

Vision

- Vision is giving sight to those who are missing out on this beautiful world we all are living. As the saying goes “we are only as blind as we want to be”, by using real time object recognition tech & YOLO technology, we are capable enough to provide this Service to those in need; This is how “The project Vision” Unfolded!

- **Technology used:**

YOLO (you only look once) is an algorithm that uses neural networks to provide real-time object detection. This algorithm is popular because of its speed and accuracy. YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects. This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously.

- **Working:**

YOLOV4 algorithm works based on the following three principles:

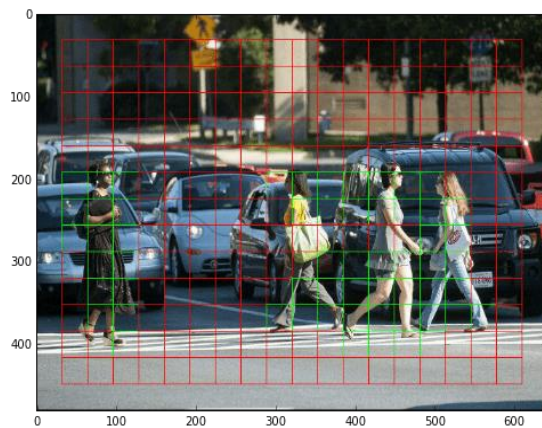
- Residual blocks
- Bounding box regression
- Intersection Over Union (IOU)

Residual blocks

First, the image is divided into various grids. Each grid has a dimension of $S \times S$. The following image shows how an input image is divided into grids

In the image below, there are many grid cells of equal dimension. Every grid cell will detect objects that appear within them. For example, if an

object centre appears within a certain grid cell, then this cell will be responsible for detecting it.



Bounding box regression

A bounding box is an outline that highlights an object in an image.

Every bounding box in the image consists of the following attributes:

- Width (bw)
- Height (bh)
- Class (for example, person, car, traffic light, etc.)- This is represented by the letter c.
- Bounding box center (bx,by)

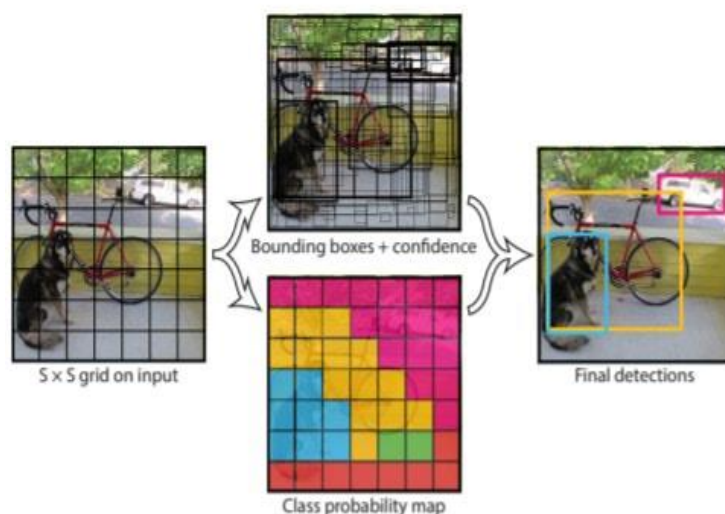
Intersection over union (IOU)

Intersection over union (IOU) is a phenomenon in object detection that describes how boxes overlap. YOLO uses IOU to provide an output box that surrounds the objects perfectly.

Each grid cell is responsible for predicting the bounding boxes and their confidence scores. The IOU is equal to 1 if the predicted bounding box is the same as the real box. This mechanism eliminates bounding boxes that are not equal to the real box.

Integrating these three principles:

The following image shows how the three techniques are applied to produce the final detection results.



First, the image is divided into grid cells. Each grid cell forecasts B bounding boxes and provides their confidence scores. The cells predict the class probabilities to establish the class of each object.

For example, we can notice at least three classes of objects: a car, a dog, and a bicycle. All the predictions are made simultaneously using a single convolutional neural network.

Intersection over union ensures that the predicted bounding boxes are equal to the real boxes of the objects. This phenomenon eliminates unnecessary bounding boxes that do not meet the characteristics of the objects (like height and width). The final detection will consist of unique bounding boxes that fit the objects perfectly.

Applications & Evolution:

1.Alzheimer's : Alzheimer's disease is a brain disorder that slowly weakens memory and thinking power ; with the help of vision they can now recognise objects Infront of them and help them remember

their loved ones ; just by opening the app on their cells, the vision will speak to them in their native language describing what's the object in real time.

2.Visually impaired : From the most recent Survey of WHO it's been expected that in the year 2026 there be an increase of Visual impairment from **34.6% in subjects 60-69 years of age** to more than 71% in subjects 70-79 years of age. With the Assistance of Vision they now can be able to be independent and Vision helps them keep track of their medical necessities & From my own experience I have seen my grandmother Mistakenly taking her morning pills at night due to less visibility .With the help of vision they now can take their pills at proper time.

3.Autonomus vehicle: Object detection is one of the key software components in the next generation of autonomous cars. Classical computer vision and machine learning approaches for object detection usually suffer from the slow response time. Modern algorithms and architectures based on artificial neural networks such as yolo can be handfull.

References:

Yolo references:

<https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection/>

Git repository:

<https://github.com/theAIGuysCode/YOLO...>

Linked in :

<https://www.linkedin.com/in/asjad-afnan-601450214>

Author:

Asjad Afnan

2nd year CSE

Guidance:

Prof. Sathish G C

REVA University

Happy Learning!