Assignment 3. Create your Project by Using Neural Networks for Object Detection and Recognition – 80 points

Tasks:

1. Study the principles of Convolutional Neural Networks (CNNs) and their applications in computer vision.
2. Explain the concepts of **validation** and **cross-validation**, their purposes, and their roles in model evaluation.
3. Review and analyze **8 popular neural network architectures** used for image recognition and object detection.
4. Prepare a dataset and implement model training with proper validation.
5. Compare model performance using metrics such as **Accuracy, Precision, Recall, F1-score, and mAP (mean Average Precision)**.
6. Develop a demonstration script or interface for visualizing detection results.

□ Theoretical Background

□ Definition: *Validation*

Validation is the process of evaluating a trained model on a separate subset of data that was not used during training.
Its goal is to assess how well the model generalizes and to prevent overfitting.

Typical validation approach:

- Split the dataset into **train / validation / test** sets (e.g., 70% / 20% / 10%).
- The validation set is used for hyperparameter tuning and monitoring model performance during training.

□ Definition: *Cross-Validation*

Cross-validation is a statistical method for evaluating model performance by dividing the dataset into *k* parts (folds), typically 5 or 10.
The model is trained on *k–1* folds and tested on the remaining one, repeating the process *k* times.

Purpose:
To ensure a more reliable and unbiased estimate of model accuracy and reduce the dependency on a single random train-test split.

□ **Neural Network Architectures for Comparison:**

№	Architecture	Type	Key Features	Application
1	**LeNet-5**	CNN	Basic early CNN with few layers	Simple image classification
2	**AlexNet**	CNN	Deep network using ReLU and Dropout	Image classification
3	**VGG16 / VGG19**	CNN	Uniform 3×3 filters and deep architecture	High accuracy, high computational cost
4	**GoogLeNet (Inception)**	CNN	Parallel convolutions with different kernel sizes	Balanced accuracy and efficiency
5	**ResNet50 / ResNet101**	CNN	Residual (skip) connections	Very deep networks without gradient vanishing
6	**MobileNetV2**	CNN	Lightweight architecture for mobile devices	Real-time recognition
7	**YOLOv5**	Object Detection CNN	Single-shot detection and classification	Fast real-time object detection
8	**Faster R-CNN**	Object Detection CNN	Two-stage model (Region Proposal + Classification)	High detection accuracy

□ **Practical Part**

1. Load and preprocess the dataset (split into train / validation / test).
2. Implement and train 2 chosen network (e.g., YOLOv5 or ResNet50).
3. Perform **validation** on a hold-out set or Apply **k-fold cross-validation** (k = 5 or k = 10) and compute the averaged results.
4. Evaluate the model using **Precision, Recall, F1-score, and mAP** metrics.
5. Visualize the results with bounding boxes and class labels.
6. Compare the obtained results with other architectures.

Report Structure:

1. Description of the dataset and a link to it
2. Theoretical Background
    - Neural Networks for Object Recognition
    - Validation and Cross-Validation
    - Overview of 8 Neural Network Architectures

3. Implementation
4. Experimental Results and Metric Comparison