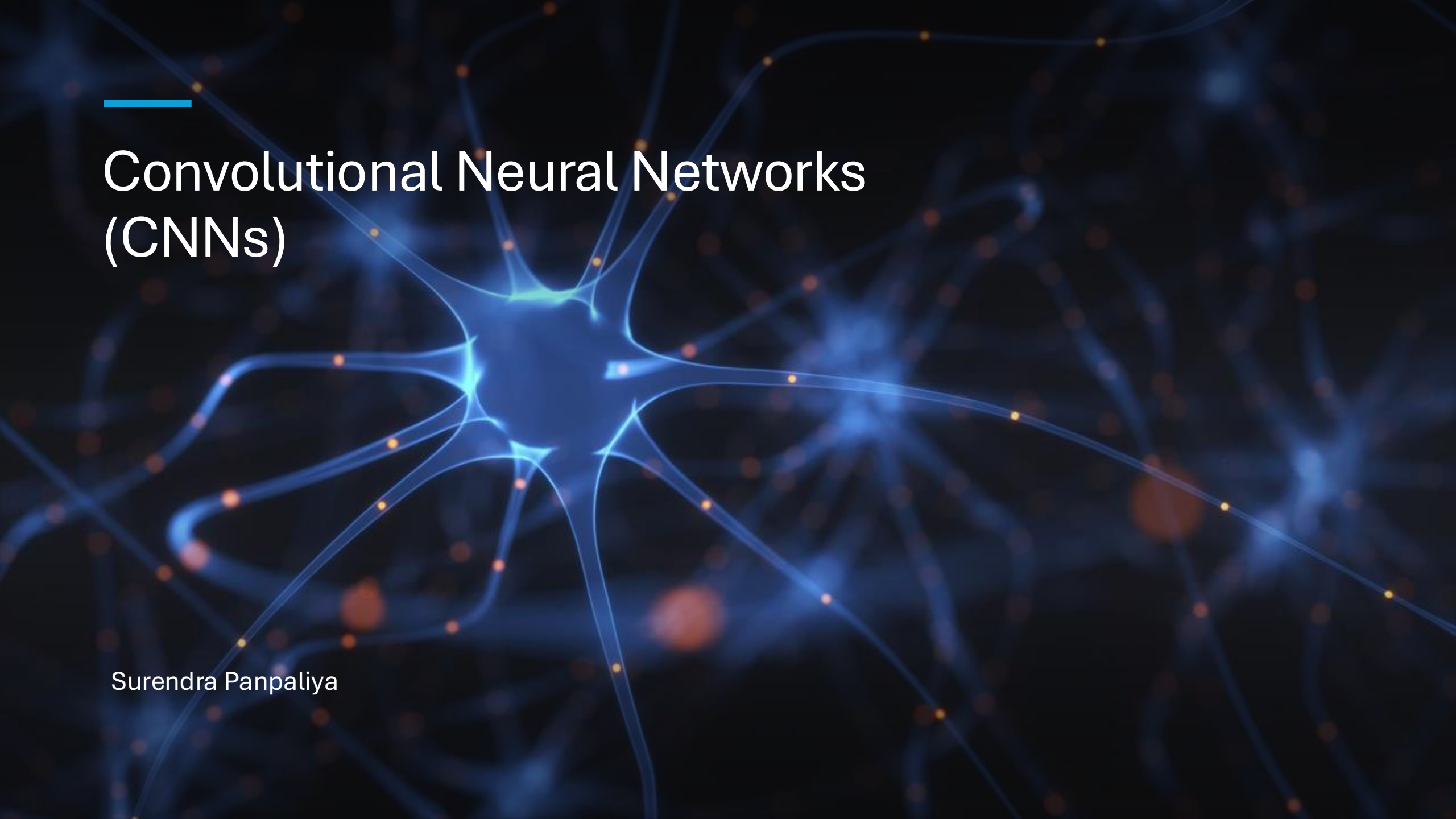


Convolutional Neural Networks (CNNs)

Surendra Panpaliya

Agenda

- Introduction to CNN
- What is CNN?
- Architecture of CNNs
- CNN Key Components
- Applications of CNNs

Introduction to CNN

CONVOLUTIONAL NEURAL NETWORKS (CNNS)

ARE SPECIALIZED DEEP LEARNING ARCHITECTURES

DESIGNED TO PROCESS GRID-LIKE DATA,

SUCH AS IMAGES AND VIDEO FRAMES.

Introduction to CNN

ENABLING MACHINES TO PERFORM TASKS

OBJECT DETECTION

IMAGE CLASSIFICATION

FACIAL RECOGNITION

WITH HIGH ACCURACY

What is CNN ?

CNN has

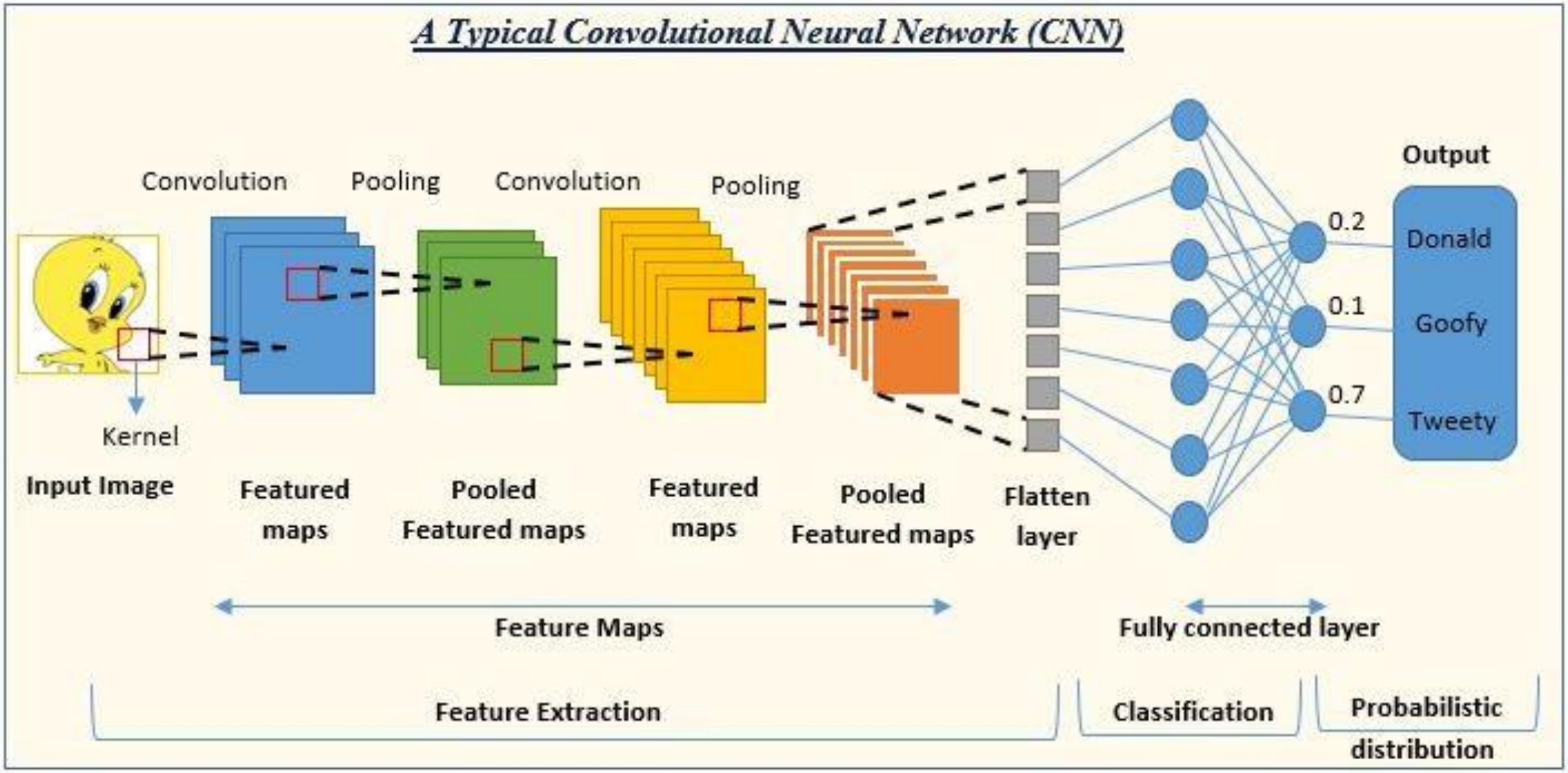
input layer, output layer,

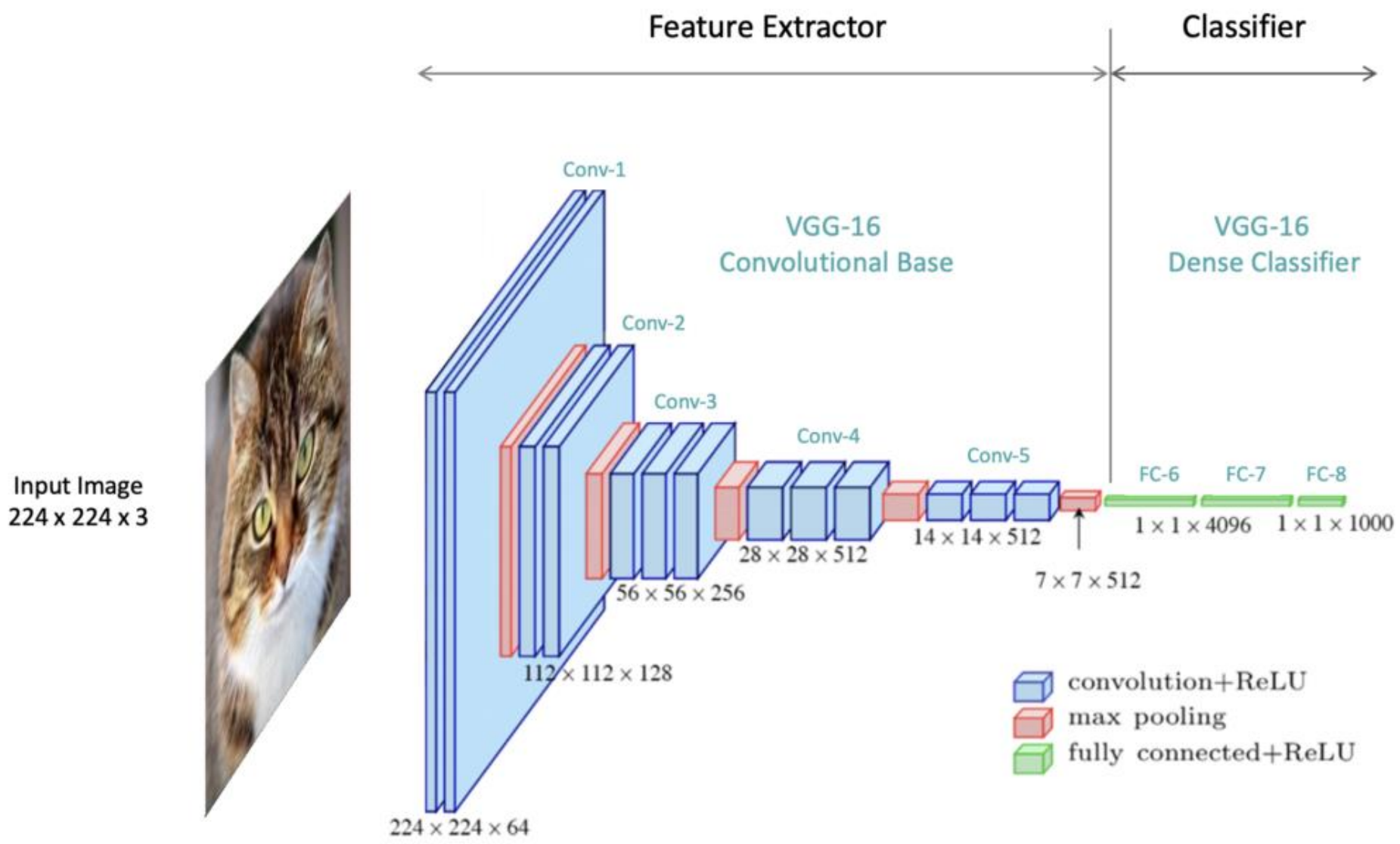
many hidden layers and

millions of parameters

that have the ability

to learn complex objects and patterns.


A Typical Convolutional Neural Network (CNN)
Convolution
Pooling
Convolution
Pooling
Output
0.2
0.1
0.7
Donald
Goofy
Tweety
Kernel
Input Image
Featured maps
Pooled Featured maps
Featured maps
Pooled Featured maps
Flatten layer
Feature Maps
Fully connected layer
Feature Extraction
Classification
Probabilistic distribution


Feature Extractor
Classifier
Input Image
224 x 224 x 3
Conv-1
Conv-2
Conv-3
Conv-4
Conv-5
FC-6
FC-7
FC-8
VGG-16
Convolutional Base
VGG-16
Dense Classifier
224 × 224 × 64
112 × 112 × 128
56 × 56 × 256
28 × 28 × 512
14 × 14 × 512
7 × 7 × 512
1 × 1 × 4096
1 × 1 × 1000
convolution+ReLU
max pooling
fully connected+ReLU

Convolution

Convolution is the process

involving combination of

two functions that produces

the other function as a result.

Convolution

Input image is subjected

to convolution with use of filters

that produces a **Feature map**.

Filters / Kernels

Filters are

randomly generated vectors

in the network consisting of

weights and biases.

Filters / Kernels

Many filters can be generated

where every filter captures

unique feature from input.

Convolution Layer

Layer generally has

Input vectors (*Image*)

Filters (*Feature Detector*)

Output vectors (*Feature map*)

Convolution Layer

Layer identify and extract

best features/patterns

from input image and

preserves the generic information

into a matrix.

Architecture of CNNs

A TYPICAL CNN CONSISTS OF

SEVERAL LAYERS,

EACH WITH A SPECIFIC ROLE

IN FEATURE EXTRACTION AND

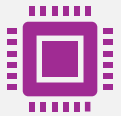
LEARNING PATTERNS.

Convolutional Layer

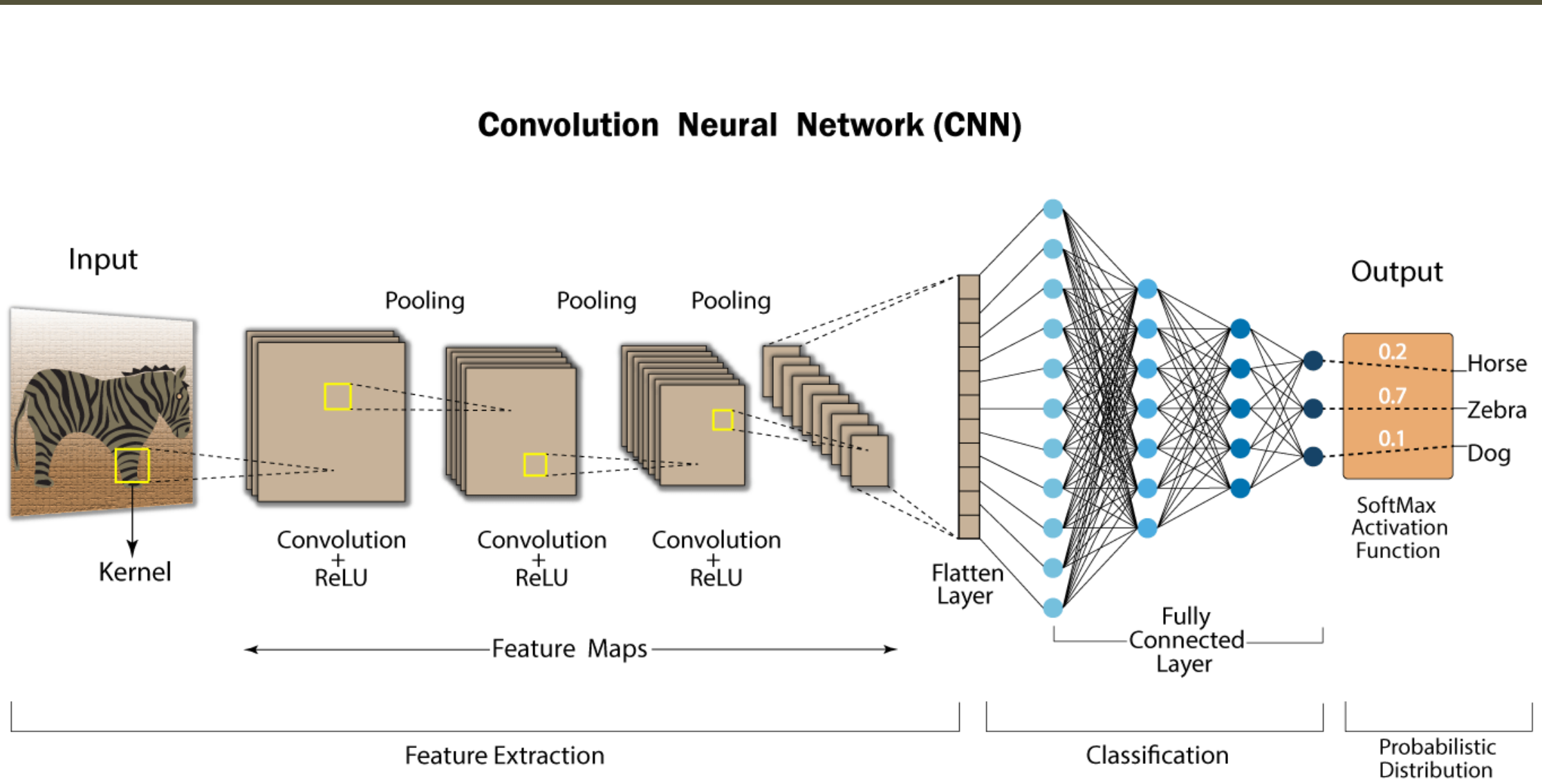
Applies convolutional filters (kernels)

to input data to extract features

such as edges, textures, and patterns.



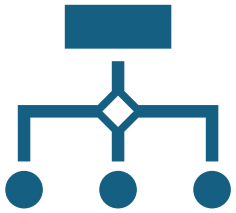
Output of this layer is called a feature map.


Convolution Neural Network (CNN)
Input
Kernel
Pooling
Pooling
Pooling
Convolution + ReLU
Convolution + ReLU
Convolution + ReLU
Feature Maps
Flatten Layer
Fully Connected Layer
Output
0.2
0.7
0.1
Horse
Zebra
Dog
SoftMax Activation Function
Feature Extraction
Classification
Probabilistic Distribution

Activation Function

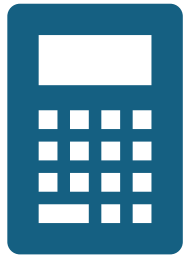
Enabling it to learn complex patterns.

Commonly used activation function



ReLU (Rectified Linear Unit)

ReLU Layer (Rectified Linear Unit)



ReLU is computed after convolution.

Allows the neural network

to account for non-linear relationships.

ReLU Layer (Rectified Linear Unit)

In a given matrix (x),

ReLU sets all negative values

to zero

all other values

remains constant.

ReLU Layer (Rectified Linear Unit)

y = max(0, x)

In a given matrix (M)

M = [[-3, 19, 5], [7, -6, 12], [4, -8, 17]]

ReLU converts it as

[[0, 19, 5], [7, 0, 12], [4, 0, 17]]

Pooling / Sub-sampling Layer

Pooling layer operates

on each feature map

independently.

Pooling / Sub-sampling Layer

REDUCES RESOLUTION

OF THE FEATURE MAP

BY REDUCING HEIGHT AND WIDTH

OF FEATURES MAPS.

Pooling / Sub-sampling Layer

Retains features of the map

required for classification

Called ***Down-sampling***

Max-pooling

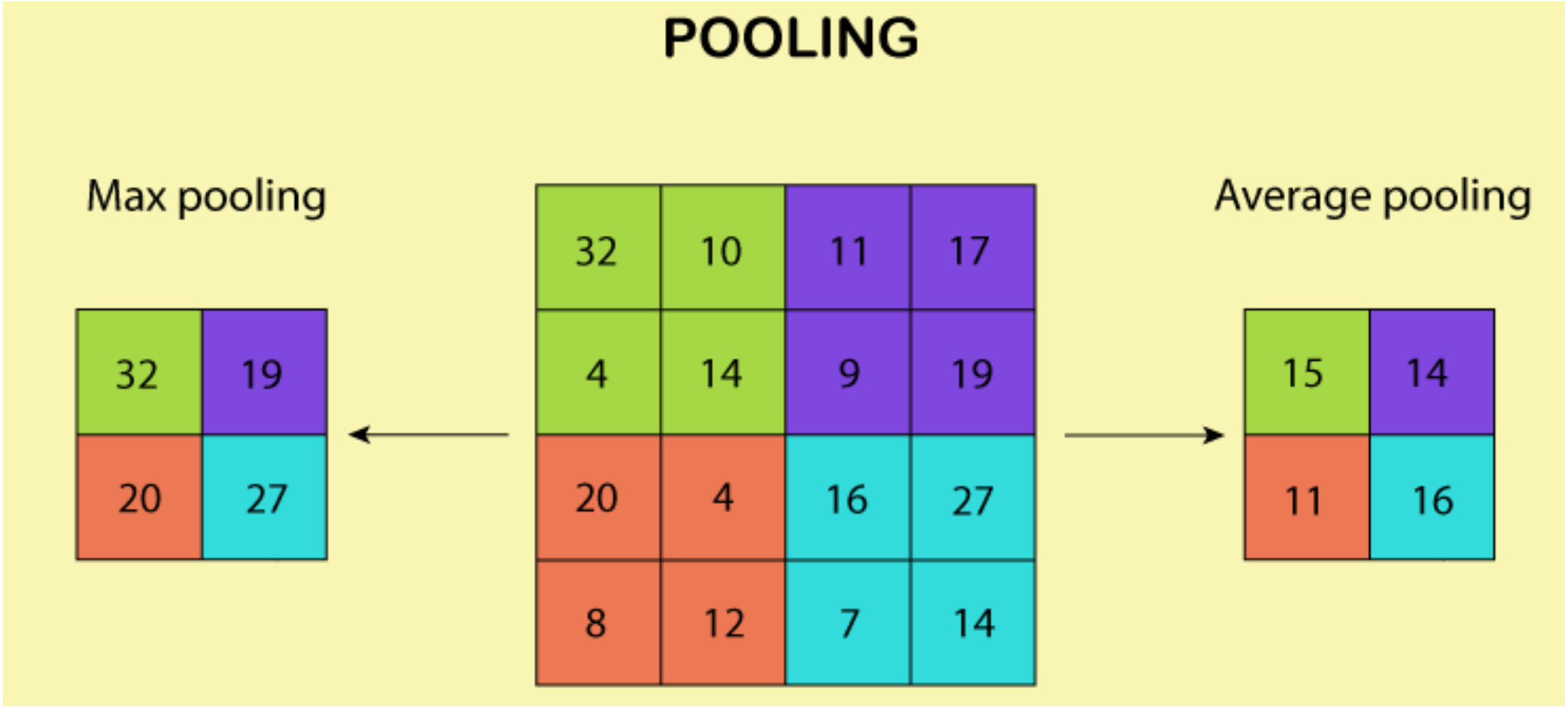
Selects maximum element

from the feature map.

Resulting max-pooled layer holds

important features of feature map.

Pooling Layer

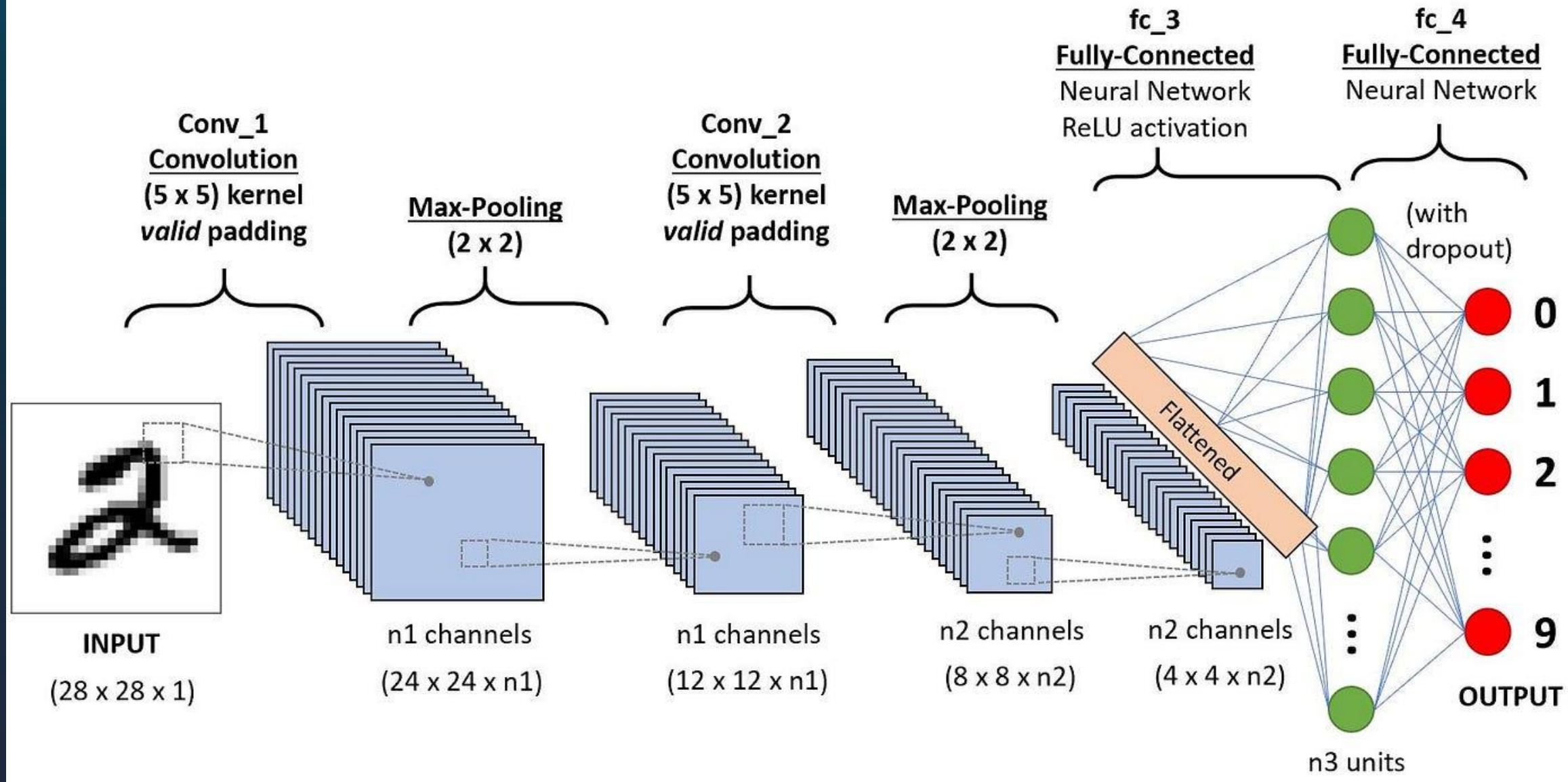


Average pooling

Involves average calculation

for each patch of

the feature map.


Conv_1
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
Conv_2
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
fc_3
Fully-Connected
Neural Network
ReLU activation
fc_4
Fully-Connected
Neural Network
(with dropout)
Flattened
INPUT
(28 x 28 x 1)
n1 channels
(24 x 24 x n1)
n1 channels
(12 x 12 x n1)
n2 channels
(8 x 8 x n2)
n2 channels
(4 x 4 x n2)
n3 units
0
1
2
9
OUTPUT

Fully Connected (Dense) Layer

Connects every neuron

from the previous layer

to the next layer,

Used for final classification or

regression tasks.

Output Layer

Produces the final prediction,

such as class probabilities in classification tasks,

often using activation functions

like Softmax for multi-class classification.

Summary

INPUT OF IMAGE DATA INTO

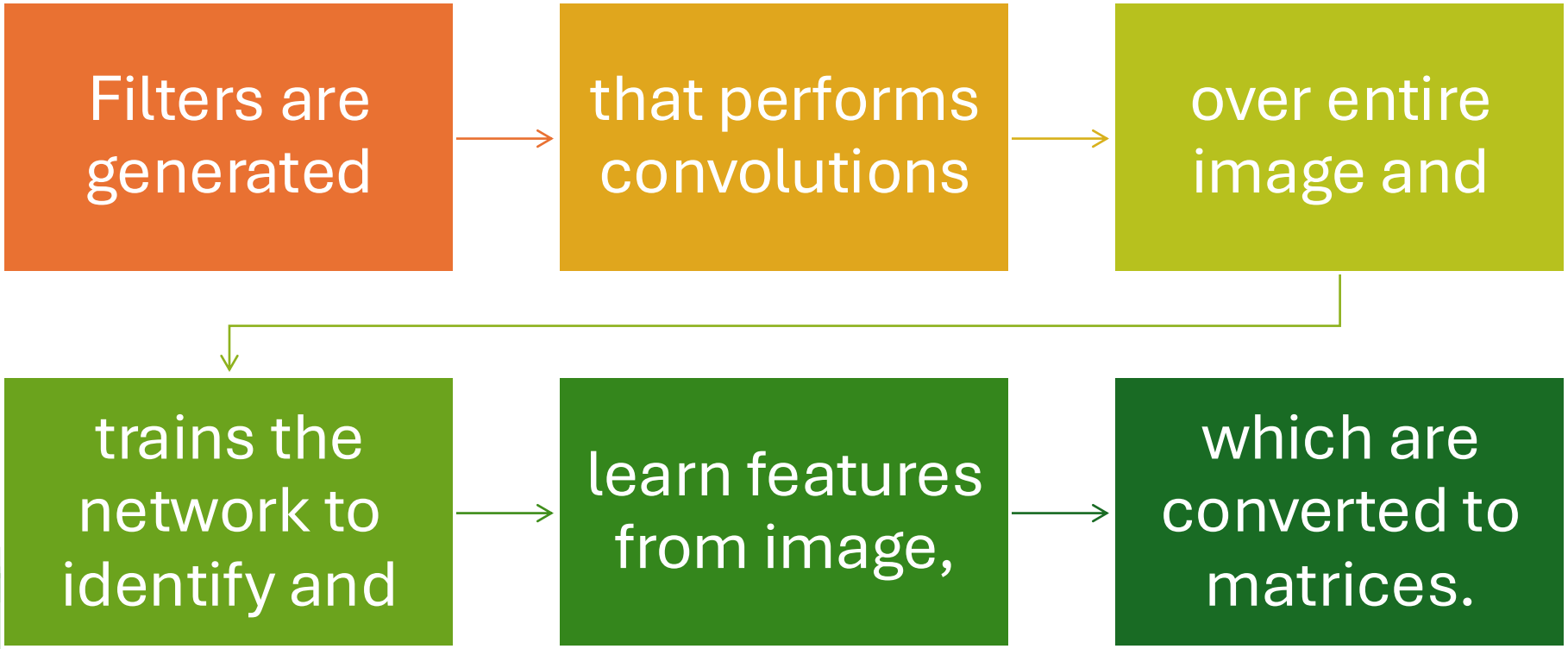
THE CONVOLUTION NEURAL NETWORK,

WHICH IS PROCESSED WITH HELP OF

PIXEL VALUES OF THE IMAGE

IN CONVOLUTION LAYER.

Summary


Filters are generated
that performs convolutions
over entire image and
trains the network to identify and
learn features from image,
which are converted to matrices.

Summary

Batch normalization of input vectors

performed at each layer,

To ensure all input vectors

are normalized.

Summary

The convolutions are performed

until better accuracy has attained

maximum feature extraction is done.

Summary

Convolutions results in

sub-sampling of image

dimensions of input gets changed

according to padding and stride chosen.

Summary

EACH CONVOLUTION FOLLOWS

ACTIVATION LAYER(RELU)

POOLING LAYER

Summary

After the final convolution,

the input matrix is converted

to feature vector.

This feature vector is

the flattened layer.

Summary

Feature vector serves as input

to next layer(fully connected layer),

where all features are collectively

transferred into this network.

Applications of CNNs in Embedded Environments

Object Detection

REAL-TIME DETECTION OF OBJECTS IN

IMAGES OR VIDEO FEEDS

AUTONOMOUS VEHICLES

SECURITY CAMERAS, ROBOTICS

Image Classification

Identifying the category

of an image

Identifying fruits or

animals in images

for smart farming

Facial Recognition

Authenticating users

based on facial features

Door security systems,

smart attendance systems

Anomaly Detection

IDENTIFYING DEFECTS OR

UNUSUAL PATTERNS IN INDUSTRIAL OR

MEDICAL IMAGES IDENTIFYING FAULTS IN

MACHINERY OR DETECTING TUMORS

IN MEDICAL IMAGING

Gesture Recognition

Recognizing hand

gestures or movements

smart home control systems,

gaming devices

Surendra Panpaliya

Founder and CEO

GKTCS Innovations

https://www.gktcs.com