Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

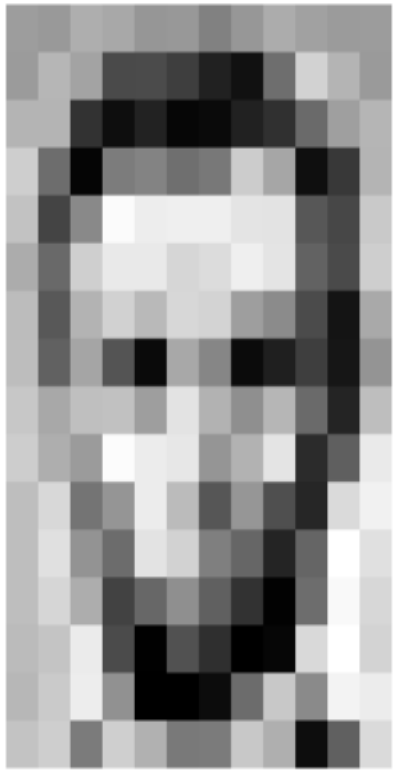
A Comprehensive Systematic Review of YOLO for Medical Object Detection

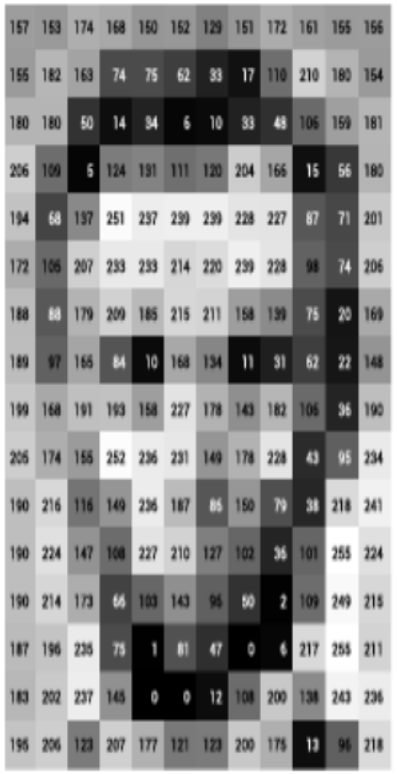
Applications, Performance, and Future Directions

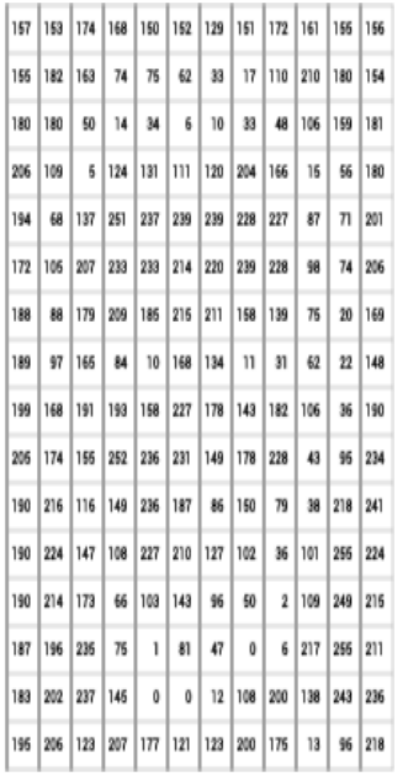
Presented by: Habib ur Rehman Owasi
04 October 2024

Introduction

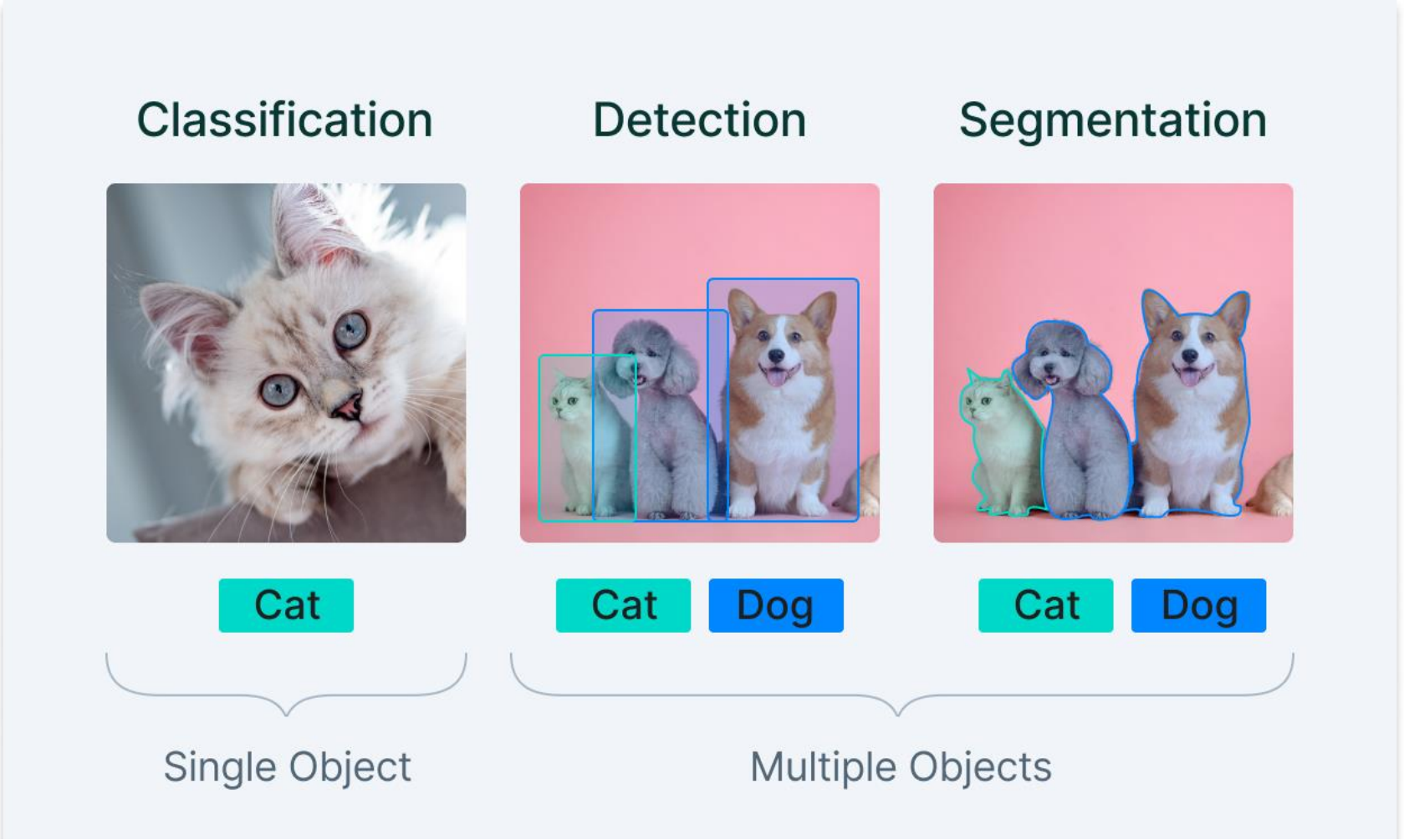
- Computers process images in the form of an array of pixels.
- While advanced tasks involve operations like convolutions (to extract features) and pooling (to reduce dimensionality), so that deep learning models can identify and classify objects effectively







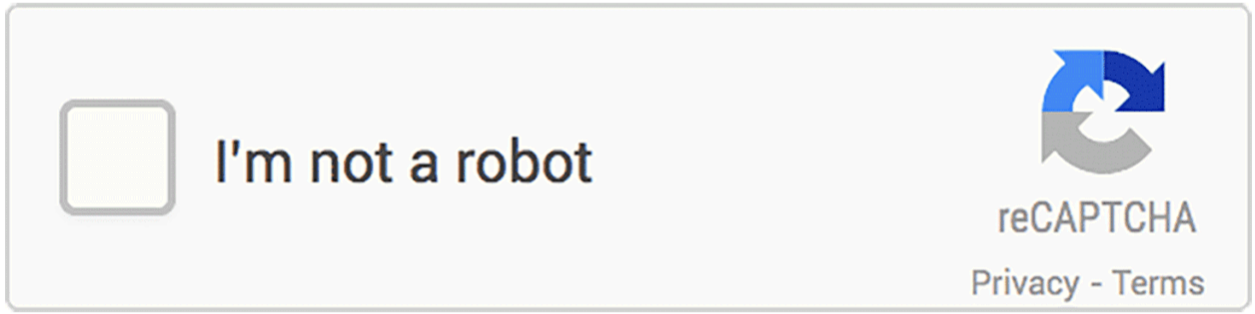
Tasks in Computer Vision

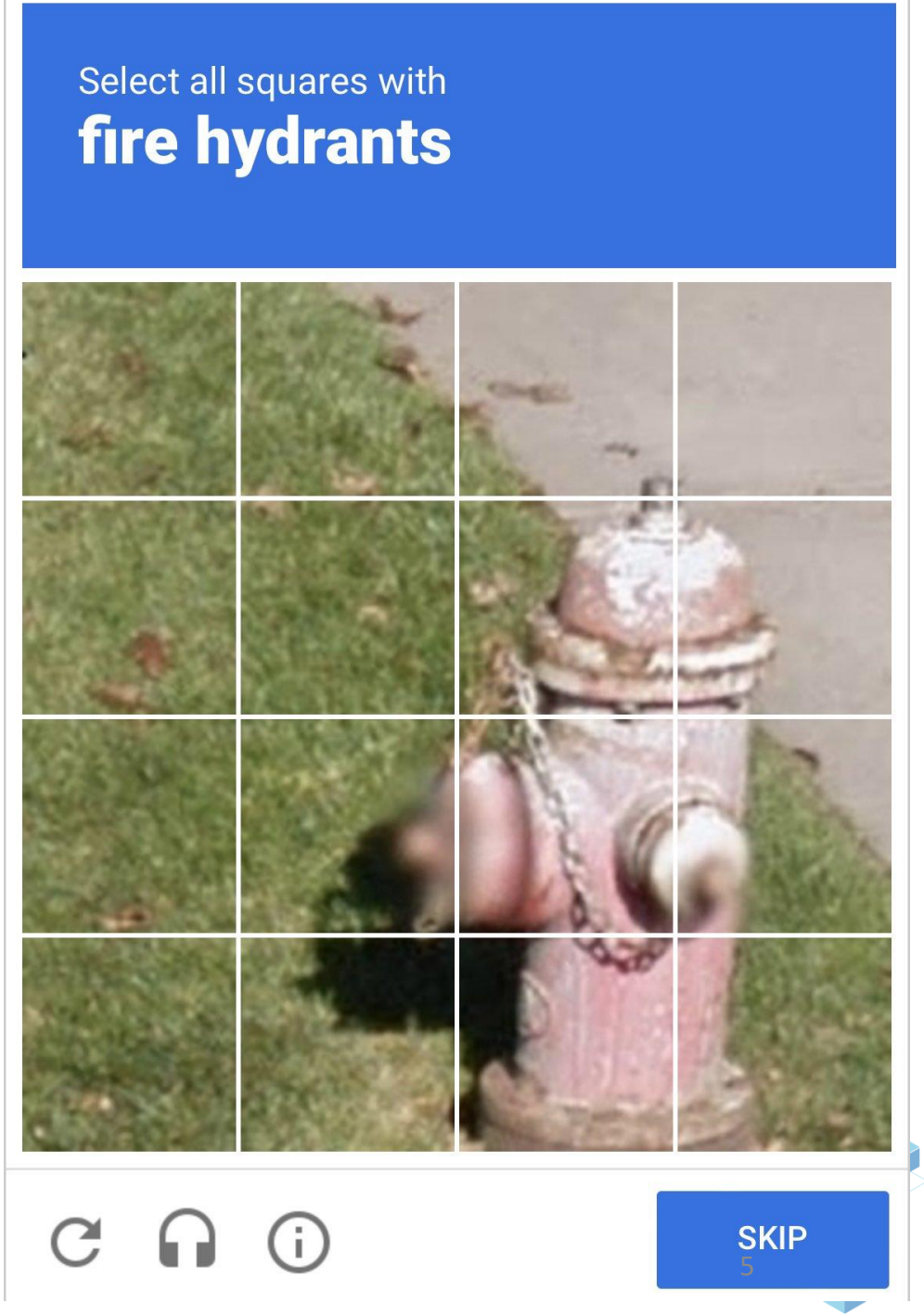


What makes us Human?

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

I am Human!



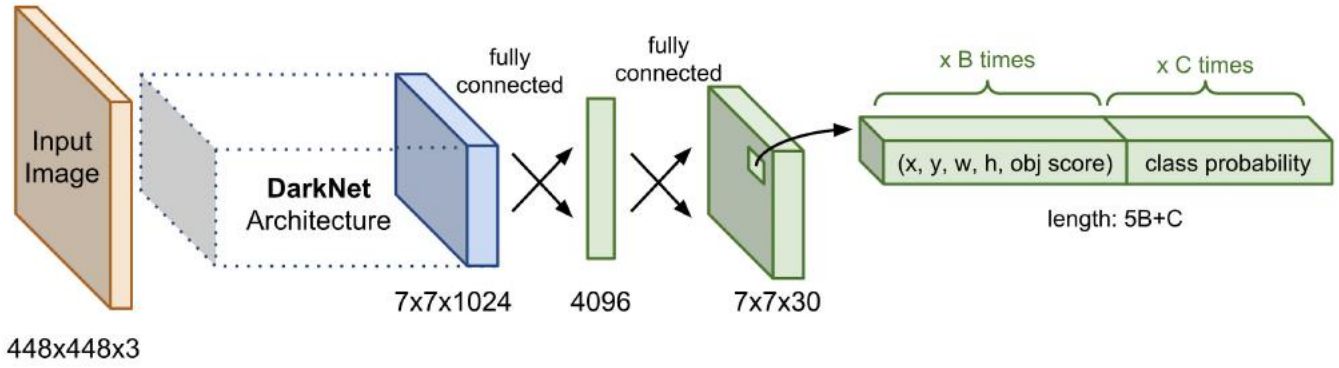


What is YOLO?

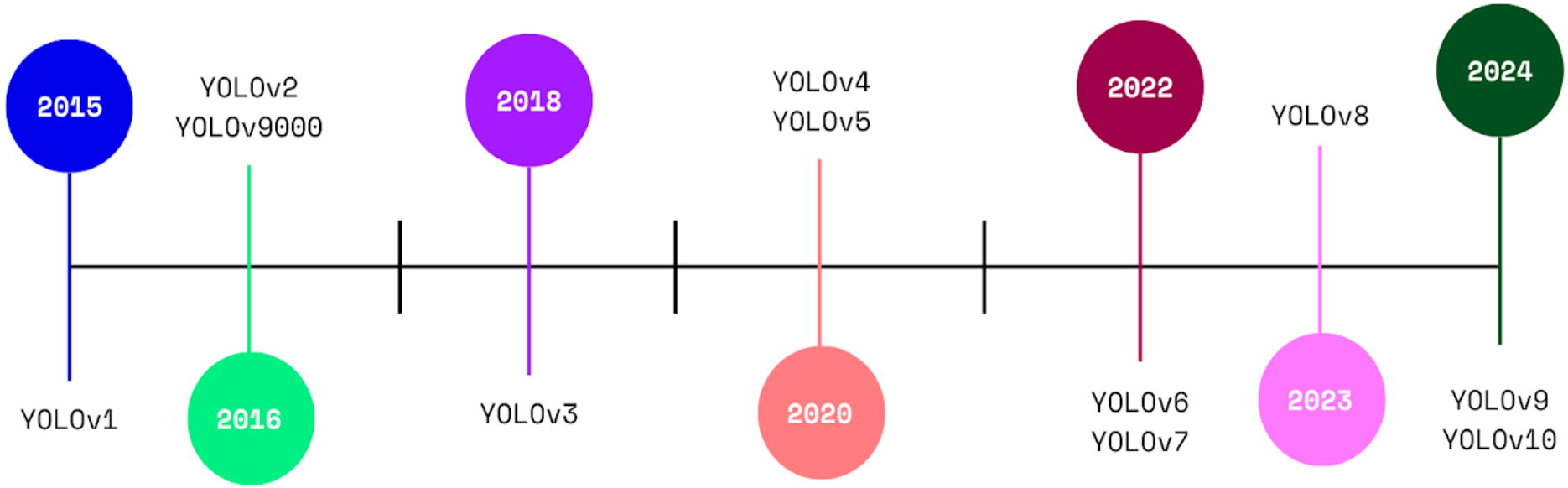
- You Only Look Once (YOLO)
- State-of-the-art, real-time, end-to-end object detection algorithm
- Single-stage method that divides the image into a grid of cells
- Each cell is responsible for detecting objects within a certain area

YOLOv1 Network Architecture

- The YOLOv1 architecture is inspired by GoogLeNet.
- Instead of the inception modules (involve multiple types of convolutions in parallel), yolo uses 1 × 1 reduction layers followed by 3 × 3 convolutional layers
- Network has 24 convolutional layers followed by 2 FC layers.



Timeline of YOLO



YOLO Architecture

- YOLO architecture consists of three main components:

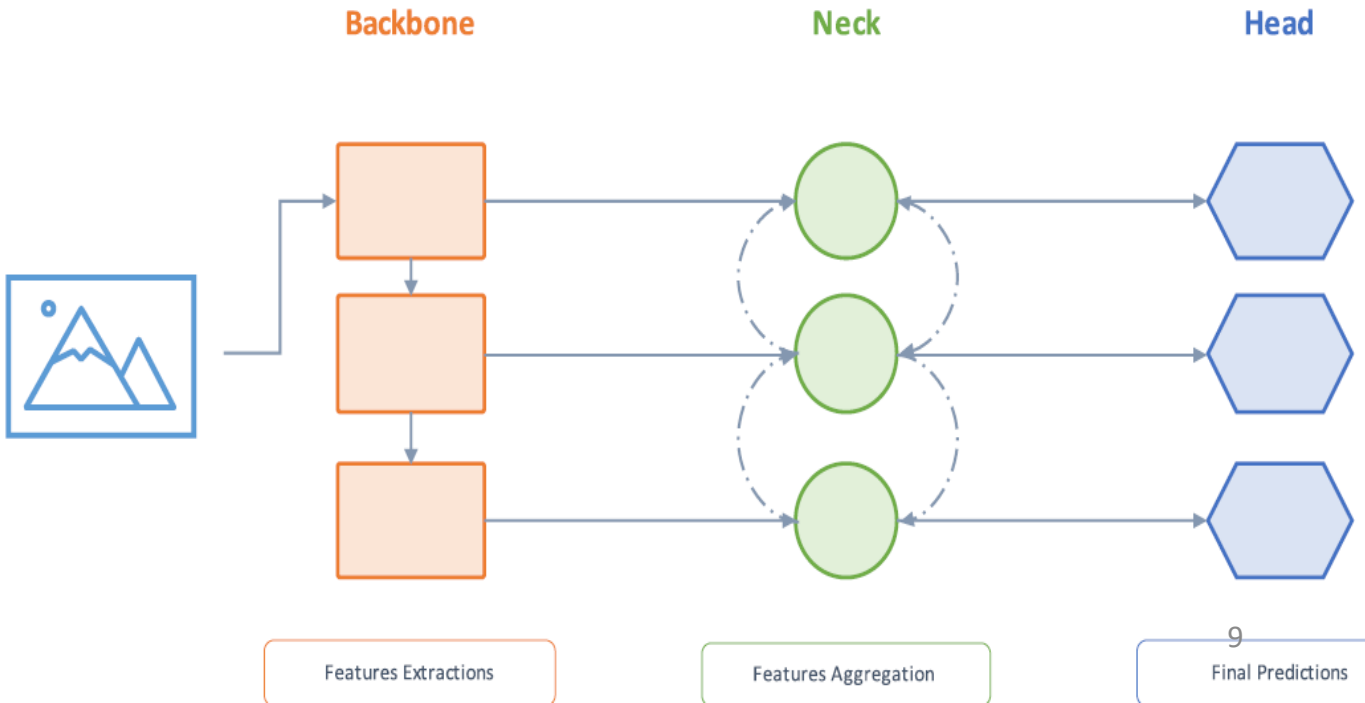
(i) the backbone

(ii) the neck

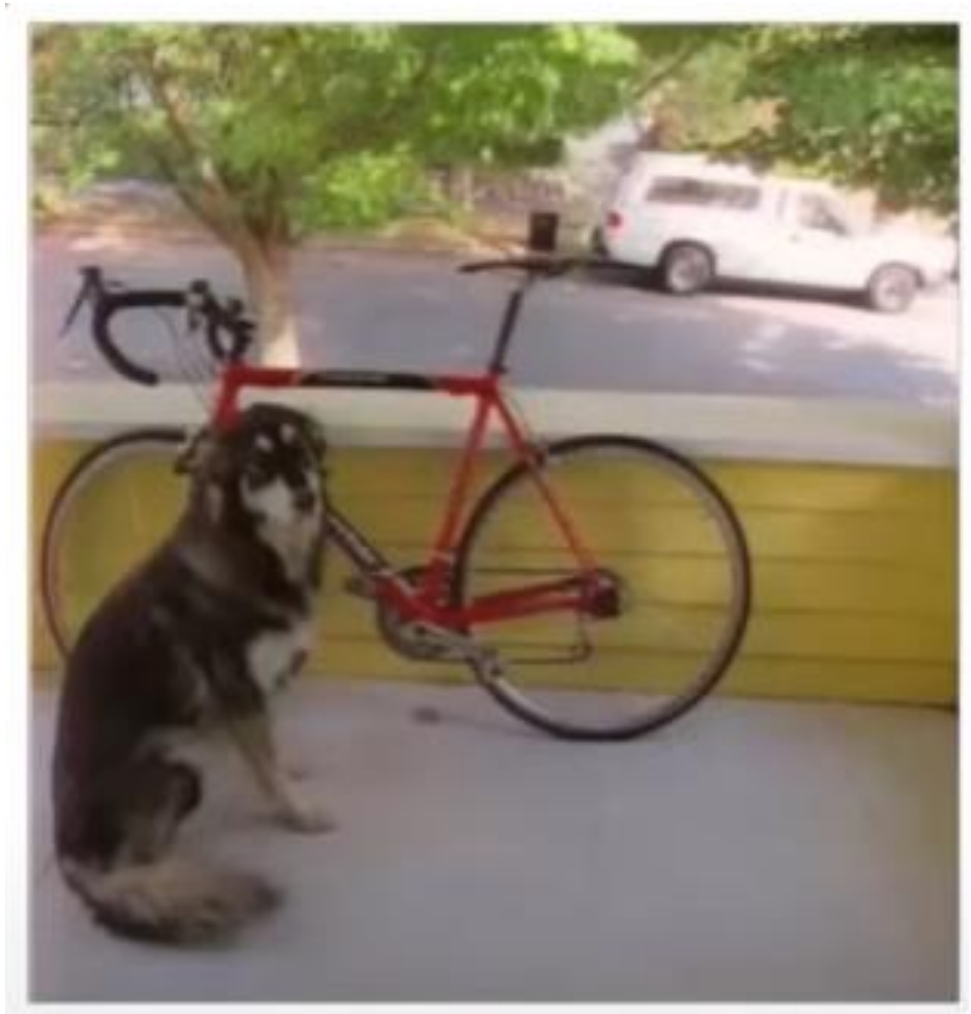
(iii) the head

The backbone, neck, and head vary in different versions of YOLO.

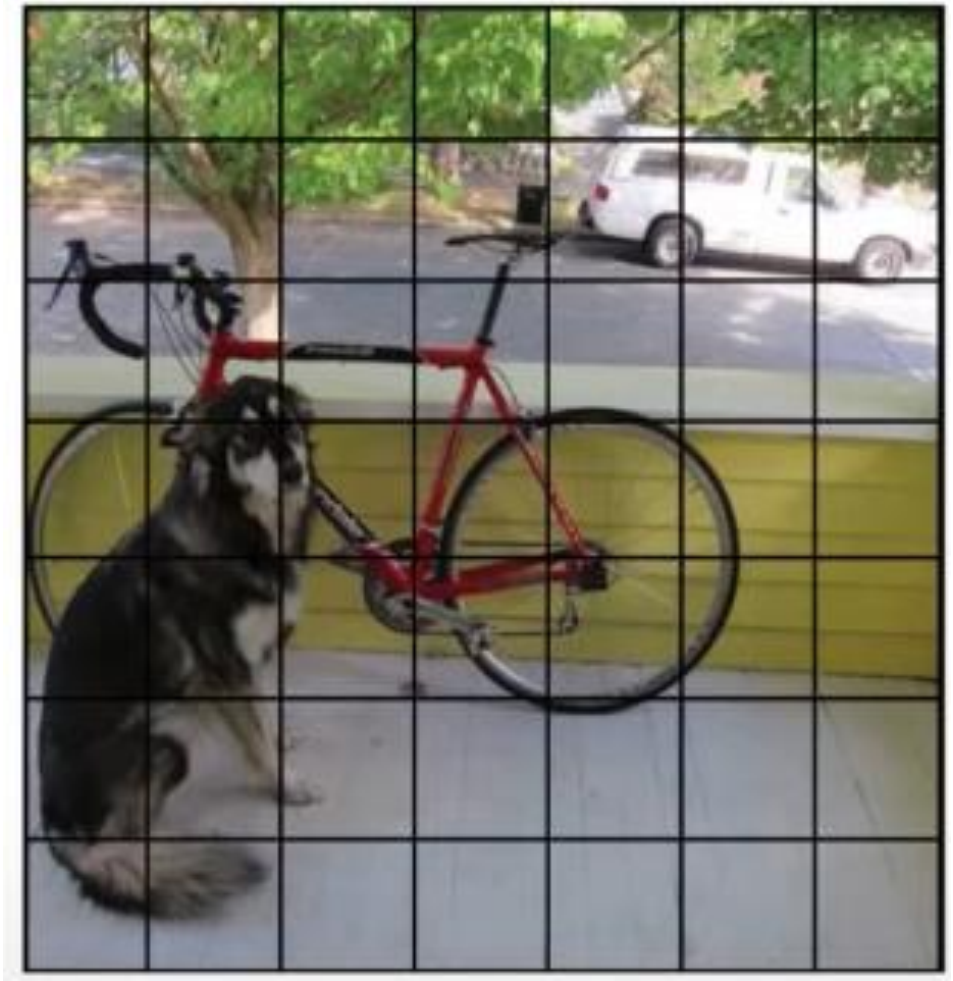
For backbone, Darknet, VGG16, or Resnet are used; for neck feature pyramid network (FPN), and for neck Densenet or sparsenet are used.



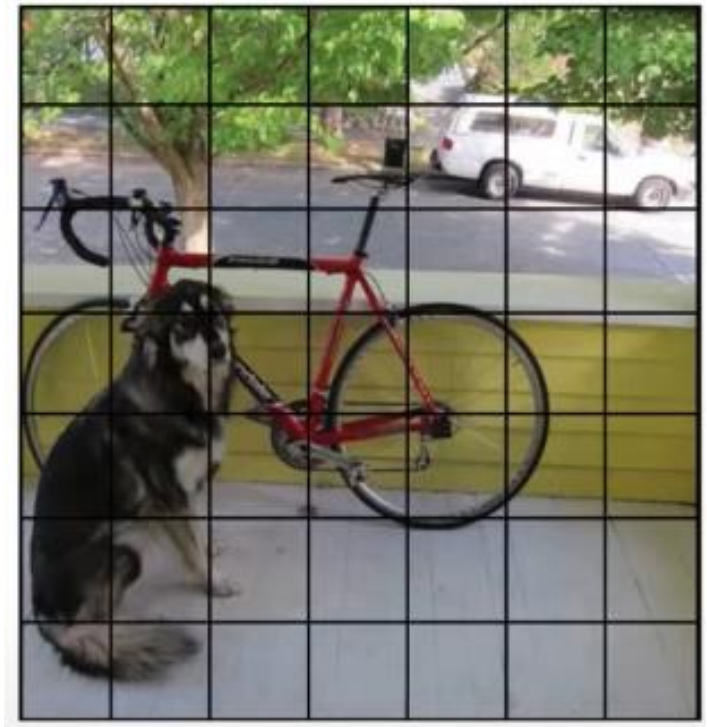
How YOLO works?



YOLO divides the image into an S × S grid

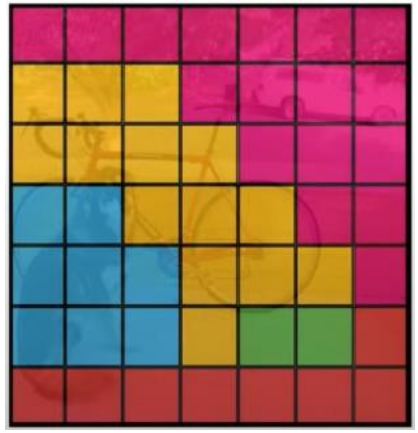


Predict bounding boxes and confidence



Bounding Boxes +

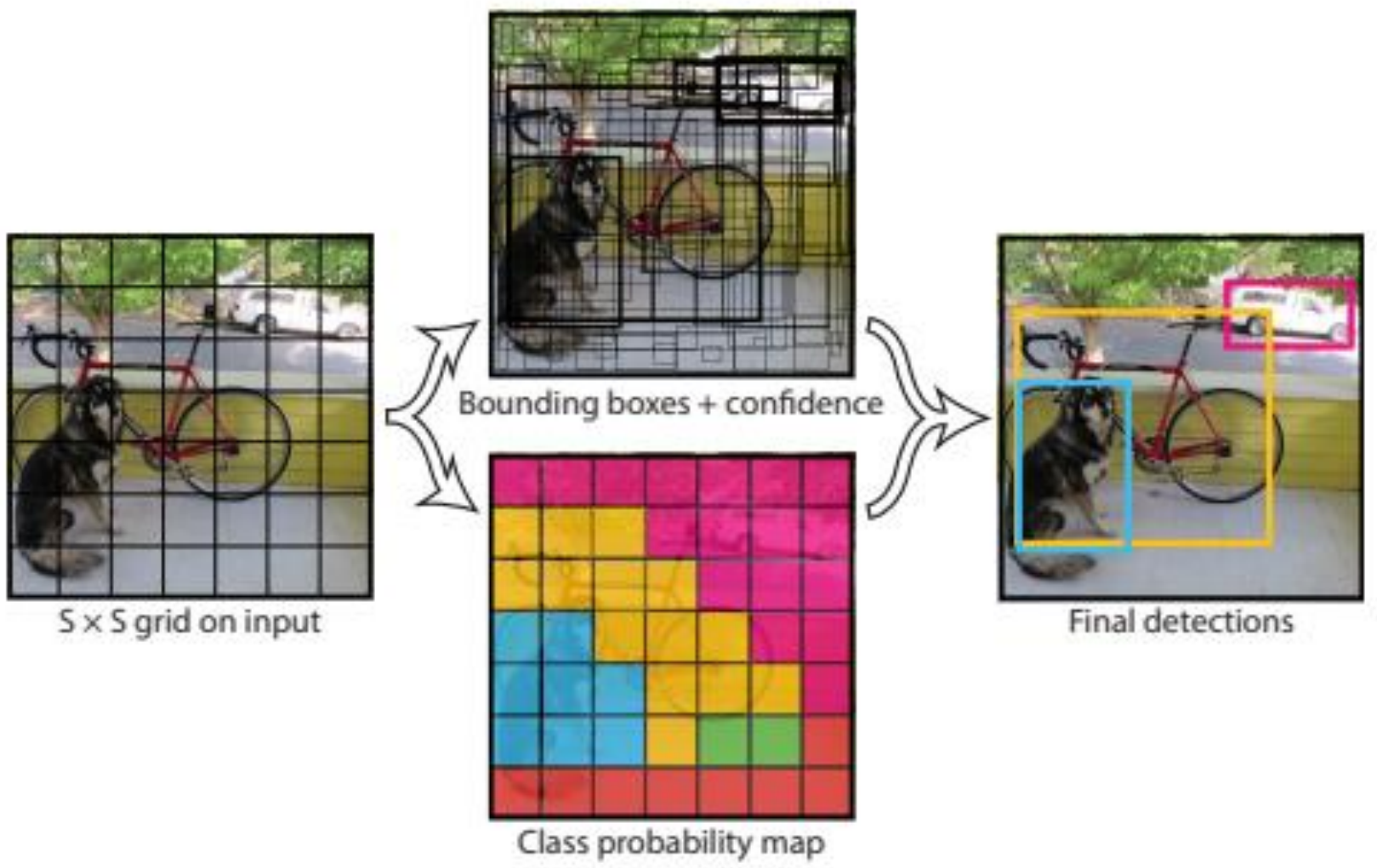
Confidence



Class Probability

map

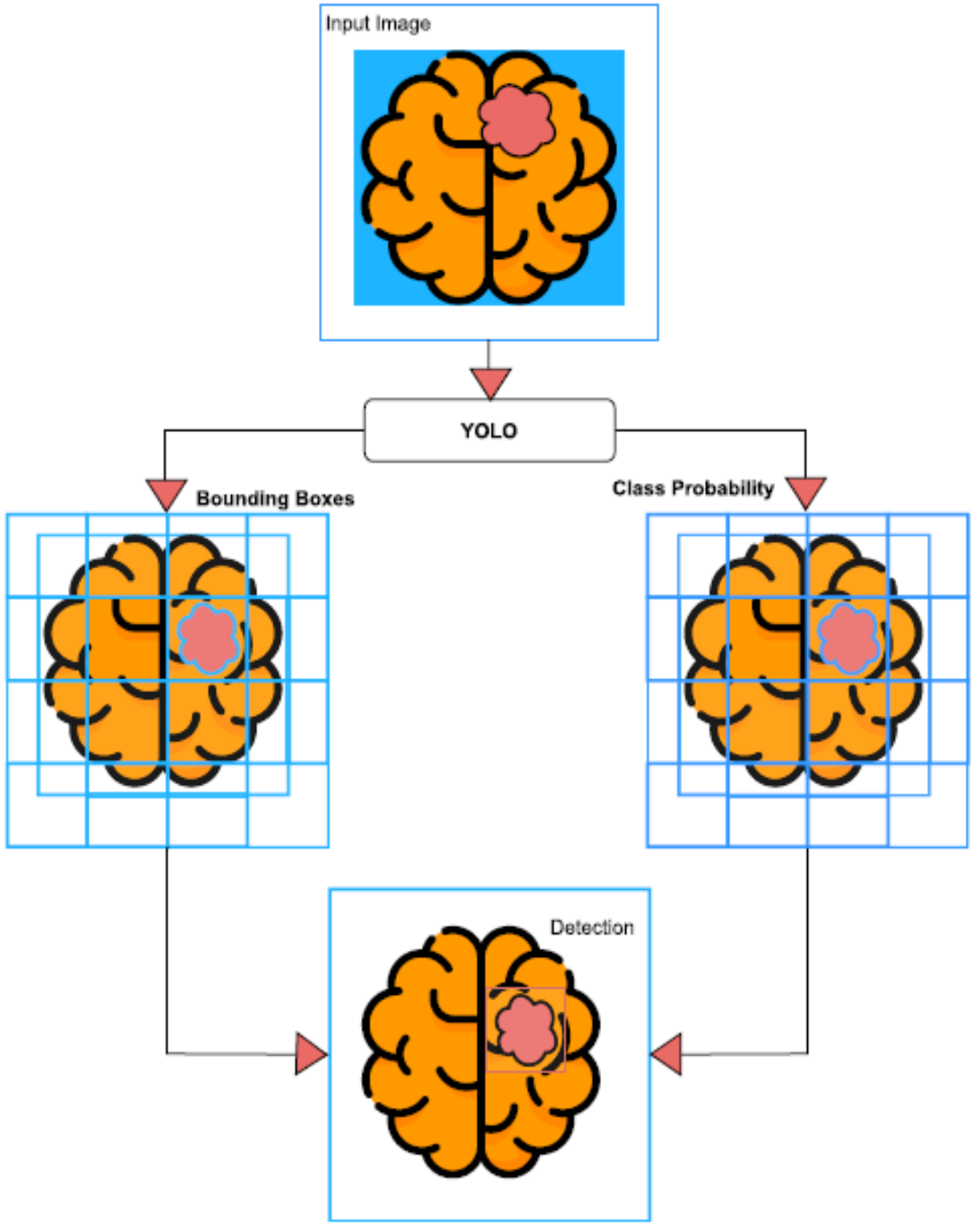
Final detection



Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

How YOLO works?

1. Divide the input image into a grid of cells.
2. Each cell predicts a bounding box, a confidence score, and the class probabilities for the objects in that cell.
3. The bounding box is a rectangle that is used to surround the object.
4. The class probabilities are the probabilities that the object in the cell belongs to each possible class.



YOLO FOR MEDICAL OBJECT DETECTION

- ADVANTAGES OF YOLO FOR MEDICAL OBJECT DETECTION:

1. Real-Time Detection
2. Efficiency
3. Accuracy
4. Multi-medical Object Detection
5. Adaptability
6. Minimal False Positives

YOLO FOR MEDICAL OBJECT DETECTION

- DRAWBACKS OF YOLO FOR MEDICAL OBJECT DETECTION:

1. Large Dataset Requirement
2. Sensitivity to Object Scale
3. Difficulty with Small Objects
4. Issues with Occluded Objects
5. Limited Object Diversity

Creating New Dataset

Following things needs to be considered:

1. Data Ethics and Privacy

2. Annotating Images

3. Class Imbalance

4. Complexity and Variability

5. Clinical Relevance

Image Annotation

- Process of labeling Regions of Interest in image.
- It involves adding metadata or annotations to images to provide additional information about the objects.
- The purpose is to create a labeled dataset that serves as training data for learning algorithms, particularly for tasks like object detection, object recognition, and image segmentation.

Image Annotation

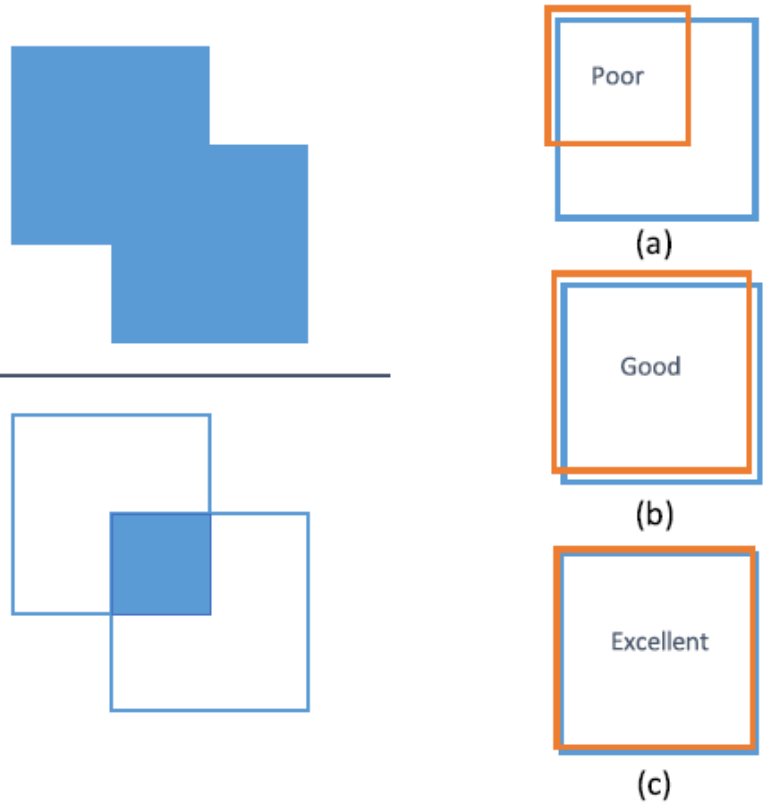
Computing the Intersection over Union

(a) poor detection performance,

(b) good detection performance,

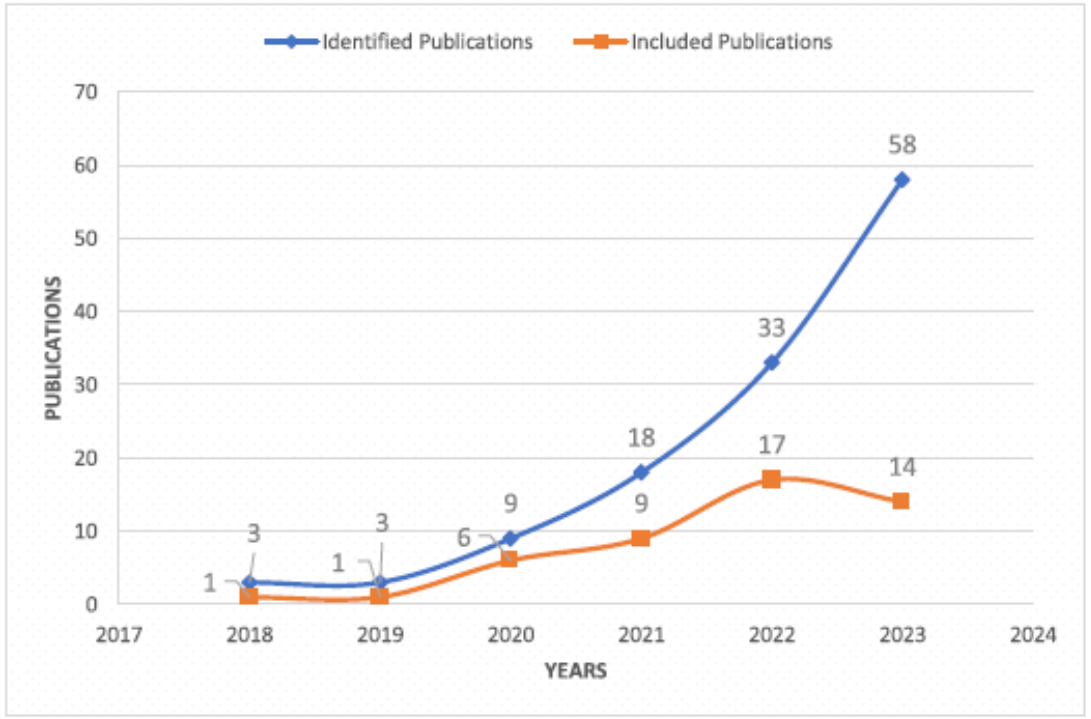
(c) excellent detection performance.

$$IoU = \frac{Area\ Overlap}{Area\ of\ Union}$$

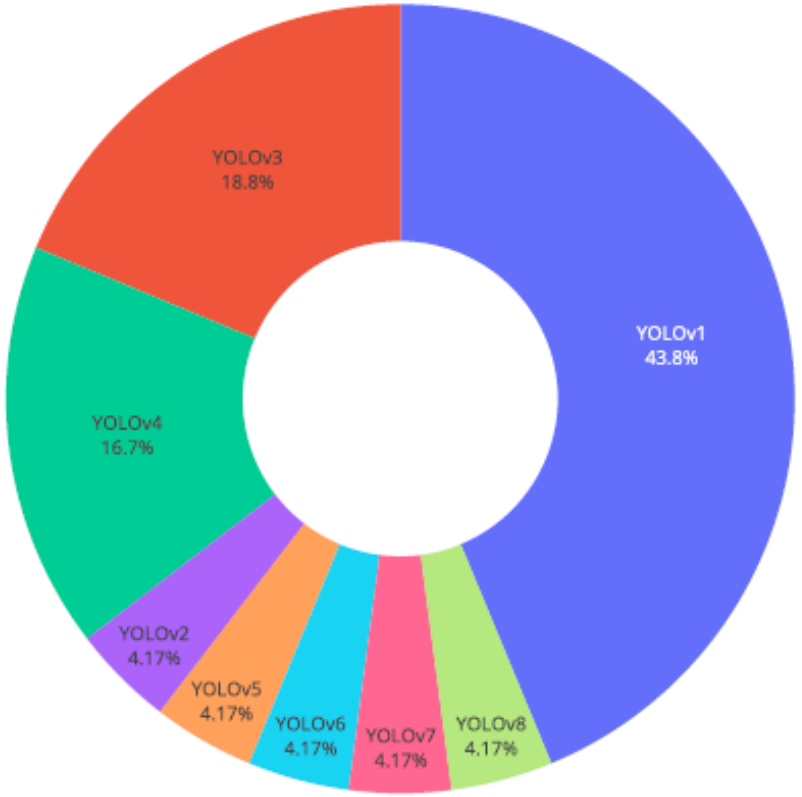


Trend of YOLO in Medical Imaging

Trend of publications on the YOLO



Distribution of YOLO versions



Major domain of YOLO in healthcare

Medical Imaging

- Radiology
- Oncology
- Pathology

Surgical Procedures

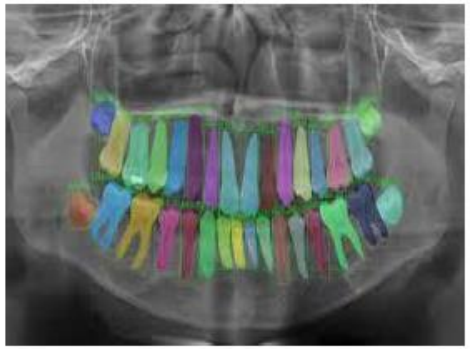
- Detection of surgical instruments.
- Segmentation of surgical instruments

Personal Protective Equipment

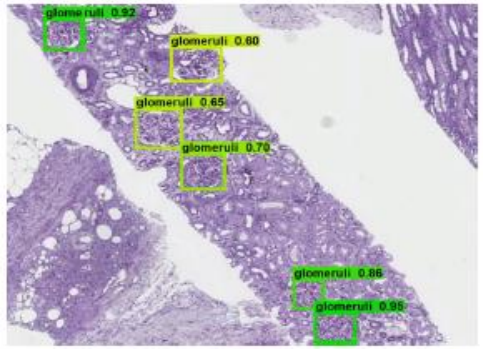
- Patients Safety
- Patients' vital signs and behavior
- Protective equipment detection

YOLO in different medical images applications

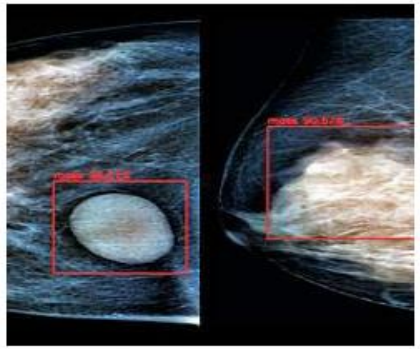
(a) periodontitis bone loss diagnosis, (b) glomerular detection, (c) breast cancer detection (d) lung normal and abnormal detection, (e) brain tumor detection, (f) white and red blood cells detections.



(a)

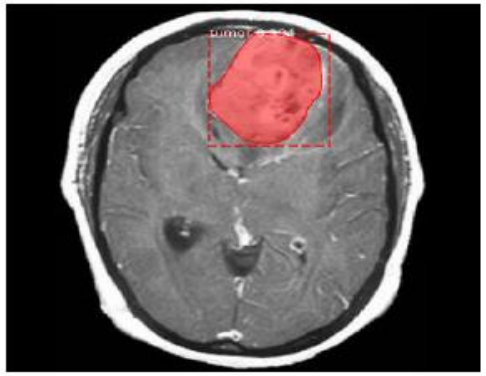


(b)

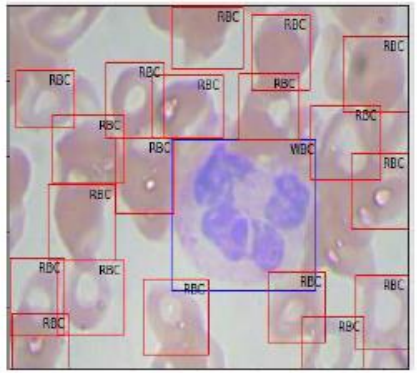


(c)

(d)



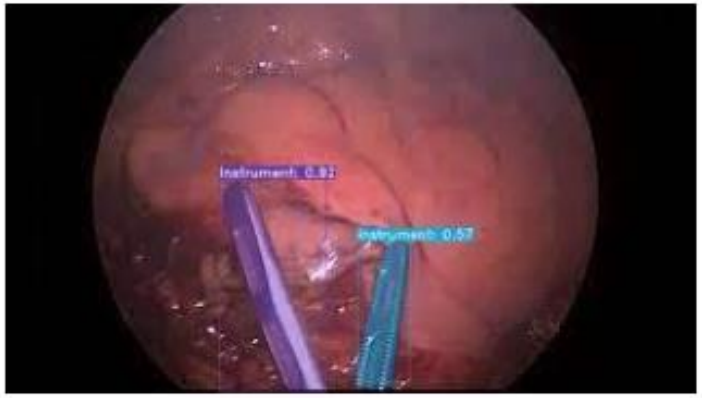
(e)



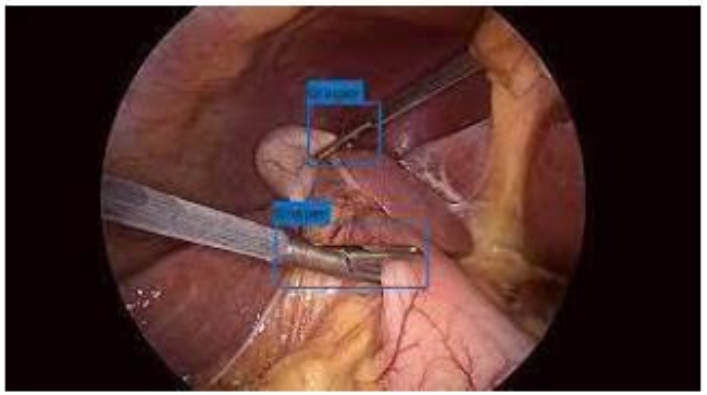
(f)

YOLO surgical procedure applications

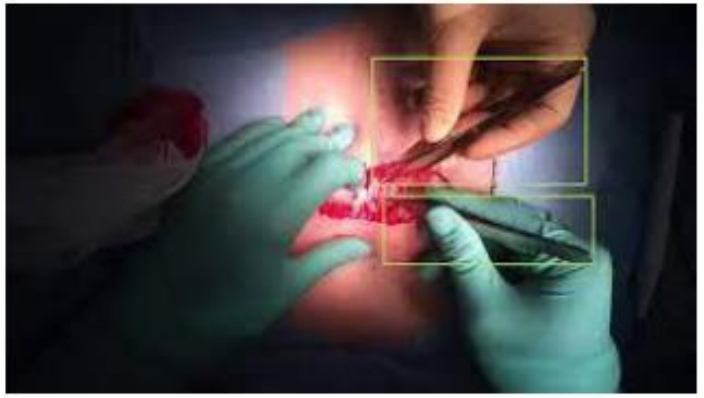
(a)surgical tool detection in open surgery videos, (b) surgical instruments, (c) real-time instance segmentation of surgical instruments.



(a)



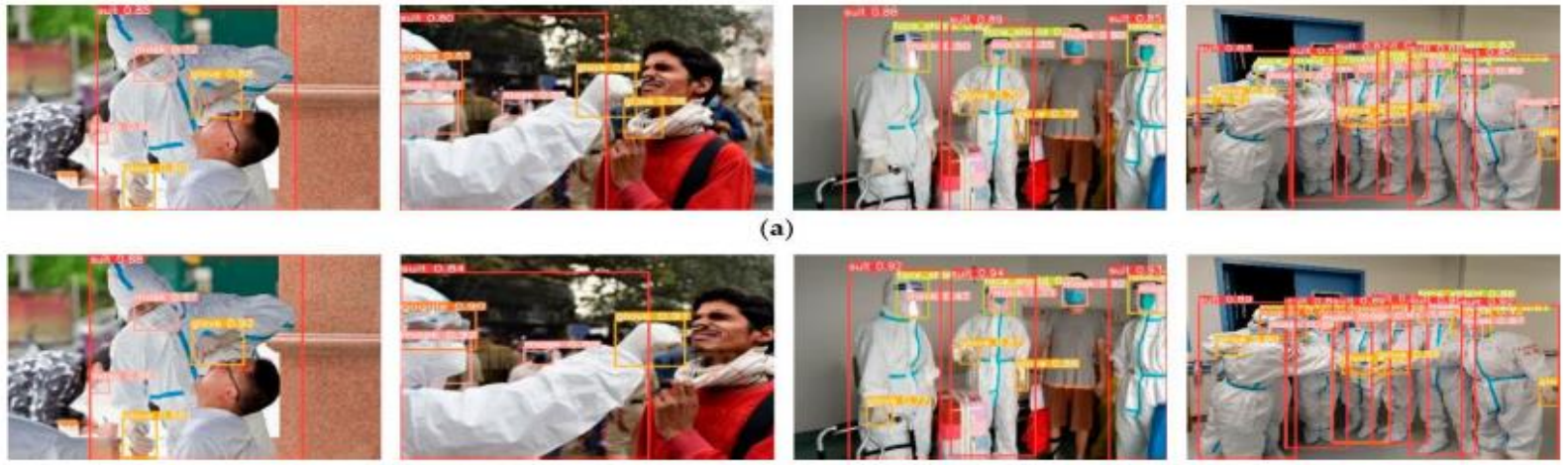
(b)



(c)

Medical personal protective equipment

(a) suit, (b) face shield;(c) goggles, (d) mask, and (e) glove



TRANSFER LEARNING

Transfer learning can be leveraged to improve YOLO's performance

1. Pre-Trained Models
2. Fine-Tuning
3. Transfer of Knowledge
4. Improved Convergence
5. Enhanced Feature Extraction

Conclusion

- YOLO has significantly improved patient outcomes while reducing diagnosis and treatment times.
- Despite the remarkable successes of YOLO, its deployment is not without challenges
- These include its sensitivity to object scale, difficulty detecting small or occluded objects, and considerable computational resource requirements.
- Ethical considerations like data privacy and algorithmic bias need to be considered in the development of YOLO-based systems.

"If AI systems like YOLO can detect medical anomalies more accurately and efficiently than humans, does this bring us closer to a 'Turing Test' for medical diagnosis? Can machines eventually surpass human abilities in critical medical decision-making?"

Thank You

C&B