Digit Recognition From BGMI Result Through Segmentation And OCR Using Yolov8

Badal Barahate, Mosin Ibrahim Hasan, Bhavesh Ashokbhai Tanawala

Computer Engineering Department, Birla Vishvakarma Mahavidyalaya Engineering College

Abstract- In recent years, numerous efforts have been made to shorten the time needed for doing manual work by automating with the help of machine learning and Artificial Intelligence, which are related to postcard, making tables and many other aspects. Esports is growing at a rocket speed and tournaments are organised for which result of matches needs to be made to evaluate points of individual players or teams. Our research focuses on automating table making through feature extraction for which the model extracts the score of players from the screenshot of current most popular game in India, BGMI. We prepared the dataset related to the games having 400 screenshots from the players results. We extracted ROI from the screen shot using Yolo. Once we separate the needed part out, we utilize the OCR that is also based with YOLO. Once we get the needed digits from the image, we will get it as a table having 20 team's score. The project hence would be able to generated a result in 98.3% times the manual labour.

Index Terms- BGMI, OCR, Segmentation and Yolo

I. INTRODUCTION

The Battle royale being an innovative and realistic game mode has brought together a lot of people indulge in online games where many teams drops from a single plane gathering loot escaping enemies moving into play areas and defeating other teams to claim the ultimate title of the last man standing. One such game which changed the battle royale seen is PUBGMOBILE / BATTLEGROUNDS MOBILE INDIA which was released on 18 march 2018 allowing people to communicate with their teammates and enemies through voice chat feature. Giving a very comfortable environment for player to come together with friends or people online to tea up and play together hence creating a great platform for esports scenario for the game.

Current point table making of the results in esports community is done manually and it takes 4 to 5 minutes to create a single match result considering more than 10000 matches takes place on daily basis in BGMI esports community it is a waste of human hours

Figure 1 Annotation for model training through CVAT

which can be automated to an extent which will help in making those results pretty quickly. Since this area of industry is a bit separate from main stream research and development so to find data set for the model was quite a task so created a whole dataset on my two devices which included iPhone12 and iPad mini-5. The project will go with two major steps one where we will divide the image into section and in the second part, we will predict the numbers having labels "killpoint" and "rank" were both the model will work with Machine learning models. Killpoint is referred to as the eliminations the players got and rank is the rank the stood in the match as a team (Nashshar, Iswara, Liliana, & Nalawati, 2023), they get varying points as shown in figure30.

For the annotation part to get the model trained for segmentation CVAT is used as shown in figure1. CVAT provides a variety of annotation tools to meet the needs of different tasks, including bounding boxes, polygons, tags, and labels. A collaborative annotation workflow: CVAT supports a collaborative annotation workflow, allowing multiple users to work on the same dataset simultaneously. A variety of export formats: CVAT supports a variety of export formats, making it easy to share your annotated data with others.

Deep learning is a part of Machine learning where the aim is to develop models that resemble human behaviour. Here models of deep learning are trained to locate the ROI in the image with the custom dataset and once the segmentation model provides ROI it is further passed on to OCR model which is based on deep learning too to read the text in ROI.

II. SIMILAR RESEARCH

In dynamic traffic regulation system members of school of Engineering and technology, Bangalore explained how using computer vision traffic could be monitored and SMS could be sent directly to the owner by extracting the number plate through computer vision techniques OCR YOLO where the used traffic cameras to get the information from the cars number plate deduced the personal information of the vehicle owner and send a direct SMS describing the violation owner performed and the fine associated with it. (Shreyas, Kumar, Adithya, Sunil, & Padmaja, 2017)

Id card extraction by members of Padjadjaran University and Telkom University another recent OCR based project and research work developed a residence identity card detection model through PYOCR, PYTesseract and TesseOCR and Natural language processing to improve the accuracy of the model to detect the information in 4510miliseconds to help reduce fraud in the area since a lot of fake identity was made and were used to person frauds in the area. (Irawan, Adhiguna, & Rusli, 2021)

III. Tools and Technology used

CVAT

CVAT, or Computer Vision Annotation Tool, is a free, open-source, web-based image and video annotation tool for computer

vision develop by Cavitation Technologies, Inc. It supports a wide range of annotation tasks, including object detection, image classification, and image segmentation. CVAT is

designed to be user-friendly and efficient, making it ideal for both beginners and experienced users. The tool would be used to create our own custom data set where we would be annotating the region of interest using CVAT as our tool. (CVAT, 2023)

YOLO

YOLOv8 is the latest version of YOLO by Ultralytics. As a cutting-edge, state-of-the-art (SOTA) model, YOLOv8 builds on the success of previous versions, introducing new features and improvements for enhanced performance, flexibility, and efficiency. YOLOv8 supports a full range of vision AI tasks, including detection, segmentation, pose estimation, tracking, and classification. This versatility allows users to leverage YOLOv8's capabilities across diverse applications and domains. A region proposal network is commonly used by well-known two-step algorithms like Fast-RCNN and Faster-RCNN to suggest regions of interest that may contain objects. (Nashshar, Iswara, Liliana, & Nalawati, 2023)
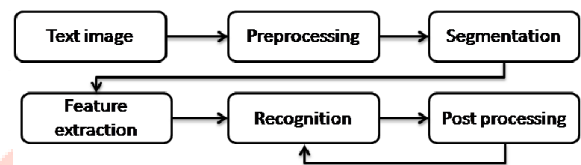

Figure 2 YOLO Work flow

IV. System

The project comprises of three major steps which includes getting data set ready to be used to train the model, training YOLOv8n model on our custom dataset to get ROI trained model, once the YOLO is ready to seperate the ROI from the image we tend to send those cropped image to the OCR model to read the text.

Figure 3 Work flow of model

IV.1 Dataset

There are not many datasets available for this purpose of object detection. There are several implementations for the object detection available and they use different datasets. Dataset for this project isn't available since it is a bit off to research purposes so it was made by us on two devices iPhone 12 and iPad mini-5 as shown in figure4 and figure5 respectively. Pixel size varying from 2400X1080 in iPhone and 2048X1536. Resolution in devices is 16:9 and 4:3 respectively. The dataset consists of 400 images out of which 350 are used for training, 25 are for testing and 25 images are for validation purpose. Sample os avavailable on Roboflow more data would be added as project progresses. (Roboflow, 2024)

Figure 4 Dataset image taken from iPhone12

Figure 5Dataset image taken from iPad mini5

IV.1.1 Annotation

Creating custom dataset requires annotation to be done on custom images. ROIs are being selected in rectangular shape from the image of different labels with the help of CVAT an innovational annotation tool which also provides AI based annotations. CVAT, or Computer Vision Annotation Tool, is a free, open-source, web-based image and video annotation tool for computer vision. It supports a wide range of annotation tasks, including object detection, image classification, and image segmentation. (Guillermo, et al., 2020)
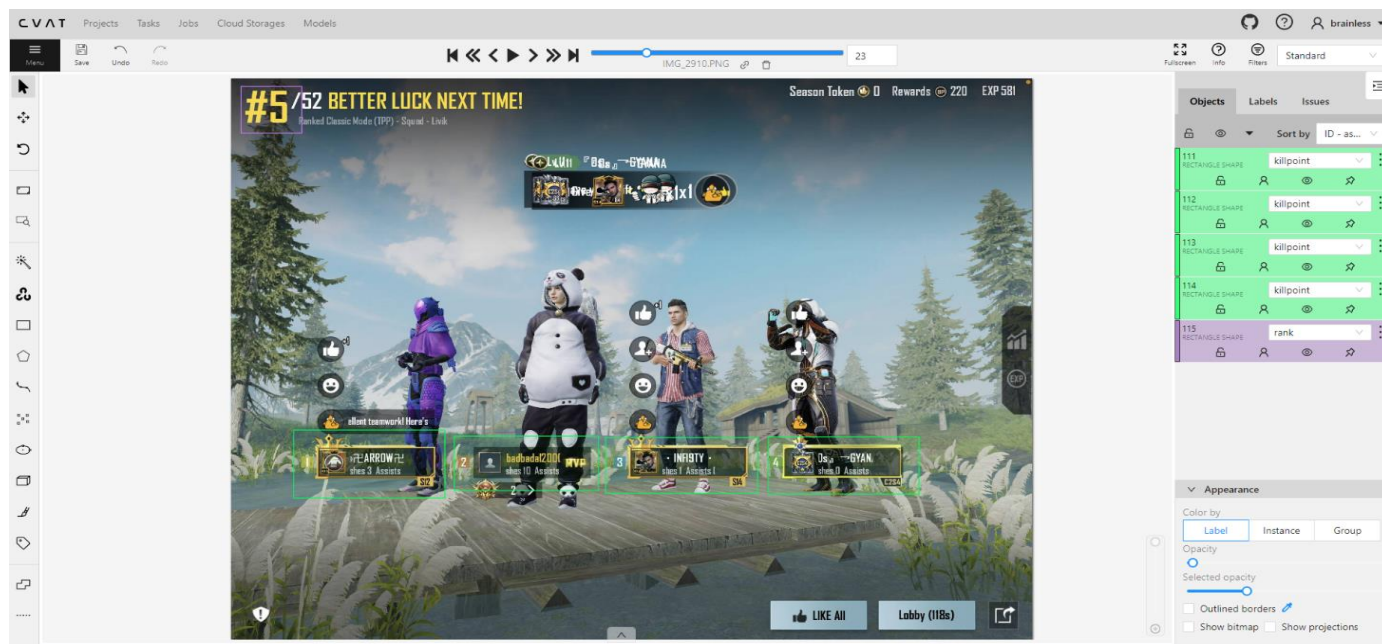
Figure 6 CVAT annotation

Figure 6 labeled information of location of ROI in the image

IV.2 Segmentation

Yolov8n model is feed with custom dataset with 2 labels "killpoint" and "rank" to learn where the ROI might be present in the image provided with details of labels to train the model.

Through the details in figure8 YOLO generates a dynamic filter based on the location of ROI to segment out the ROI. YAML file plays a vital role in models like YOLO since a lot of configurations are already set in YOLO and users are unable to make changes to them hence through YAML custom configuration could be applied based on our custom dataset. (Fu, Si, Liu, & Ma, 2022)

Figure 7 YAML configurations

Figure 8 Dataset sample for segmentation

Figure 9 Filter generated by the model

The YOLO model segments out the ROI to be used further as shown in the figure 2.

Figure 10.1 Rank region

Figure 11 segmented Rank point region

Figure 12 Kill point region

Figure 13 Segmented killpoint region

Figure 14 Segmented rank section

Figure 15 Segmentation for kill point for player1

Figure 16 Segmentation for kill point for player2

Figure 17 Segmentation for kill point for player3

Figure 18 Segmentation for kill point for player4

IV.3 OCR

OCR also referred to as Optical Character Recognition takes the ROI segmented by the model to read the text from it. Figure below shows the segmented images from our segmentation model from a different sample of our dataset.

From the segmented image as shown in above figure the OCR runs based on the label/class it has been specified and we get the result in text form as shown below with respect to their labels/classes with some minor changes in configurations. (Nashshar, Iswara, Liliana, & Nalawati, 2023)
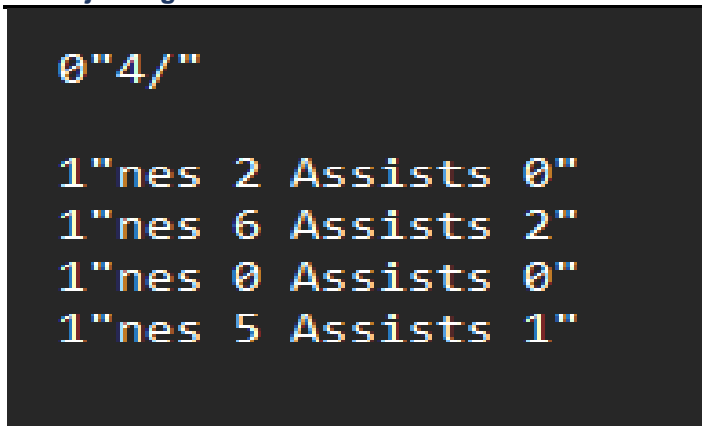
Figure 19 OCR result

V. Implementation

The flow of implementation starts with annotation of region of interest (ROI) in the image with the tool CVAT allowing us to train our model on the annotated images. Once we have completed training the model on our two classes "killpoint" and "rank" then we take 16 images at random decided by the model take a mirror image of it, then taking it upside down, a tilted image at 25 degree and placing the image slightly out of the frame to test the segmentation accuracy of the model trained and we get 100% accuracy.

Figure 20 Sample image IMG_2874.PNG

Figure 21 Result from segmentation through labels

Figure21 and Figure 22 shown how the segmentation model separates out the ROI from the given image based on labels which we have provided as 0 for rank and 1 for killpoint.

If we don't specify the labels to OCR model the result shown the test of the whole image results are shown below in figure 24.

But we only need the model to extract text from the region of our interest hence we specify the labels that we have segmented out from the image the OCR model provides us with the following result. Hence for the same image in figure 23 when we add our configurations specifying the labels we get the following result as shown in figure 25.
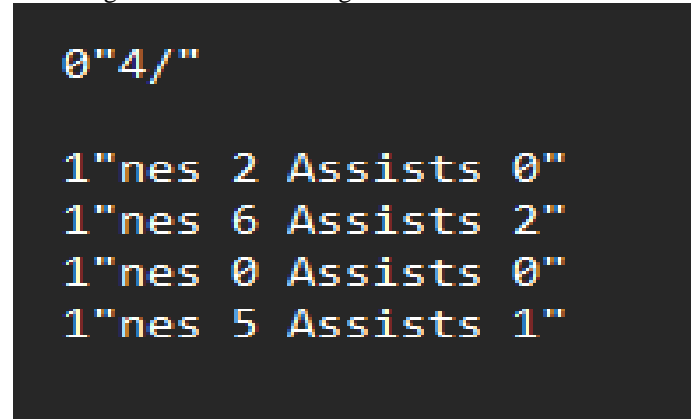

Figure 25 OCR results with configurations specified

Real time implementation of the OCR model is shown below where 6 team came to play a match and based on their screen shot a table has been generated by the system.

The result in figure27 is calculated based on the point system provided by the games publishers itself which is "Krafton"as shown in figure 28. Here we see that the players of team "Badesports" has uploaded their screenshot where they were placed #1 hence according to the point system 15 points has been allotted for the position and kill points amongst the team came out to be a total of 5 hence summing up to 20 as their teams score.

Figure 23 Data sample to test on OCR model with no configurations provided

Figure 24 OCR result when no label specific configurations are provided

Figure 26 Result screenshot from Badesports

No	Team Name	Kill Pts	Position Pts	Overall
1	Badesports	5	15	20
2	Yakuza	0	6	6
3	NeX	0	8	8
4	Bad	0	0	0
5	S7	0	0	0
6	Botteam	0	0	0

Figure 227 Final result after calculating results from each team

Rank	New
1	15
2	12
3	10
4	8
5	6
6	4
7	2
8	1
9	1

Figure 28 Point system provided by crafton in 2021

VI. Conclusion

There are major sectors were computer vision and computer optics could be used to improve the functioning of cause. YOLOV8n and OCR both model shows 100% present accuracy since this custom dataset has computer generated results.
The model will decrease the time consumed for making results in BGMI esports community from 5 mins to 5 secs. Considering it on daily basis on average 10000 matches takes place in BGMI esports daily considering 5 minutes to make result for each match manually it would take 50000 minutes which our model brings down to 50000 sec or 833.33 minutes hence reducing the time by 98.3%. YOLOV8n and OCR both model shows 100% present accuracy since this custom dataset has computer generated results.
We might see some issues with the model proposed like the players not providing relevant screenshot or a forged screenshot for which in future we can train the same model with the dataset of images from spectator's side of the match and we can cross check the information from both sides of images to eliminate any misconduct. Due to lack images from spectators the model couldn't be trained on this date but would go in future work.

VII. Acknowledgment

VII. References

[1] Dynamic Traffic Rule Violation Monitoring System Using Automatic Number Plate Recognition with SMS Feedback Sunil M P, R Shreyas, Pradeep Kumar B V, Adithya H B and Padmaja B from the Department of Electronics and Communication Engineering School of Engineering and Technology, Jain University, Bangalore, India. 2021 9th International Conference on Information and Communication Technology (ICoICT)

[2] Indonesian ID Card Extractor Using Optical Character Recognition and Natural Language Post-Processing Firhan Maulana Rusli School of Computing, Telkom University, Bandung City, Indonesia, Kevin Akbar Adhiguna Faculty of Mathematics and Natural Science, Padjadjaran University, Indonesia, Hendy Irawan School of Computing, Telkom University, Bandung City, Indonesia. 2021 9th International Conference on Information and Communication Technology (ICoICT)

[3] Vehicle Number Plate Recognition in Intelligent Transportation System using YOLO v8 and EasyOCR Rizki Elisa Nalawati; Dewi Yanti Liliana;Ratna Widya Iswara; Muhammad Dzaky Nashshar;M.S. Rumika Damayanti;Rahma Maulida Shaliha 2023 11th International Conference on Cyber and IT Service Management (CITSM)

[4] PSA-YOLO: License Plate Detection Method Based on Pyramid Segmentation Attention in Complex Scenes Tian Ma;Zhen Liu;Yuge Si;Chenhui Fu 2022 7th International Conference on Image, Vision and Computing (ICIVC)

[5] Implementation of Automated Annotation through Mask RCNN Object Detection model in CVAT using AWS EC2 Instance Marielet Guillermo; Robert Kerwin Billones; Argel Bandala; Ryan Rhay Vicerra; Edwin Sybingco; Elmer P. Dadios; Alexis Fillone 2020 IEEE REGION 10 CONFERENCE (TENCON)

[6] Implementation of Automated Annotation through Mask RCNN Object Detection model in CVAT using AWS EC2 Instance Marielet Guillermo; Robert Kerwin Billones; Argel Bandala; Ryan Rhay Vicerra; Edwin Sybingco; Elmer P. Dadios; Alexis Fillone 2020 IEEE REGION 10 CONFERENCE (TENCON)

[7] https://app.roboflow.com/badal-dandv/bgmiresult/1

[8] https://app.cvat.ai