**IMAGE RECOGNITION WITH IBM CLOUD VISUAL RECOGNITION**

PHASE 2 :INNOVATION

Image recognition is the process of identifying an object or a feature in an image or video. It is used in many applications like defect detection, medical imaging, and security surveillance.using deep learning machine learning to create image recohnition with IBM cloud visual recognition.

**1.Integrating the intial steps:**

Image recognition can speed up tedious tasks and process images faster or more accurately than manual image inspection. Image recognition is a crucial technique in many applications, and is the main driver in deep learning applications like:

* **Visual Inspection**: Identifying parts as defective or non-defective in manufacturing can quickly inspect thousands of parts on an assembly line.
* **Image Classification**: Categorizing images based on the image content. This is especially useful in applications such as image retrieval and recommender systems in e-commerce.
* **Automated Driving**: The ability to recognize a stop sign or a pedestrian in an image is crucial to autonomous driving applications.
* **Robotics**: Image recognition can be used by robots to identify objects and enhance autonomous navigation by identifying locations or objects on their path.

### 2.Data collection

To achieve image recognition, machine vision artificial intelligence models are fed with pre-labeled data to teach them to recognize images they’ve never seen before.

Some of the massive publicly available databases include Pascal VOC and ImageNet. They contain millions of labeled images describing the objects present in the pictures—everything from sports and pizzas to mountains and cats.

* **Variation in the viewpoint of the image.**The images can be aligned at different angles or vary in dimension, which can lead to inaccurate prediction of the machine learning model. The system fails to understand the effect of changing the alignment and viewport of the image.
* **Deformation.** Generally, training data gives a biased perception that a particular object can only have a specific shape.
* **Occlusion.** Some objects may obstruct the full view of an image and result in partial information being fed to the system. The neural network should acknowledge these variations as a part of the training process.

3.Pre-processing of the image data

Once the dataset is ready, there are several things to be done to maximize its efficiency for model training.

**DATA ANNOTATION**

* Once the dataset is ready, there are several things to be done to maximize its efficiency for model training.
* The objects in the image that serve as the regions of interest have to labeled (or annotated) to be detected by the computer vision system. In other words, labels have to be applied to those frames or images.
* Annotations for segmentation tasks can be performed easily and precisely by making use of V7 annotation tools, specifically the polygon annotation tool and the [auto-annotate tool](https://www.v7labs.com/automated-annotation). A label once assigned is remembered by the software in the subsequent frames.

4.Model architecture and training process

* Due to their unique work principle, convolutional neural networks (CNN) yield the best results with deep learning image recognition.
* The complete pixel matrix is not fed to the CNN directly as it would be hard for the model to extract features and detect patterns from a high-dimensional sparse matrix. Instead, the complete image is divided into small sections called feature maps using filters or kernels.
* The convolution layers in each successive layer can recognize more complex, detailed features—visual representations of what the image depicts.
* The corresponding smaller sections are normalized, and an activation function is applied to them. Rectified Linear Units (ReLu) are seen as the best fit for image recognition tasks.
* The objects in the image that serve as the regions of interest have to labeled (or annotated) to be detected by the computer vision system. In other words, labels have to be applied to those frames or images.

5.Traditional machine learning algorithms for image recognition

Before the development of parallel processing and extensive computing capabilities required for training deep learning models, traditional machine learning models had set standards for image processing.

Let us quickly walk through some of the most learned machine learning models:

**Support Vector Machines**

SVMs describe features by making histograms of images. They use a sliding detection window technique by moving around the image. The algorithm then takes the test picture and compares the trained histogram values with the ones of various parts of the picture to check for close matches.

**Bag of Features**

Bag of Features models like Scale Invariant Feature Transformation (SIFT) does pixel-by-pixel matching between a sample image and its reference image. The trained model then tries to pixel match the features from the image set to various parts of the target image to see if matches are found.

Some other machine learning models widely used in computer vision include:

* Regression Algorithms
* Instance-based Algorithms
* Regularization Algorithms
* Decision Tree Algorithms
* Bayesian Algorithms
* Clustering Algorithms

6.Popular deep learning models for image recognition

**Single-shot detector (SSD)**

Single-shot detectors divide the image into a default number of bounding boxes in the form of a grid over different aspect ratios. The feature map that is obtained from the hidden layers of neural networks applied on the image is combined at the different aspect ratios to naturally handle objects of varying sizes.

These types of object detection algorithms are flexible and accurate and are mostly used in face recognition scenarios where the training set contains few instances of an image.

Other machine learning algorithms include Fast RCNN (Faster Region-Based CNN) which is a region-based feature extraction model—one of the best performing models in the family of CNN.