Karuna sree
(TASK-0)

Object detection in Games

ABSTRACT

In this report, we explore the expediency of object detection as applied in games. Object detection has many practical applications concerning games, ranging from promising AI design to gathering data on user/player habits or in-game characters for competitive intents or advancement updates. Since the advent of Machine Learning, computers have even been able to identify patterns and make decisions with minimal intervention from humans. Ideally, machines increase accuracy and efficiency and remove (or greatly reduce) the possibility of human error.

Computer Vision has lately been earning a lot of popularity. Image classification and object detection are major subsets of Computer Vision that are being acclimated to fields such as medicine, engineering, and many more. Computer Vision research in games is presently lacking. There are a few studies that seem to focus particularly on image classification combined with reinforcement learning or Q-learning and creating agents that can play games. Other subsets of Computer Vision such as object detection don't seem to have been recognized by the gaming community much as a useful area of study hence in this report I choose to focus on object detection as a whole with several core classification concepts.

Problem Statement

The problem statement is to sample the practicality of object detection as used in games through a convolutional neural network detection model.
This supports the implementation of a real-time detection model to classify and chase video game characters. With this, one will be able to assemble a simplistic bot, qualified for movement established on tracked locations of a secondary character on screen. This uses labeled frame data from the game to perform basic actions in real-time.

Market/Customer/Business needs Assessment

This research will supply the gaming industry with some insights on object detection models to enhance their gaming capability. A study suggests that with the recent advent of the e-sports industry and the rise of competitive gaming, the live streaming of games will expand rapidly in the upcoming years. The study also signifies an expansion concerning the viewing audience for games. This heightened growth of competitive gaming unlocks new routes for Computer Vision study and the gaming industry. Instances of possible benefits of this kind of analysis contain: gathering data on player habits to find flaws or modifications, designing more complicated game AI to assist players to learn faster, or executing perfect balance updates.

Target Specification

The presented object detection model has several practical applications:
• It can be utilized in producing an adaptive AI. Regardless of casual or competitive gaming, it can assist the player to rehearse their fundamentals and improve.
• Detection blended with classification can be utilized to follow poses of the character in games permitting new players to understand button combinations.
• With gaming competitions on the rise, tracking objects in real-time authorizes for information extraction about characteristics of the player, the game, or even the character to find priceless improvements or defects in gameplay.

External Search

The sources I have used as references for analyzing the capability of a CNN to achieve higher accuracy predictions when using large datasets to learn from are mentioned below :

https://www.fritz.ai/object-detection/

https://paperswithcode.com/task/object-detection

https://machinelearningmastery.com/object-recognition-with-deep-learning/

https://developers.google.com/ml-kit/vision/object-detection

https://www.linkedin.com/pulse/how-gaming-industry-using-machine-learning-ai-ajay-pathak/

Benchmarking

CNNs are qualified to perform higher precision predictions when using enormous datasets to learn from. Image classification and object detection are well-known subsets of Computer Vision

that are being adapted to domains such as medicine and engineering. In the automotive industry, research has been suggested utilizing this in applications like pedestrian detection and road detection.
This method can execute a real-time detection model to compartmentalize and track video game characters. A simplistic bot can also be constructed which is qualified for movement established on tracked locations of a secondary character on screen. This bot can be trained utilizing labeled data to successfully perform basic actions in real-time.

Applicable Patents

Patent 1: Systems and methods for end-to-end object detection

Our day-to-day lives abound with instances of object detection. Region-based CNN (R-CNN) methods further improve the accuracy of object detection beyond FCN-based methods. Object detection is one of the core tasks in computer vision applications. Presented in this patent are systems and methods that provide a unified end-to-end detection pipeline for object detection that achieves impressive performance in detecting very small and highly overlapped objects. This patent can be significantly considered while designing and executing the model.

Patent 2: GAMING STATE OBJECT TRACKING

The above-presented invention relates generally to gaming systems, apparatus, and methods and, more particularly, to image analysis and tracking of physical objects in a gaming environment.
This patent can be referred to while analyzing the objects in a competitive game.

Countless patents can be examined. but since the patents mentioned on the top relate most to the Object Tracking in Games using Convolutional Neural Networks, I decided to mentioned them.

Applicable Constraints

● A dataset of labeled images should be readily available so no one has to manually compose it.

● Constant data collection and maintenance should be performed.

● Lack of technical knowledge for the players or end-users.

● Lack of computer vision research in the game industry.

● Imperfect detections and Liability.

● Requires a lot of research to build a promising algorithm in-order to accurately detect or track objects in images or games.

● Difficulty to choose among several applicable algorithms, therefore, trading the quality of the model.

Applicable Regulations

● Data security and privacy regulations(end users).

● Patents on ML algorithms designed.

● Patents on computer vision research developed in the gaming industry.

● Laws regulating data collection.

Business Opportunity

The global video-game industry has been on the rise ever since the invention of smartphones and video game consoles like the Xbox and the PlayStation. The gaming market is growing extensively in recent years and will most likely continue to advance. Experts anticipate that by 2025 it may become a $300 billion industry. As the number of gaming aficionados is boosting, companies functioning in the game market can utilize this opportunity to develop products to enhance customer experience. That is why AI and machine learning solutions are becoming popular topics in the gaming industry. Therefore, there is a golden chance of this technology potentially becoming a great business opportunity.

Final Product Prototype

The final product is a service that provides a custom convolutional neural network detection model that can be enhanced through classification weights developed from the pre-training on the suitable datasets. Then it can be trained on annotated frames from the competitive games to track the coordinated locations of the characters in real-time.

Object localization:

Object localization is equivalent to that of classification. Localization produces the classification task and generates a bounding box founded on the location of the classified object. Localization can be performed in two probable ways - with classification or classification and regression.

1) With the pure classification system, a sliding window would need to be used to travel across the image and classify each region.

2) classification and regression, This concept utilizes a CNN with a classification head and a regression head to anticipate the object's class and the bounding box of the object respectively.

When training this CNN, two types of input data are sent: images and ground-truth bounding box coordinates. The output result ends up being a classification vector and coordinates for the bounding box.

Detection:

The fundamental issue with localization is that it can only classify and predict boxes for one object at a time. In contrast, the goal of detection is to classify all object instances of every class in an image.
several algorithms can be used such as,
> R-CNN
> Faster R-CNN

YOLO Object Detection:

The YOLO algorithm outperforms Faster R-CNN in speed and accuracy. The fundamental concept behind the YOLO algorithm is that the input images are resized to a fixed size and then fed into a CNN

IOU:

IOU or intersection over union is a way to estimate the bounding box prediction for the ground-truth box. It is used in both object localization and detection.

Anchor Boxes:

Anchor boxes are predefined box shapes that are used to help grid cells detect multiple different objects. Each defined anchor box can detect one object per cell.
If an object is detected to be of a certain class, that value is set to 1 in the anchor box output for that grid cell, otherwise, it is a 0.

Loss Function:

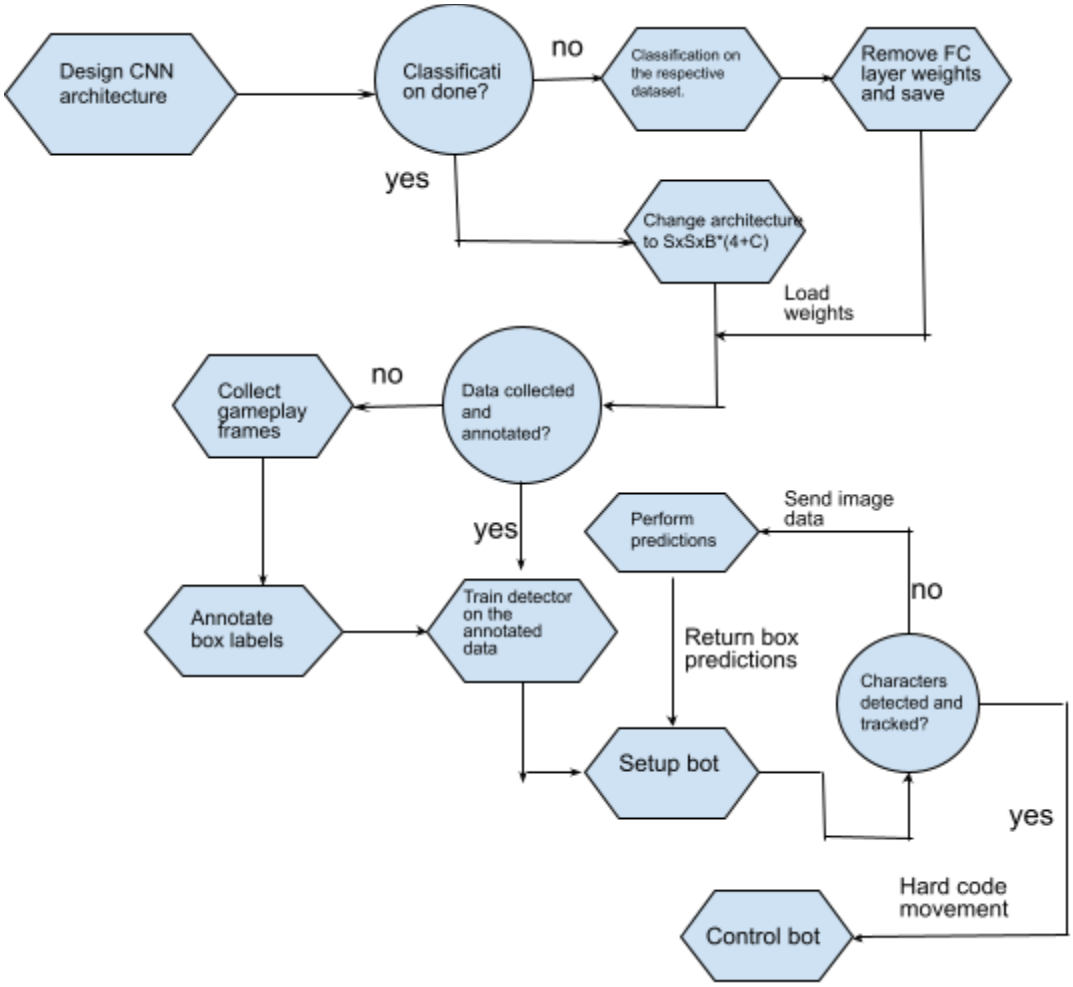
YOLO utilizes its own custom loss function the YOLO loss function contains the coordinates of the bounding box at the respective grid cells.
While other coordinates represent the confidence of an object present in a particular grid cell and the predicted class of the object.

Evaluation Metrics

The common evaluation metric for object detection is Mean Average Precision (MAP). MAP is the mean across the average precision of all classes. Average precision or AP gives an idea of how well the detection system is performing by analyzing the precision to the recall of the detections of each class.

system pipeline is divided into 3 chronological stages:

classification, detection, and the proof-of-concept bot. Figure 4.1 depicts a diagram of our various pipeline stages



CNN Architecture:

Choose a CNN architecture that meets the goals such as Portability and Parameter and Computation Limitations.

Classification Pre-Training:

Classification is the beginning of the first stage of our pipeline. We can execute this training to construct high-level feature extracting weights that are used during object detection.

Training Process:

The individual images and labels can be split into training and validation data for the objectives of training and evaluating the implementation of the model. Then the optimization algorithm is used permitting the model to learn from the data. Once training is completed, the final saved weights with the lowest validation loss are utilized in detection.

Detection:

As the input to the detection model, select the sized images. The prevalent goal is to detect the distinctive characters from competitive games, which can be classified into different classes. The purpose of detection as with classification is to decrease training loss and validation loss as the model learns

Data Collection:

A dataset of labeled images should be readily available

Training Process:

The input takes is a blend of images alongside the ground truth bounding box coordinates for each character present in the image. After the model learns from the batch of images, the YOLO loss function is used to calculate the loss, and the weights are updated by backpropagation.

Warm-up Training:

First is the pre-training step which interests loading the classification weights into the modified detection model and then training the model.

Post-Warm-up Training:

This training can be considered the final step for object detection predictions. The weights collected from this training are utilized during prediction for just standard images, pre-recorded video, and live video for real-time bot responses.

Conclusion

In this report, we explored the development of a real-time character tracking system utilizing a custom CNN model. We considered Integrated classification pre-training and object detection training as parts of a 3-stage pipeline to develop this system.
More and more gaming companies are discovering ways of using machine learning and computer vision to gain useful insights into building much more interactive games. But object tracking using CNNs can be a great opportunity to improve the overall quality and features of competitive gaming.