missinglink.ai

Neural Networks for Image Recognition

Methods, Best Practices, Applications

Computer Vision

Neural Networks for Image Recognition:

thousands of companies and millions of consumers every day. Under the hood, image recognition is powered by deep learning, specifically Convolutional Neural Networks (CNN), a neural network architecture which emulates how the visual cortex breaks down and analyzes image data. CNN and neural network image recognition is a core component of deep learning for computer vision, which has many applications including e-commerce, gaming, automotive, manufacturing, and education.

In this article you will learn

☰ What is Image Recognition?   ↓

☰ How Does Image Recognition Work?   ↓

☰ Image Data Pre-Processing Steps for Neural Networks   ↓

![missinglink.ai]

What is Image Recognition?
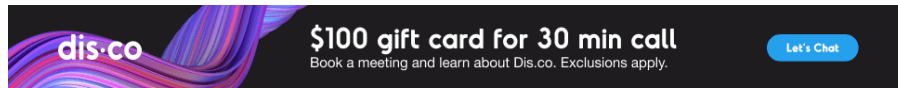
Image recognition uses artificial intelligence technology to automatically identify objects, people, places and actions in

Image recognition is natural for humans and animals but is an extremely difficult task for computers to perform. Over the past two decades, the field of Computer Vision has emerged, and tools and technologies have been developed which can rise to the challenge.

The most effective tool found for the task for image recognition is a deep neural network (see our guide on artificial neural network concepts), specifically a Convolutional Neural Network (CNN). CNN is an architecture designed to efficiently process, correlate and understand the large amount of data in high-resolution images.

How Does Image Recognition Work?

The human eye sees an image as a set of signals, interpreted by the brain's visual cortex. The outcome is an experience of a scene, linked to objects and concepts that are retained in memory. Image recognition imitates this process. Computers

depict physical objects and features. Computer vision systems can logically analyze these constructs, first by simplifying images and extracting the most important information, then by organizing data through feature extraction and classification.

Finally, computer vision systems use classification or other algorithms to make a decision about the image or part of it – which category they belong to, or how they can best be described.

Image Recognition Algorithms

such as dog, cat or table. The algorithm needs to be trained to learn and distinguish between classes.

In a simple case, to create a classification algorithm that can identify images with dogs, you'll train a neural network with thousands of images of dogs, and thousands of images of backgrounds without dogs. The algorithm will learn to extract the features that identify a "dog" object and correctly classify images that contain dogs. While most image recognition algorithms are classifiers, other algorithms can be used to perform more complex activities. For example, a Recurrent Neural Network can be used to automatically write captions describing the content of an image.

Image Data Pre-Processing Steps for Neural Networks

Neural network image recognition algorithms rely on the quality of the dataset – the images used to train and test the

missinglink.ai

- **The number of images**—the more data you feed to a model, the more accurate it will be, but ensure the training set represents the real population.

- **The number of channels**—grayscale image have 2 channels (black and white) and color images typically have 3 color channels (Red, Green, Blue / RGB), with colors represented in the range [0,255].

- **Aspect ratio**—ensure the images have the same

- **Image scaling**—once all images are squared you can scale each image. There are many up-scaling and down-scaling techniques, which are available as functions in deep learning libraries.

- **Mean, standard deviation of input data**—you can look at the 'mean image' by calculating the mean values for each pixel, in all training examples, to obtain information on the underlying structure in the images.

- **Normalizing image inputs**—ensures that all input parameters (pixels in this case) have a uniform data distribution. This makes convergence speedier when you train the network. You can conduct data normalization by subtracting the mean from each pixel and then dividing the outcome by the standard deviation.

- **Dimensionality reduction**—you can decide to collapse the RGB channels into a gray-scale channel. You may want to reduce other dimensions if you intend to make

neural network to a variety of variations. This way this neural network is less likely to identify unwanted characteristics in the data-set.

Building a Predictive Model for Images with Neural Networks

Once training images are prepared, you'll need a system that can process them and use them to make a prediction on new,

(see our in-depth article on classification and neural networks). Neural networks are an interconnected collection of nodes called neurons or perceptrons. Every neuron takes one piece of the input data, typically one pixel of the image, and applies a simple computation, called an activation function to generate a result. Each neuron has a numerical weight that affects its result. That result is fed to additional neural layers until at the end of the process the neural network generates a prediction for each input or pixel.

Perceptron Input And Output

Input Layer

Hidden Layer

Output Layer

co missinglink.ai

≡
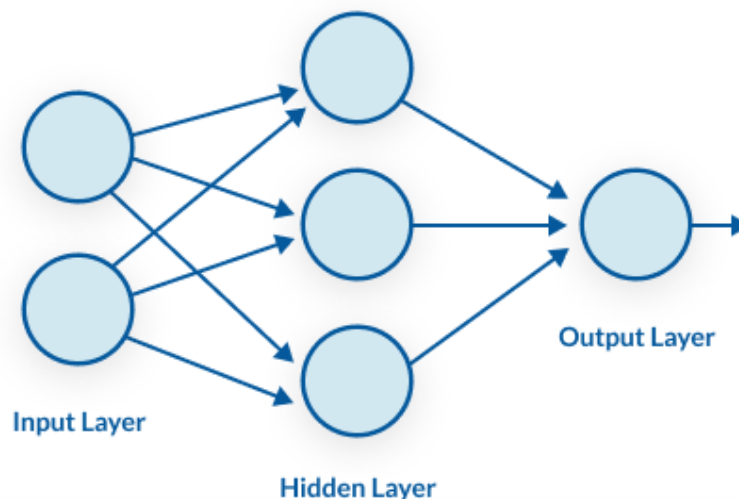
model can be used to classify real-world images.

Limitations of Regular Neural Networks for Image Recognition

Traditional neural networks use a fully-connected architecture, as illustrated below, where every neuron in one layer connects to all the neurons in the next layer.


Fully-Connected Neural Network
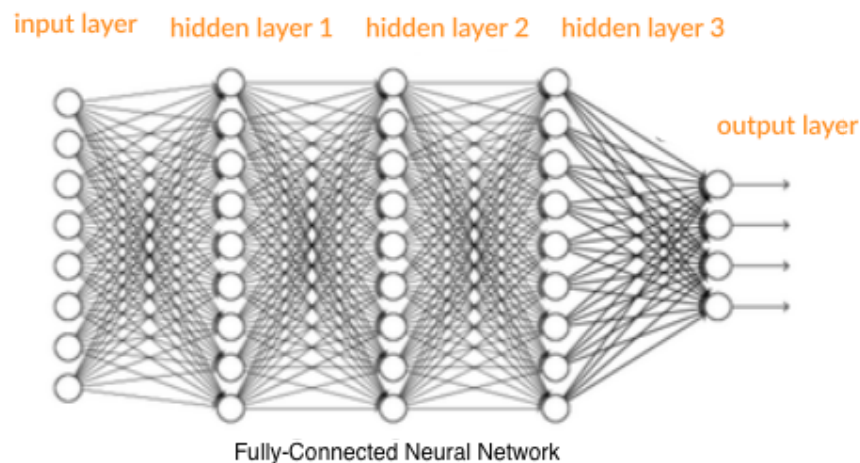
A fully connected architecture is inefficient when it comes to processing image data:

- For an average image with hundreds of pixels and three channels, a traditional neural network will generate millions of parameters, which can lead to overfitting.

- The model would be very computationally intensive.

- It may be difficult to interpret results, debug and tune the model to improve its performance.

Convolutional Neural Networks and Their Role in Image Recognition

only analyzed in relation to pixels nearby), making the training process computationally achievable. In a CNN each group of neurons focuses on one part of the image. For example, in a cat image, one group of neurons might identify the head, another the body, another the tail, etc. There may be several stages of segmentation in which the neural network image recognition algorithm analyzes smaller parts of the images, for example, within the head, the cat's nose, whiskers, ears, etc. The final output is a vector of probabilities, which predicts, for each feature in the image, how likely it is to belong to a class or category.

A traditional neural network structure (left)   VS.   3-dimensional CNN structure (right)
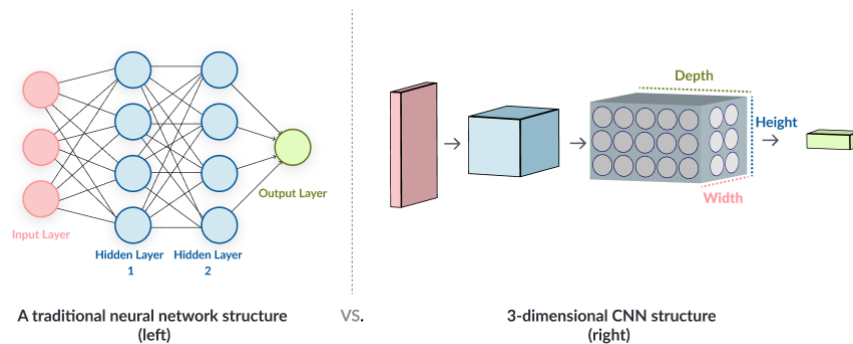
To learn more about how CNNs work, see our in-depth Convolutional Neural Networks Guide.

Effectiveness and Limitations of Convolutional Neural Networks

A CNN architecture makes it possible to predict objects and faces in images using industry benchmark datasets with up to 95% accuracy, greater than human capabilities which stand at 94% accuracy. Even so, convolutional neural networks have their limitations:

- **Require high processing power**. Models are typically trained on high-cost machines with specialized Graphical Processing Units (GPUs).

missinglink.ai

Image Recognition Applications

Implementations of image recognition include security and surveillance, face recognition, visual geolocation, gesture recognition, object recognition, medical image analysis, driver assistance, and image tagging and organization in websites or large databases. Image recognition has entered the mainstream. Face, photo, and video frame recognition is used in production by Facebook, Google, Youtube, and many other high profile consumer applications. Toolkits and cloud services have emerged which can help smaller players integrate image recognition into their websites or applications.

Use of Image Recognition Across Industries

- **E-commerce industry**—image recognition is used to automatically process, categorize and tag product images, and enable powerful image search. For example, consumers can search for a chair with a particular armrest and receive relevant results.

- **Gaming industry**—image recognition can be used to transpose a digital layer on top of images from the real world. Augmented reality adds details to the existing environment. Pokemon Go is a popular game that relies on image recognition technology.

- **Automotive industry**—autonomous vehicles are in testing phases in the United States and are used for public transport in many European cities. To facilitate autonomous driving, image recognition is taught to identify objects on the road, including moving objects, vehicles, people and pathways, as well as recognize traffic lights and road signs.

- **Education**—image recognition can help students with learning difficulties and disabilities. For example, applications powered by computer vision provide image-to-speech and text-to-speech functions, which can read out materials to students with dyslexia or impaired vision.

In this article we explained the basics of image recognition, and how it can be achieved by Convolutional Neural Networks. When you start working on CNN projects, using deep learning frameworks like TensorFlow, Keras and PyTorch to process and classify images, you'll run into some practical challenges:

Tracking experiments

Tracking experiment source code, configuration, and hyperparameters. Convolutional networks can have many parameter and structural variations. You'll need to run hundreds or thousands of experiments to find hyperparameters that provide the best performance. Organizing, tracking and sharing experiment data and results can be a challenge.

missinglink.ai
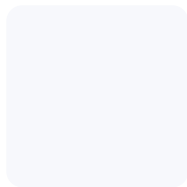
experiments between them, is time-consuming.

Manage training data

Computer vision projects involve rich media such as images or video, with large training sets weighing Gigabytes to Petabytes. Copying data to each training machine, then re-copying when you change training sets, can be time-consuming and error-prone.

MissingLink is a deep learning platform that can help you automate these operational aspects of CNNs and computer vision, so you can concentrate on building winning image recognition experiments. Sign up for free to see how easy it is.

Learn More About Computer Vision