolutional blocks.

REFERENCES

1. Vishnu, C., Singh, D., Mohan, C.K. and Babu, S., "Detection of motorcyclists without helmet in videos using convolutional neural network." In *2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 3036-3041. IEEE, 2017.
2. Boonsirisumpun, N., Puarungroj, W. and Wairotchanaphuttha, P., "Automatic detector for bikers with no helmet using deep learning." In *2018 22nd International Computer Science and Engineering Conference (ICSEC)*, pp. 1-4. IEEE, 2018.
3. Maliye, S., Oza, J., Rane, J., and Pathak, N., "Mask and Helmet Detection in Two-Wheelers using YOLOv3 and Canny Edge Detection." In *International Research Journal of Engineering and Technology (IRJET)*, vol.8, pp. 1225-1230. 2021.
4. www.kaggle.com/andrewmvd/helmet-detection