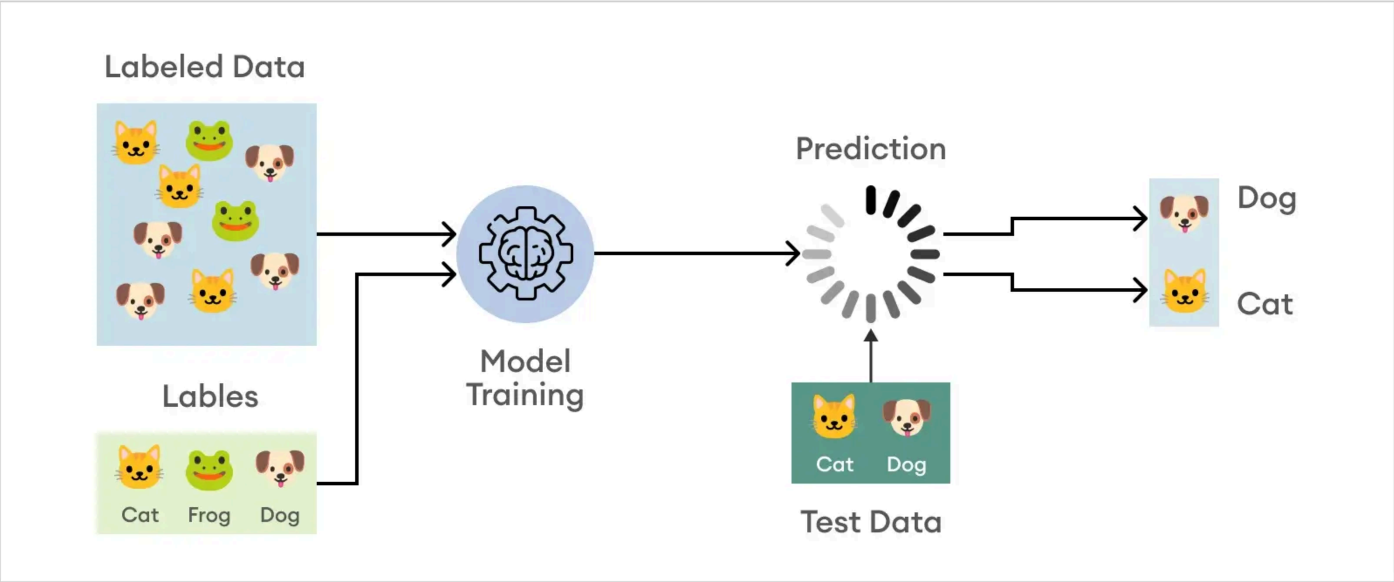
Index

Image classification....2
- Introduction....2
- Types of image classification....2
- How image classification works....3
  - Image pre-processing....3
  - Feature extraction....3
  - Object classification....4
- Image Classification using ML....5
  - Unsupervised Classification....5
  - Supervised Classification....5
- Deep neural networks for image classification....6
  - Components of CNN:....6
  - Benefits of CNN:....7
- Applications of image classification....8
- Image classification vs. Object detection....9
- Github Repo....9

Image classification

Introduction

Image classification is a crucial computer vision task that involves categorizing entire images into predefined classes. It's a fundamental process that relies on training data to learn patterns and assign appropriate labels. Image classification requires a deep analysis of each pixel to determine the most accurate overall label.



Types of image classification

- **Binary:** Classifies images into two categories (e.g., dog or cat). The output layer for binary classification in a neural network typically consists of a single neuron with a sigmoid activation function. (Binary classification-github)

- **Multiclass:** Categorizes images into more than two classes (e.g., fruits). The output layer consists of more than 1 neuron which is designed to produce a probability distribution of the classes. Softmax Activation Function is used.

- **Multilabel:** Assigns multiple labels to a single image (e.g., classifying image as both beach and sunset). The output layer typically consists of multiple single neurons with sigmoid activation function.

- **Hierarchical:** Organizes classes into a hierarchical structure (e.g., classifying a cat as animal->mammal->cat). The output layer varies depending on the specific structure of the hierarchy, commonly used architecture are multiple output layers or a single output layer with a specialized activation function.

- **Zero-Shot:** Classifying images that the model has not seen in training phase (e.g., classifying image of panda as animal). The output layer varies based on the model architecture.

How image classification works

Image classification involves breaking down an image into its individual pixels and analyzing them to determine the overall image label. The pixels are treated as a matrix, and the chosen algorithm extracts key features from the image to aid in classification.
The general pipeline includes:

- Image pre-processing
- Feature extraction
- Object classification.

Image pre-processing

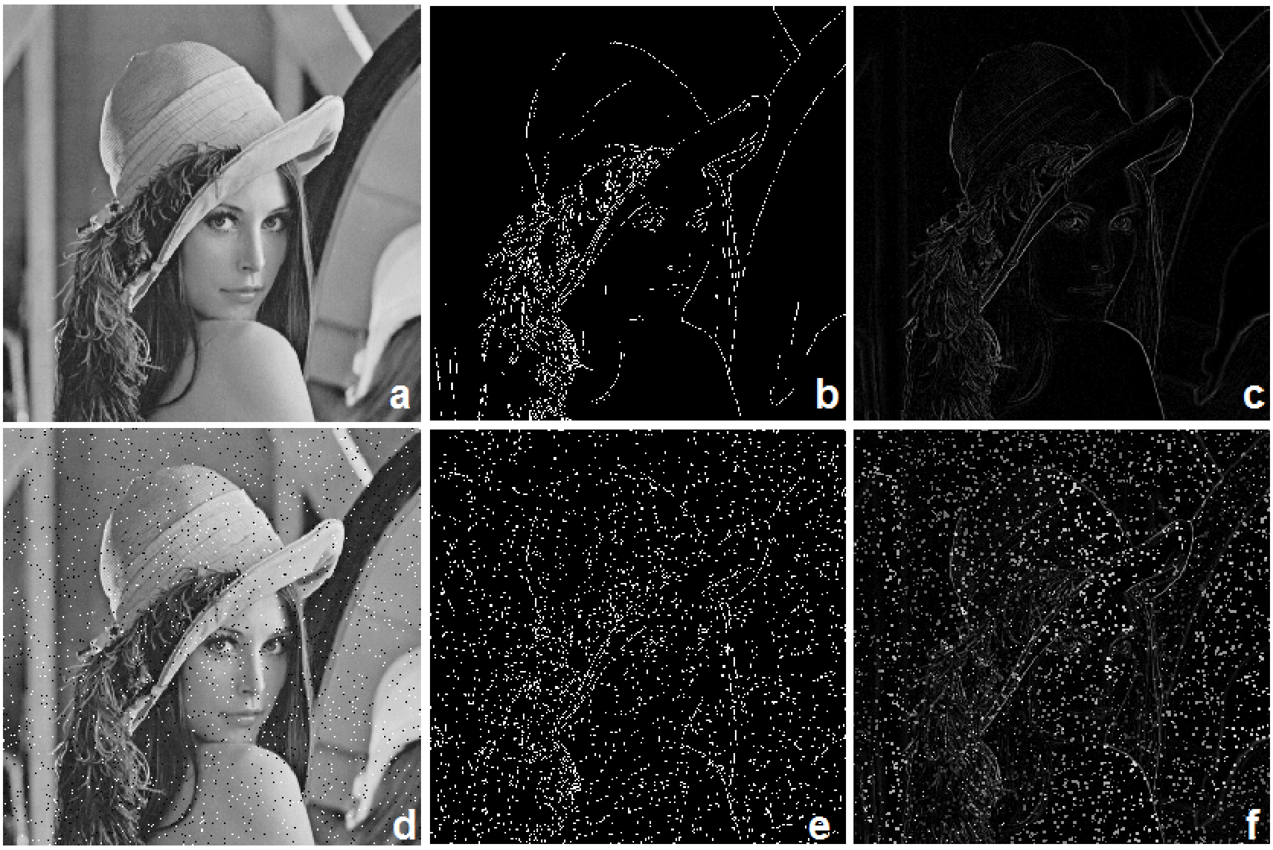
Image preprocessing is a crucial step in image classification that involves enhancing image quality and preparing it for further analysis. Techniques like resizing, cropping, normalization, noise reduction, and data augmentation are commonly used to improve image data quality.

- **Resizing:**
  - Adjusts image dimensions to a standard size for computational efficiency and consistency.
  - Avoid excessive resizing that may lead to image degradation.
- **Cropping:**
  - Removes irrelevant or unnecessary parts of the image that may hinder model performance.
  - Focuses on the region of interest for better classification.
- **Normalization:**
  - Standardizes pixel values to a specific range (e.g., 0-1) to ensure consistent input for the model.
  - Improves model convergence and stability.
- **Noise reduction:**
  - Removes noise (e.g., random pixel values) that can affect model accuracy.
  - Improves image clarity and reduces the impact of noise-induced errors.
- **Data augmentation:**
  - Creates new variations of the images by applying transformations like rotation, zooming, flipping, and changing brightness/contrast.
  - Increases the dataset size and diversity, improving model generalization and robustness.

Feature extraction

Feature extraction is a key process in image classification that involves identifying unique visual patterns to distinguish different objects. These patterns, such as fur texture, ear shape, and body size, are learned through model training. Feature extraction improves machine learning models' performance by focusing on the most relevant aspects of data.

There are multiple features that can be extracted from an image Color Features, Texture Features, Shape Features. One commonly used technique for shape feature extraction is **edge detection**, which involves identifying boundaries between regions in an image.



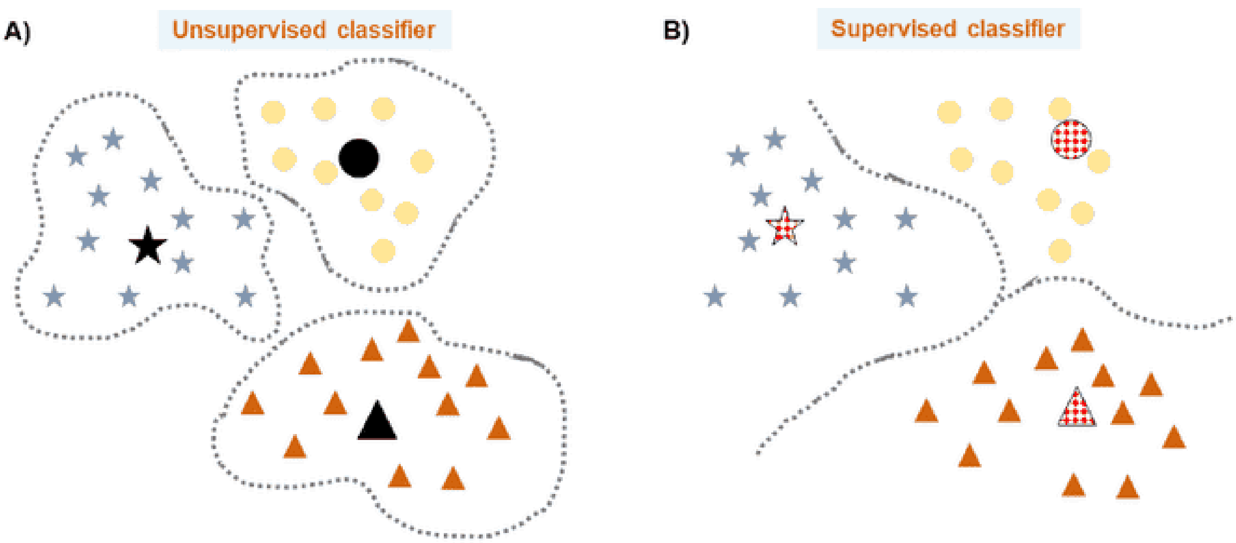
Results of proposed edge detector operator in comparison with the Sobel operator, here (a) original Lena image, (b) result of the Sobel operator (c) result of proposed edge detector operator (d) Lena image corrupted from SNP noise (e) result of the Sobel operator for image of (d) and (f) result of the proposed edge detector operator for image of (d).

Object classification

Object classification involves training a machine learning model on extracted features and corresponding labels. Once trained, the model's performance is evaluated using metrics like accuracy, precision, recall, and F1-score. Finally, the trained model can be deployed to classify new, unseen images.

Image Classification using ML

Image classification is a multifaceted task with no single definitive approach. Two prominent ML methods are **supervised** and **unsupervised** classification, and have emerged as leading strategies in the field.



Unsupervised Classification

- **K-Means**: While K-Means and KNN share a similar name, they are fundamentally different techniques. K-Means is an unsupervised clustering algorithm that groups data points based on their inherent similarities. It operates by selecting initial centroids, assigning data points to their nearest centroid, recalculating centroids, and iterating until convergence. This process helps identify patterns and insights within the data without relying on external labels.
- **Gaussian Mixture Models:** GMMs offer a more sophisticated approach to clustering compared to K-Means. By assuming that data points are drawn from a mixture of Gaussian distributions, GMMs can capture more complex data patterns and handle overlapping distributions. This flexibility makes GMMs a popular choice for image classification tasks.

Supervised Classification

- **Logistic regression**: A binary classification technique, is a valuable tool in image classification. It models the relationship between input image features and the probability of an image belonging to a specific category. By constructing a logistic function, logistic regression assigns a probability value to each input image, enabling a binary classification decision based on a predefined threshold.
- **K-Nearest Neighbors:** Knn is a simple yet effective classification algorithm that operates by memorizing the entire training dataset. Unlike other algorithms that actively learn patterns, KNN relies on comparing new data points to the closest neighbors in the training set to make predictions. This approach can be computationally expensive for large datasets, especially during prediction time.
- **Support Vector Machines:** SVMs are a powerful classification algorithm that aims to find the optimal hyperplane to separate data points into different classes. By maximizing the margin between the hyperplane and the closest data points of each

class, SVMs ensure robust classification performance, even in the presence of noise or outliers.
- **Decision trees:** Decision trees are a simple yet effective classification technique that can be visualized as a flowchart. They make decisions by asking a series of yes/no questions about the features of the data. Similar to guessing a fruit based on its characteristics, decision trees progressively narrow down the possibilities until a prediction can be made. This approach makes them highly interpretable, as the decision-making process can be easily visualized and understood.

Deep neural networks for image classification

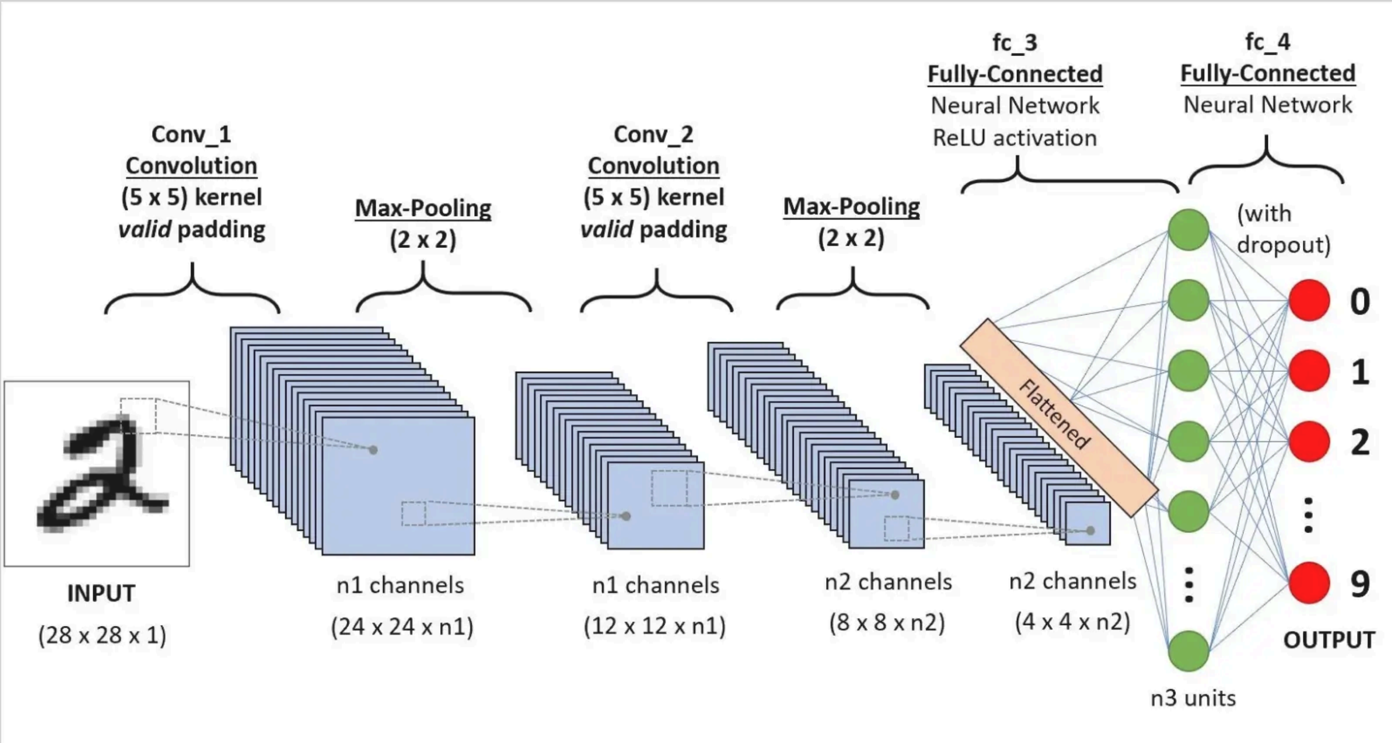
Convolutional Neural Networks (CNNs) are a powerful type of deep learning architecture that have significantly improved the performance of computer vision tasks.
By mimicking the human brain's neural networks, CNNs can process images and extract relevant features with minimal human intervention. CNNs are part of supervised learning algorithms.

Components of CNN:

- **Input Layer:**
    - This is the initial layer where the input image is fed into the network.
    - The image is often resized to a standard size to ensure consistency.
- **Convolutional Layer:**
    - Applies filters to the input image to extract features.
    - Each filter slides across the image, detecting and activating specific patterns.
    - The output is a feature map, which represents the presence or absence of these features in different regions of the image.
- **Activation Function:**
    - Introduces non-linearity into the network, allowing it to learn more complex patterns.
    - Commonly used activation functions include ReLU (Rectified Linear Unit), Leaky ReLU, and sigmoid.
- **Pooling Layer:**
    - Downsamples the feature maps to reduce computational cost and improve generalization.
    - Common pooling techniques include max pooling, average pooling, and global average pooling.
- **Fully Connected Layer:**
    - Connects all neurons in the previous layer to all neurons in the current layer.
    - Combines the extracted features to produce a final classification.
- **Output Layer:**
    - The final layer of the network that produces the classification result.
    - The number of neurons in the output layer corresponds to the number of possible classes.

Additional Layers:

- **Dropout:** Prevents overfitting by randomly dropping neurons during training.
- **Dense:** Combines the features from the previous layer, reducing the dimensionality and preparing the data for the final output.
- **Batch normalization:** Helps stabilize training and improves convergence speed.
- **Residual connections:** Improve the flow of information through the network, especially for deep architectures.
- **Softmax layer:** Converts logits into probabilities for multi-class classification.



Benefits of CNN:

- **Hierarchical feature learning:** CNNs learn features at different levels of abstraction, from simple low-level features (e.g., edges, lines) to complex high-level features (e.g., objects, scenes).
- **Spatial invariance:** CNNs are translationally invariant, meaning they can recognize objects regardless of their position within an image.
- **Efficient computation:** CNNs use shared weights, reducing the number of parameters and computational cost compared to fully connected networks.
- **Adaptability:** CNNs can be adapted to various image-related tasks, such as image classification, object detection, semantic segmentation, and image generation.

Applications of image classification

Image classification has a wide range of applications across various industries and domains. Here are some of the most common uses:

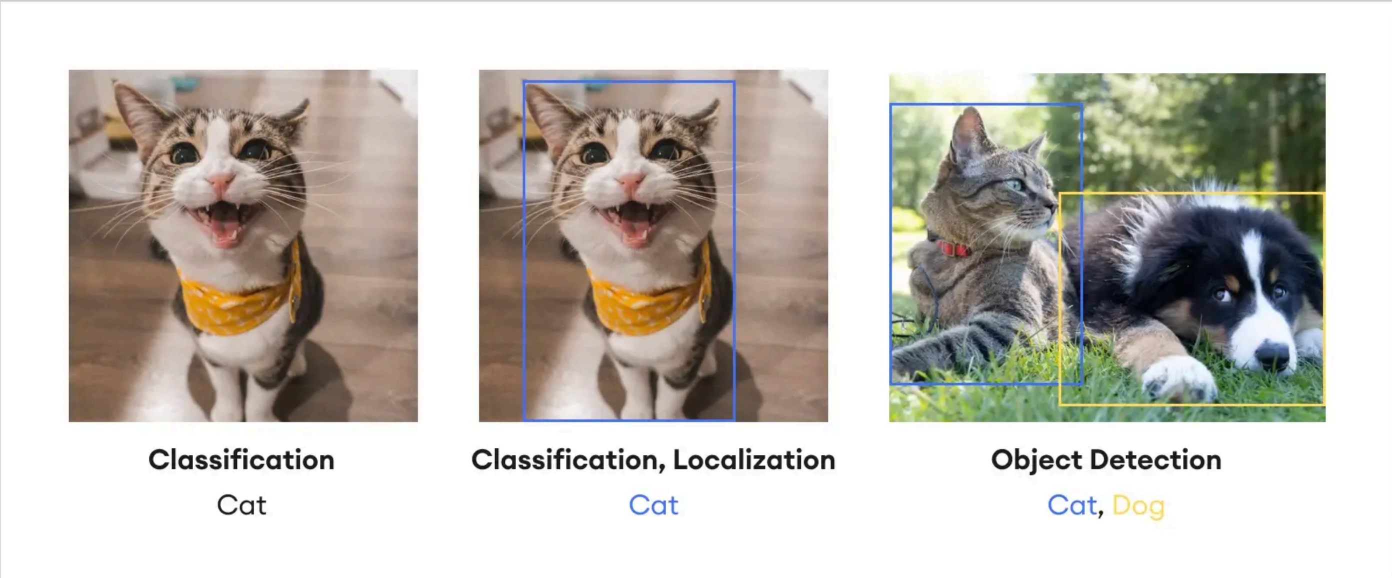
- **Medical Imaging**
    - **Disease Diagnosis:** Identifying diseases based on X-rays, CT scans, MRIs, and other medical images.
    - **Tumor Detection:** Locating and classifying tumors in images.
    - **Anomaly Detection:** Identifying abnormalities in medical images.
- **Retail and E-commerce**
    - **Product Search:** Enabling users to search for products based on images.
    - **Visual Inventory Management:** Automatically categorizing and managing product images.
    - **Personalized Recommendations:** Suggesting products based on visual preferences.
- **Autonomous Vehicles**
    - **Object Detection:** Identifying objects on the road, such as pedestrians, vehicles, and traffic signs.
    - **Scene Understanding:** Understanding the surrounding environment for safe navigation.
- **Security and Surveillance**
    - **Facial Recognition:** Identifying individuals based on their facial features.
    - **Object Tracking:** Tracking objects in videos for surveillance purposes.
    - **Anomaly Detection:** Identifying unusual activities or objects in images and videos.
- **Agriculture**
    - **Crop Monitoring:** Assessing crop health and yield using aerial imagery.
    - **Pest and Disease Detection:** Identifying pests and diseases in crops.
    - **Yield Prediction:** Predicting crop yields based on image analysis.
- **Manufacturing**
    - **Quality Control:** Inspecting products for defects using visual inspection.
    - **Defect Detection:** Identifying defects in manufactured goods.
- **Entertainment**
    - **Image Editing:** Enhancing images using various filters and effects.
    - **Content Moderation:** Filtering inappropriate content from images and videos.
- **Art and Design**
    - **Style Analysis:** Analyzing artistic styles and techniques.
    - **Image Generation:** Creating new images based on existing styles and patterns.

Image classification vs. Object detection

Image classification labels the entire image as a single class.

Object localization identifies and locates specific objects within the image by drawing bounding boxes around them.

Object detection is a more advanced task than image classification. It classifies multiple entities in a single image by combining image classification and object localization.



Github Repo

https://github.com/RahulM-3/Object-Detection