KLE Technological University
Creating Value
Leveraging Knowledge

KLETech CEVI Enabling AI
Center of Excellence in Visual Intelligence

CEVI WORKSHOP PROJECT 2022

TEAM NO : 20

TEAM MEMBERS		
1	VINOD PATIL	01FE20BEC143
2	ALLABAKASH	01FE20BEC261

OVERVIEW

- Problem statement.
- Dataset used.
- Dataset description.
- Dataset description analysis.
- Methodology.
- RESNET-50 architecture.
- Analysis of model-1.
- Analysis of model-2.
- Validation.
- Observations
- Efficient model.
- References.

LITERATURE SURVEY.

RESEARCH PAPER:Object Detection Based on VGG with ResNet Network.

- Introduction to an improved very deep convolutional network for accurate and significant object detection.

- It extracts high-level features that help to achieve tremendous performance to classify the image and detect objects

ARCHITECTURE.

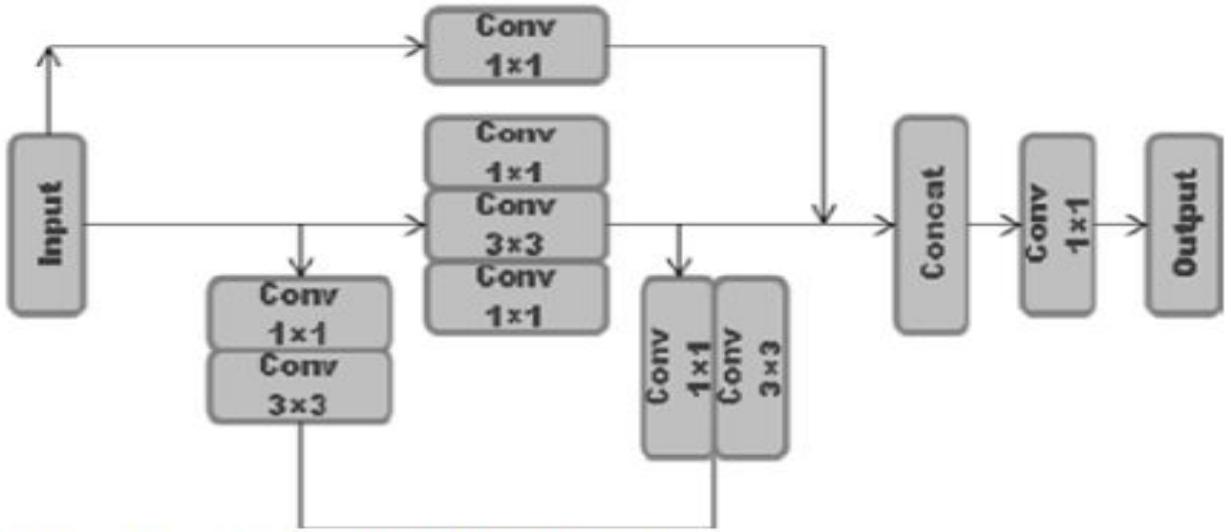

Fig. 2. The proposed architecture of ResNet

PROBLEM STATEMENT

Detection of Non-Helmet Riders.

OBJECTIVES

- Detect helmet from the image.
- Predicting correct bounding box.

figure a: input image

figure b: predicted output

1.DATASET USED

HELMET DETECTION DATASET.

figure 1.a: Image1

Filename	width	height	xmin	ymin	xmax	ymax	label
Image1	576	460	289	5	361	125	Without helmet
Image1	576	460	196	10	282	158	Without helmet

figure 1.b:Image1 annotation's

figure 1.c:Image2

Filename	width	height	xmin	ymin	xmax	ymax	label
Image2	576	460	276	78	381	190	With helmet
Image2	576	460	336	138	441	250	With helmet

figure 1.d:Image2 annotation's

2. DATASET DESCRIPTION

- **1200 –** Train images.
- **300 –** Test images.
- The dataset contains both Helmet and Non-helmet riders images and their corresponding annotation’s in a single csv .

3. DATASET DESCRIPTION ANALYSIS

- we have all image annotations in a single CSV file. reading the contents from CSV, filename, width, height, label, xmin, ymin, xmax, ymax.
- Splitting the lists like 80% for training and 20% for testing using train_test_split().
- Resizing images to 224x224(because resnet50 takes input images of size 224x224).
- Converting the lists in the form of tensors from the torch library.

4. METHODOLOGY

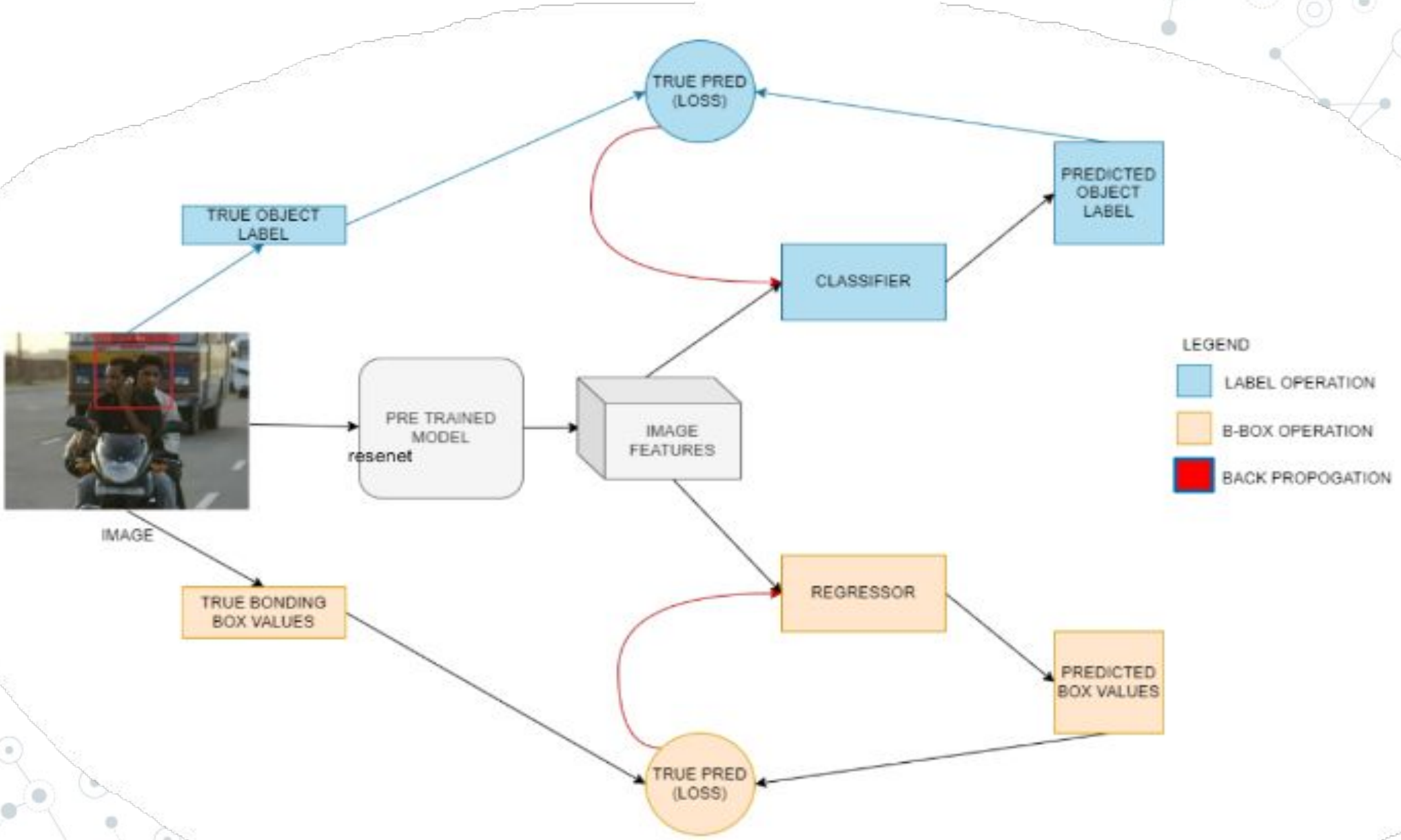

figure 3 : Methodology

5.RESNET-50 ARCHITECTURE

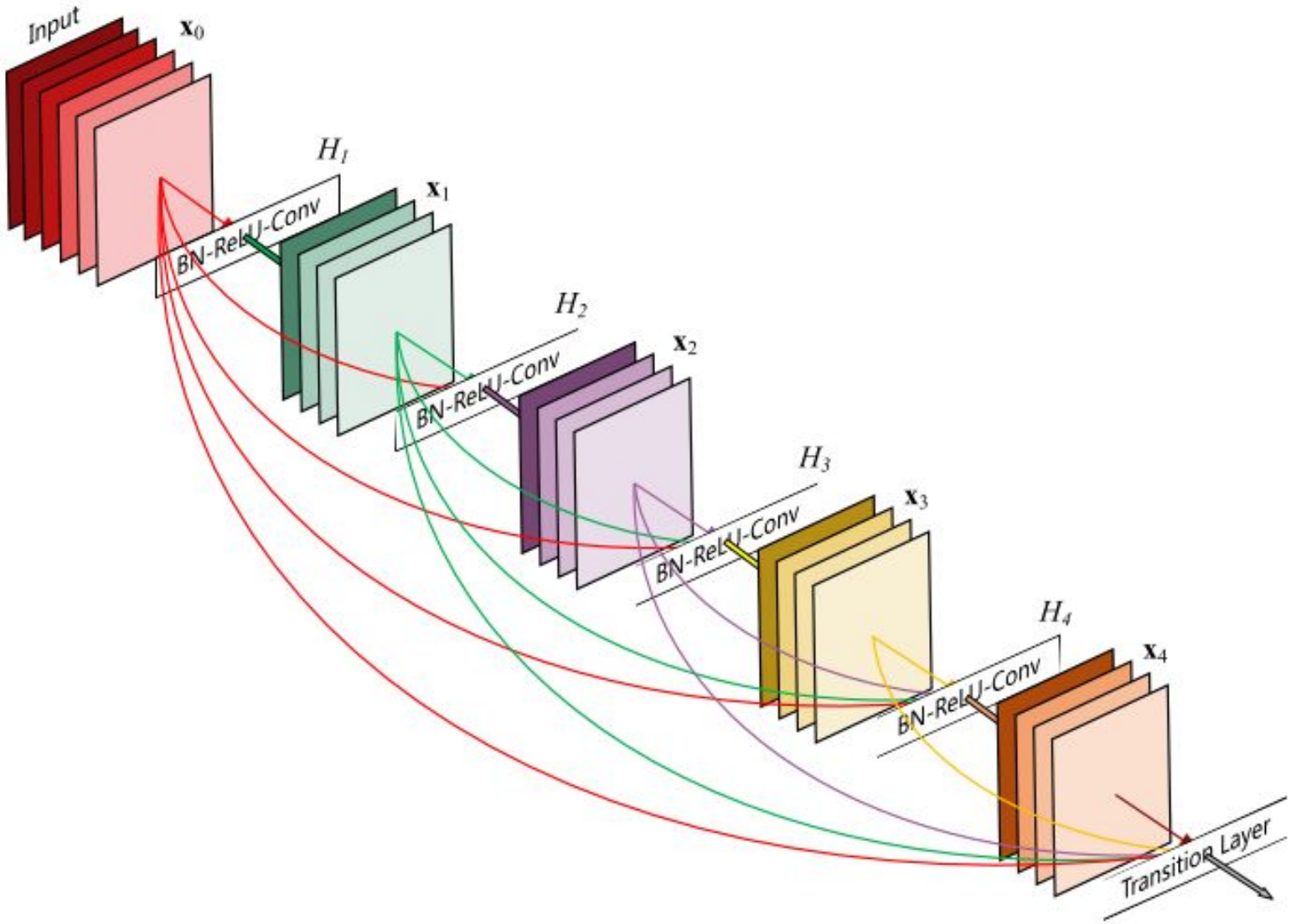

figure 5 : resnet50 architecture

6.MODEL-1.

- **Pretrained model:** RESNET-50.
- **Number of convolution layers for regressor:** 4.
- **Number of epochs trained:** 20.
- **Loss function for bounding box:** I1Loss.
- **Loss function for labels:** CrossEntropyLoss.
- **Optimizer:** Stochastic gradient descent.

RESULTS OF MODEL-1

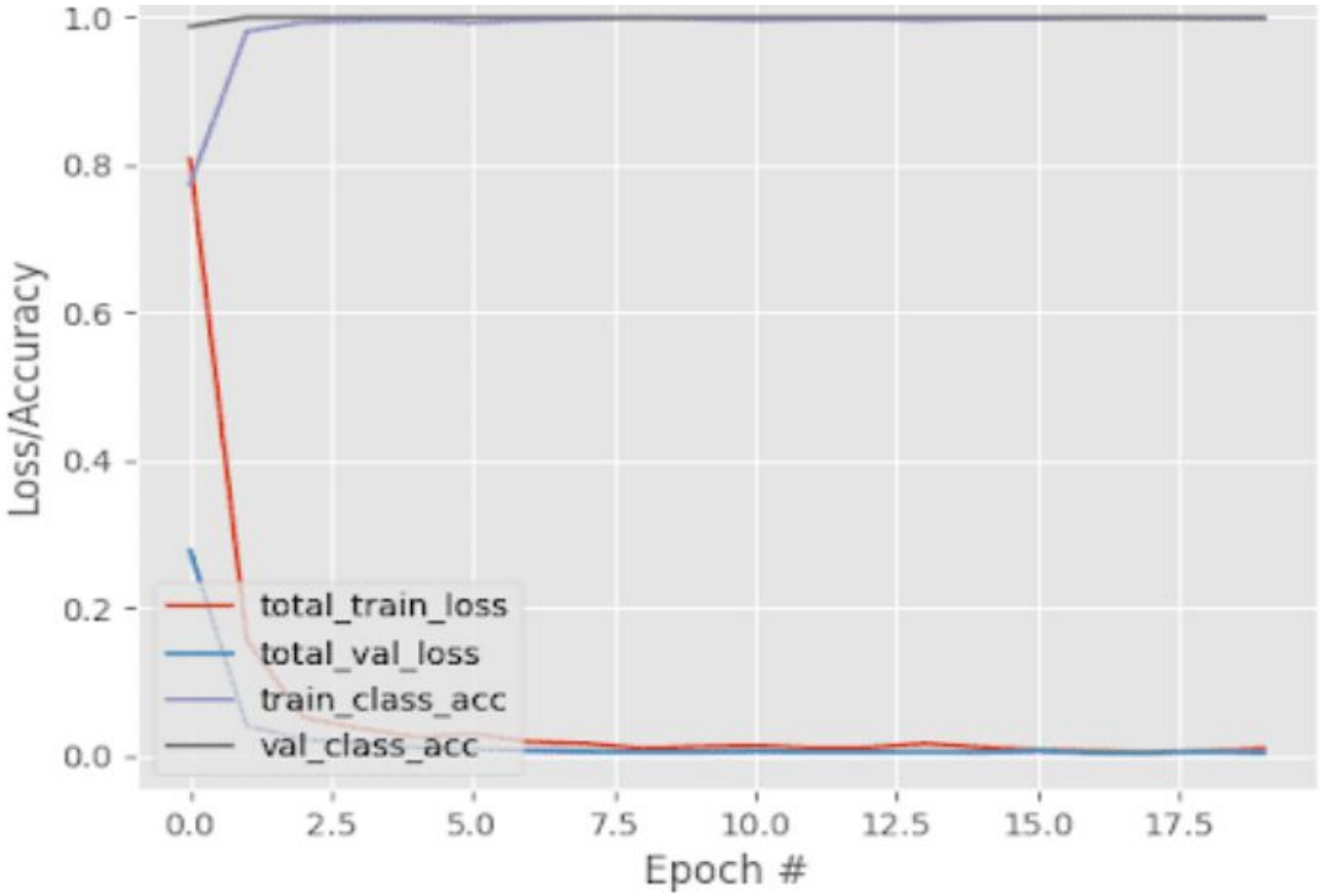

figure 6: training loss/accuracy

6.1 PREDICTION OF MODEL-1

figure 6.1.a

figure 6.1.b

figure 6.1.c

When model-1 was trained for 200 epoch's the results were not that satisfactory.

6.2 METRIC EVALUATION

 Evaluating model-1 using intersection over union.

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{\text{Intersection}}{\text{Ground truth box} + \text{Detected box} - \text{Intersection}}$$

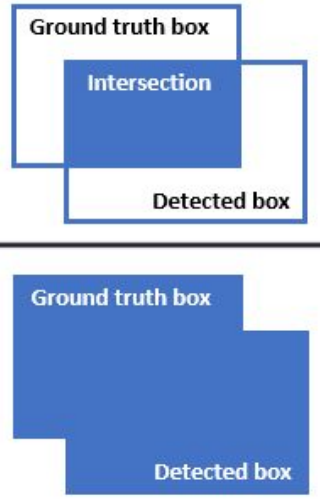

figure 6.2.a: formula for intersection over union.

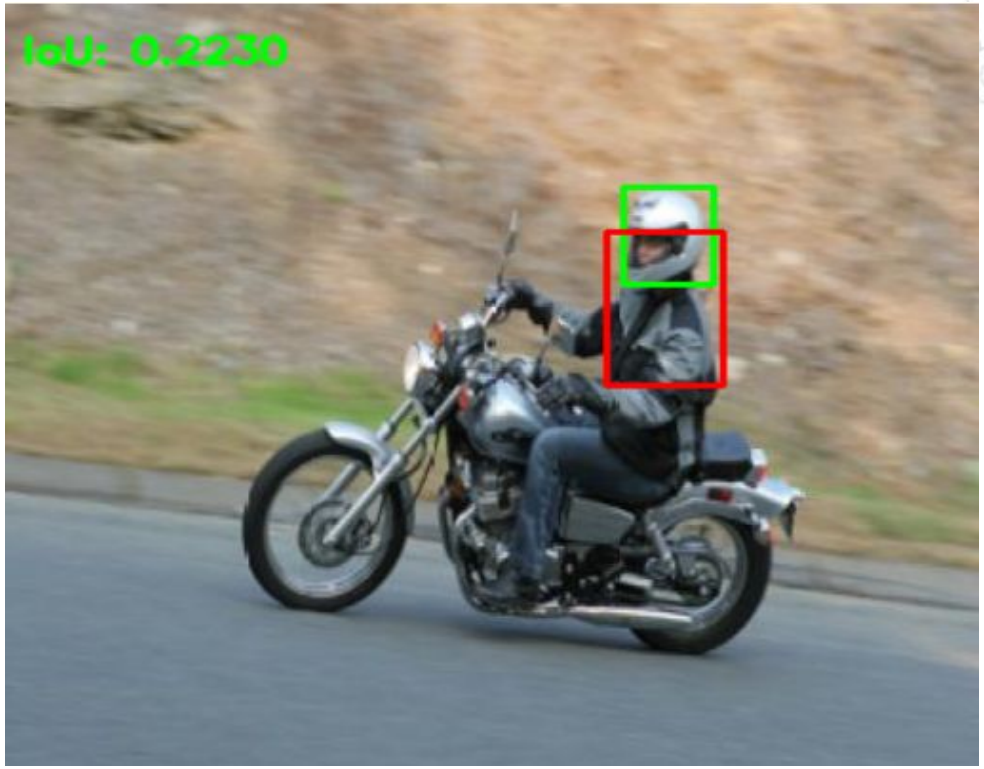

figure 6.2.a: IOU of the output image.

7.MODEL-2

- **Pretrained model:** RESNET-50.
- **Number of convolution layers for regressor:** 8.
- **Loss function for bounding box:** Mean squared error loss.
- **Loss function for labels:** CrossEntropyLoss.
- **Optimizer:** Adam optimizer.

7.1 EPOCHS: 70

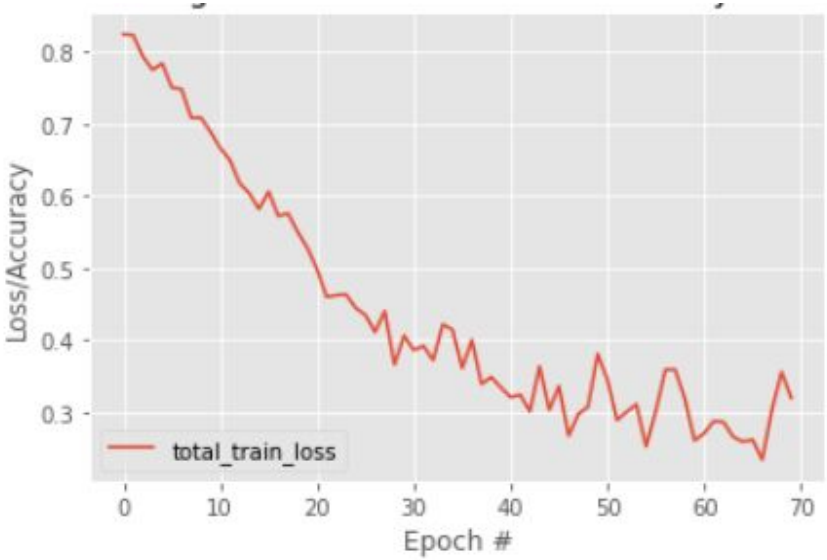

figure 7.1.a:total training loss

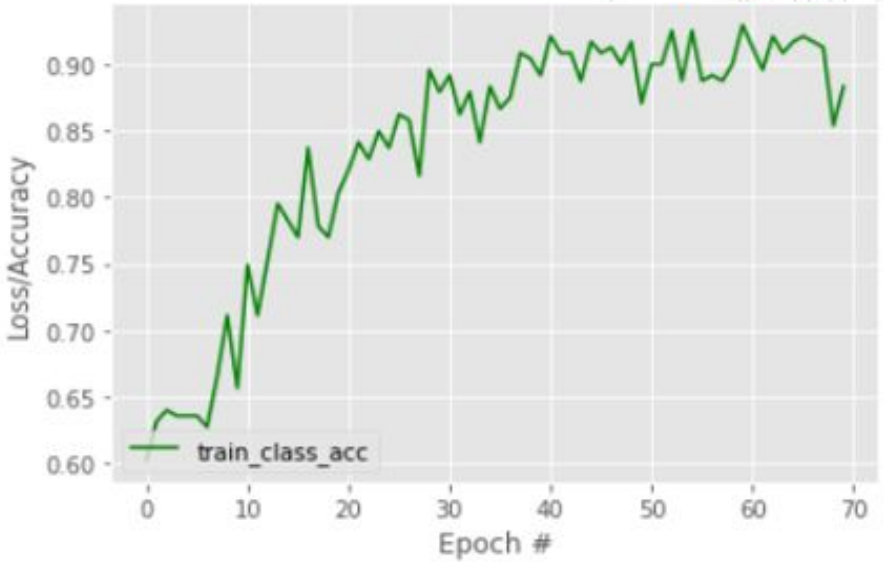

figure 7.1.b:total training accuracy

figure 7.1.c : input image

figure 7.1.d: predicted output

7.2 EPOCHS: 200

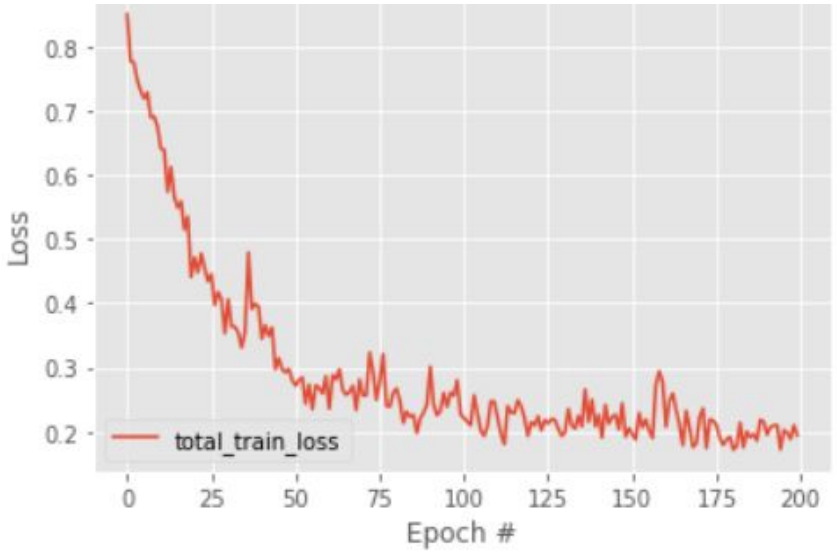

figure 7.2.a:total training loss

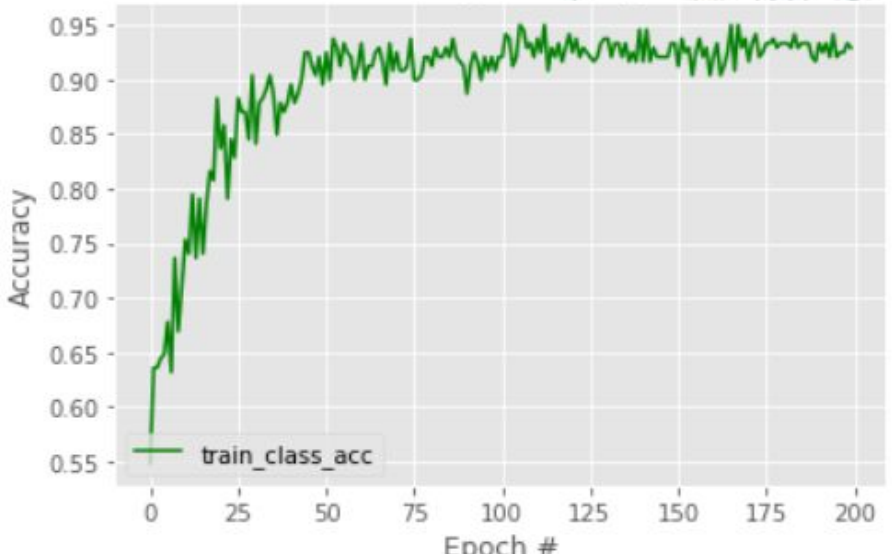

figure 7.2.b:total training accuracy

figure 7.2.c : input image

figure 7.2.d: predicted output

7.3 Learning rate: 0.001

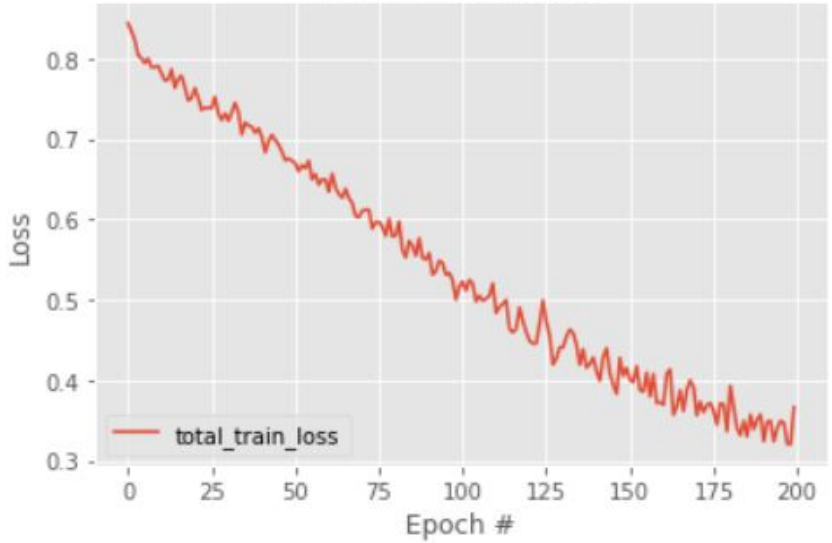

figure 7.3.a : total training loss

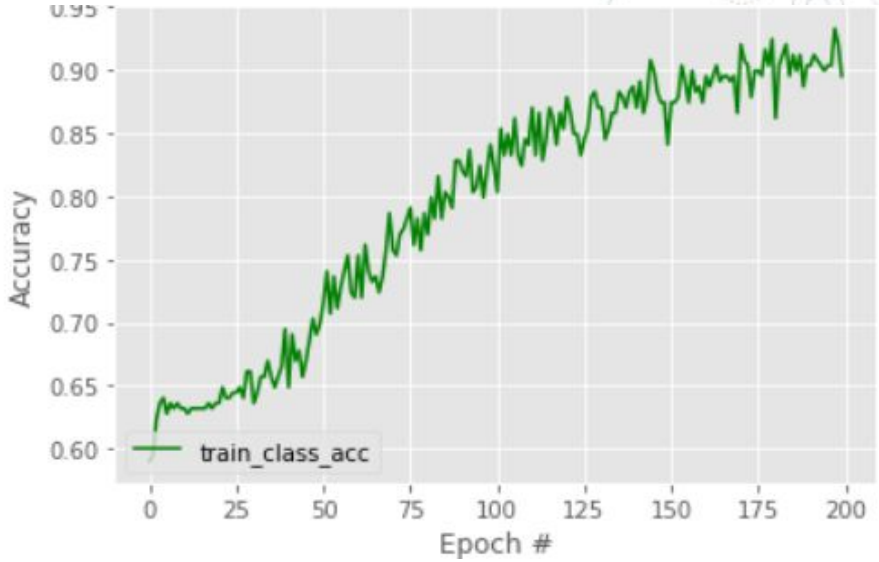

figure 7.3.b : training accuracy

figure 7.3.c : input image

figure 7.3.d : predicted output

7.4 Learning rate : 0.01

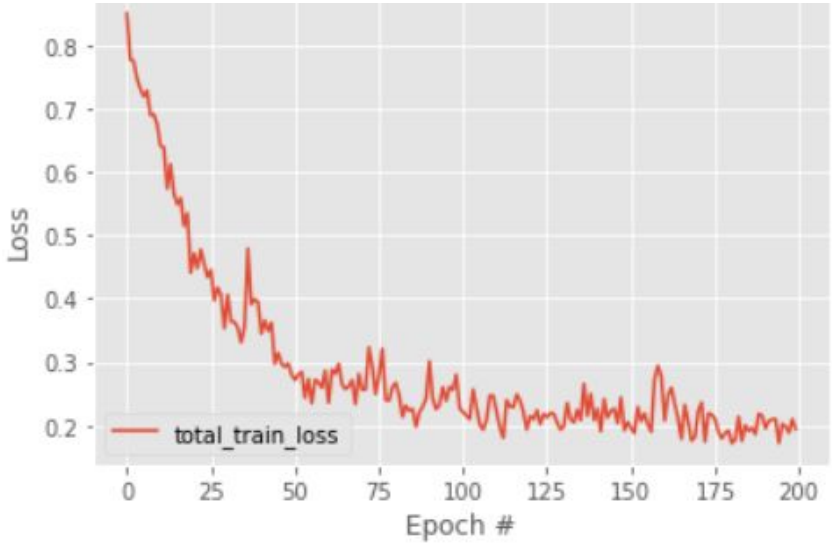

figure 7.4.a : total training loss

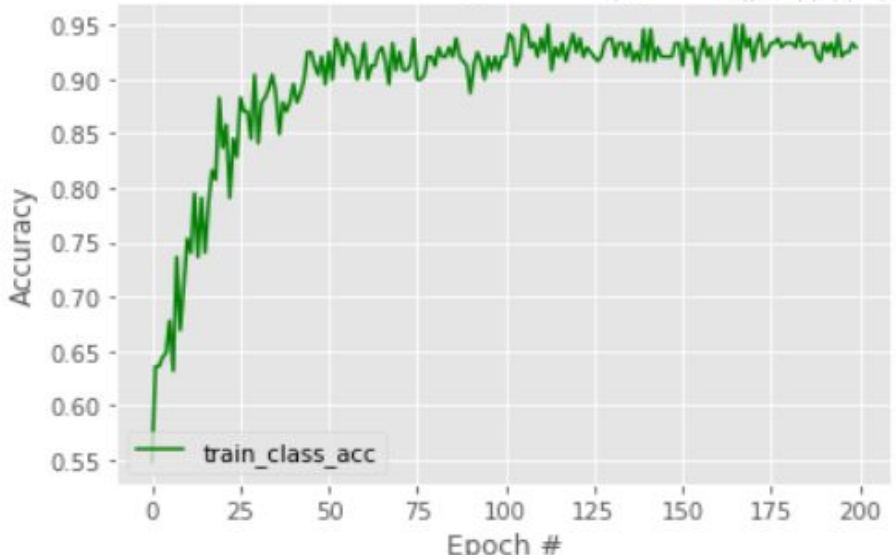

figure 7.4.b : total accuracy loss

figure 7.4.c : input image

figure 7.4.d : predicted output

7.5 VALIDATION OF MODEL-2

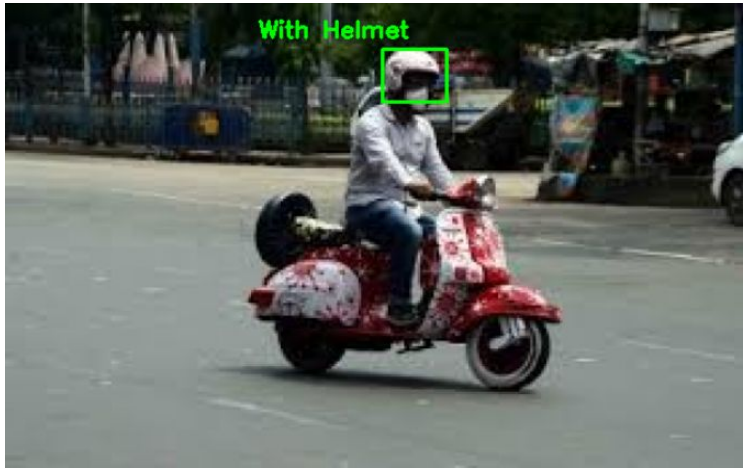

figure 7.5.a

figure 7.5.b

figure 7.5.c

figure 7.5.d

7.6 BONUS

figure 7.6

7.6 METRIC EVALUATION

 Evaluating model-2 using intersection over union.

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{\text{Intersection}}{\text{Ground truth box} + \text{Detected box}}$$

figure 5.2.a: formula for intersection over union.

figure 7.2.a: IOU of the output image.

8 OBSERVATION.

Model	Highest Train Accuracy	Highest Validation Accuracy	Iou
Model-1	85.46%	89.96%	0.22
Model-2	91.21%	90%	0.37

9. EFFICIENT MODEL

After training our model with different loss functions, different optimizers, training with different numbers of epochs, and adding more convolution layers.
We conclude that the best-suited parameters for our model are:

- **Loss function for bounding box:** MSELoss.
- **Loss function for labels:** CrossEntropyLoss.
- **Optimizer:** Adam optimizer.
- **Pretrained model:** RESNET-50.
- **Learning rate:** 0.01
- **Number of convolution layers:** 8.

10. REFERENCES

- https://www.kaggle.com/datasets/andrewmvd/helmet-detection
- Training an object detector from scratch in PyTorch – PyImageSearch
- https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/
- https://www.irjmets.com/uploadedfiles/paper//issue_5_may_2022/25059/final/fin_irjmets1654316725.pdf
- https://www.researchgate.net/publication/333232428_Object_Detection_Based_on_VGG_with_ResNet_Network
- https://pytorch.org/
- https://stackoverflow.com/


THANK
YOU!