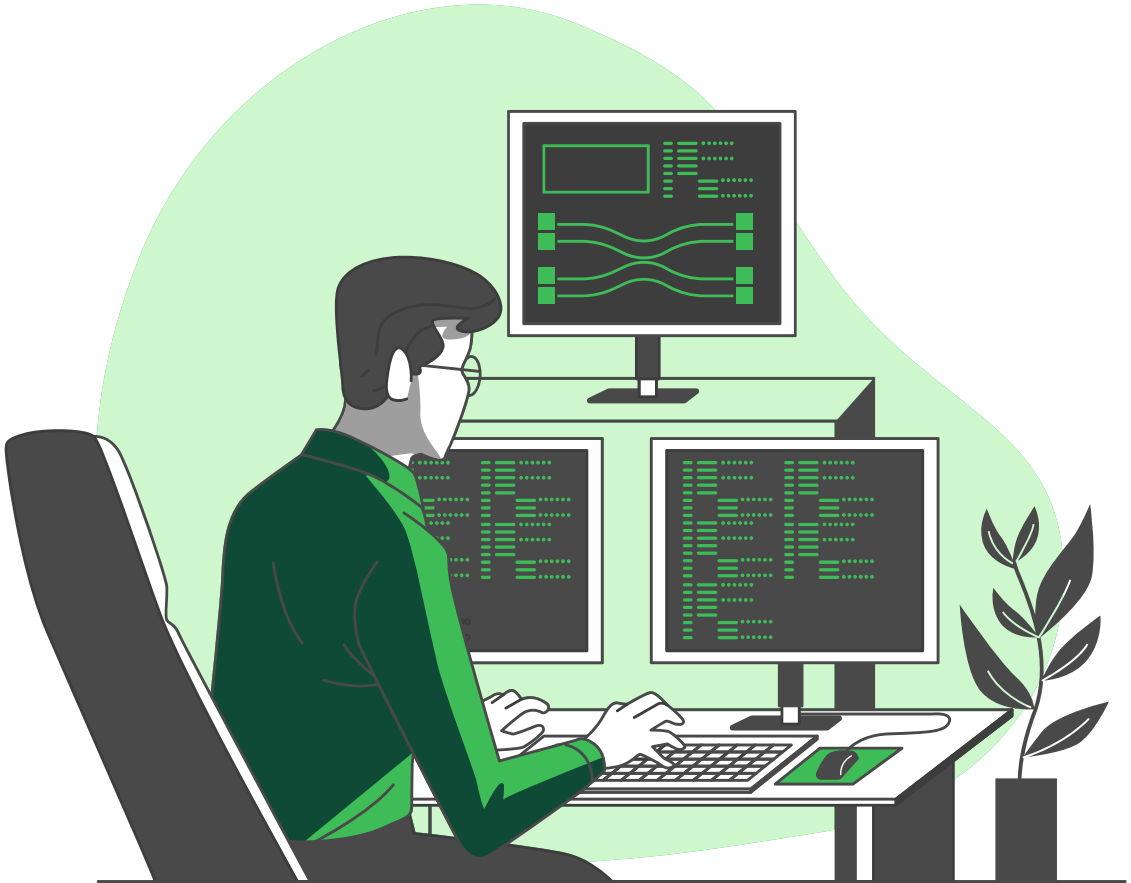
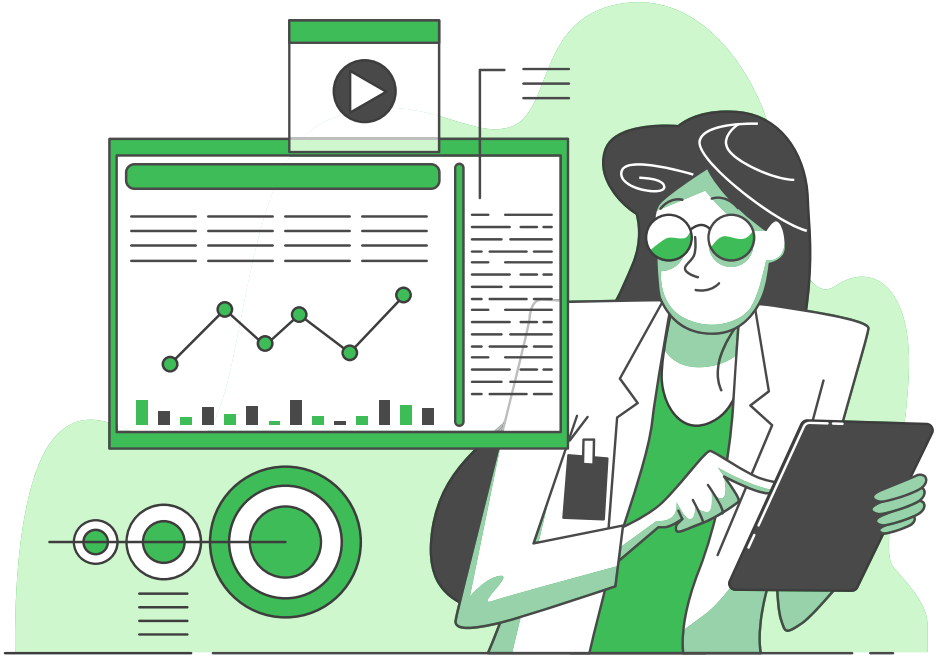


Summary : Intermediate Data Analysis For Image Data

by : Dhea Fajriati Anas

iykra

Table of Contents



01

Type of Computer Vision

Type of Computer Vision

01 Image Classification

Given a set of images that are all labeled with a single category.

02 Object Detection

Define objects within images usually involves outputting bounding boxes and labels for individual objects.

03 Semantic Segmentation

Process of segmentation, which divides whole images into pixel groupings which can then be labelled and classified.

04 Instance Segmentation

Instance Segmentation segments different instances of classes, such as labelling 5 cars with 5 different colors.

Type of Computer Vision

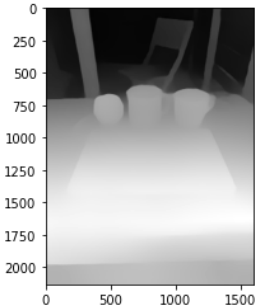
05

Panopic Segmentation

Combination of Semantic Segmentation and Instance Segmentation.

06

Depth Perception



You can also estimate the 3D depth of the objects and the scenes with this neural network.

07

Object Tracking

Object tracking refers to the process of following a specific object of interest, or multiple objects, in a given scene.

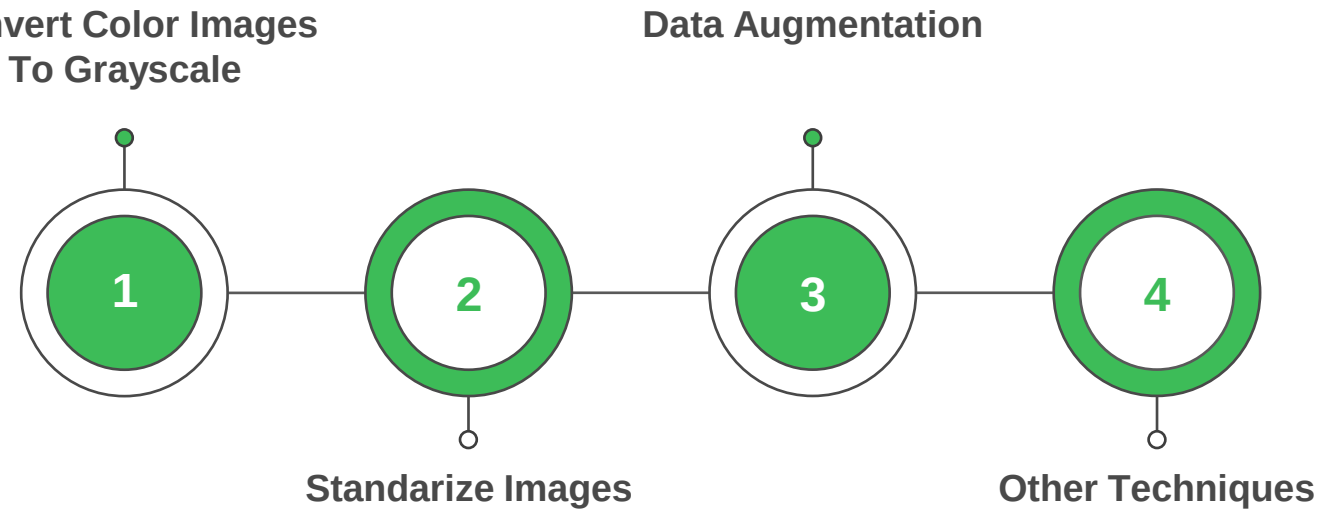
08

Image Captioning

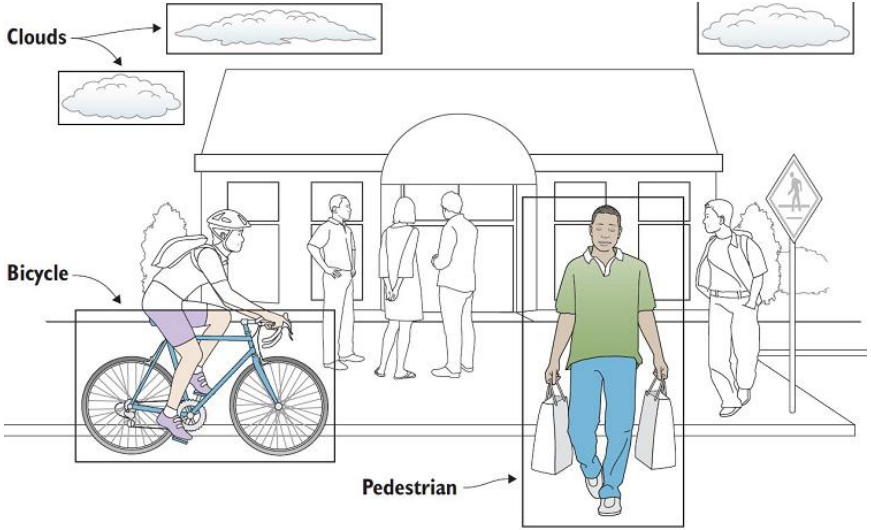
When you give the neural network an image, it creates a caption for you describing the image.

02
Preprocessing for Image Data

Image Data Preprocessing Techniques



1. Convert Color Images To Grayscale



Convert color images to grayscale to reduce computation complexity: in certain problems you'll find it useful to lose unnecessary information from your images to reduce space or computational complexity.

For example, converting your colored images to grayscale images. This is because in many objects, color isn't necessary to recognize and interpret an image.

In the example beside, you can see how patterns in brightness and darkness of an object (intensity) can be used to define the shape and characteristics of many objects. In other applications, color is important to define certain objects. Like skin cancer detection which relies heavily on the skin colors (red rashes).

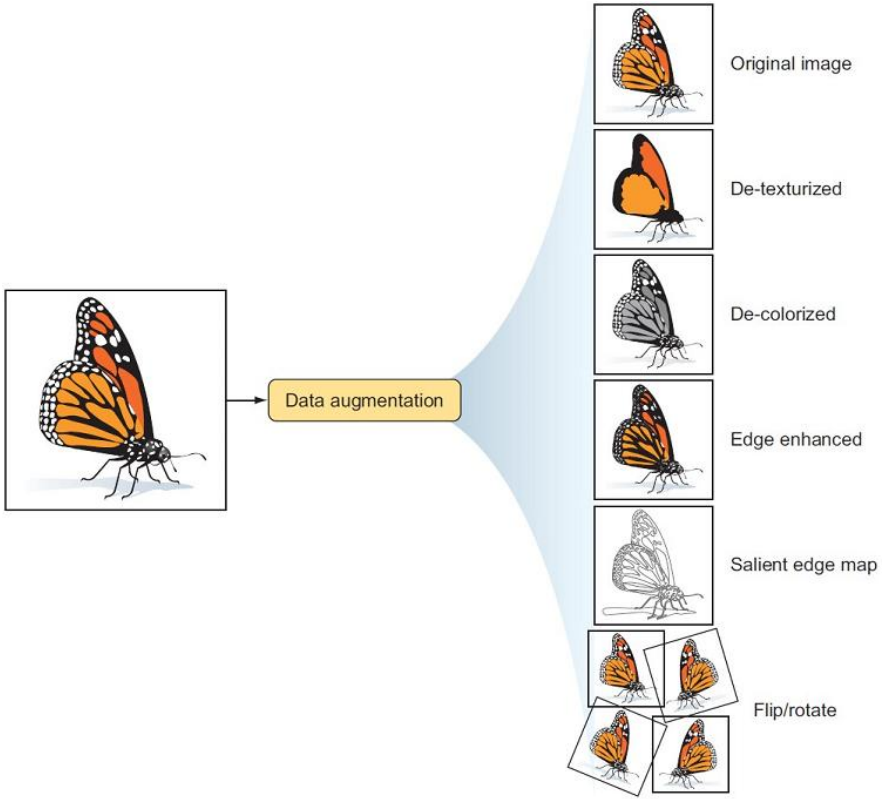
Source: https://freecontent.manning.com/the-computer-vision-pipeline-part-3-image-preprocessing/

2. Standarize Images

One important constraint that exists in some machine learning algorithms, such as CNN, is the need to resize the images in your dataset to a unified dimension. **This implies that our images must be preprocessed and scaled to have identical widths and heights before fed to the learning algorithm.**

Source: https://freecontent.manning.com/the-computer-vision-pipeline-part-3-image-preprocessing/

3. Data Augmentation



Another common pre-processing technique involves augmenting the existing dataset with perturbed versions of the existing images. **Scaling, rotations and other affine transformations are typical**. This is done to enlarge your dataset and expose the neural network to a wide variety of variations of your images. This makes it more likely that your model recognizes objects when they appear in any form and shape.

Source: https://freecontent.manning.com/the-computer-vision-pipeline-part-3-image-preprocessing/

4. Other Techniques

Many preprocessing techniques can be used to get your images ready to train the machine learning model

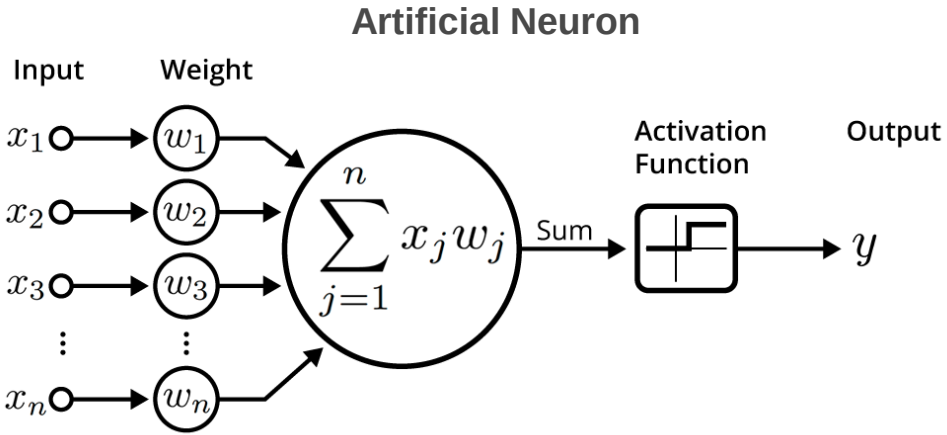
- In some projects, you might need to remove the background color from your images to reduce the noise.
- Other projects might require that you brighten or darken your images.

In short, any adjustments that you need to apply to your dataset are considered a sort of preprocessing. And you'll select the appropriate processing techniques based on the dataset at hand and the problem you're solving. That builds your intuition of which ones you need when working on your own projects.

03

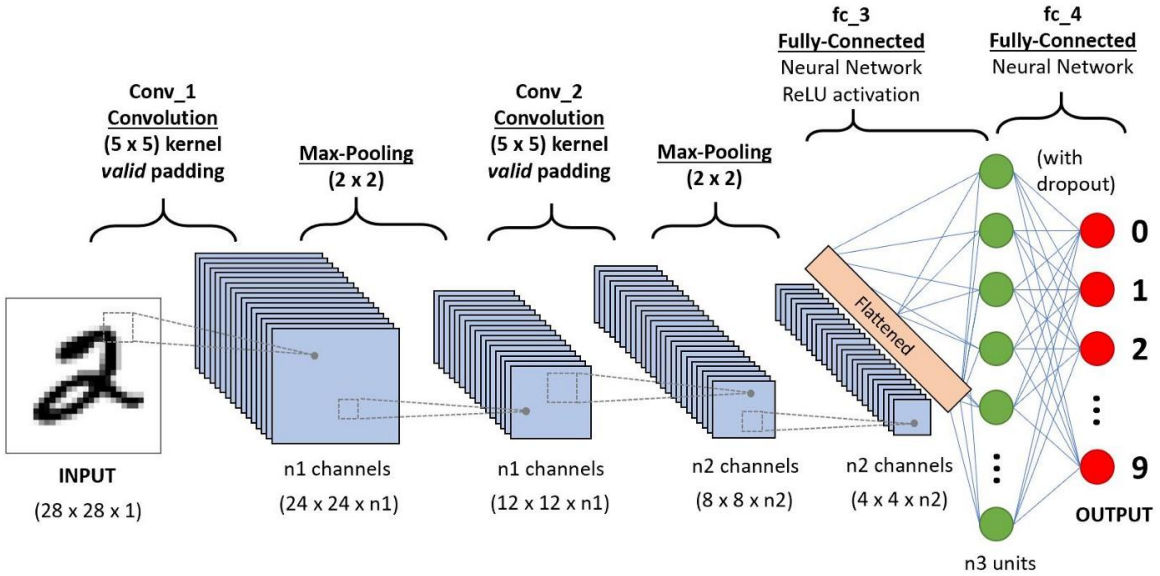
Convolutional Neural Nerwork (CNN)

Introduction



- Once a neuron receives its inputs from the neurons in the preceding layer of the model, it adds up each signal multiplied by its corresponding weight and passes them on to an activation function.
- The activation function calculates the output value for the neuron. This output value is then passed on to the next layer of the neural network through another synapse.

Introduction



A **Convolutional Neural Network (ConvNet/CNN)** is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.

Feature Learning

01

Filter

Component in the convolutional layer which contains several weight matrices or commonly called the kernel. For example: The convolutional layer uses 6 filters, each filter has 3 kernels with a size of 5x5.

02

Stride

The number of filter shifts in the convolution process. If the stride value is 1, then the filter will shift by 1 pixel horizontally and then vertically.

03

Padding

The addition of pixel size with a certain value around the input data so that the results of the receptive field are not too small so that not much information is lost.

04

Pooling

The process of reducing the size of an image data. Basically the pooling layer consists of a filter with a certain size and stride that will alternately shift over the entire feature map area.

Classifier

01

Softmax Classifier

Another form of Logistic Regression algorithm that can be used to classify more than two classes.

02

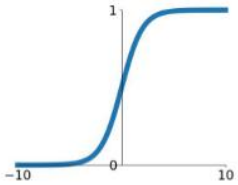
Adaptive Moment Estimation or ADAM

Type of optimizer that can be used on Convolutional Neural Networks. The optimizer functions to update the weights, so that a weight is obtained that can minimize the loss function.

Activation Function

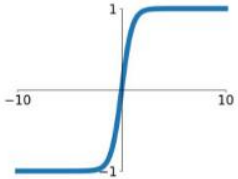
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



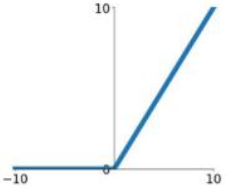
tanh

$$\tanh(x)$$



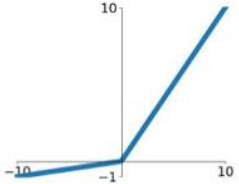
ReLU

$$\max(0, x)$$



Leaky ReLU

$$\max(0.1x, x)$$

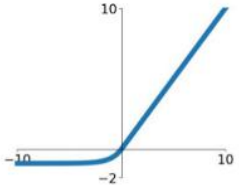


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



Architectures of CNNs

- LeNet
- AlexNet
- VGGNet
- GoogleNet
- ResNet
- ZFNet

Thanks!

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, infographics & images by Freepik and illustrations by Stories