TASK 3 Report
Based on the research study,

1. Comparison of CNN and Yolo
2. Try implementing CNN and Yolo with built-in functions for any dataset
   Note down the results and accuracy
3. Different layers in CNN and your understanding

Both CNNs and YOLO have made significant contributions to deep learning and computer vision. CNNs are widely used for image classification, segmentation, and feature extraction, while **YOLO excels in real-time object detection tasks**

You Only Look Once (YOLO): YOLO is a real-time object detection algorithm that operates on entire images, aiming to detect objects and their bounding boxes in a single pass. Unlike traditional object detection methods that involve sliding windows or region proposals, YOLO takes a different approach to achieve efficiency and accuracy.

- Speed
- Detection accuracy
- Good generalization
- Open-source

Key Features of YOLO:

1. Grid-based Approach: YOLO divides the input image into a grid and assigns each cell responsibility for predicting objects. Each cell predicts a fixed number of bounding boxes along with their class probabilities.
2. Single Pass Detection: YOLO performs object detection in a single pass through the neural network, making it extremely fast compared to other methods that require multiple stages or region proposals.
3. Feature Extraction: YOLO uses a convolutional neural network to extract features from the input image. These features are then used for predicting bounding boxes and class probabilities.
4. Multi-Scale Training: YOLO trains on images of different scales to improve its ability to detect objects at various sizes. It uses anchor boxes to handle objects of different aspect ratios.

Benefits of YOLO:

- Real-Time Detection: YOLO's single-pass architecture allows it to process images or videos in real-time, making it suitable for applications that require low latency.
- Simplicity: YOLO's approach is conceptually simple and straightforward compared to more complex methods. It directly predicts bounding boxes and class probabilities without the need for additional post-processing steps.
- Contextual Information: Since YOLO operates on the entire image, it can capture rich contextual information that aids in accurate object detection, especially in cluttered scenes.
- End-to-End Learning: YOLO is trained end-to-end, optimizing both the feature extraction and object detection components simultaneously, leading to better overall performance.

CNN Implementation :

Code :
https://colab.research.google.com/drive/1VCaUU4_zTvTdrp7AIEtxCi5RAxKed-Rp#scrollTo=VZmdgRudcMcP

References :

- https://towardsdatascience.com/building-a-convolutional-neural-network-cnn-in-keras-329fbbadc5f5
- https://www.datacamp.com/tutorial/convolutional-neural-networks-python
- https://www.geeksforgeeks.org/applying-convolutional-neural-network-on-mnist-dataset/

YOLO Implementation :

Code :

- https://blog.paperspace.com/train-yolov5-custom-data/#set-up-the-code
- https://www.datacamp.com/blog/yolo-object-detection-explained
- https://www.v7labs.com/blog/yolo-object-detection

References :

- https://www.datacamp.com/tutorial/object-detection-guide
- https://blog.paperspace.com/train-yolov5-custom-data/
- https://towardsdatascience.com/implementing-yolo-on-a-custom-dataset-20101473ce53