Shape

Description automatically generated with medium confidence  


**Essay / Assignment Title:** Enhancing Environmental Monitoring through Advanced Object Detection in Satellite Imagery

**Programme title: MSc of Data Analytics**

**Name: Turan Can Gün**

**Year: 2024 (2023 October Intake)**

[INTRODUCTION 3](#_Toc1244399413)

[CHAPTER ONE Problem Formulation 4](#_Toc1793461002)

[1.1 Limitations of Traditional Environmental Monitoring 5](#_Toc2081839043)

[1.2 Role of Computer Vision Models in Environmental Monitoring 5](#_Toc617451982)

[CHAPTER TWO Data Preparation 6](#_Toc1544790601)

[2.1 Collecting Relevant Dataset and Explanation of Features 7](#_Toc626114741)

[2.2 Preprocessing Data 9](#_Toc1734558544)

[CHAPTER THREE Model Implementation 11](#_Toc1086063732)

[CHAPTER FOUR Model Training and Evaluation 14](#_Toc1416874173)

[4.1 Model Training 14](#_Toc1854134619)

[4.2 Model Evaluation 15](#_Toc107675288)

[4.3 State of Art and Comparisions 18](#_Toc2044395030)

[CHAPTER FIVE Practical Application 19](#_Toc827592723)

[5.1 Snippets of Python Code 19](#_Toc896931764)

[5.2 Potential Impact 21](#_Toc547607622)

[CONCLUDING REMARKS 22](#_Toc80046115)

[BIBLIOGRAPHY 23](#_Toc307182121)

[APPENDIX (if necessary) 24](#_Toc462197051)

**Statement of compliance with academic ethics and the avoidance of plagiarism**

I honestly declare that this dissertation is entirely my own work and none of its part has been copied from printed or electronic sources, translated from foreign sources and reproduced from essays of other researchers or students. Wherever I have been based on ideas or other people texts I clearly declare it through the good use of references following academic ethics.

(In the case that is proved that part of the essay does not constitute an original work, but a copy of an already published essay or from another source, the student will be expelled permanently from the postgraduate program).

Name and Surname (Capital letters): TURAN CAN GÜN

.......................................................................................................................................

Date: 31 / 10 /2024

# **INTRODUCTION**

Around the world many researcher is studying different object detection techonologies at different instutions , universities and companies.Most of them prefers to use Convolutional Neural Network (CNN) to handle projects.Moreover object detection is main area within communities of computer vision area then usage areas of object detection can be classified as autonomous driving , remote sensing , military operations and so on (K. Tong et al. , 2020) , ( G. Cheng et al. , 2022). Although modern systems are created object detection tasks for real time frame mainly focus on speed and cost of computational resources but they are suffered from low detection accuracy.As a result of this situation making new improvements in this area is benefitical to use in real world life more broadly.Moreover from general aspects deepl learning algorithms in terms of object detection are built up from two types.First type is two stage version which aims to higher accuracy than one stage.However recent studies started to involve one stage systems such as YOLOv5.In this study object detection project is handled by utilizng YOLOv5 model in terms of from installing packages to evaluation the model performance.

# **CHAPTER ONE Problem Formulation**

# **1.1 Limitations of Traditional Environmental Monitoring**

Traditional environment methods requires high intensive resource then not have real time flowing and data is coming from basic sensors.At the same time these traditional system aren’t capable to adapt rapid changes such as environemntal pollutions , recent construction areas or other recent events.However new generation monitoring systems are different becuase their structures are based on deep learning.This learning type is useful for especially gathering recent update from geographical areas continously by using automated systems.This recent update is able to collected from combination of sateilites and cameras within urban areas.One of the critical usage area for this feature is environmental disasters since which can track the last situaiton during any disaster.

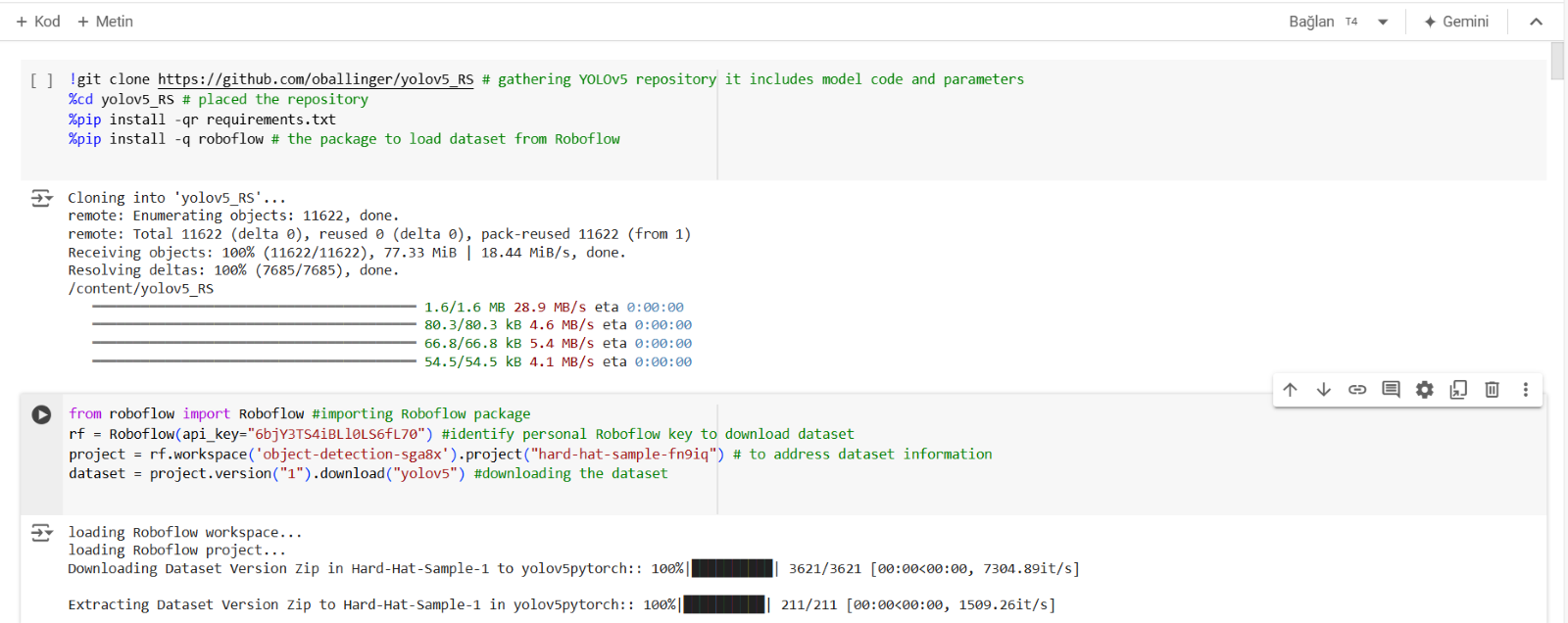
## **1.2 Role of Computer Vision Models in Environmental Monitoring**

Computer vision models such as Faster R-CNN and YOLO has a crucial role to observe environmental events coontinously by automated systems.Faster R-CNN is suitable to detect detailed objects within specific areas with high accuracy.On the other side YOLO (You Only Look Once) is appropraite for real time processing becuase which aims to track conitnous environment events like marine pollution , forest fire and so on.Both models are built up based on deep learning algorithm then when the time to selectin of models which is related to monitoring needs and priority of benchmark score like importance of either speed or accuracy.For example if there is a need across to identify aircrafts within a specified area in that case Faster R-CNN is a well selection but if there is a need towards to identify a continous event such as earthquake in this situation YOLO is a good option.

# **CHAPTER TWO Data Preparation**

Data preparation is critical stage to handle any machine larning (ML) projects because to be ensure dataset should have enough quality to obtain trustful results also dataset shouldn’t have any missing points that is called as “ Consistency “.Moreover normalization and scaling is essential for perfromance becuase to get as high as possible benchmark score then also this part is importtant maintainance in terms of not excessive hardware capacity.

## **2.1 Collecting Relevant Dataset and Explanation of Features**

Figure 2.1 Collecting Required Dataset

In this study YOLOv5 model is implemented to handle object detection.Firstly on the contrary of traditional applications of YOLOv5 in this study the mdoel is gathered from GitHub by using just oen line code with model code and parameters.Moreover the dataset is collected from RoboFLow then RoboFLow is a plotfrom across machine learning applications and computer vision implementations at the same time this platform is able to trained model by selected dataset within the inside of the platform.Before download the dataset from RoboFlow second part of the code indicates that an account must be created to get personal key to download the dataset.

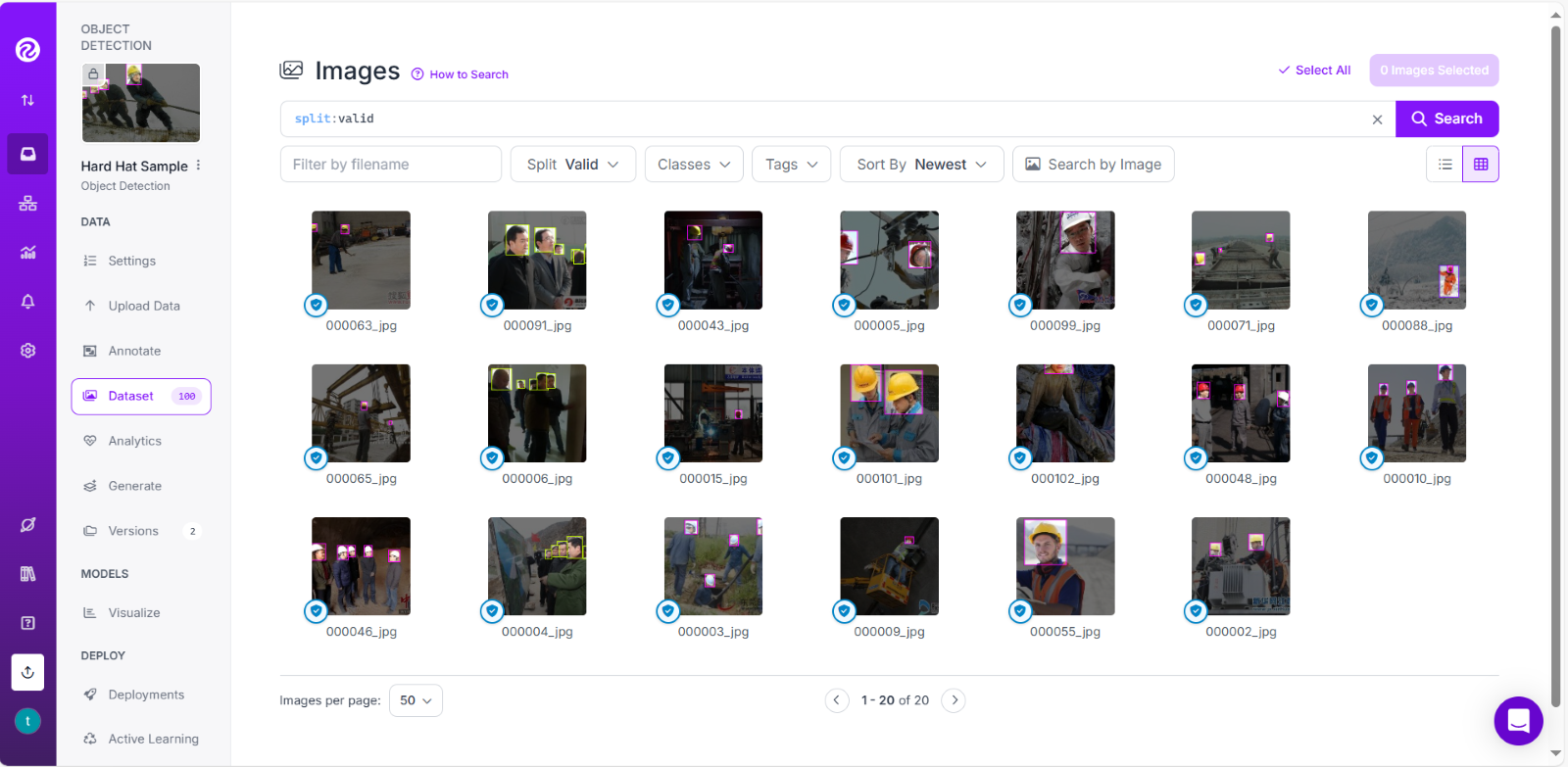
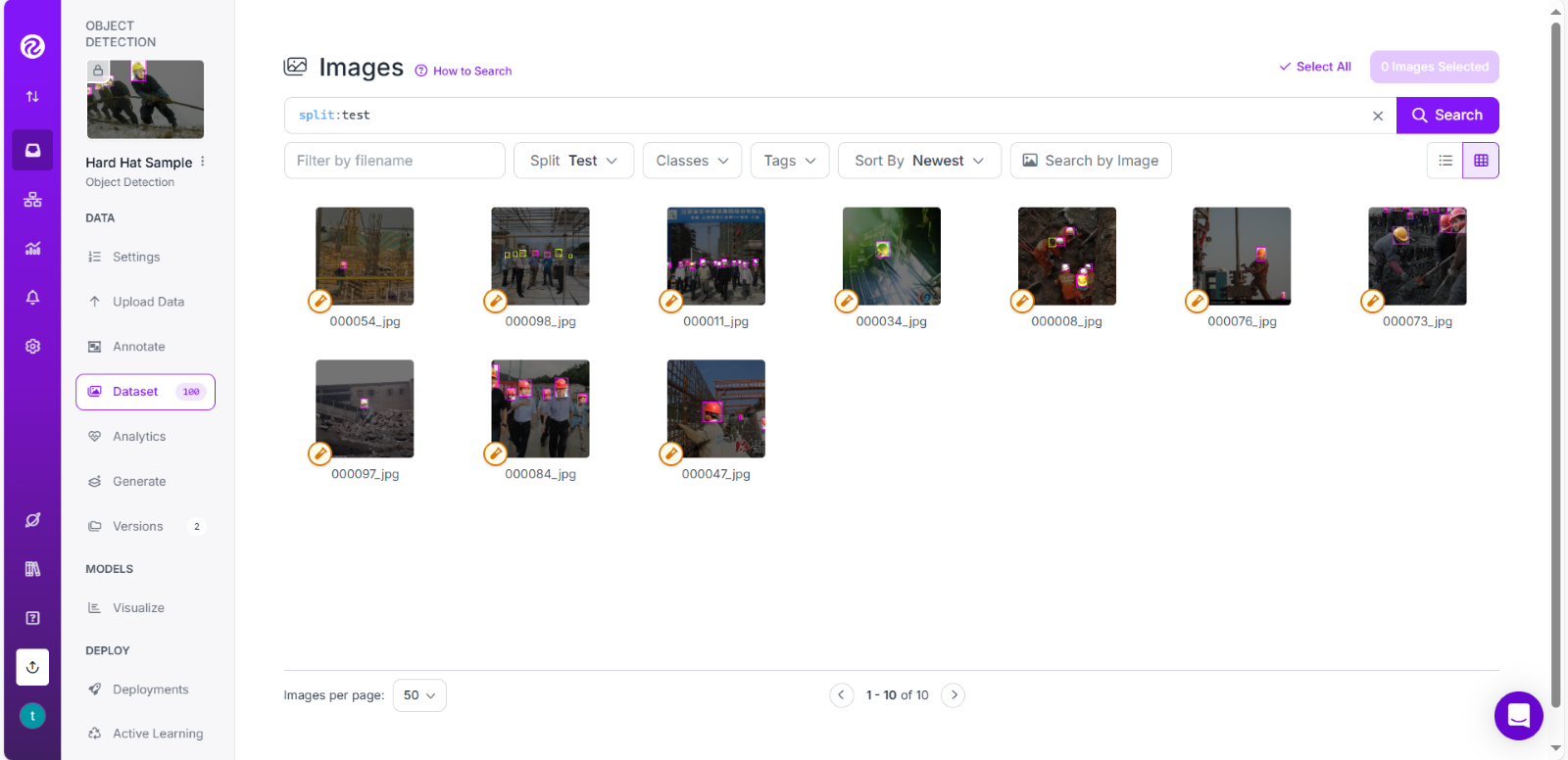
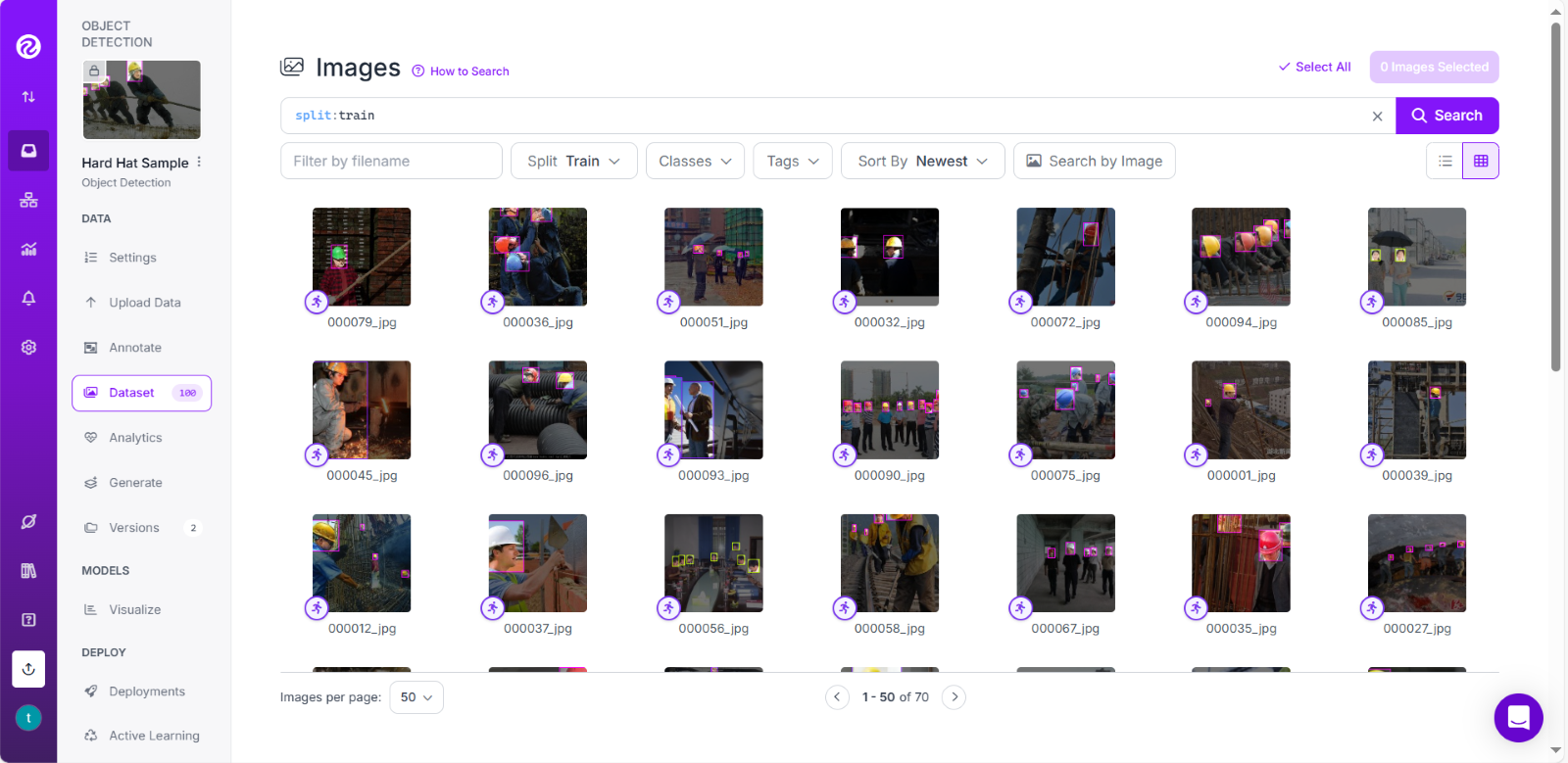


Figure 2.2 Review of Test Data and Validaton

About splitting of the datast as training , test and validation part which identifies by itself of structure of the dataset and as Fig 2.3 indicates that %70 of the dataset is train , %20 is validation part and %10 of the dataset is test.

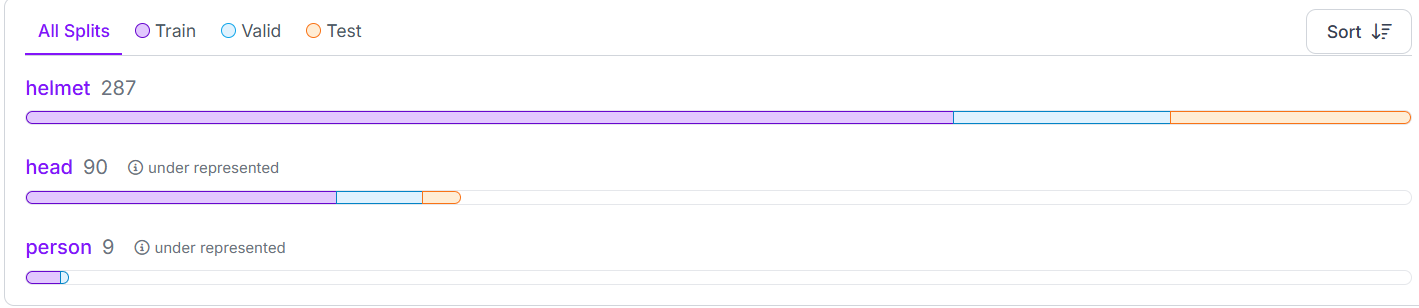


Figure 2.3 Overview of Classes

There are 3 classes which are helmet person and head.That means the model aims to identify to these three specific classes among others.Moreover also there are other objects like 911 sign , construction , climbing , shoes , body , boots , hat , house etc.Moreover as mentioned before YOLOv5 is appropraite solution to track any disaster.This dataset confirms that adive because of its classes.In this study the model is helpful identify civilian people , workers for mining accident or landslide in construction areas.One of challenge in that stage is collecting data from RoboFlow since create personal key might be difficult.

## **2.2 Preprocessing Data**

Data preprocessing is important pahse to handle object detection models which is able to provide higher durability across different real life scenario..It contains data augmentation and transformation to increase model performance as well as provide robustness.In this part annotations are carefully processed , resized and normalized.Data augmentation plays critical role to enchance variability of the dataset also that has capability of preventing overfitting.

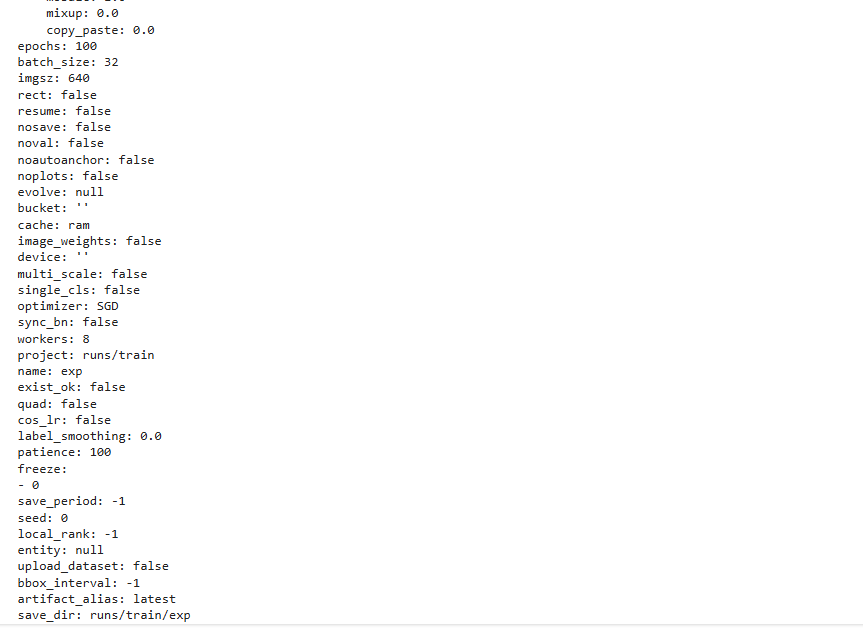
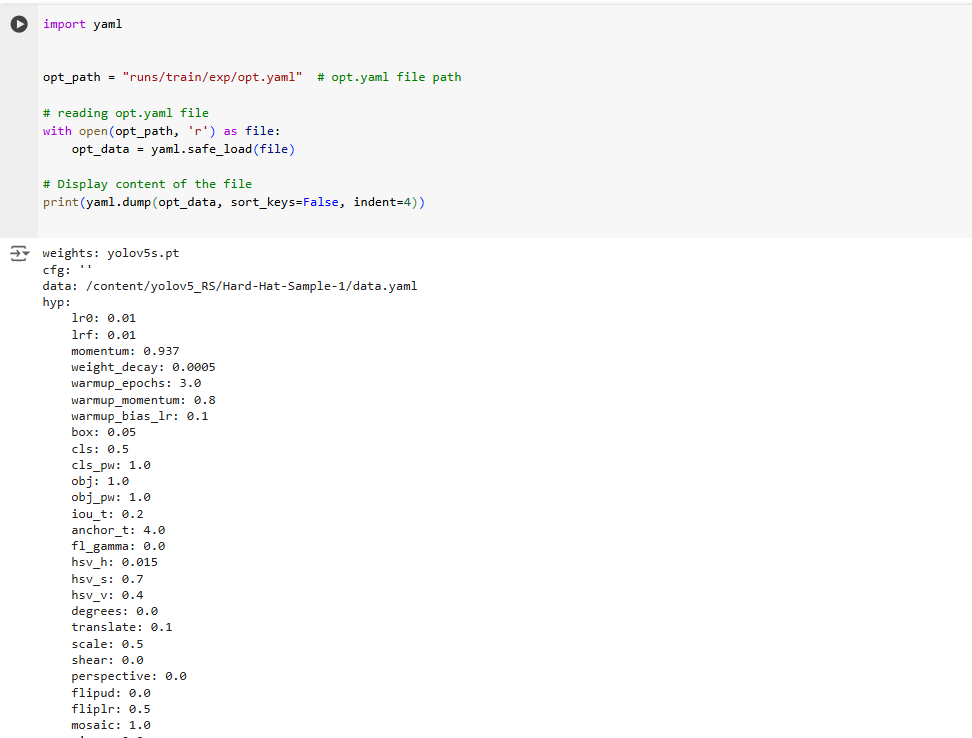


Figure 2.4 Pre Processing of The Data

As Figure 2.4 indicates that there are many hyperparameters of YOLOv5 but in this some of important parameters are elobrated.

|  |  |  |  |
| --- | --- | --- | --- |
| Basic Settings | Hyperparameters | Training Settings | Optimizaiton |
| Weights:YOLOv5 | lr0: 0.01- lrf: 0.01 (Inital and final learning rate) | Epochs:100  Batch size:32  Imgsz: 640 x 640  rect: false (Rectangular training) | Optimizer:SGD (Stochastic gradient descent) |
| cfg: '' '' | weight\_decay: 0.0005 (Regulating strenght to avoid overfitting) | Imgsz: 640 x 640  rect: false (Rectangular training) | Cos\_lr:False (cosine learning rate schedule) |
| data: /content/yolov5\_RS/Hard-Hat-Sample-1/data.yaml | box: 0.05 (loss gain for box accuracy) | Cache:RAM (using RAM for faster access | patience=100 (early stopping place for no improvements |
|  | cls: 0.5 (loss gain for class accuracy) cls\_pw: 1.0 (wieght to process imbalanced classes |  | Freeze= - 0 ( freezing layers during training) |

|  |  |  |  |
| --- | --- | --- | --- |
|  | degrees:0.0 (angle of rotation for augmentation)  translate:0.1 range pf translation for augmentation  Scale:0.5 Scaling range for size augmentation |  |  |
|  | flipud: 0.0 (Possiblity of vertical flip )  fliplr: 0.5 (Possibility of horizontal flip) |  |  |

Table 2.1 Parameters of the Model

# **CHAPTER THREE Model Implementation**

In this study YOLOv5 model used then this model is popular for speed and accuracy.Also which seperates images then predicts bounding boxes with calculating their probabilities so this situation makes it high efficiently method.The model set up performance balance by using CSPDDarknet for feature extraction and PANet path – aggregation enhance feature fusion so YOLOv5 is able to detect objects at various scales.Furthermore YOLOv5 has capability track events based on real time frame and its pre defined weights are able to provide strength across various difficult conditions at the same time the model has an advantages related too compatibility with different kind of framewroks such as Pytorch.

More specifically YOLOv5 was released evolution of YOLOv3 and Pytorch implementation of YOLO has been realized after released of YOLOv4 ( Z. , Yao et al. , 2021) and YOLOv5 has gained its popularity during transitioning YOLOv3 models from DarkNET to Pytorch for prodcution distribution ( Khanam , R. , 2024 ).During the trainig step for YOLOv5 model there two main points should be considered.

First point is related data augmentation which is able to provide to the model robustness towards to different real life scenario by enchancing model generalizaiton.Second point is loss function which is a metric calculated from 3 primary components.These components are called as Generalized Intersection over Union (GIoU), objectness, and classification loss.Main goal of these three components is creating optimize mean average precision (mAP) carefully a broadly adopted evaluation score for object detection.

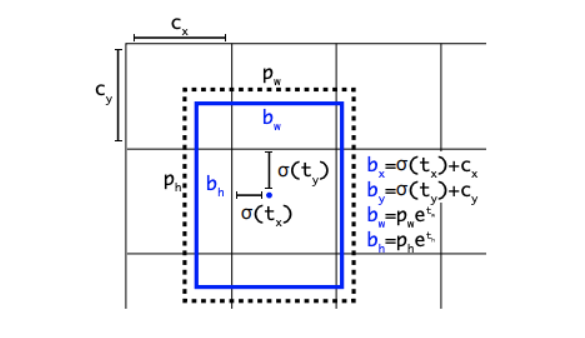


Figure 3.1 Bounding Box Prediction ( Solawetz , J. , 2020)

YOLOv5 architecture tries to predict coordinates of bounding box across pre defined anchor box diemnsions.Size of anchor box is necessary to begin the model traninig also the size has crucial role for model accuracy and performance.

Loss=λ1⋅Lcls+λ2⋅Lobj+λ3⋅Lloc *Loss=𝜆1⋅Lcls+𝜆2⋅Lobj+𝜆3⋅Lloc*

Loss function is an essential parameter and which is computes from three parameters. Binary Cross-Entropy (BCE) is used for class prediction , objectness and Complete Intersection over Union (CIoU) for localization. The overall loss is evaluated as a composition of weighted sum of these three losses. Lcls , Lobj and Lloc parts are related to BCE is for class prediction BCE loss for objectness, and CIoU loss for localization respectively ( Khanam , R. , 2024 ) on the other side   
*λ1*, *λ2*, and *λ3* help to make balance to each loss components for optimizing of overall process.

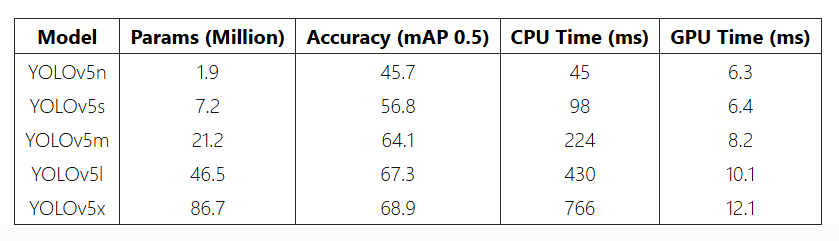


Figure 3.2 Comparision Performance Results of YOLOv5 Types ( Khanam , R. , 2024 )

Figure 3.2 is able to provide copmherensive results regarding to perfomance outlook of different versions of YOLOv5 then which indicates that YOLOv5n and YOLOv5s is a good choice for speed and accuracy, on the other hand YOLOv5l and YOLOv5x are appropraite solution when accuracy is critical. Selection among these models is depend on requirements of defined project and hardware limitations.

# **CHAPTER FOUR Model Training and Evaluation**

## **4.1 Model Training**

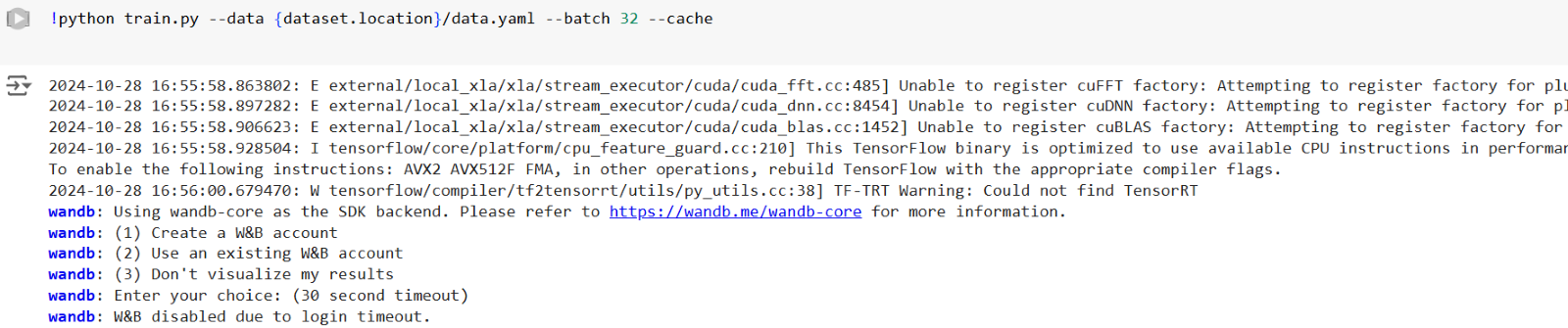


Figure 4.1 Python Command to Start Traning

Figure 4.1 implies to begin traning step of the model based on YOLOv5 architecture one of the advantage this model is actuaaly everything (including parameters) are pre defined that’s why which is very user friendly also details of parameter and share of traning , testing and validaiton part are elobrated at the previous chapters.Moreover the figure displays “wandb” section which can be ignored when implemented becuase it is optional to demonstrate performance results also performance results can be extracted with python script as implemented further steps.

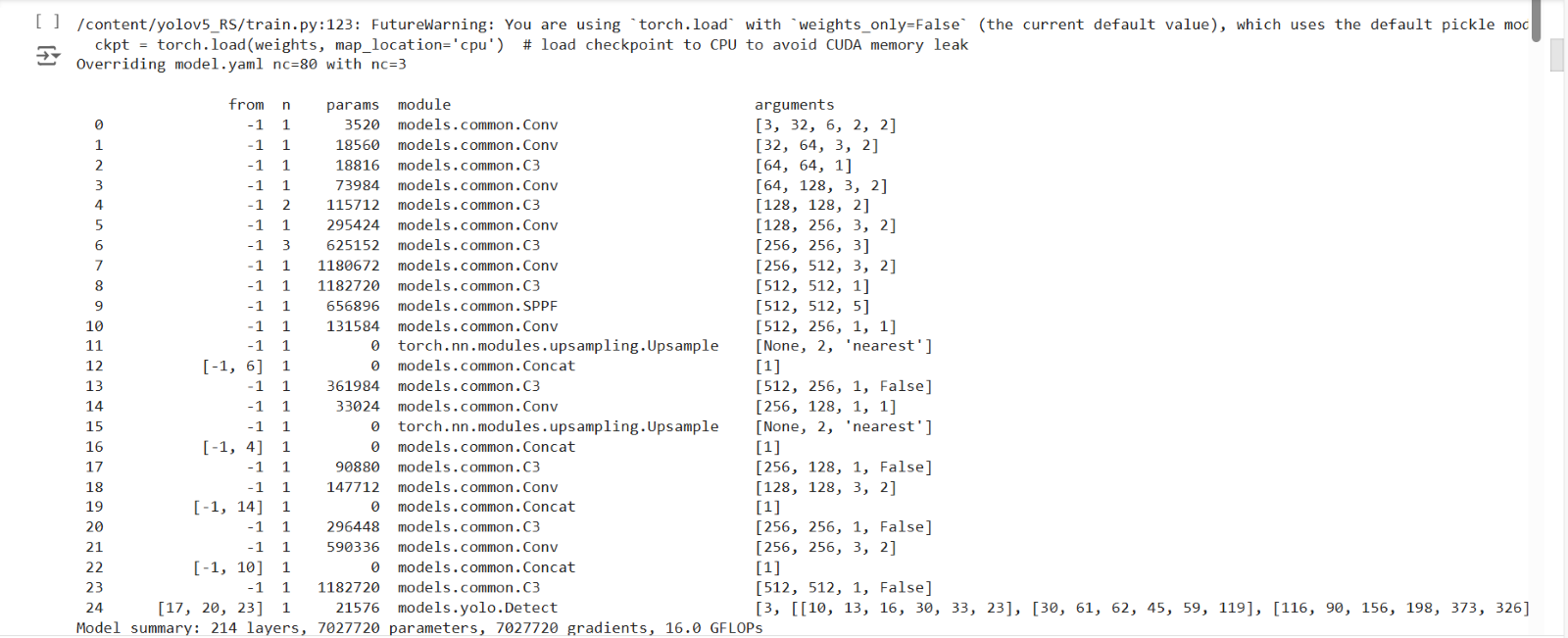
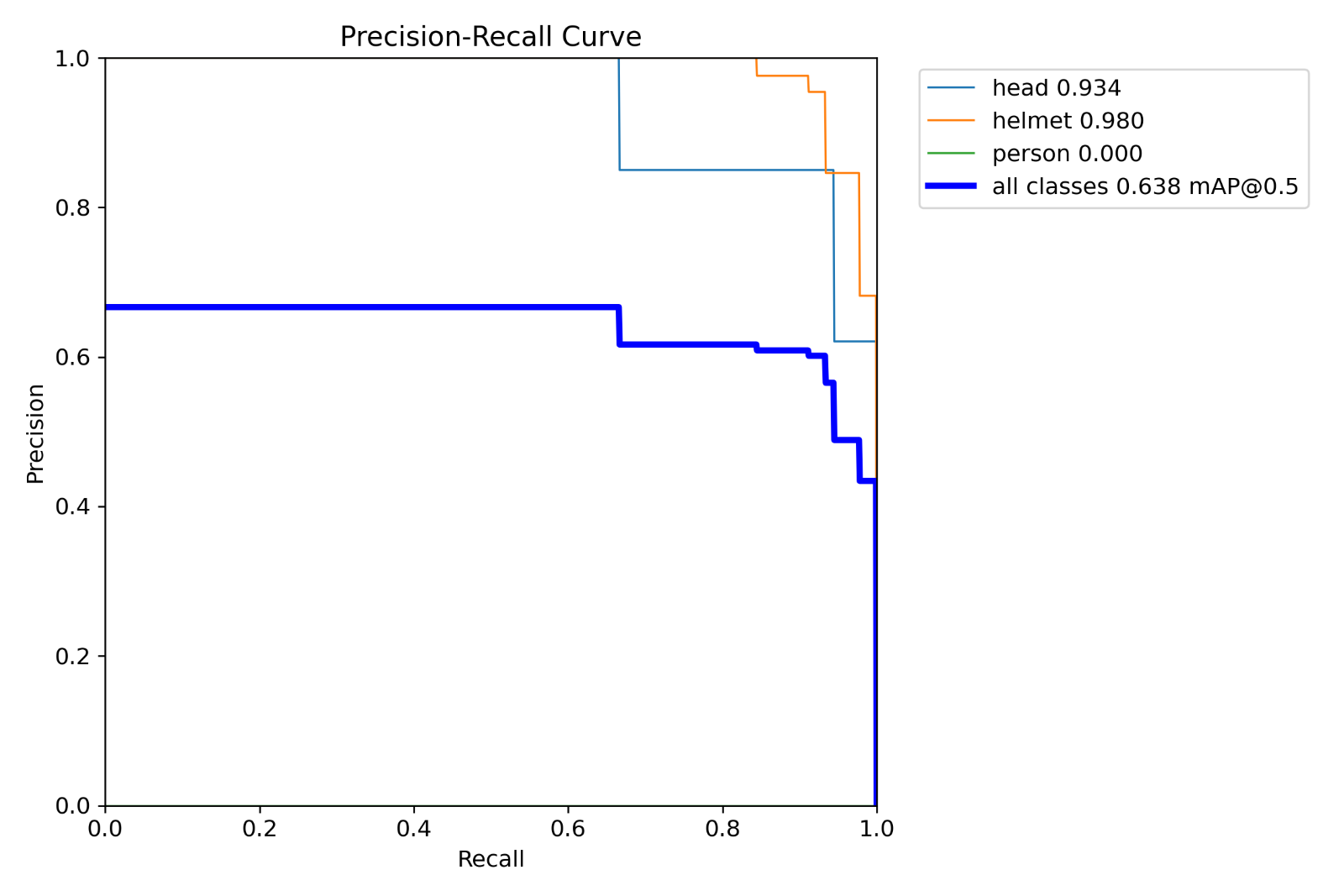
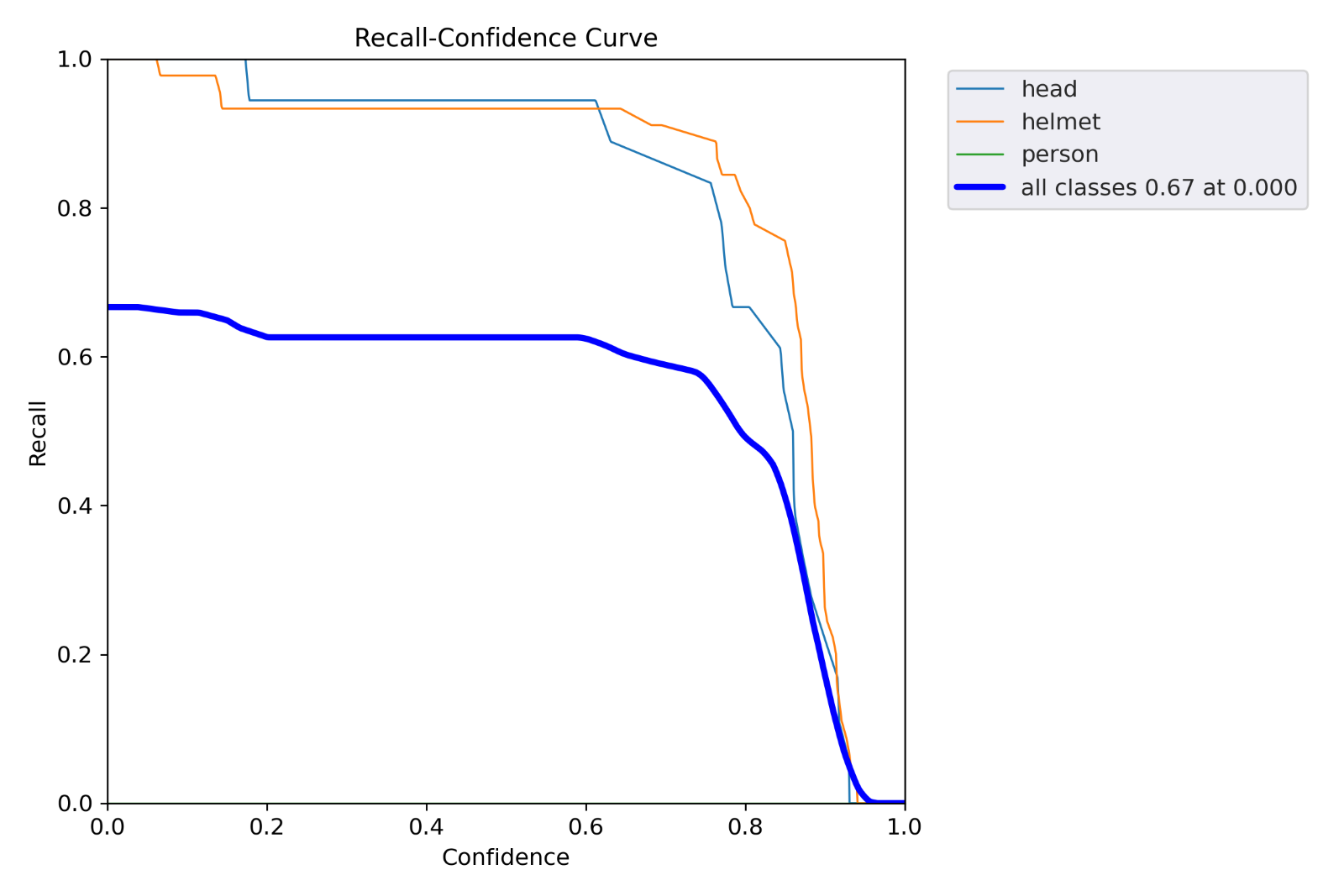
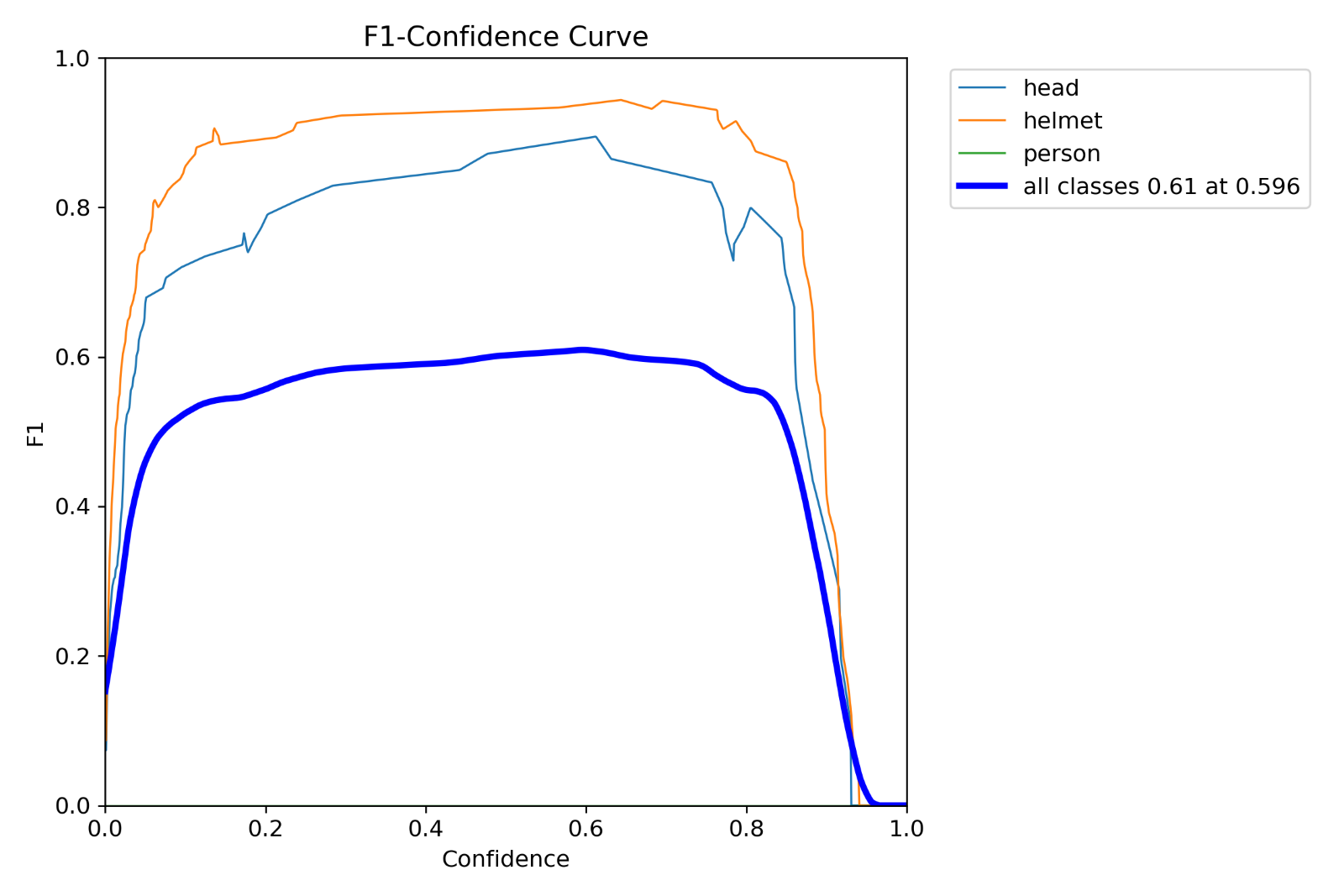


Figure 4.2 Parameters and Packages of the Training Stage

Figure 4.2 review packages and parameters for training stage there are 214 layers , 7027720 parameters and 7027720 gradients are available and 16.0 GFLOPs indicates that how many billions of floating-point operations are required by a model to make a prediction across specific task.

## 4.2 Model Evaluation

Figure 4.3 Performance Results of the Model

* Box loss , obj loss and cls loss are decreasing as epcoh increase which means that the model is able to learn effectively.
* Validation demosntrate decreasing trend over time this situation elobrates that the model has capability to generalize without occuring serious overfitting issue.
* Precision and recall parameters have uptrend that means the model is able to maximize true positive as well as minimize false positive cases.  
   Figure 4.4 Evaluation Metrics of the Model
* F-1 – Confidence curve shows higher values for helmet and head class.On the contrary for person the value is relatively lower than others.
* Recall – Confidence curve demonstrates high recall across helmet but for other classes this case isn’t valid.
* Precision - Recall curve displays high precision for helmets and heads which means that this curve is reliable for them but for person the curve isn’t reliable enough.

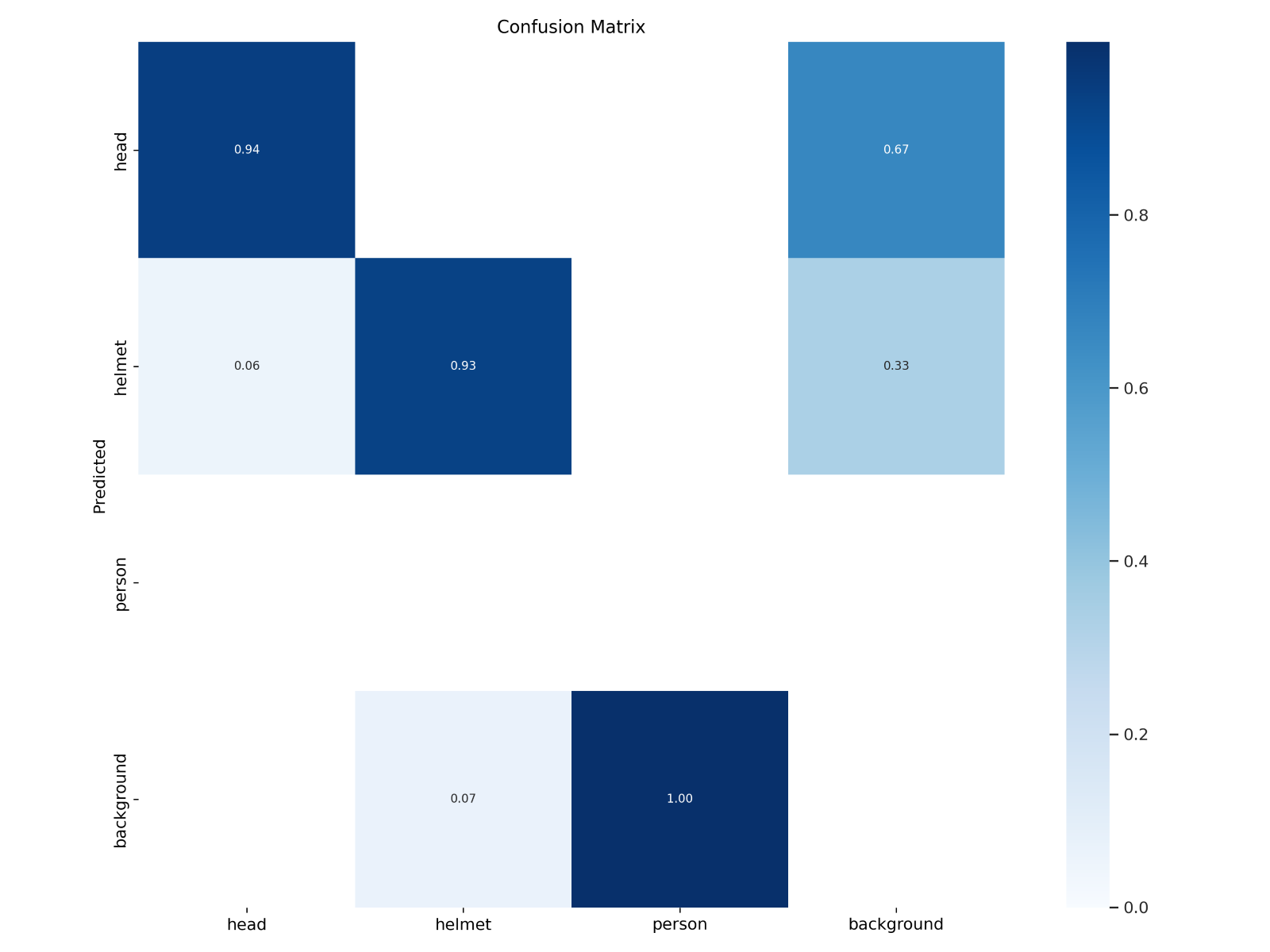


Figure 4.5 Confusion Matrix of the Model

* Confusion matrix of the model demonstrates relatively higher correlation for head and helmet classes than others.

## **4.3 State of Art and Comparisions**

State of arts indicates that the latest and most sophisticated situation in itechnology , art or science.Moreover which is very benefitical for understand the progress of any technology based on chronologically as well as compare results between the specified study and others.

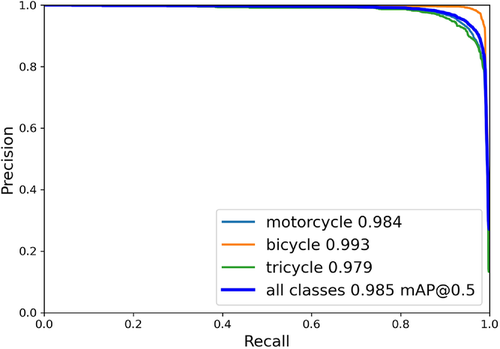


Figure 4.6 Precision and Recall Results ( Jia , W. et al. , 2021 )

* Jia , W. et al. have been introduced to a study which aims to identify weared helmet persons among motorcycle drivers during urban traffich then they are able to obtain high well performance results to detect helmet or non helmet situation among dirvers.

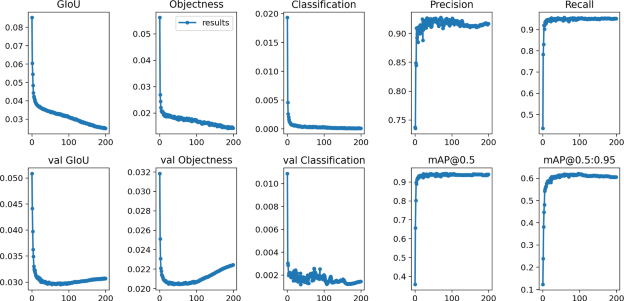
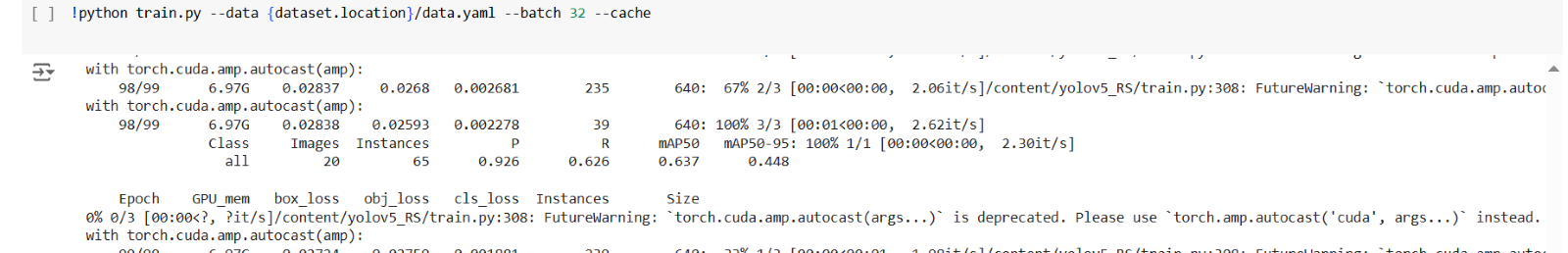
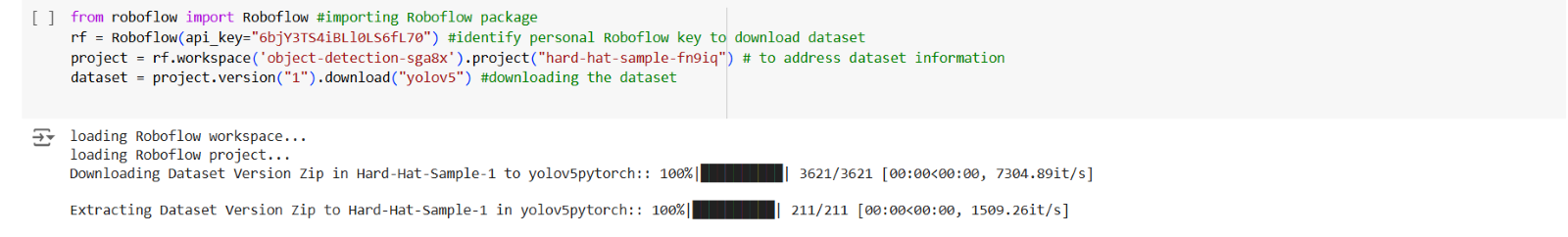
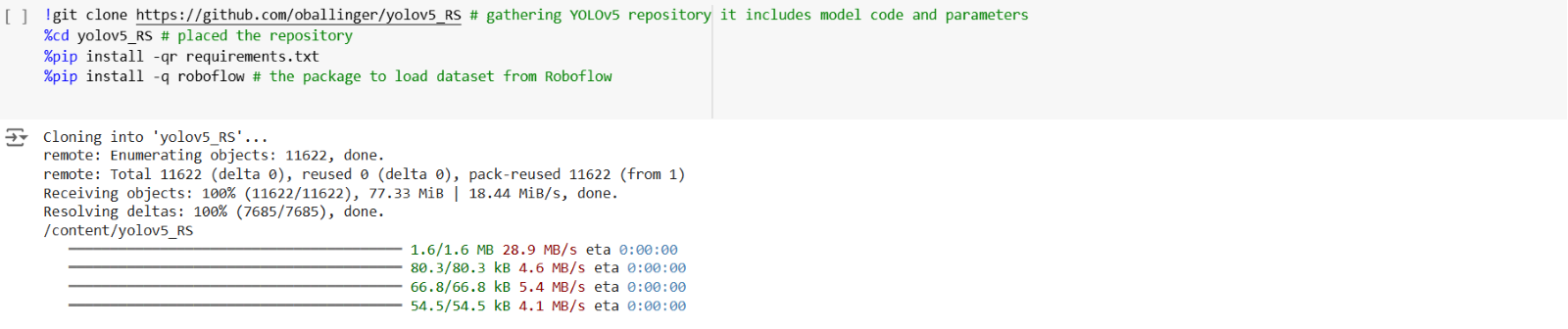


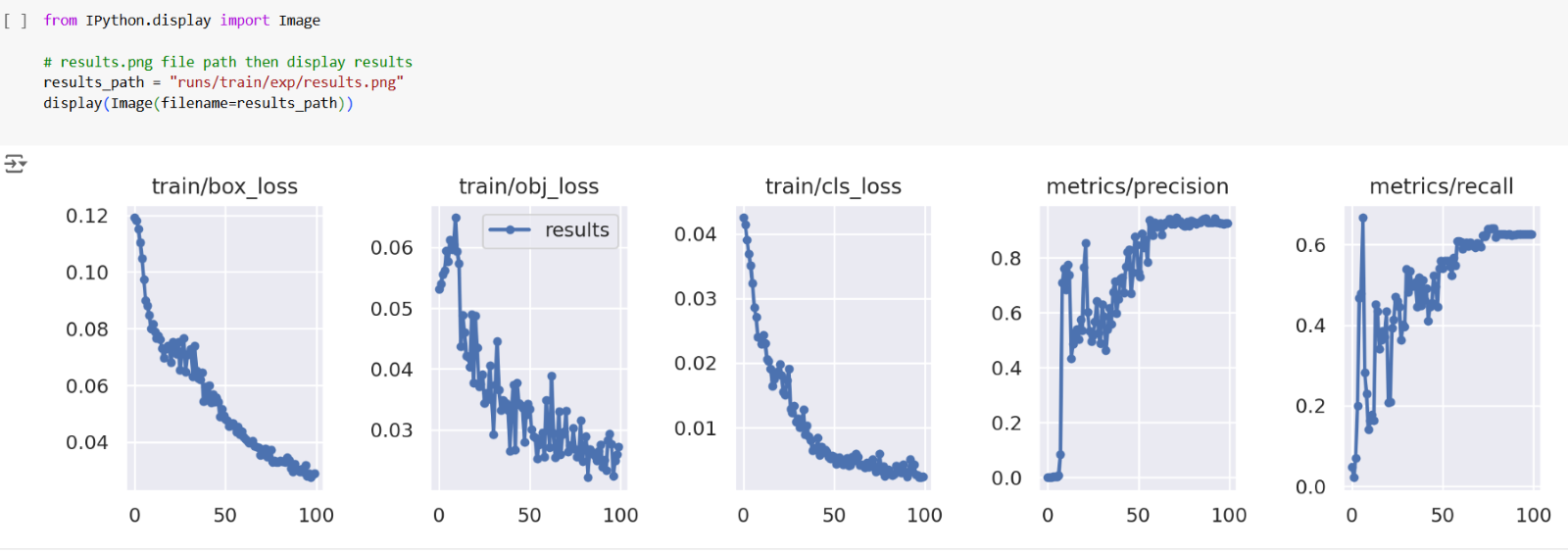
Figure 4.7 Benchmark Results ( Zhao , L. et al. , 2023 )

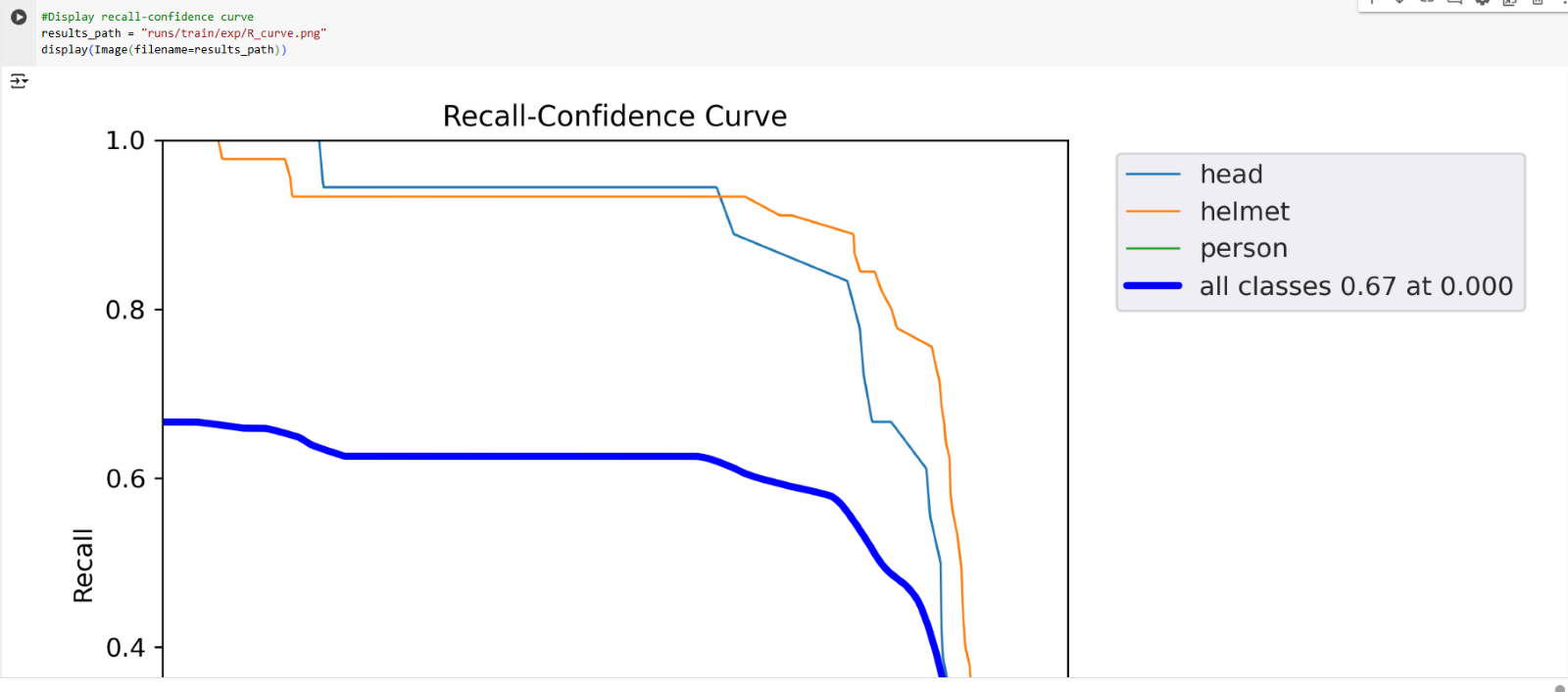
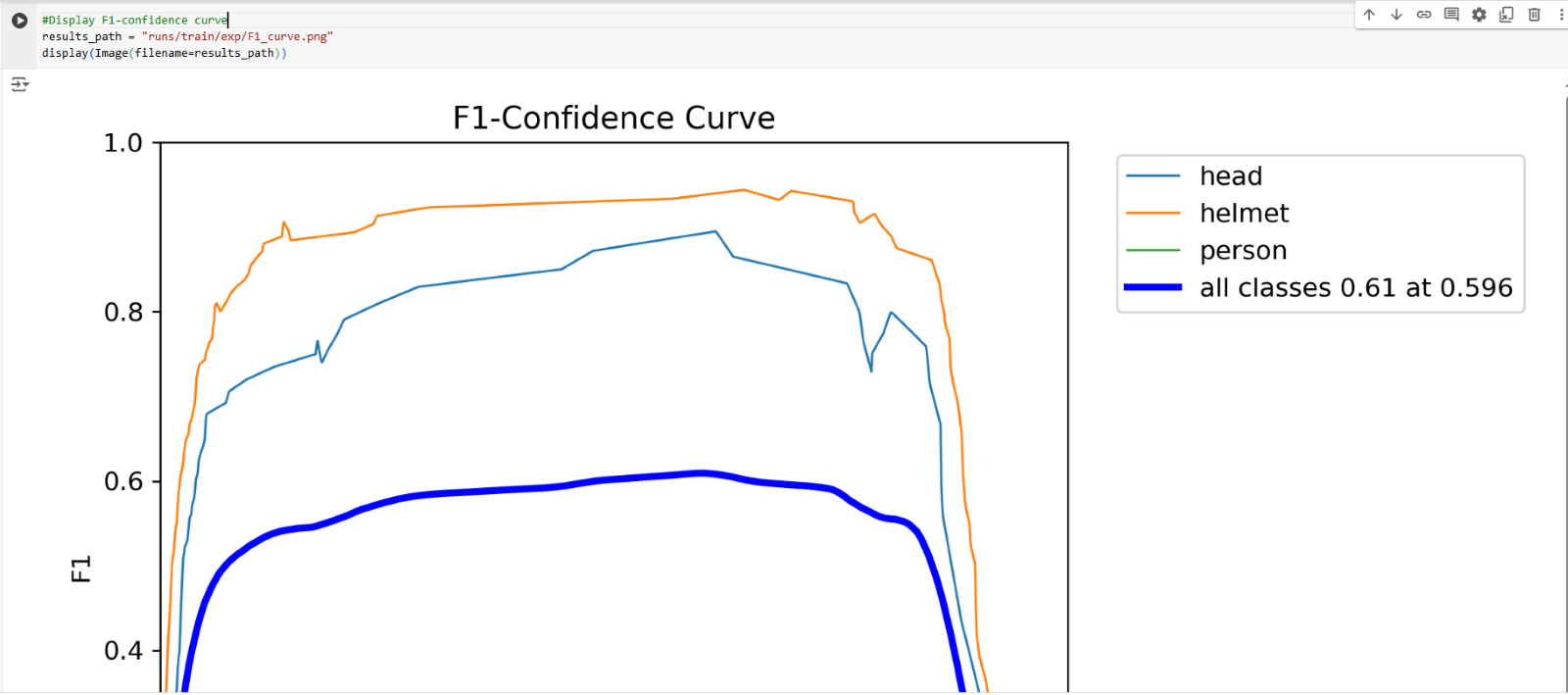
* Zhao , L. et al. have been introduced to a study which aims to identify weared helmet persons who are working in construction site.Also this study argues that exisitng studies have lack of robustness across complex real life scenarios to detect helmets then overcome to this issue they provide two stage solution.First one is adding one more layer into YOLOv5 architecture after that second stage is using Bidirecitonal Feature Pyramid to reducing detection error rate of the model.According to results scientist are able gather satisfactory results based on their aim becuase precision and recall are able to increase over time as continous as possible.
* Comparision between other two studies and this study others are able to demonstrate better results because this situation might be happen regarding to dataset quality and hardware capacity.

# **CHAPTER FIVE Practical Application**

## **5.1 Snippets of Python Code**

Figure 5.1 Snippets of Python Code





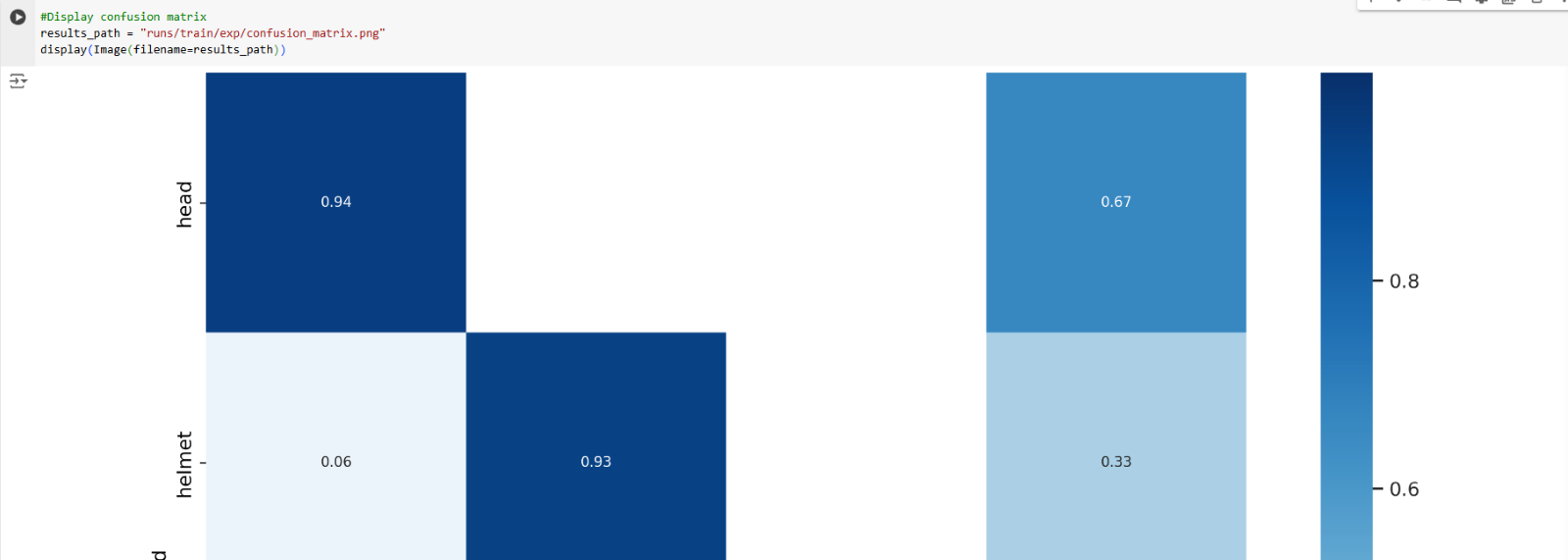


Figure 5.2 Snippets of Python Code

## **5.2 Potential Impact**

Potential usage area of this study can be placed in mining , firefightining , construction , supply chain , drivers (bicycles or motorbicycles) , industrial zones and so on.For instance this type studies can be enchance to work safety especially for third world countries.On the other side this technology might be high cost for corporations at the same time they are life saving in terms of legal sanction and people life since the system can proof the situation when accident happened to courts.Another giant impact might be happen in environmental disaster as understnad from the study the system is able to identify helmets as well as people.In that case differentiating civiliant people and rescue teams during any disaster is importnat for time saving.

# **CONCLUDING REMARKS**

In this study an object detection project is handled by utilized YOLOv5 architecture.In this context firstly limitations traditional computer vision models and role of Faster R-CNN – YOLOv5 architecture are handled then the selected dataset collected – analyzed and preprocessed which includes normalizaiton , augmentation and resizing for training.Before traning part YOLOv5 architecture and its capability are explained.Afte that training step is realized next at end of the day performance results are gathered extracted and the model is able to demonstrate satisfactory results.Afterwards the evaluation metrics are compared to toher studies in terms of accuracy and robustness.Moreover at final stage potentical effect in real life is discussed elobrated for corporaitons and people.

In the future these type systems are able to developed for higher accuracy and faster verison but additionally some automated systems can be integrated into computer vision architecture.For example when a person don’t use helmet within in a construction site in that case the system must have capabiliity to inform executives of the company , specialist of work safety and police or in that case the system is able to prevent to use any mechanical device within the construction site.

# **BIBLIOGRAPHY**

Cheng, G., Harmel, R.D., Ma, L., Derner, J.D., Augustine, D.J., Bartling, P.N.S., Fang, Q.X., Williams, J.R., Zilverberg, C.J., Boone, R.B. and Yu, Q., 2022. Evaluation of the APEX cattle weight gain component for grazing decision-support in the Western Great Plains. Rangeland Ecology & Management, 82, pp.1-11.

Khanam, R. and Hussain, M., 2024. What is YOLOv5: A deep look into the internal features of the popular object detector. arXiv preprint arXiv:2407.20892.

Jia, W., Xu, S., Liang, Z., Zhao, Y., Min, H., Li, S. and Yu, Y., 2021. Real‐time automatic helmet detection of motorcyclists in urban traffic using improved YOLOv5 detector. IET Image Processing, 15(14), pp.3623-3637.

Mahaur, B. and Mishra, K.K., 2023. Small-object detection based on YOLOv5 in autonomous driving systems. Pattern Recognition Letters, 168, pp.115-122.

Tong, K., Wu, Y. and Zhou, F., 2020. Recent advances in small object detection based on deep learning: A review. Image and Vision Computing, 97, p.103910.

Redmon , J. et.al. , 2015 You Only Look Once: Unified, Real-Time Object Detection Dataset available at <<https://bellingcat.github.io/RS4OSINT/C5_Object_Detection.html> > [Accessed 23 October 2024]

Yao, Z., Cao, Y., Zheng, S., Huang, G. and Lin, S., 2021. Cross-iteration batch normalization. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 12331-12340).

Zhao, L., Tohti, T. and Hamdulla, A., 2023. BDC-YOLOv5: a helmet detection model employs improved YOLOv5. Signal, Image and Video Processing, 17(8), pp.4435-4445.

<https://github.com/TuranCANG/Turan-Can-G-n-BSBI-Projects.git> (codes are available there)

# **APPENDIX (if necessary)**

BCE : Binary Cross Entropy

CIoU : Complete Intersection over Union

CNN : Convolutional Neural Netowrk

DL : Deep Learning

GıOU : Generalized Intersection over Union

mAP : Mean Average Precision

ML : Machine Learning

YOLO : You Look at Once