Brand Detection using Yolo

-Team BASA

TABLE OF CONTENTS

OBJECTIVE

BRAND DETECTION USING YOLO :
Train a model that is able to detect various brands and keep adding more brands to be learned by the model.

INTRODUCTION TO DEEP LEARNING

Deep learning is a subfield of machine learning that uses artificial neural networks with multiple layers to learn and solve complex problems. These neural networks are inspired by the structure and function of the human brain, with interconnected nodes that process and transmit information

Deep learning algorithms use large amounts of data to train neural networks and learn patterns in the data. The more data the algorithm is trained on, the better it can recognize and classify new data. This approach has been very successful in tasks such as image and speech recognition, natural language processing, and autonomous driving.

Deep learning has revolutionized many fields, including computer vision, speech recognition, natural language processing, and robotics, among others. It has enabled the development of more sophisticated and accurate models that can process large volumes of complex data and make predictions and decisions in real-time.

SIGNIFICANCE OF YOLO

YOLO (You Only Look Once) is a real-time object detection system that uses a neural network to detect and classify objects in images or video frames. YOLO was developed by Joseph Redmon and his team in 2016 and is widely used in computer vision applications.

Unlike other object detection systems, YOLO processes the entire image in one pass, which makes it very fast and efficient. YOLO divides the input image into a grid of cells and predicts the bounding boxes and class probabilities for each cell. This allows YOLO to detect multiple objects in a single image and produce real-time results.

YOLO has several versions, with the latest being YOLOv5, which was released in 2020. YOLOv5 uses a more efficient architecture and is faster and more accurate than its predecessors. YOLO is widely used in applications such as self-driving cars, security cameras, and robotics.

WIDE RANGE OF APPLICATIONS OF YOLO

YOLO (You Only Look Once) has many applications in computer vision and image processing. Here are some examples of how YOLO is used:

1. Object detection in autonomous vehicles: YOLO is used to detect objects such as pedestrians, cyclists, and other vehicles in real-time, which is essential for the safe operation of self-driving cars.
2. Security and surveillance: YOLO is used in security cameras and systems to detect and track people, vehicles, and other objects in real-time.
3. Robotics: YOLO is used in robotic systems to detect and classify objects in the environment, which is useful for tasks such as navigation, manipulation, and object recognition.
4. Medical imaging: YOLO is used in medical imaging to detect and classify abnormalities in images, such as tumors or lesions.
5. Retail and marketing: YOLO is used to track and analyze customer behavior in retail stores, such as counting the number of people in a store or tracking customer movements to optimize product placement.
6. Sports analysis: YOLO is used to track and analyze player movements in sports such as basketball, soccer, and football, which is useful for training and game analysis.

Overall, YOLO is a versatile and powerful tool for object detection and classification in a wide range of applications.

SOCIAL MEDIA ANALYSIS

Brand detecting AI can monitor social media platforms to identify mentions of a brand or a specific product. This can help businesses to understand how customers are talking about their brand and products and to identify potential issues or opportunities.

By analyzing social media conversations, a brand detecting AI can identify trends and patterns in customer behavior, such as changes in preferences or the emergence of new competitors.

By tracking the number of mentions of a brand or product on social media, a brand detecting AI can help businesses to measure their brand awareness and to identify opportunities to increase visibility.

WHAT DOES OUR PROJECT DO?

With the rise of social media and e-commerce, brands need to monitor their presence and reputation online. This is where brand detection comes in, as it allows companies to identify and track the usage of their brand assets (such as logos, slogans, and product images) on various digital platforms.
Our project aims to detect the names of the brand logos that appear in the input video.
YOLO (You Only Look Once) is a popular deep learning algorithm for object detection tasks because of several key features that make it stand out from other object detection algorithms.
YOLO is incredibly fast and efficient, making it suitable for real-time object detection tasks. Unlike other algorithms that use region-based approaches and multiple passes over an image, YOLO applies a single neural network to the entire image and makes predictions for all objects at once. This makes it much faster than other algorithms and enables it to achieve real-time performance, even on low-end hardware.

Challenges involved and how our project aims to overcome them :

Brand detection using YOLO can be a challenging task as it requires large amounts of annotated data in order to train the algorithm to recognize specific brand assets, such as logos or product packaging. Collecting and annotating this data can be a time-consuming and expensive process, and the quality of the data can have a significant impact on the accuracy of the algorithm.

Brand assets such as logos and product packaging can vary significantly in terms of size, orientation, color, and other visual features. This means that the algorithm needs to be trained on a wide range of variations in order to be able to recognize brand assets accurately and reliably.

In some cases, YOLO may detect objects that are not actually brand assets, leading to false positives. This can be particularly problematic if the false positive detection is frequent and causes the algorithm to produce inaccurate or unreliable results.

Brand assets in images or videos may be partially or fully occluded by other objects or background elements. This can make it difficult for the algorithm to accurately detect and recognize brand assets, particularly if the occlusion is significant or occurs frequently.

Brand detection using YOLO is often applied to new and previously unseen images or videos. This means that the algorithm needs to be able to generalize from its training data in order to accurately detect brand assets in new contexts. Achieving good generalization can be challenging, particularly if the training data is limited or unrepresentative.

We prepared a custom yolov5 object detection for training. We took a deep dive into the yolov5 data set format and learned how to prepare a data set for training. How many images do we need? If the classes we try to detect are very different, we need a few hundred images. On the other hand, if we are trying to detect similar objects, we need a few thousand images of each class. After this, we need to create one file per image containing the location data of the objects. This text file is named the same as the image file but has a dot txt extension. If there are no objects in the image, for example, if we want to include some background images to reduce false positives, then we do not need these text files; we can only add the image file, which is sufficient. Inside this file, we have one row per object. In this row, there are five numbers. The first one is the object's id, the second and third numbers are the entries of the object center, and the fourth and fifth are for the width and height of the object, respectively. These numbers are in normalized coordinates. To get the actual image coordinates, we need to multiply the x coordinate with the width of the image and the y coordinate with the height of the image, respectively.
While creating the dataset, we need to make sure of a few things.
Try to make the data set balanced. For example, we should not have 100 images for coca cola and ten images for BMW.
Label every instance of every data, and find the bounding boxes for all the images.
The dataset should have much variety, and the annotations should be consistent.
NOW we have to make this data set usable for yolov5 . We did this by creating a directory named data that consist of two sub-directories named images and labels. Images consist of 2 sub-directories: train and val, and labels consist of train and val.
The final step is to create a yaml file needed by the training script, and this yaml file consists of the path to all data relative to the training set. It also needs a path to training and validation images. The training script automatically determines the path to the ground truth metadata files. Finally, we need to specify the names of the classes, which are the brand names. That is all we need to train our object detection model using yolov5.

CONCLUSION

We have succesfully trained a model that is able to detect various brands and keep adding more brands to be learned by the model.