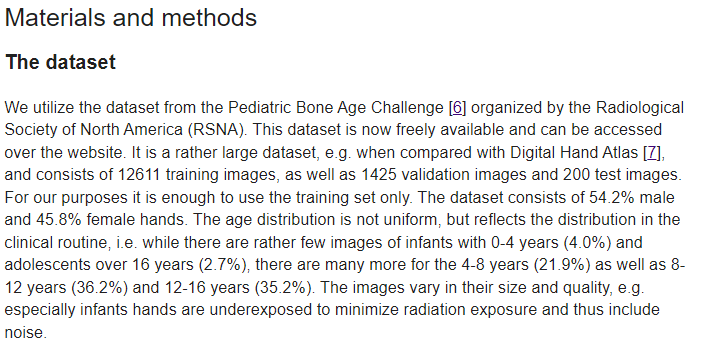
Data Analysis Report capstone project.

Introduction

This project is based on object detection where the bones and joints are detected and from that we will recognize the age of the child whether the child is healthy or not in other words the bones of the child are grown as they need to be or not.

For this purpose, we will be using yolov5 framework and link for setting everything is given below.

**Object Detection of Joints on the Bone age Dataset with YOLOv5**



This piacular paragraph is referred by the research paper

Reference of that is given below with already research which is going on and now further improvement on research.

Reference:- <https://doi.org/10.1371/journal.pone.0207496> by S. Koitka, A. Demircioglu, M.S. Kim, C.M. Friedrich, F. Nensa.

The goal of this project is object detection by using bb boxes and then applying various models to find the best model to detect the age of the bone of the child to know whether the child is healthy or not based on the bones whether the child is grown properly with this purpose first we have xrays images we will draw boxes and understand the joints properly. Trained to validate and apply different models to find the best output we can.

Input:-

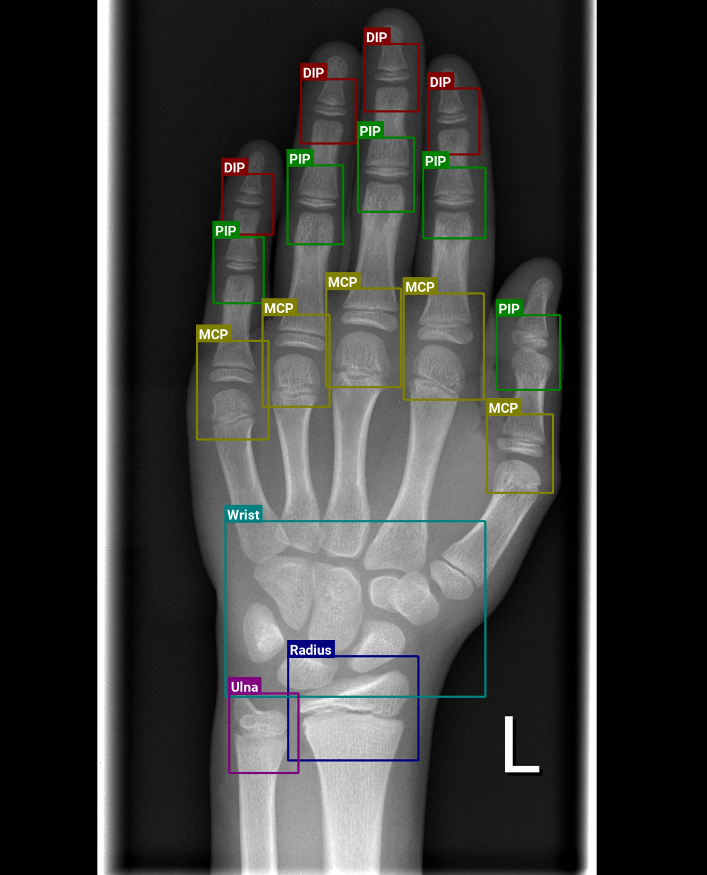


Goal:-

To have BB boxes as shown below with a good model to analyze or predict data analysis.

Conclusion and Analysis results are shared below to find the best model.

Each and every steps with the input results are shared below.

Goal

To perform the following project yolo5 framework has been used.

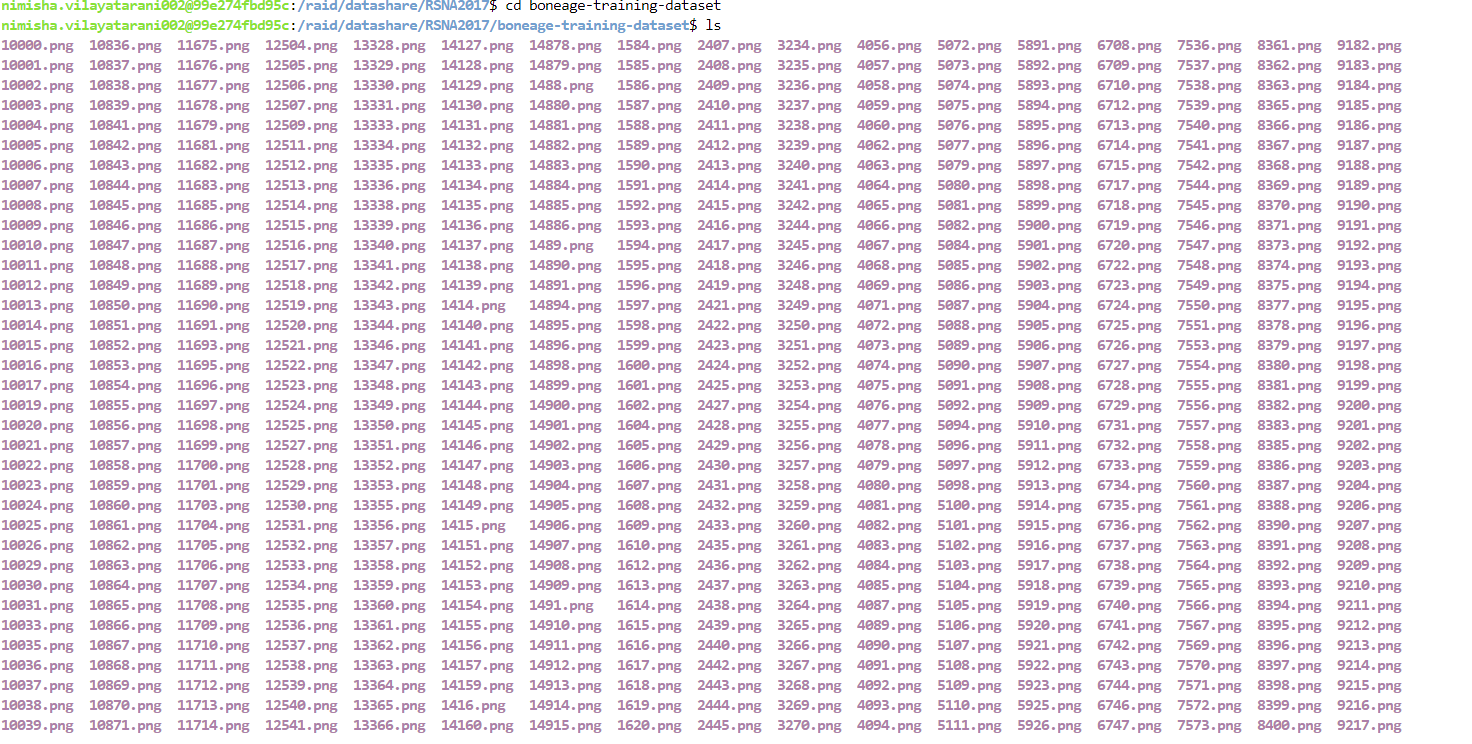
The data set for BB-Boxes has been stored in CSV file format.

The data set for validation and training has been provided in CSV format. Earlier the dataset was the Xray images from which the BB-boxes was maintained by using the yoloframework this has been stored in csv file format.

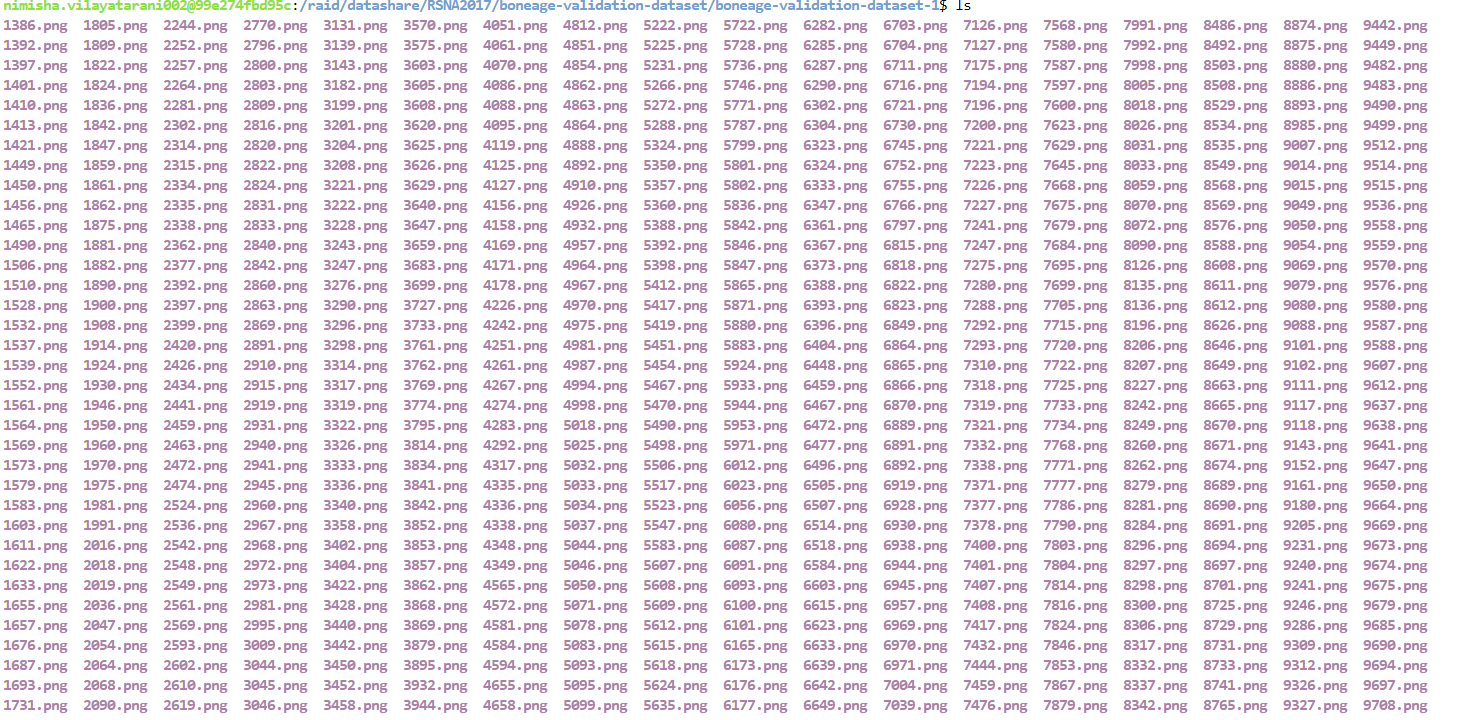
All the dataset and everything is given for the checking purpose.

**Step 1:**- Preparation for Dataset which is images.



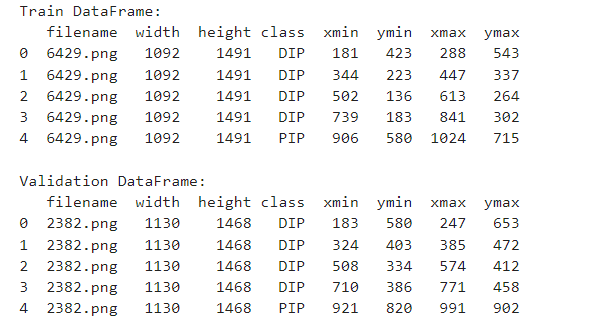


**Step2:-** Dataset for validation

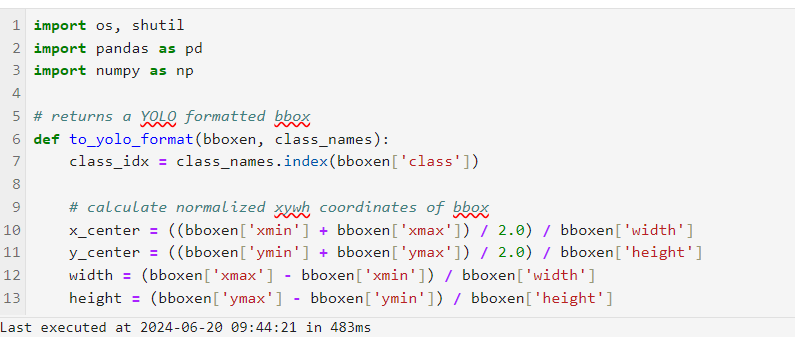


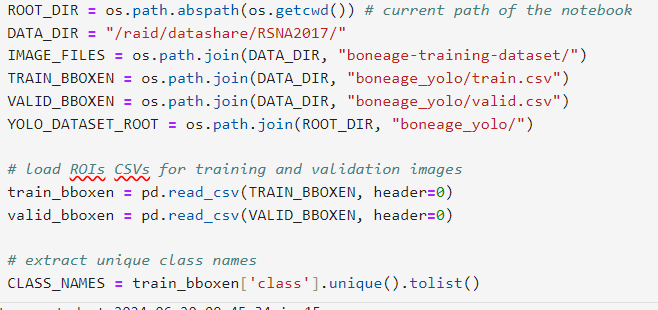
Data representation in yolo5.

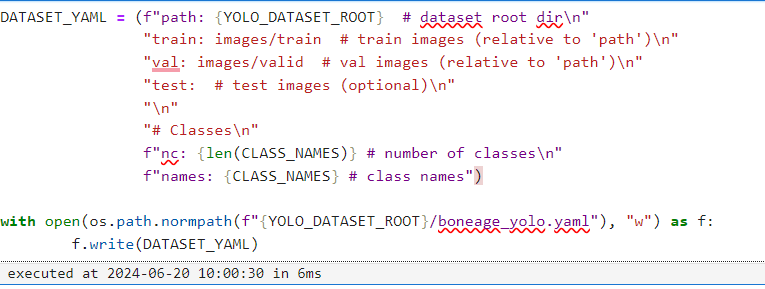
The column filename refers to the underlying X-ray image with width as image width and height as image height. The columns xmin, ymin, xmax, ymax define the corner pixels of a bounding box (BBox) of the X-ray image with class as the joint class. Multiple bounding boxes (BBox) are possible per X-ray image.



**Step 3**:- Code for conversion in yolo5 format is:-

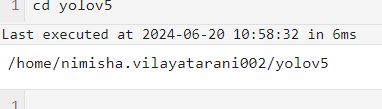




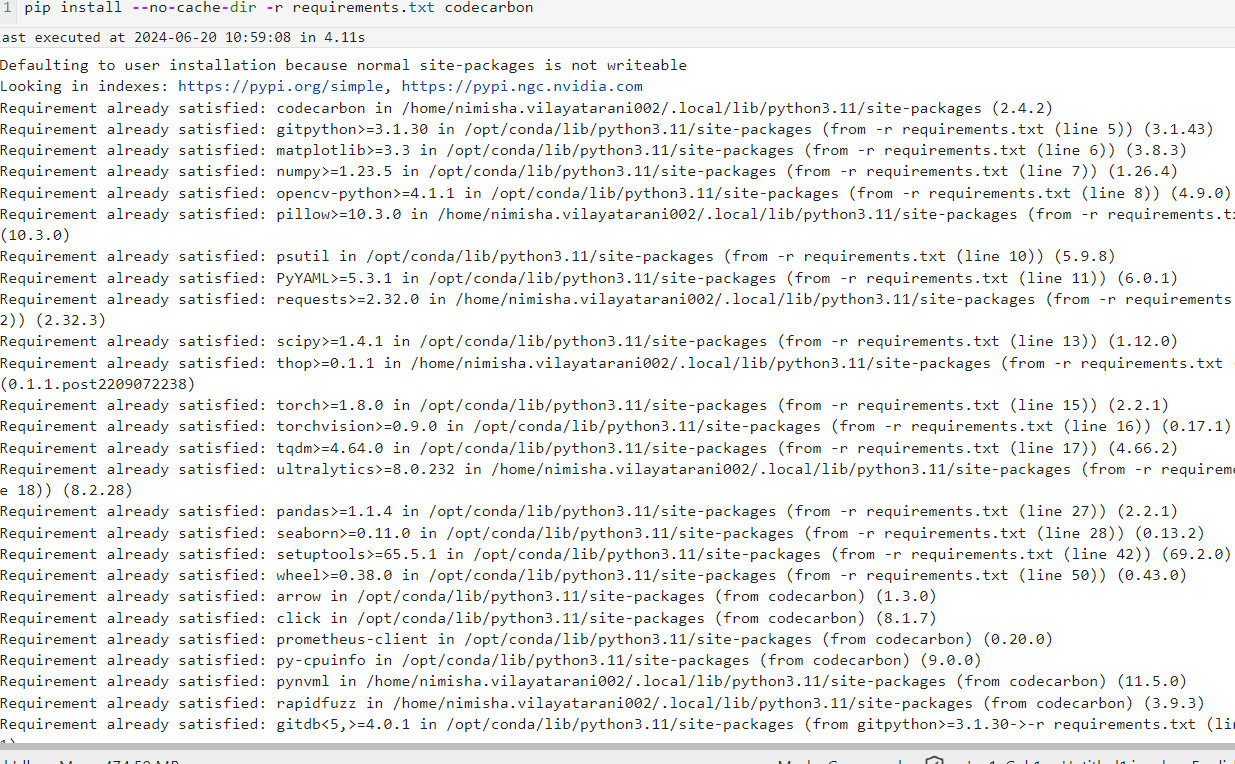


**Step 4:**- just for ease create a button to code on the terminal by just one click we can go directly go to the terminal.

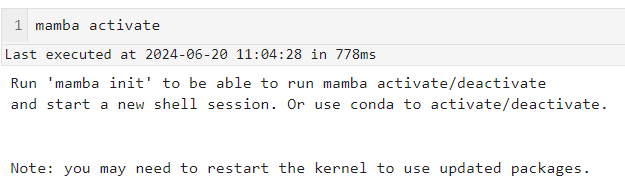




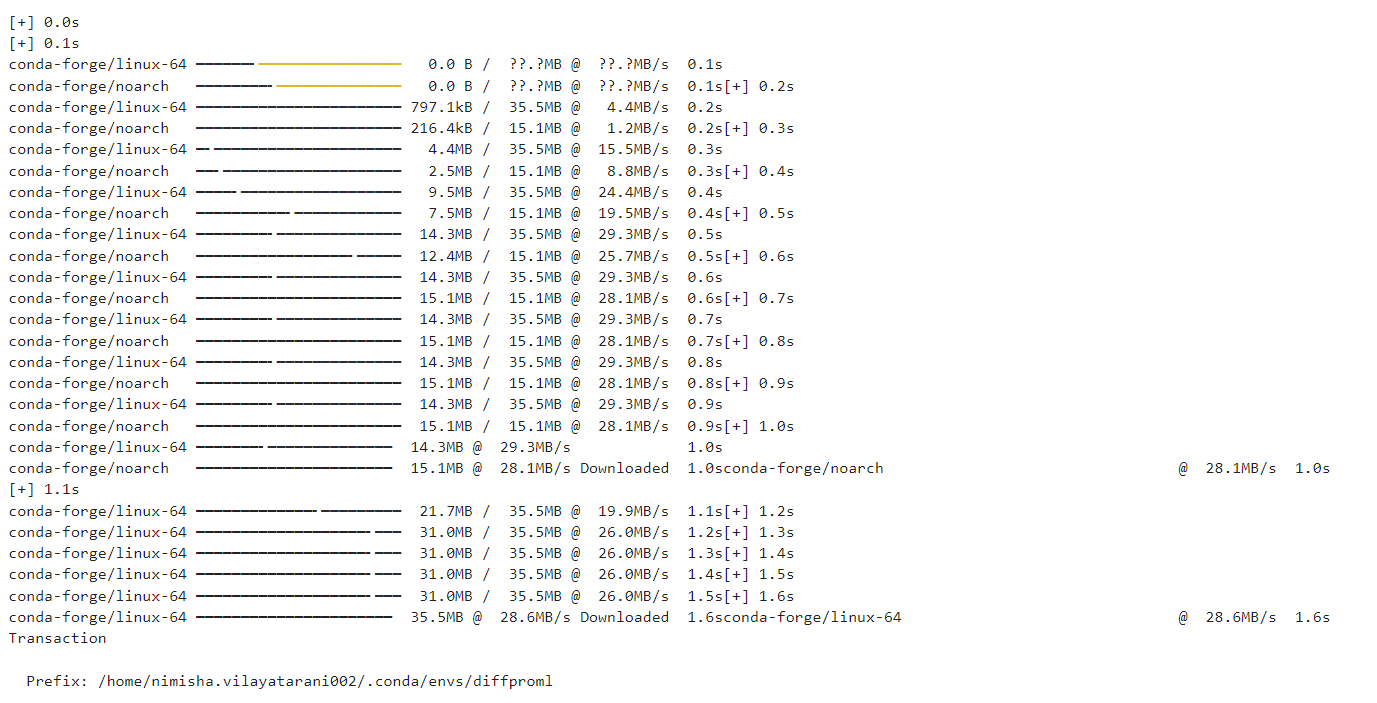
**Step:5** setup yolo5 frame work

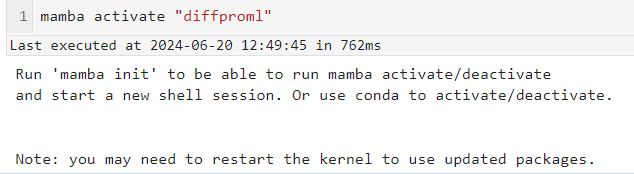


**Step:6** To install all dependencies for running it separately on the Anaconda environment.









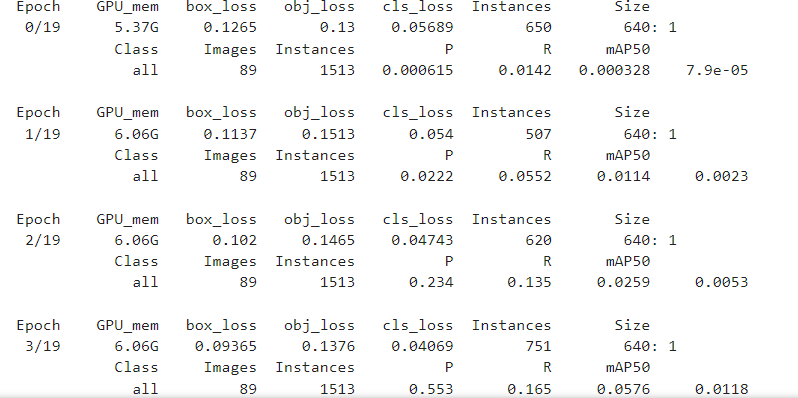
# Step 7 :- YOLOv5 Model Training

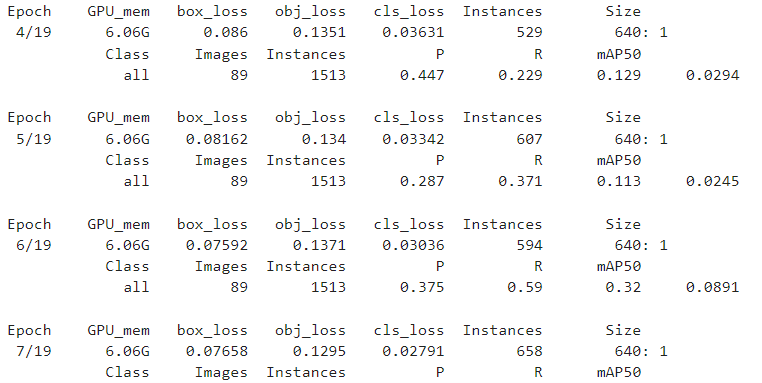
With !python train.py the training script of the YOLOv5 framework can be called. Here are some parameters to define.

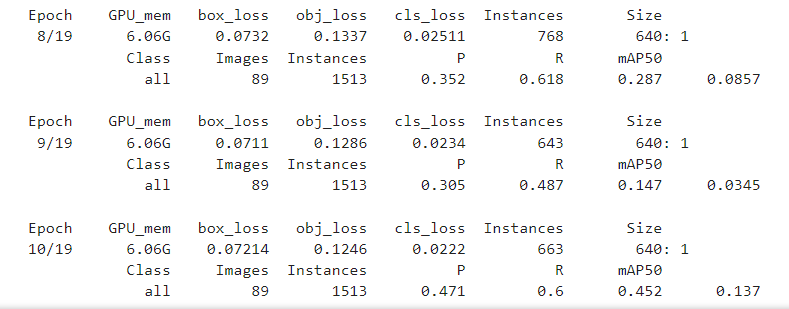
* --img: Set the image size
* --batch: batch size (please do not change, because GPU resources are shared between users)
* --epochs: number of training epochs
* --data: path to the YAML configuration file of the dataset
* --weights: path to pre-trained model weights (see picture for different model sizes)
* --hyp: path to YAML configuration file with additional training hyperparameters

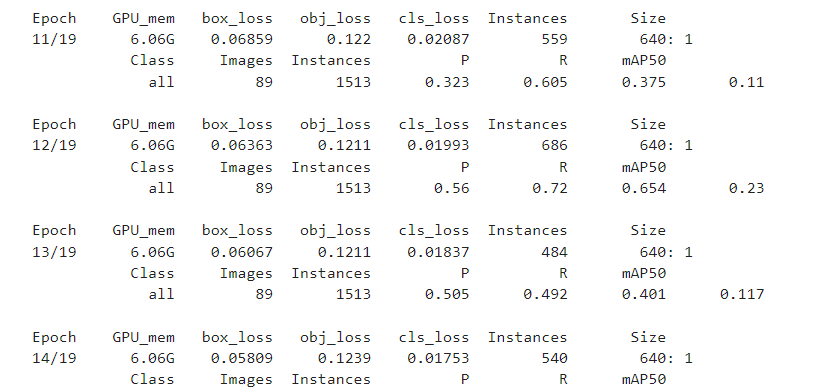
`input code for training the yolo5 

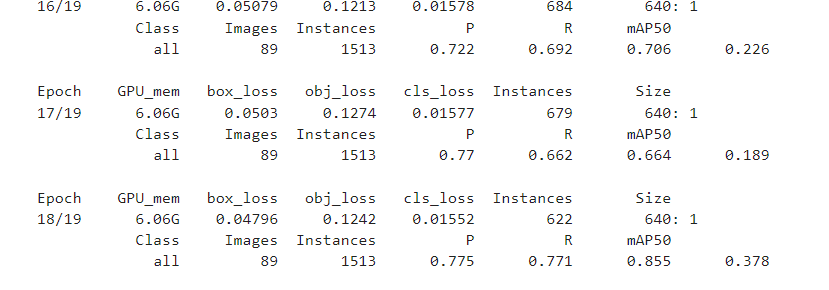
Output for different Epoch values as result of changing the number of epochs so as to provide better training to the model.







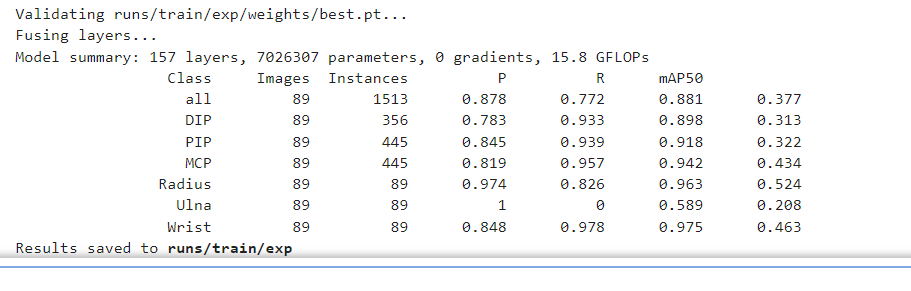




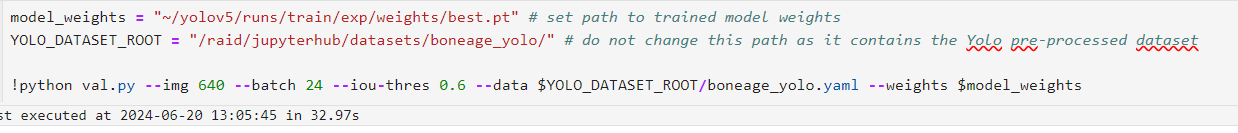
**Step 8** :- Model results

Here as with xray image we make the BB-Boxes from the xray image we train the model for providing the best we have then store the output yolo5 format which is csv the dataset which is been provided to us is .

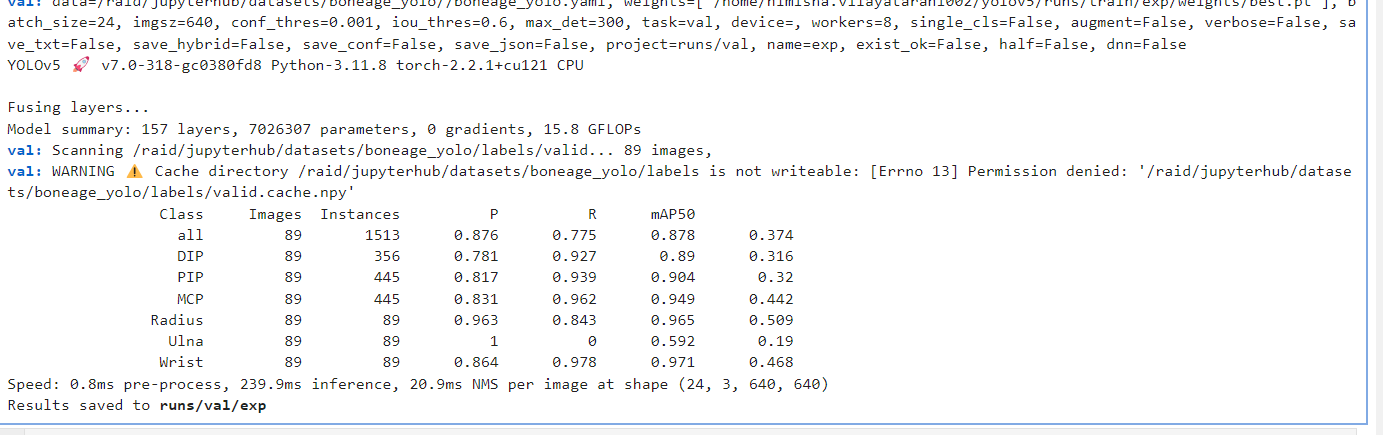
* distal interphalangeal joints (DIP, below in green).
* proximal interphalangeal joints (PIP, below in turquoise)
* metacarpophalangeal joints (MCP, below in pale green)
* wrist
* ulna
* radius

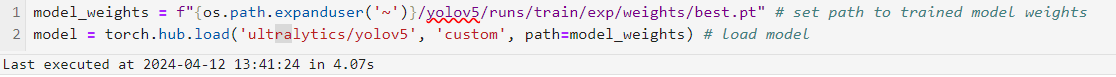


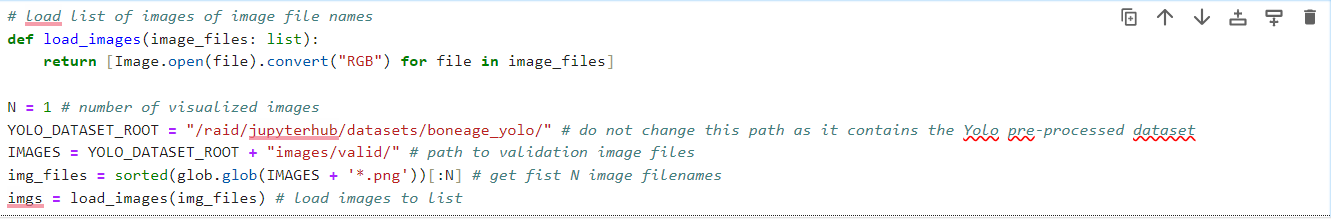
**Step 9:-** Loading some images of the validation dataset. The N parameter can be used to specify the number of validation images.

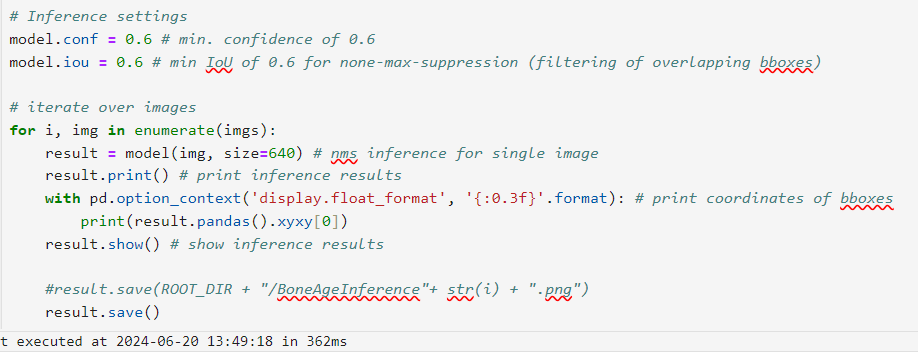


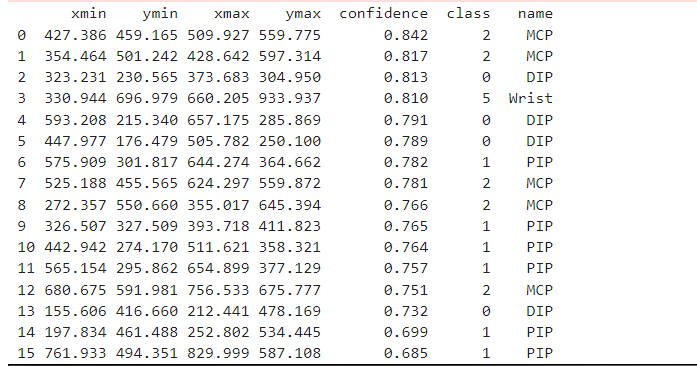
**Step 10:**- visualization we will able to see boxes on the x-ray images which would work as input for our regression model further analysis now data we will use this for further analysis and finding the best model for finding the bone age of a child to know by seeing the bone age whether the child is healthy or not is that bone grows as per the age.

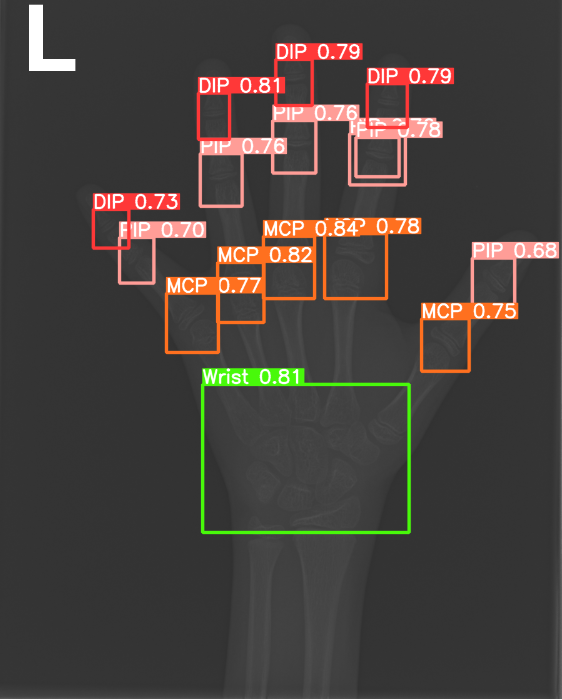




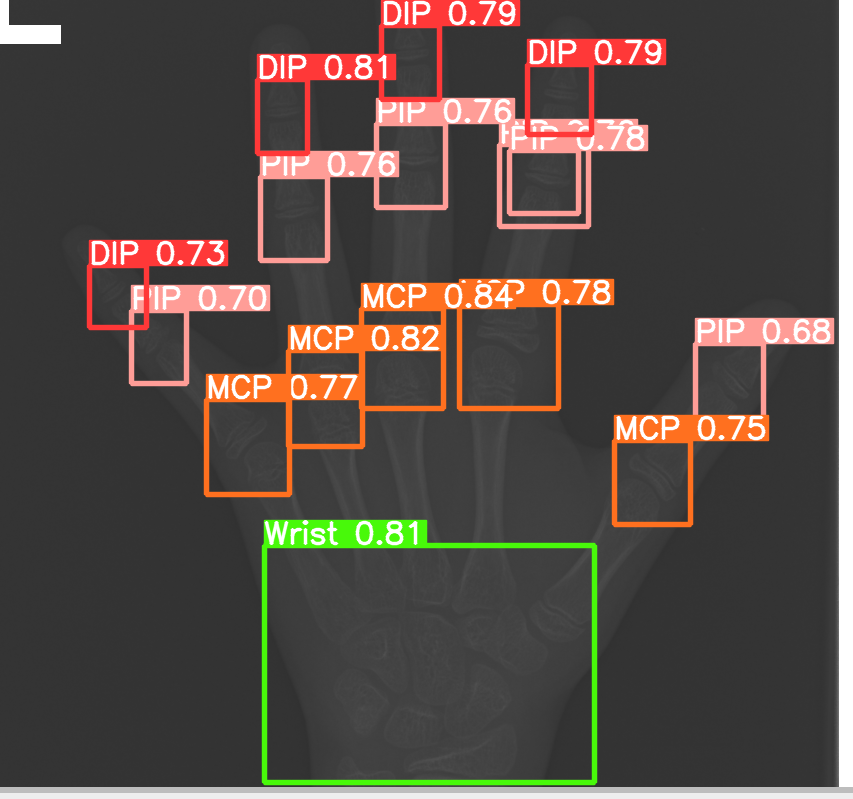








From the x-ray, we were successfully able to draw BB-boxes and convert our dataset into the yolo5 framework.

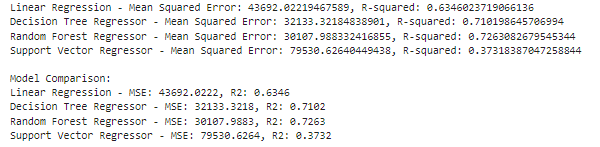


**Step 11**:- perform the four different models such as Linear Regression, Decision Tree Regressor, Random Forest Regressor, and Support Vector Regressor. And then we tried to find out which is the best model based on Mean Squared Error, R squared.





Output



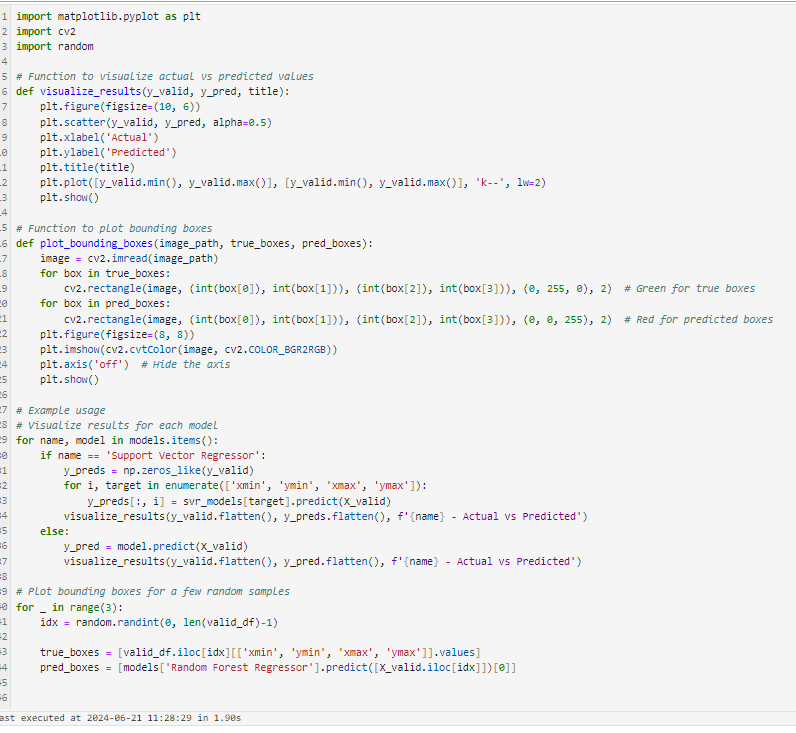
Output in tabular form

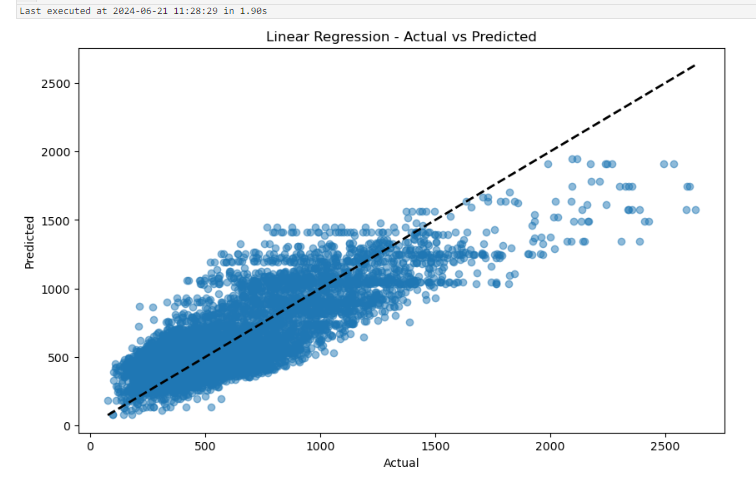
|  |  |  |
| --- | --- | --- |
| Model | Mean Squared Error | R-squared |
| Linear Regression | 43692.02 | 0.6346 |
| Decision Tree Regressor | 32345.35 | 0.7079 |
| Random Forest Regressor | 30079.01 | 0.7262 |
| Support Vector Regressor | 46071.58 | 0.6204 |

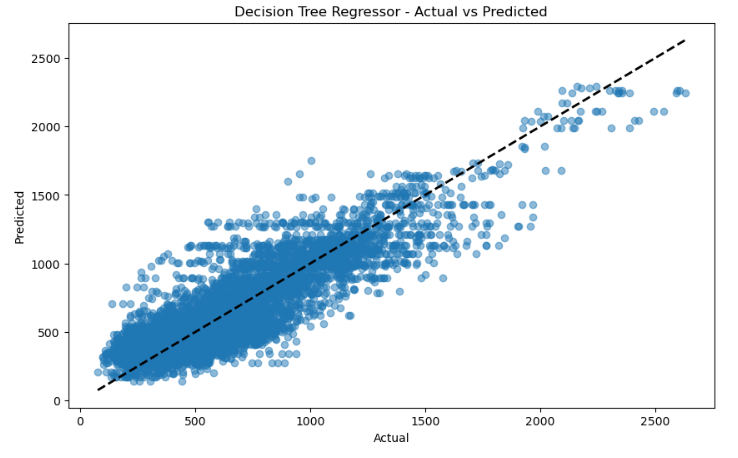
Through this, it is easily understood that data tree regression model provide the result and the reason for this is the minimum value of mean squared error which means that less possible chances to make errors and a high value R squared which is equally good.

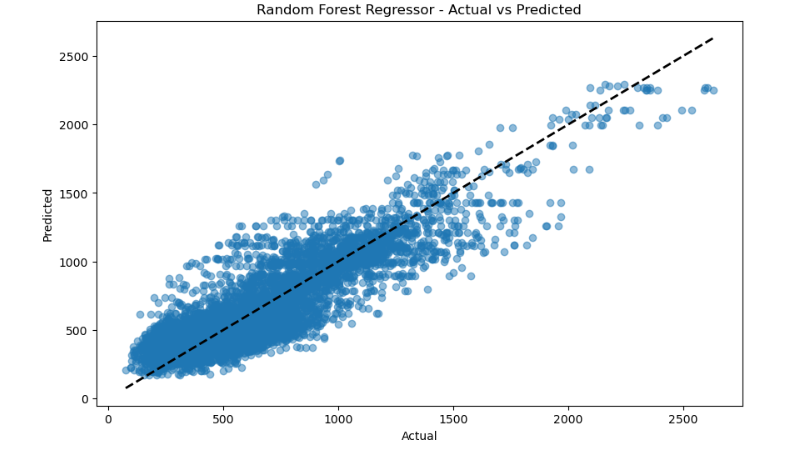
**Step 12**:- Here is the graphical representation through which we can visualize the result in this we can see that which model performs well on the basis of predicting actual value and final value.

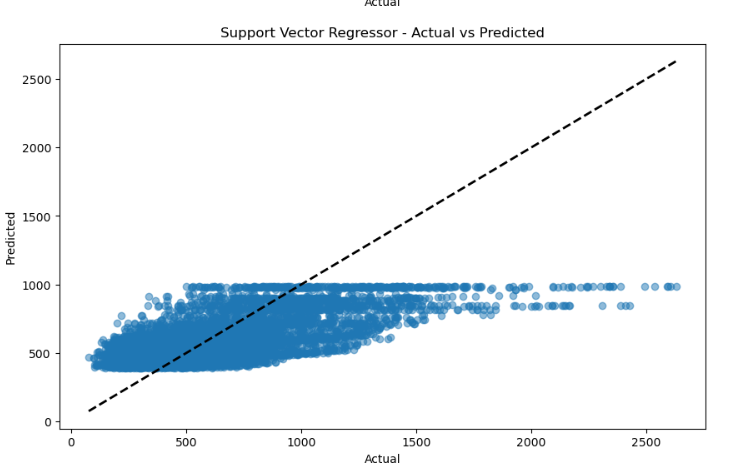
Here is the input code and output graph.











**Conclusion:-**

In this project, we explored various regression models to predict bone age from X-ray image data. The dataset consisted of annotated bounding boxes around specific bone regions, which were used to derive features for training and evaluation.

Four regression models were evaluated:

* **Linear Regression**
* **Decision Tree Regressor**
* **Random Forest Regressor**
* **Support Vector Regressor**

Each model was trained and evaluated using metrics such as Mean Squared Error (MSE) and R-squared (R2) to assess their predictive performance. Based on the results, the **Random Forest Regressor** emerged as the top-performing model, demonstrating the lowest MSE and highest R-squared among all models tested. This indicates that the Random Forest model not only minimized prediction errors but also explained a significant portion of the variance in bone age predictions.

The success of the Random Forest Regressor can be attributed to its ability to handle complex relationships and interactions in the data, which is crucial in medical image analysis where non-linear patterns may exist. Additionally, Random Forests are robust against overfitting and are known for their generalization capabilities.

Moving forward, further optimization of the Random Forest model through hyperparameter tuning and potentially exploring ensemble methods could enhance its performance even more. Moreover, deploying the model in a real-world setting would involve considerations such as scalability, interpretability, and integration with existing medical imaging systems.

Overall, this project highlights the effectiveness of machine learning techniques in predicting bone age from X-ray images, paving the way for enhanced diagnostic tools and clinical decision support systems in radiology.