FPT Education

FPT UNIVERSITY

BACHELOR OF ARTIFICIAL INTELLIGENCE THESIS

FRUIT TYPE AND WEIGHT RECOGNITION FOR PAYMENT PURPOSES USING COMPUTER VISION

TEAM MEMBERS:

NGUYỄN NGỌC BÁCH

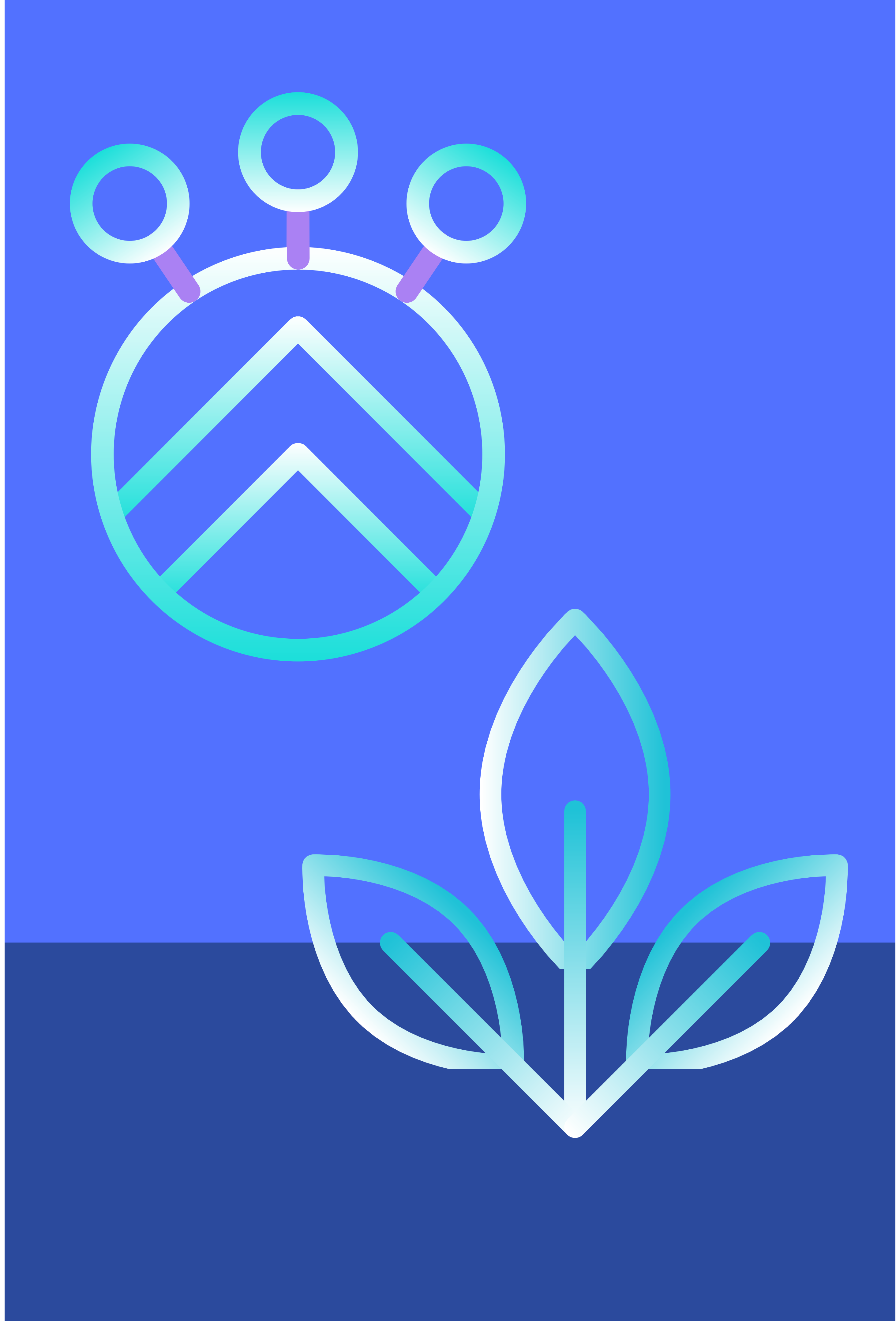
LÊ MINH CHÍ

BÙI VĂN PHÚC

SUPERVISOR:

MR. BÙI VĂN HIỆU

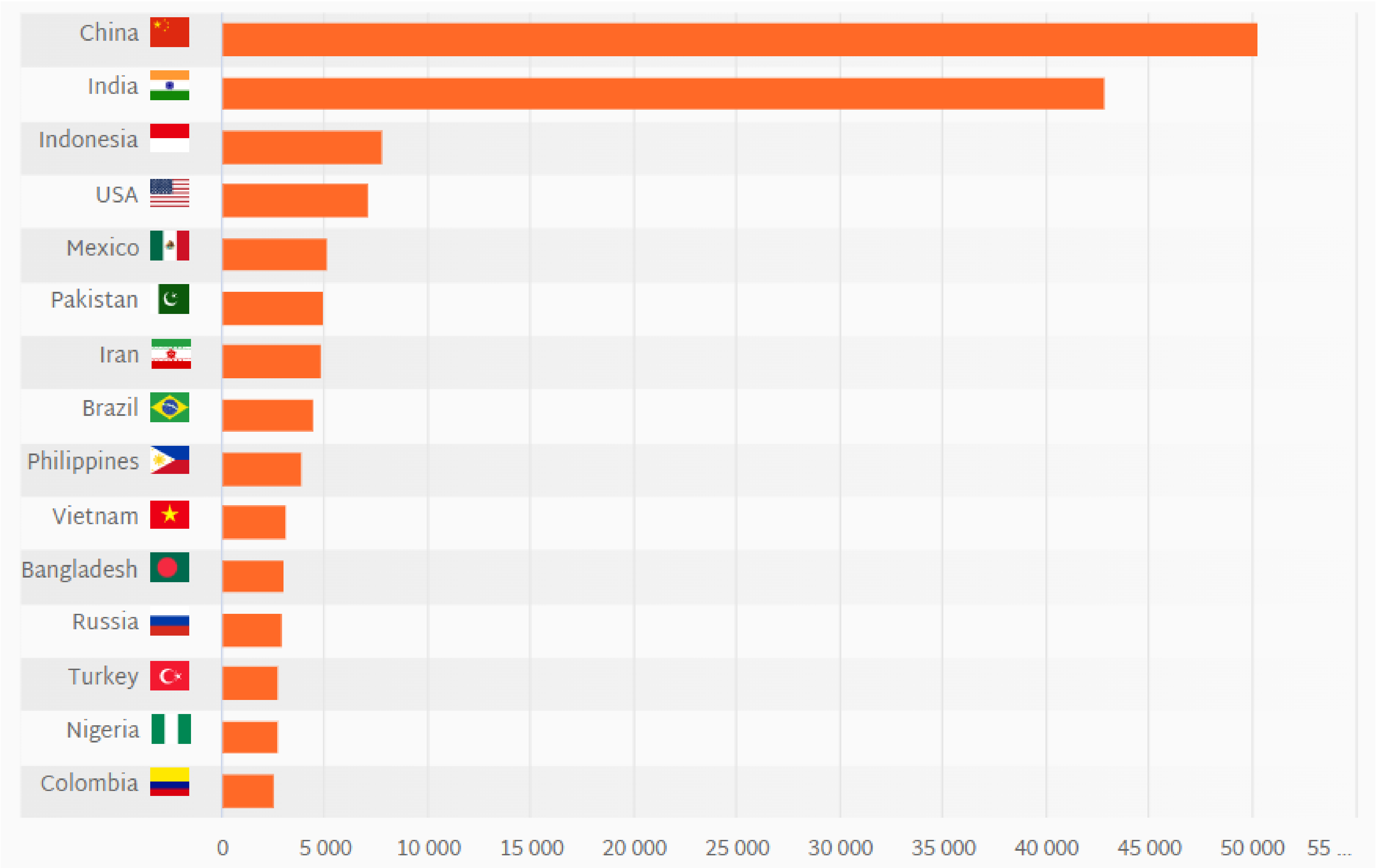
1. INTRODUCTION
2. BUILDING MODEL
3. IMPLEMENTATION
4. DEMO
5. CONCLUSION



1

INTRODUCTION

WHY CHOOSE THIS WORK?



Current fruit checkout methods

Traditional methods: Fruits are recognized and scaled manually at the time of selling.
 - Price is calculated by mental arithmetic or using a calculator.

Disadvantage:

- Take a variety of human efforts to calculate the price
- Hight risk of making mistakes: miscalculating or recognizing the wrong type of fruit or forgetting unit prices

Current fruit checkout methods

Electronic scale with the printer: Fruits are recognized manually to input code into the scale.
 - Price is calculated and shown by scale after inputting code.

Disadvantage:

- Cost the owner time and money
- Chance of workers forgetting the fruit code or misremembering the fruit type
- Take time if having to look up paper

Current fruit checkout methods

Using barcode: The fruit is scaled and labeled before selling.
 - Price is shown right after scanning the barcode.

Disadvantage:

Packed Fruits	Single fruit
• Make customers less choice in buying • Release waste to the environment	• Take a numerous amount of human effort to weigh and add stickers to every fruit • Can be harmful to human

Challenges

Fruit recognition:

- The fruit has multiple features: shape, size, color, ...
- Some fruits are similar to each other
- Effect of environmental factor

Scale:

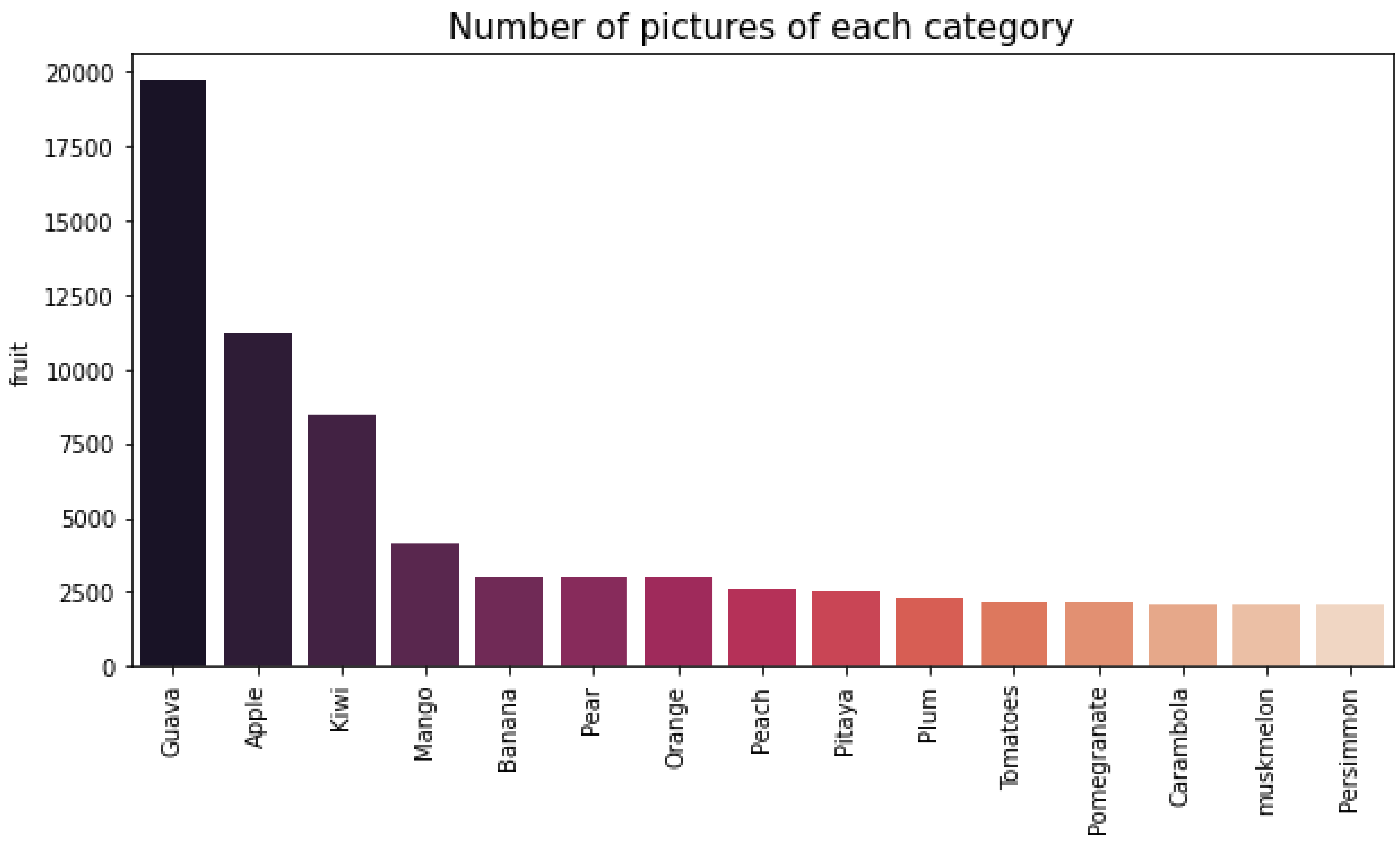
- Building a scale that has the ability to connect and give parameters to another device in real-time

2

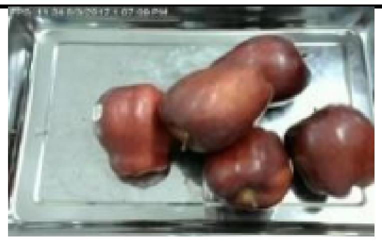
BUILDING THE MODEL

DATA PREPARATION

- CONSISTS OF 44406 IMAGES
- 15 KINDS OF FRUITS
- VARIOUS CONDITIONS LIKE POSES, ANGLES, LIGHTNING OR SHADOWS


Number of pictures of each category
fruit
20000
17500
15000
12500
10000
7500
5000
2500
0
Guava
Apple
Kiwi
Mango
Banana
Pear
Orange
Peach
Pitaya
Plum
Tomatoes
Pomegranate
Carambola
muskmelon
Persimmon

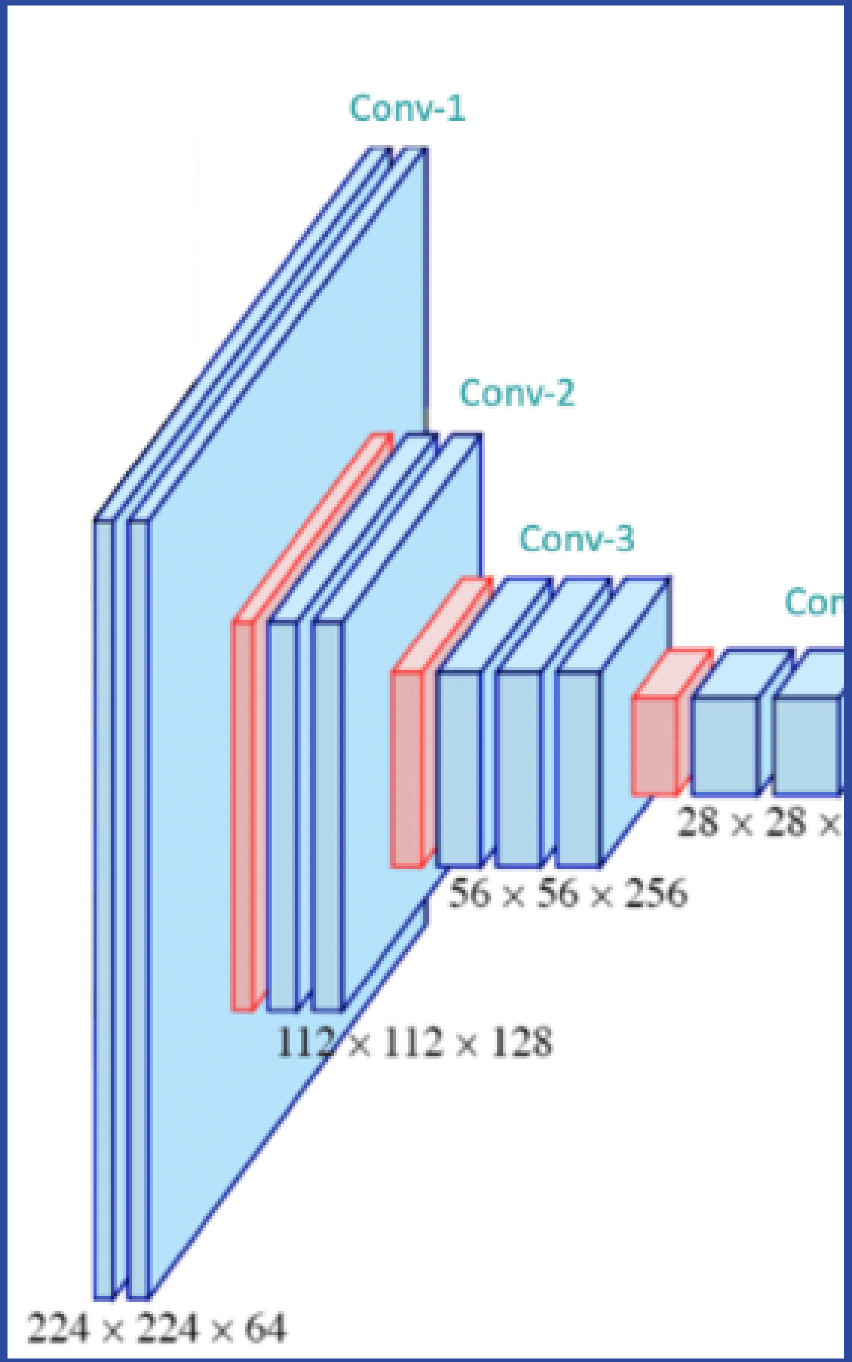
Apple Fruits

Fruit Name	Images				
Apple A	957				
Apple B	740				
Apple C	870				
Apple D	1033				
Apple E	664				
Apple F	1338				

WHY USE A CNN?

EFFICIENCY

- EXTRACTED THE MOST CORE FEATURES
- HIGH ACCURACY RATE
- ROBUST TO NOISE



EASE OF USE

- AUTOMATED, REQUIRED LITTLE HUMAN SUPERVISE
- SUITABLE FOR TRANSFER LEARNING

A BASIC CNN

What We See

What Computers See

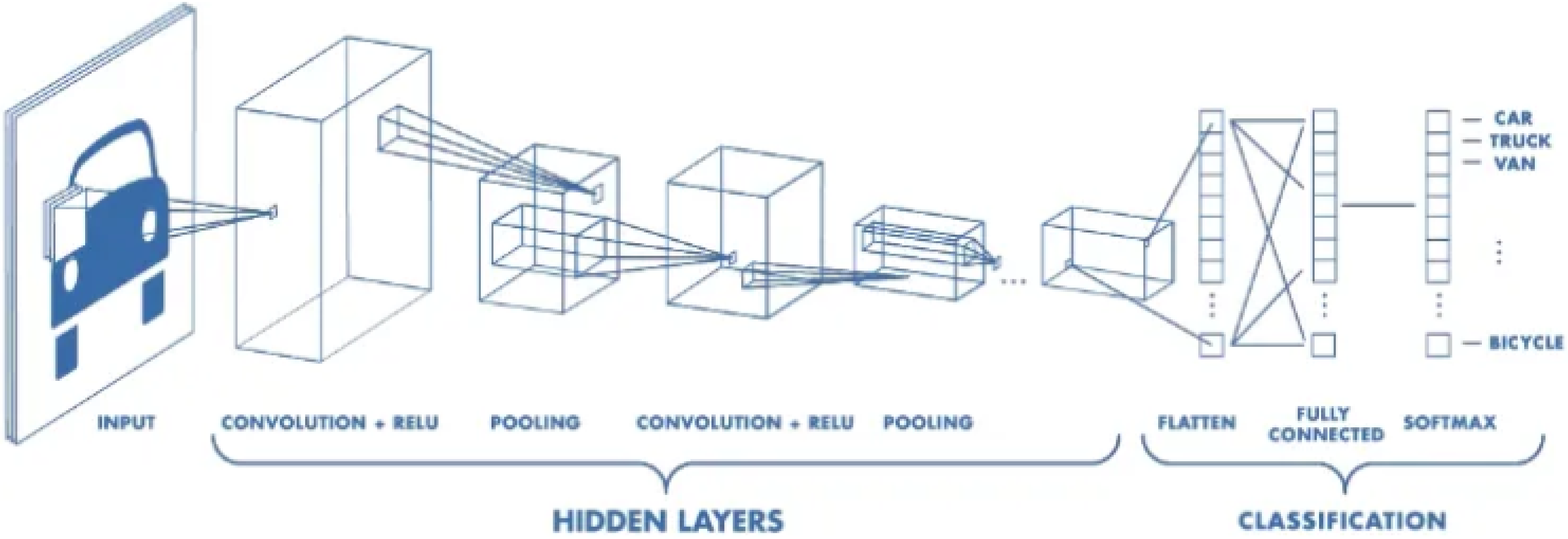
A BASIC CNN

What We See

What Computers See

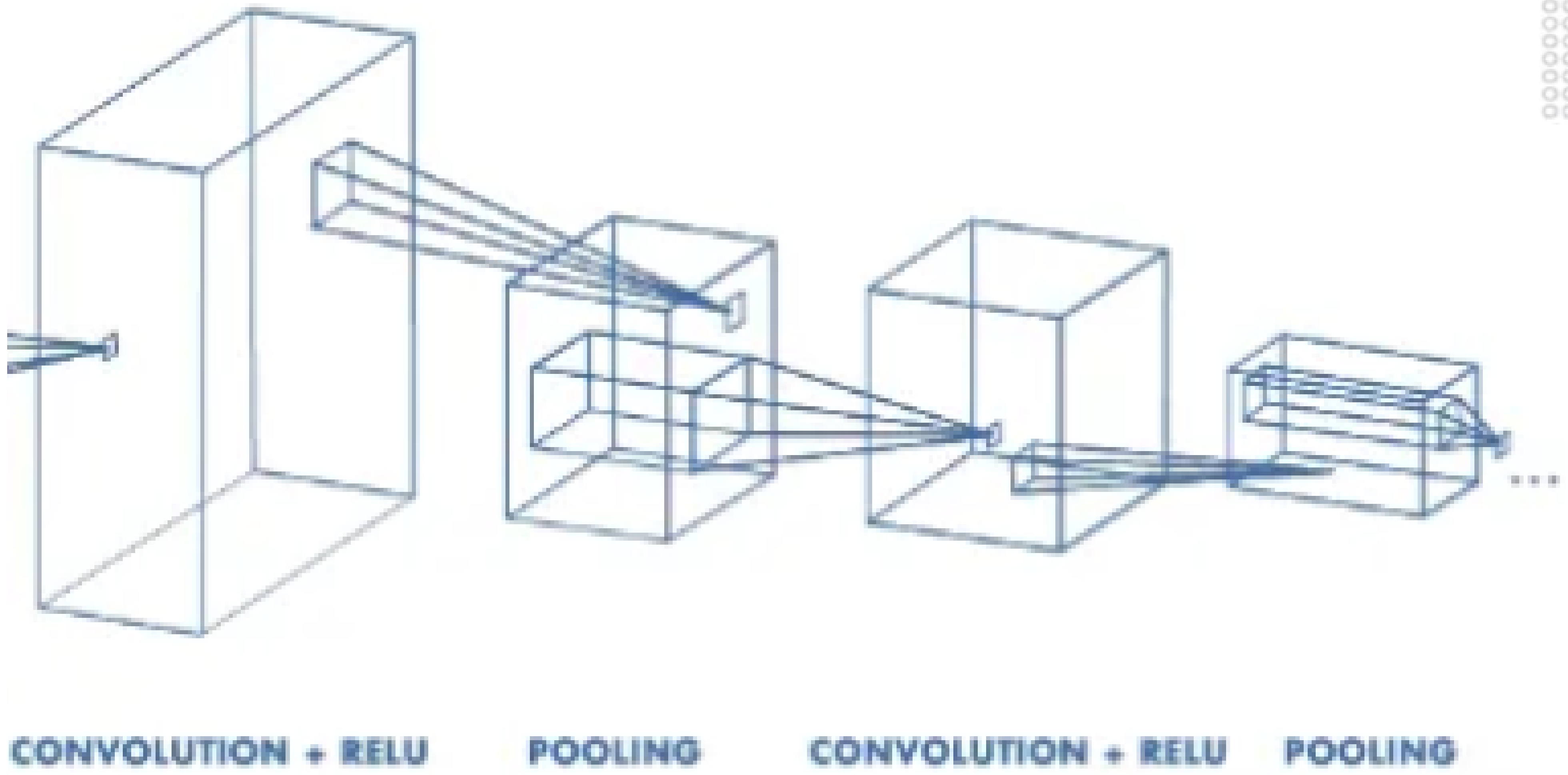
- FOCUS ON RELEVENT HIGH-LEVEL FEATURES.
- USE FILTERS TO SELECT LOW-LEVELED FEATURE LIKE EDGES AND CURVES, FROM THAT DEVELOP TO A MORE SOPHISICATED DETAILS.

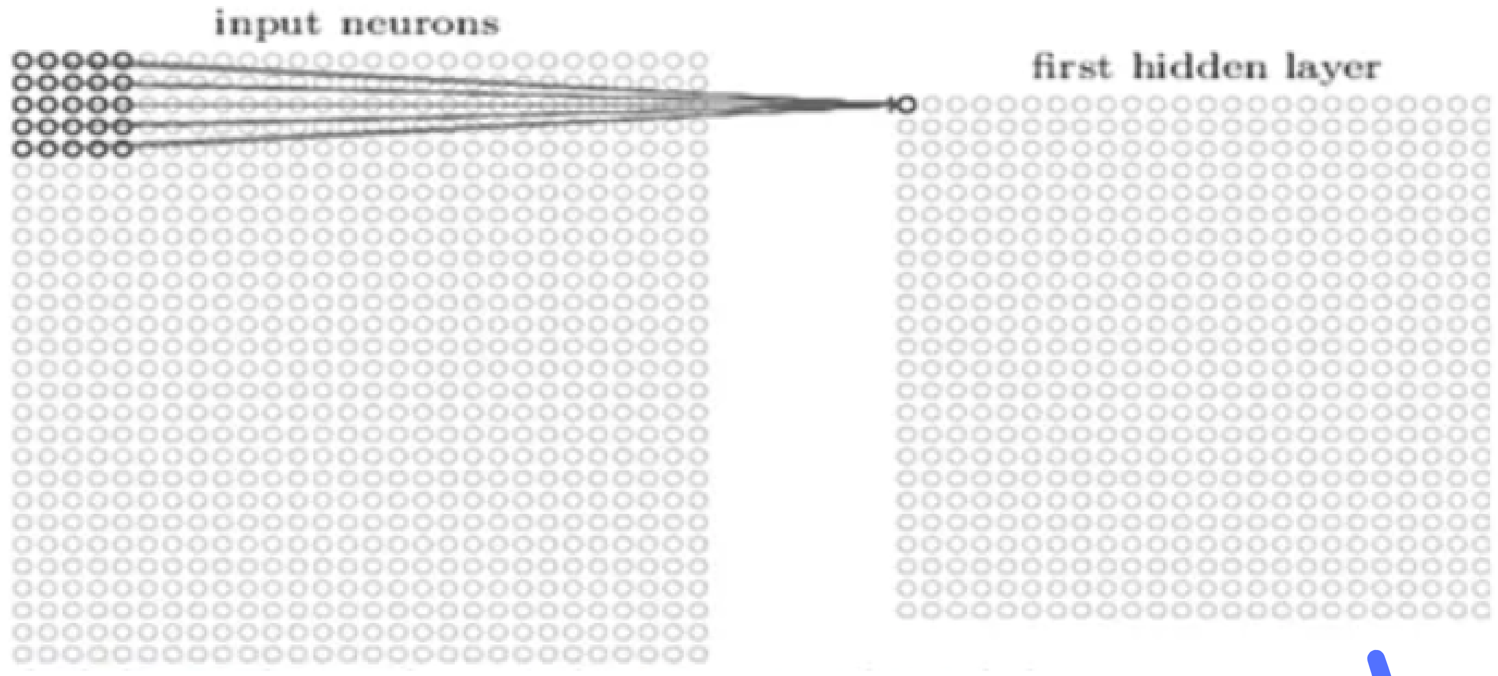
A BASIC CNN

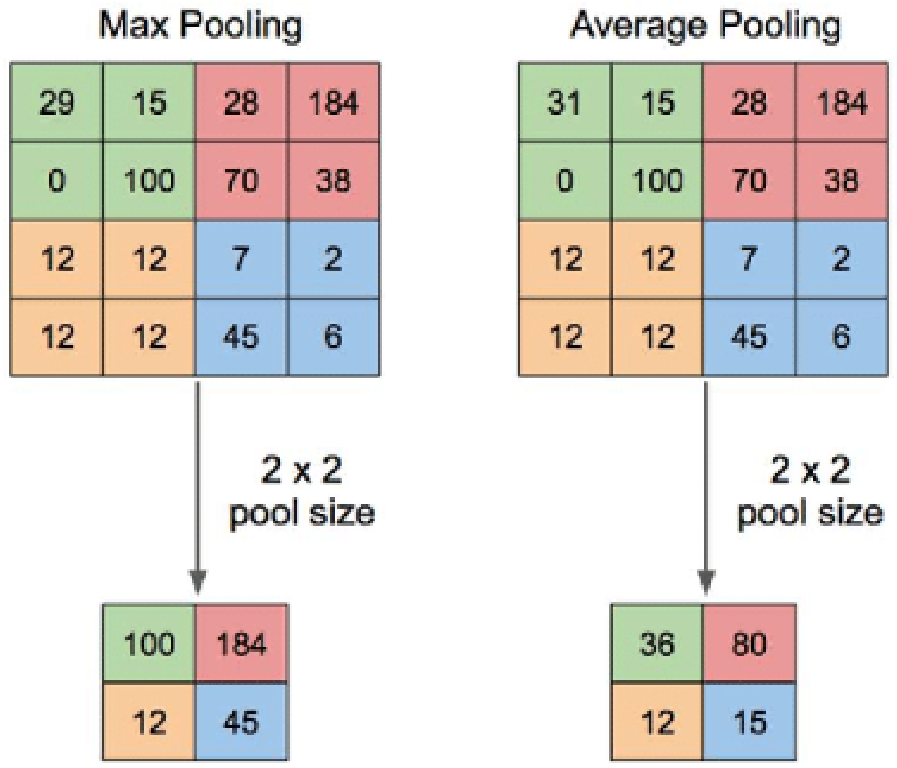


CAR
TRUCK
VAN
BICYCLE
INPUT
CONVOLUTION + RELU
POOLING
CONVOLUTION + RELU
POOLING
FLATTEN
FULLY CONNECTED
SOFTMAX
HIDDEN LAYERS
CLASSIFICATION

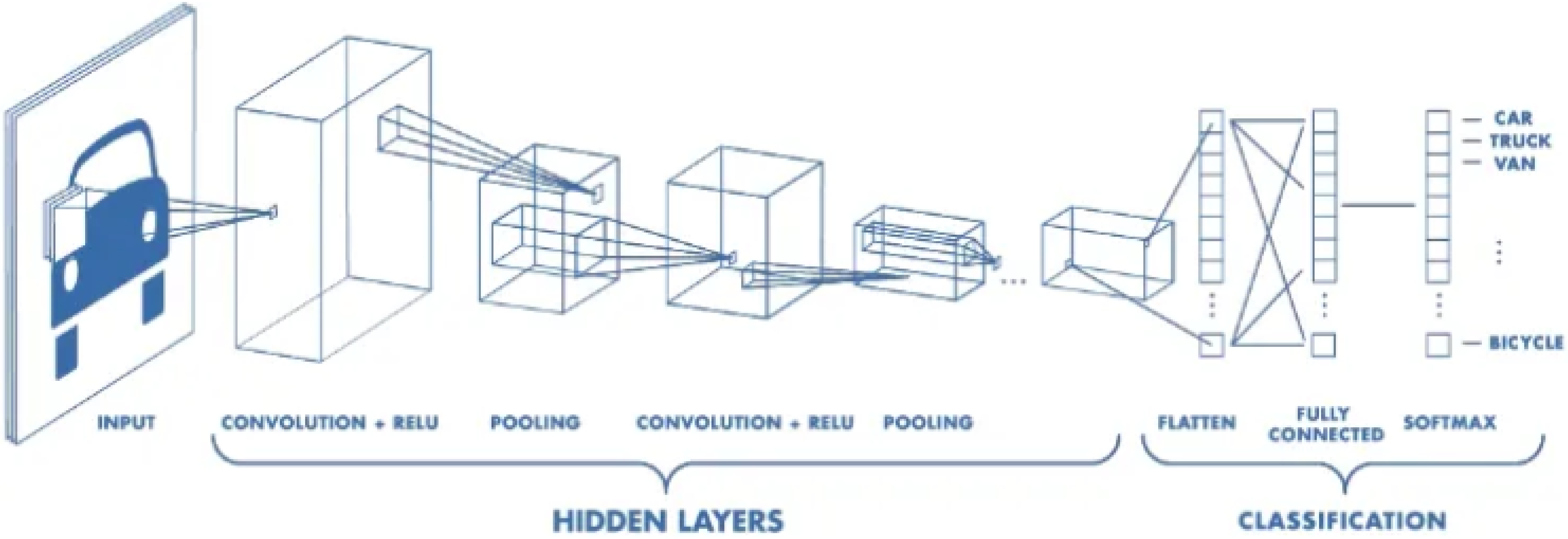
A BASIC CNN


CONVOLUTION + RELU
POOLING
CONVOLUTION + RELU
POOLING


input neurons
first hidden layer


Max Pooling
29 15 28 184
0 100 70 38
12 12 7 2
12 12 45 6
2 x 2
pool size
100 184
12 45
Average Pooling
31 15 28 184
0 100 70 38
12 12 7 2
12 12 45 6
2 x 2
pool size
36 80
12 15

A BASIC CNN



CAR
TRUCK
VAN
BICYCLE
INPUT
CONVOLUTION + RELU
POOLING
CONVOLUTION + RELU
POOLING
FLATTEN
FULLY CONNECTED
SOFTMAX
HIDDEN LAYERS
CLASSIFICATION

A BASIC CNN

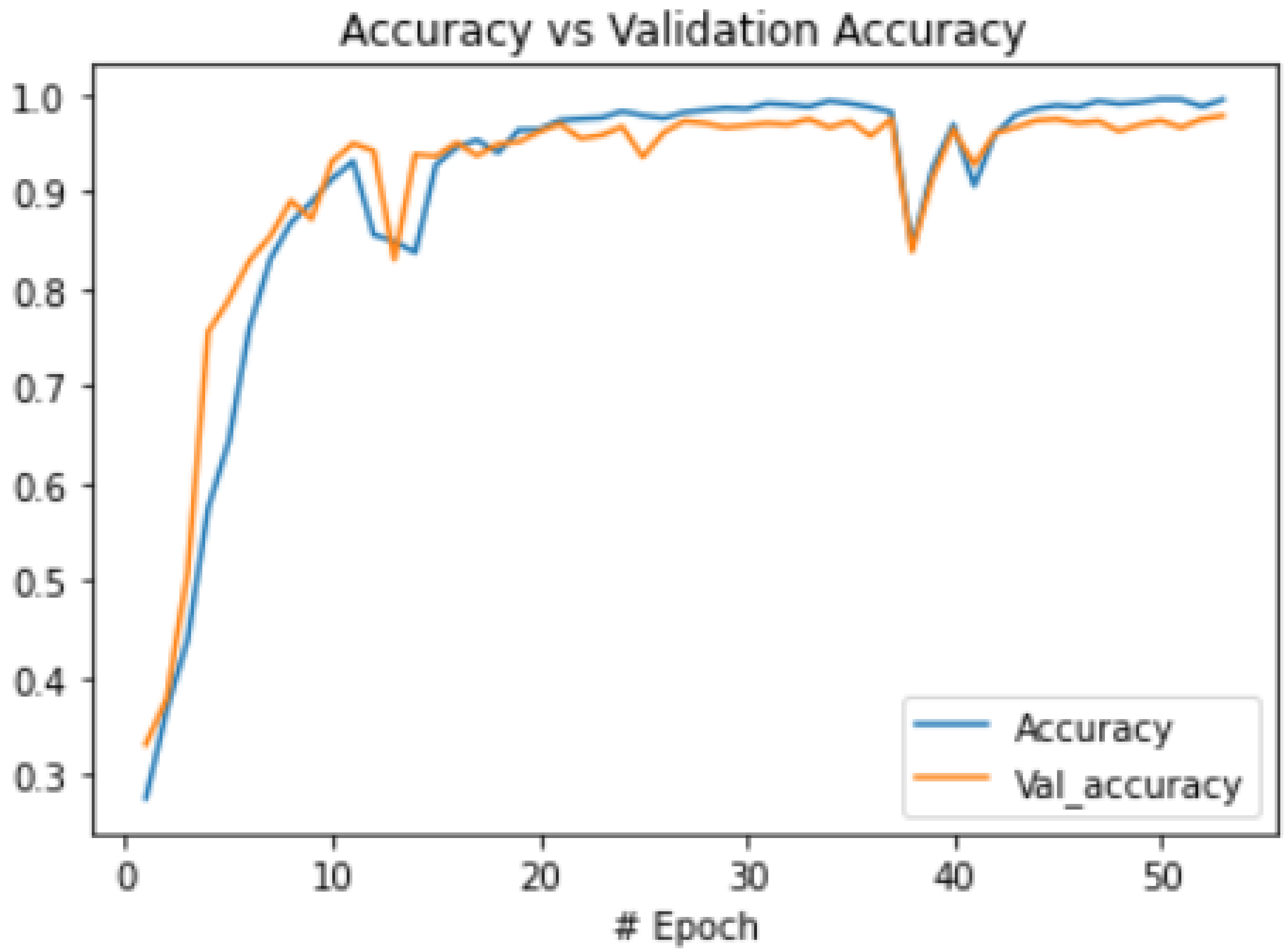
Layer	Details
Input	(150,150,3)
Convolution + Relu	number of filers: 32; kernel: 3x3
Max Pooling	2x2
Convolution + Relu	number of filers: 64; kernel: 3x3
Max Pooling	2x2
Convolution + Relu	number of filers: 64; kernel: 3x3
Max Pooling	2x2
Convolution + Relu	number of filers: 64; kernel: 3x3
Max Pooling	2x2
Convolution + Relu	number of filers: 64; kernel: 3x3
Max Pooling	2x2
Convolution + Relu	number of filers: 64; kernel: 3x3
Max Pooling	2x2
Flatten	
Dense layer	size: 256 + reLu
Dense layer	size: 15
Softmax	
Output	

A BASIC CNN

- 6 CONVOLUTION LAYERS DEEP
- USING 1/5 OF DATASETS TO TRAIN
- ABOUT 100 EPOCHS

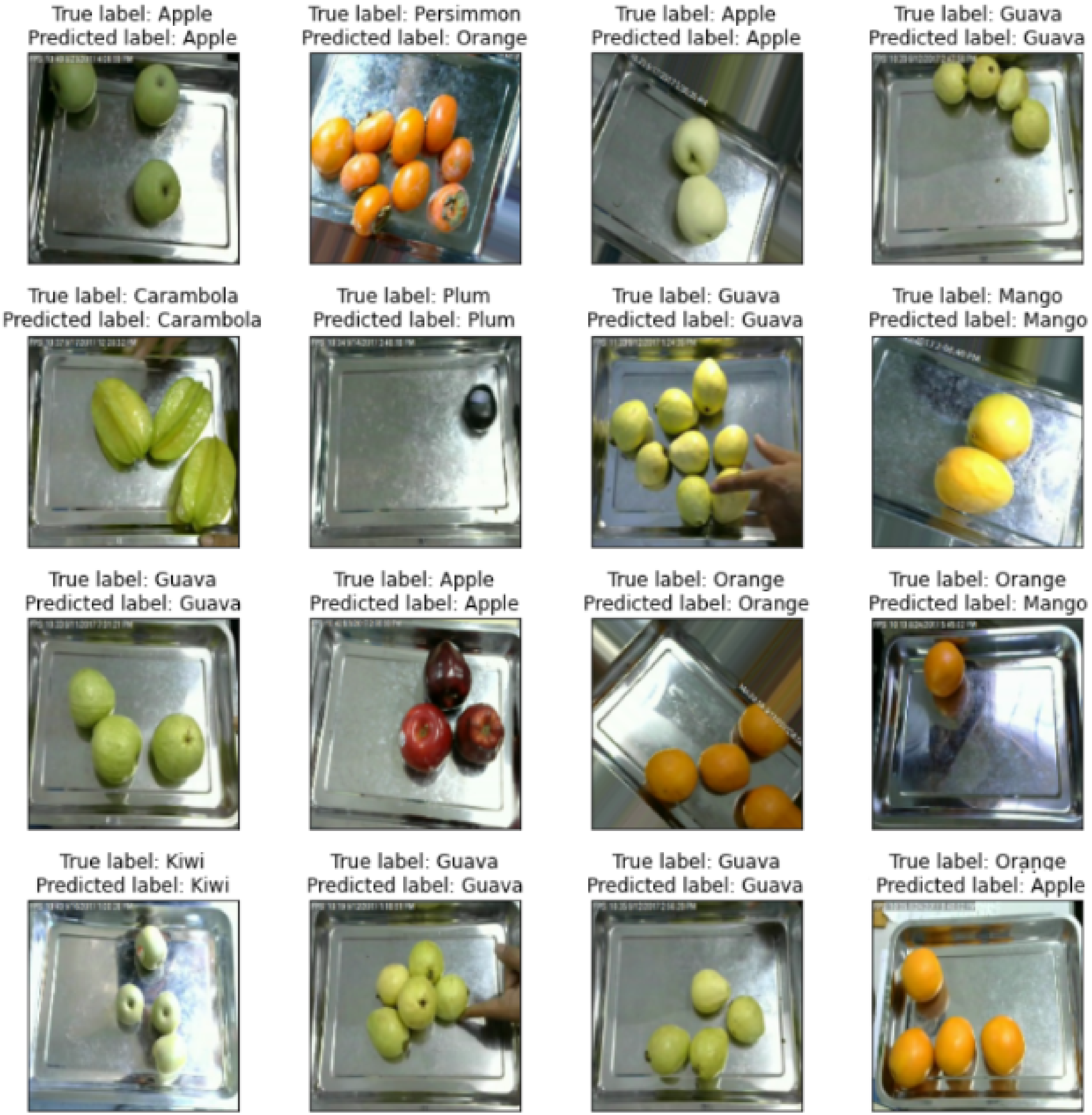
Layer	Details
Input	(150,150,3)
Convolution + Relu	number of filers: 32; kernel: 3x3
Max Pooling	2x2
Convolution + Relu	number of filers: 64; kernel: 3x3
Max Pooling	2x2
Convolution + Relu	number of filers: 64; kernel: 3x3
Max Pooling	2x2
Convolution + Relu	number of filers: 64; kernel: 3x3
Max Pooling	2x2
Convolution + Relu	number of filers: 64; kernel: 3x3
Max Pooling	2x2
Convolution + Relu	number of filers: 64; kernel: 3x3
Max Pooling	2x2
Flatten	
Dense layer	size: 256 + reLu
Dense layer	size: 15
Softmax	
Output	

RESULT OF THE BASIC CNN



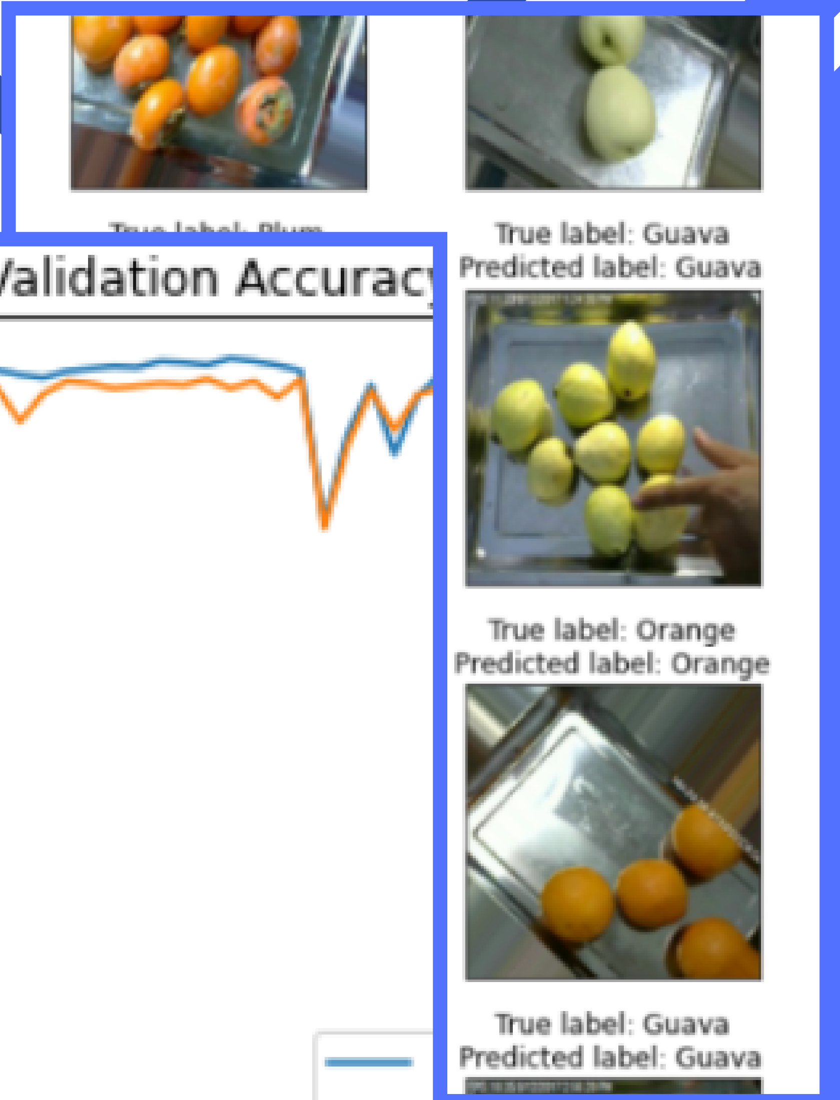
	Apple	Banana	Carambola	Guava	Kiwi	Mango	Orange	Peach	Pear	Persimmon	Pitaya	Plum	Pomegranate	Tomatoes	muskmelon
Apple	523	0	0	1	0	0	0	12	1	0	1	0	0	5	2
Banana	1	120	0	5	1	0	0	0	1	0	0	0	1	0	0
Carambola	0	0	109	0	0	0	0	0	0	0	0	0	0	0	0
Guava	0	0	0	1026	1	0	0	0	0	0	0	0	0	0	0
Kiwi	0	0	0	0	394	0	0	0	1	0	0	0	0	0	0
Mango	0	1	0	0	0	208	0	0	0	0	0	0	0	0	1
Orange	0	0	0	0	0	0	150	0	0	1	0	0	0	0	0
Peach	1	0	0	0	0	0	0	126	0	0	0	0	0	0	0
Pear	1	0	0	0	0	0	0	1	129	0	0	0	0	0	0
Persimmon	0	0	0	0	0	0	2	0	0	102	0	0	0	0	0
Pitaya	0	0	0	0	0	0	0	1	0	0	130	0	0	0	0
Plum	0	0	0	0	0	0	0	0	0	0	0	107	0	0	0
Pomegranate	0	1	0	0	0	1	0	0	2	0	0	0	125	0	0
Tomatoes	12	0	0	0	0	0	0	0	0	0	0	0	0	117	0
muskmelon	13	0	0	0	0	0	0	0	0	0	0	0	0	0	90

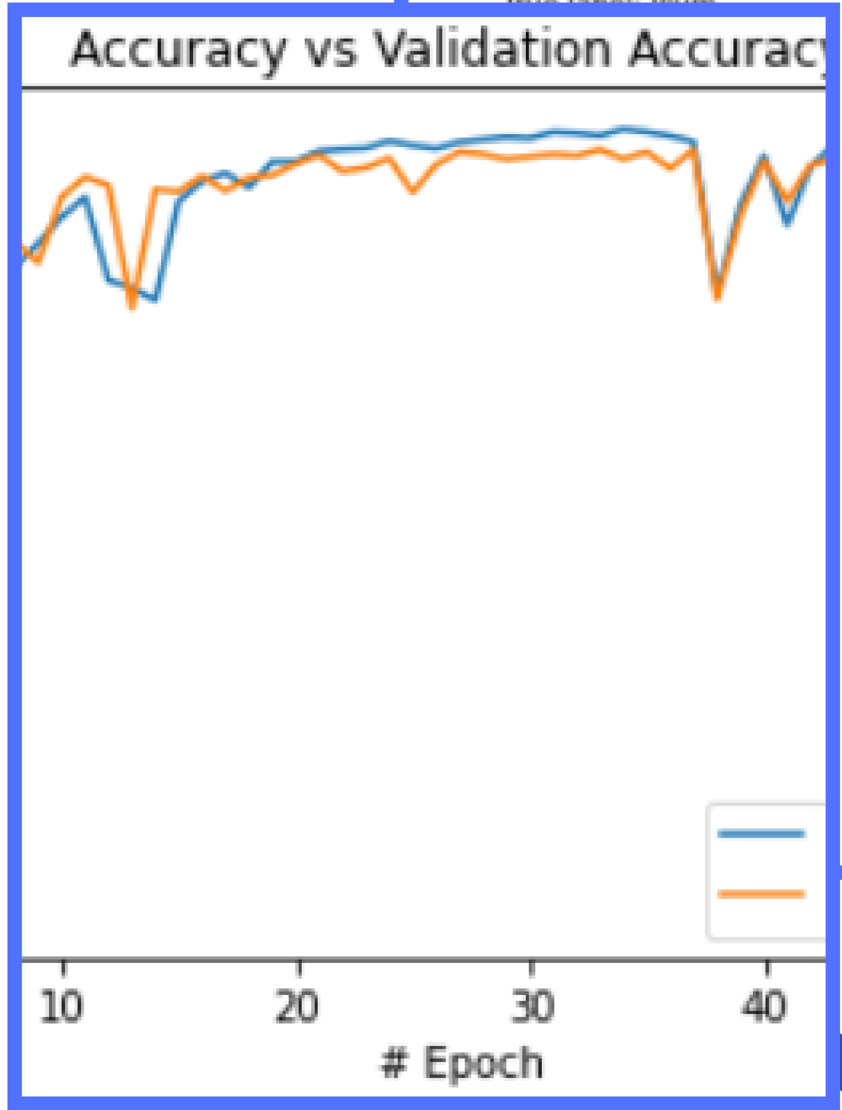
	Apple	Banana	Carambola	Guava	Kiwi	Mango	Orange	Peach	Pear	Persimmon	Pitaya	Plum	Pomegranate	Tomatoes	muskmelon
Apple	523	0	0	1	0	0	0	12	1	0	1	0	0	5	2
Banana	1	120	0	5	1	0	0	0	1	0	0	0	1	0	0
Carambola	0	0	109	0	0	0	0	0	0	0	0	0	0	0	0
Guava	0	0	0	1026	1	0	0	0	0	0	0	0	0	0	0
Kiwi	0	0	0	0	394	0	0	0	1	0	0	0	0	0	0
Mango	0	1	0	0	0	208	0	0	0	0	0	0	0	0	1
Orange	0	0	0	0	0	0	150	0	0	1	0	0	0	0	0
Peach	1	0	0	0	0	0	0	126	0	0	0	0	0	0	0
Pear	1	0	0	0	0	0	0	1	129	0	0	0	0	0	0
Persimmon	0	0	0	0	0	0	2	0	0	102	0	0	0	0	0
Pitaya	0	0	0	0	0	0	0	1	0	0	130	0	0	0	0
Plum	0	0	0	0	0	0	0	0	0	0	0	107	0	0	0
Pomegranate	0	1	0	0	0	1	0	0	2	0	0	0	125	0	0
Tomatoes	12	0	0	0	0	0	0	0	0	0	0	0	0	117	0
muskmelon	13	0	0	0	0	0	0	0	0	0	0	0	0	0	90


True label: Apple
Predicted label: Apple
True label: Persimmon
Predicted label: Orange
True label: Apple
Predicted label: Apple
True label: Guava
Predicted label: Guava
True label: Carambola
Predicted label: Carambola
True label: Plum
Predicted label: Plum
True label: Guava
Predicted label: Guava
True label: Mango
Predicted label: Mango
True label: Guava
Predicted label: Guava
True label: Apple
Predicted label: Apple
True label: Orange
Predicted label: Orange
True label: Orange
Predicted label: Mango
True label: Kiwi
Predicted label: Kiwi
True label: Guava
Predicted label: Guava
True label: Guava
Predicted label: Guava
True label: Orange
Predicted label: Apple

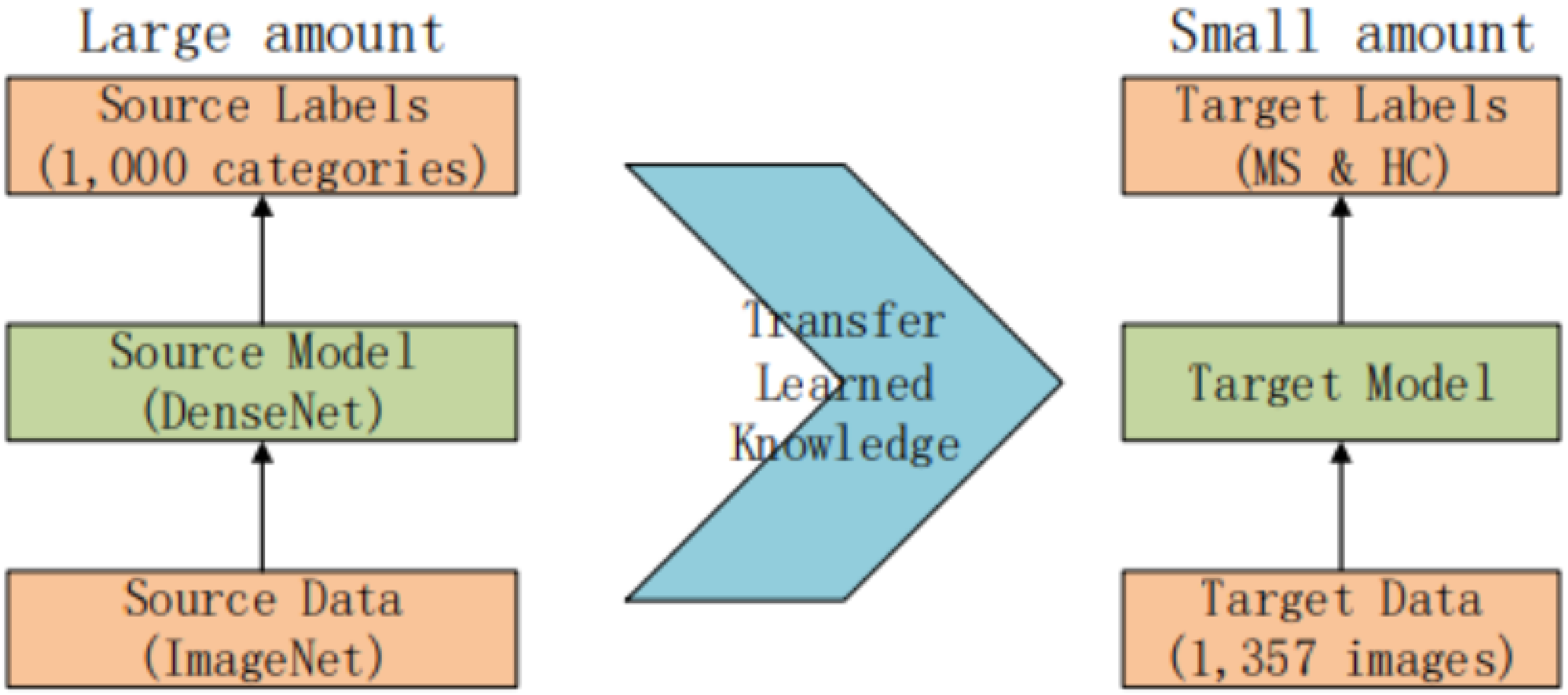
DISCUSION

- IMPRESSIVE RESULT: 97,6 %
- PERFORMS WELL ENOUGH ON MOST PARTS EVEN WITH ONLY 20% OF THE DATASET
- SEEMS TO BE STRUGGLE TO RECOGNIZE FRUITS WITH SIMILAR SHAPE AND COLOR LIKE ORANGE, APPLE, PERSIMMON OR TOMATO.





TRANSFER LEARNING



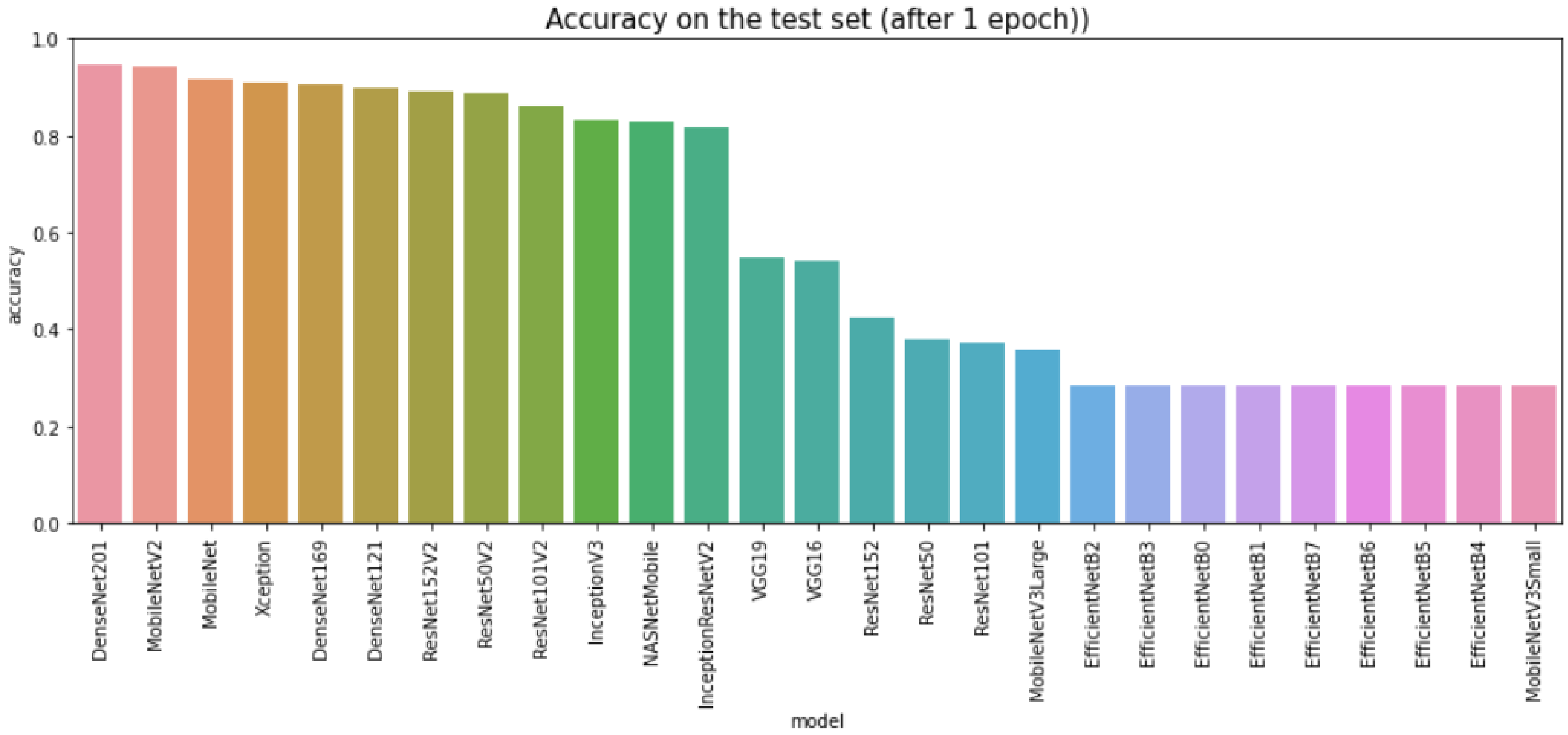
PRETRAINED MODEL

- 27 MODELS AVAILABLE
- USING 1/10 OF THE DATASET
- TRAIN AND TEST THE ACCURACY AFTER 1 EPOCHS

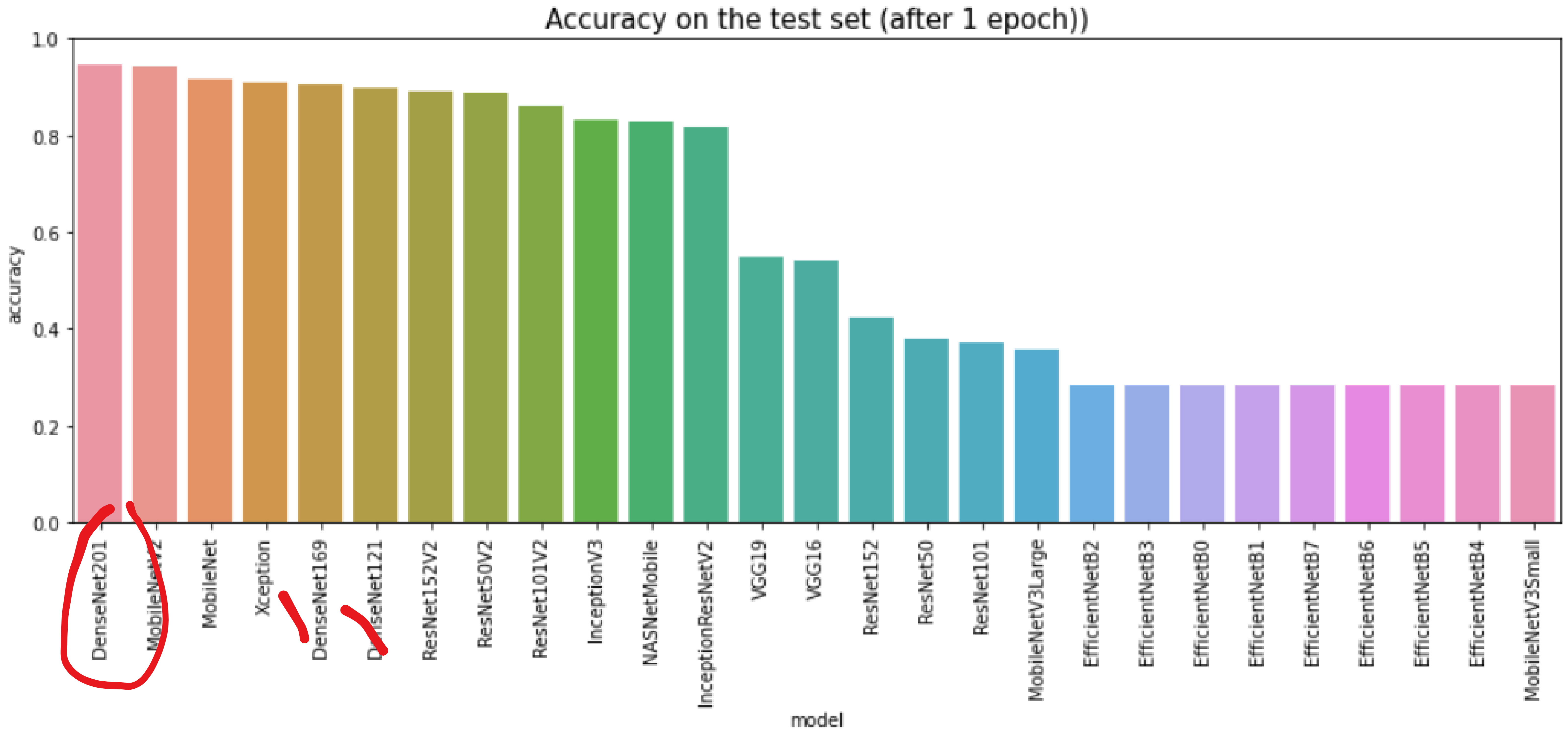
TensorFlow

Keras

PRETRAINED MODEL

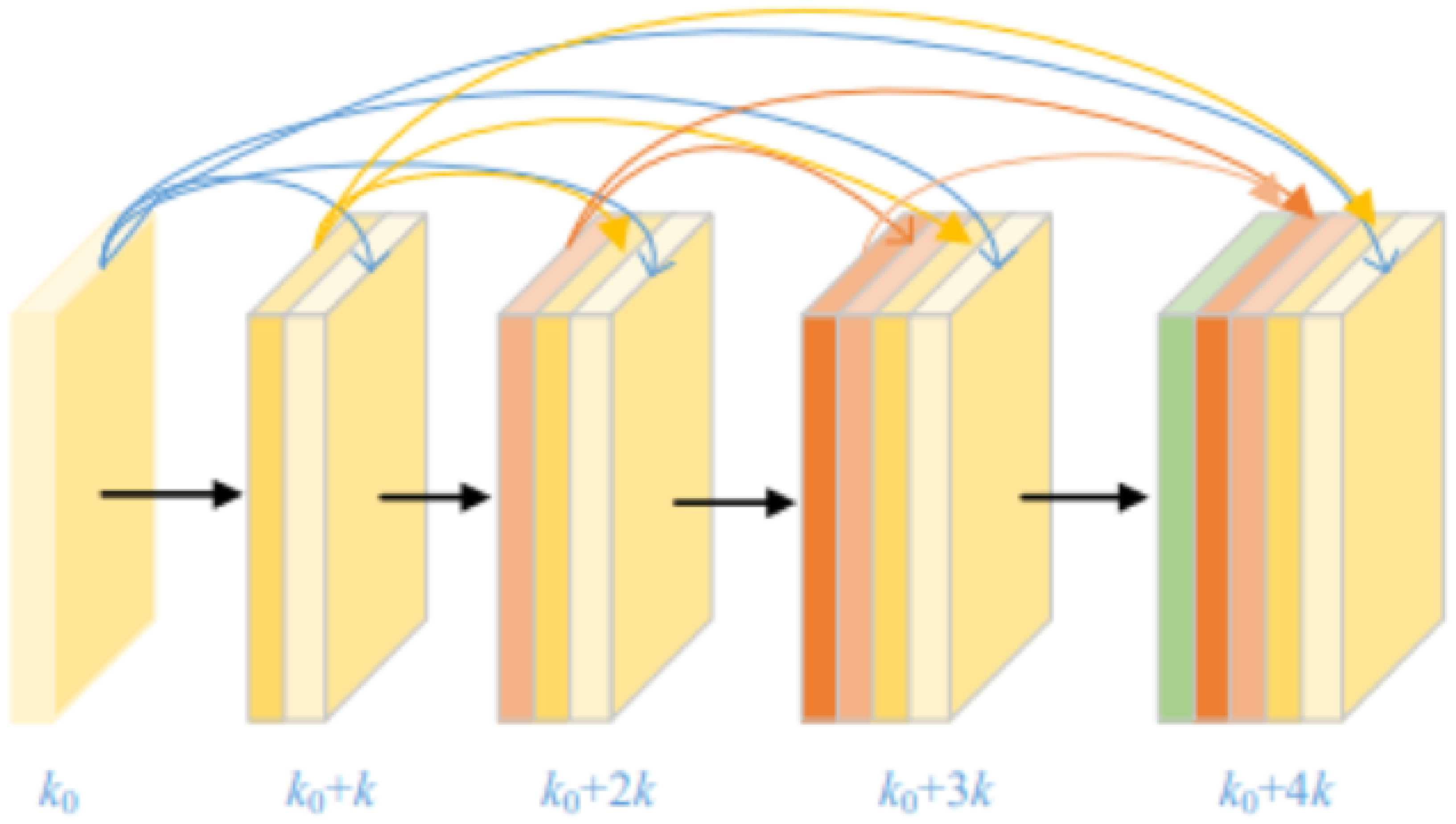


PRETRAINED MODEL

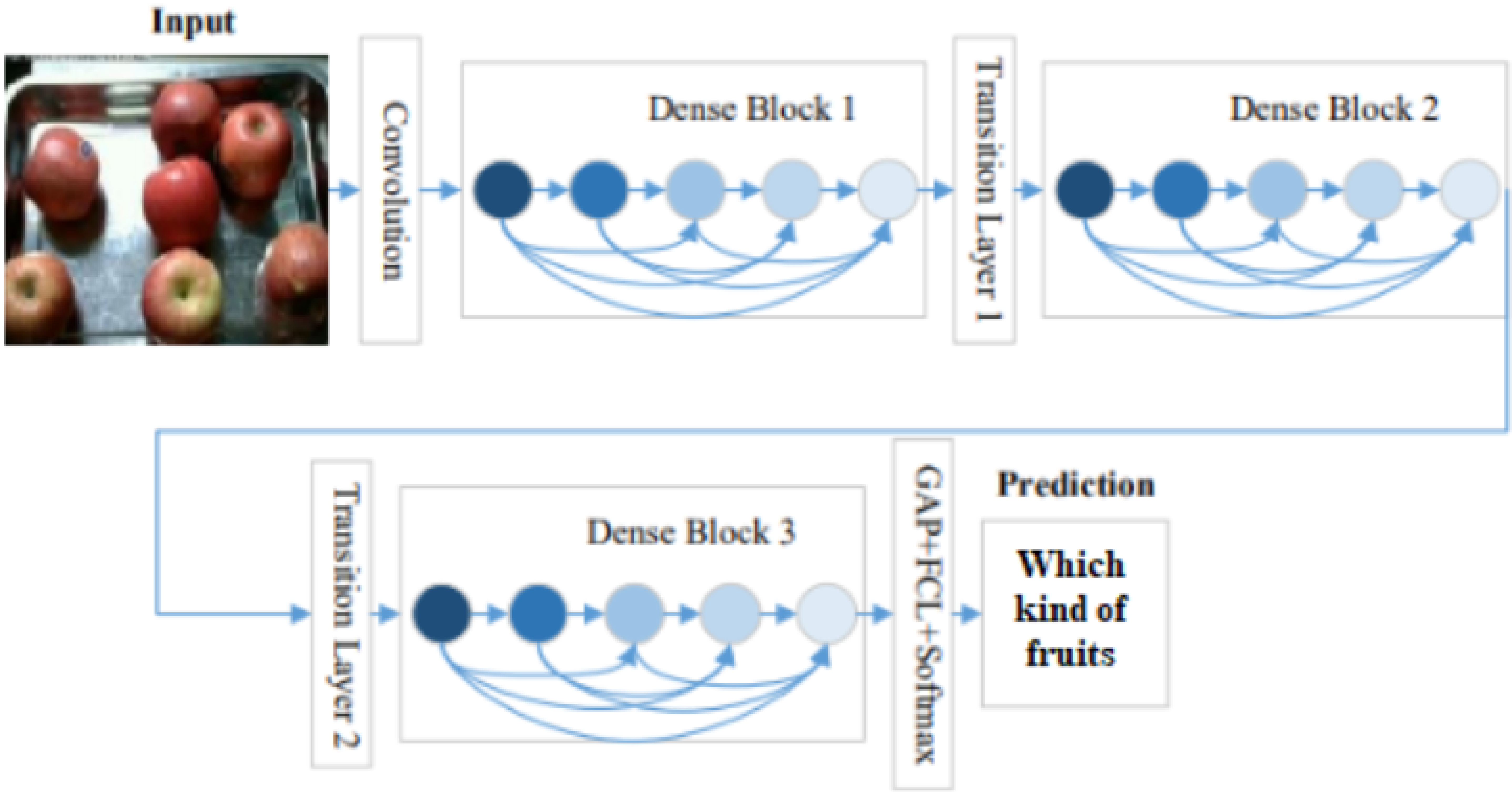


DENSENET

- THE DEFINED FEATURE: THE DENSE BLOCKS.
- CONSISTS OF ALL FEATURE MAPS OF ALL PREVIOUS CONVOLUTIONAL PROCESS



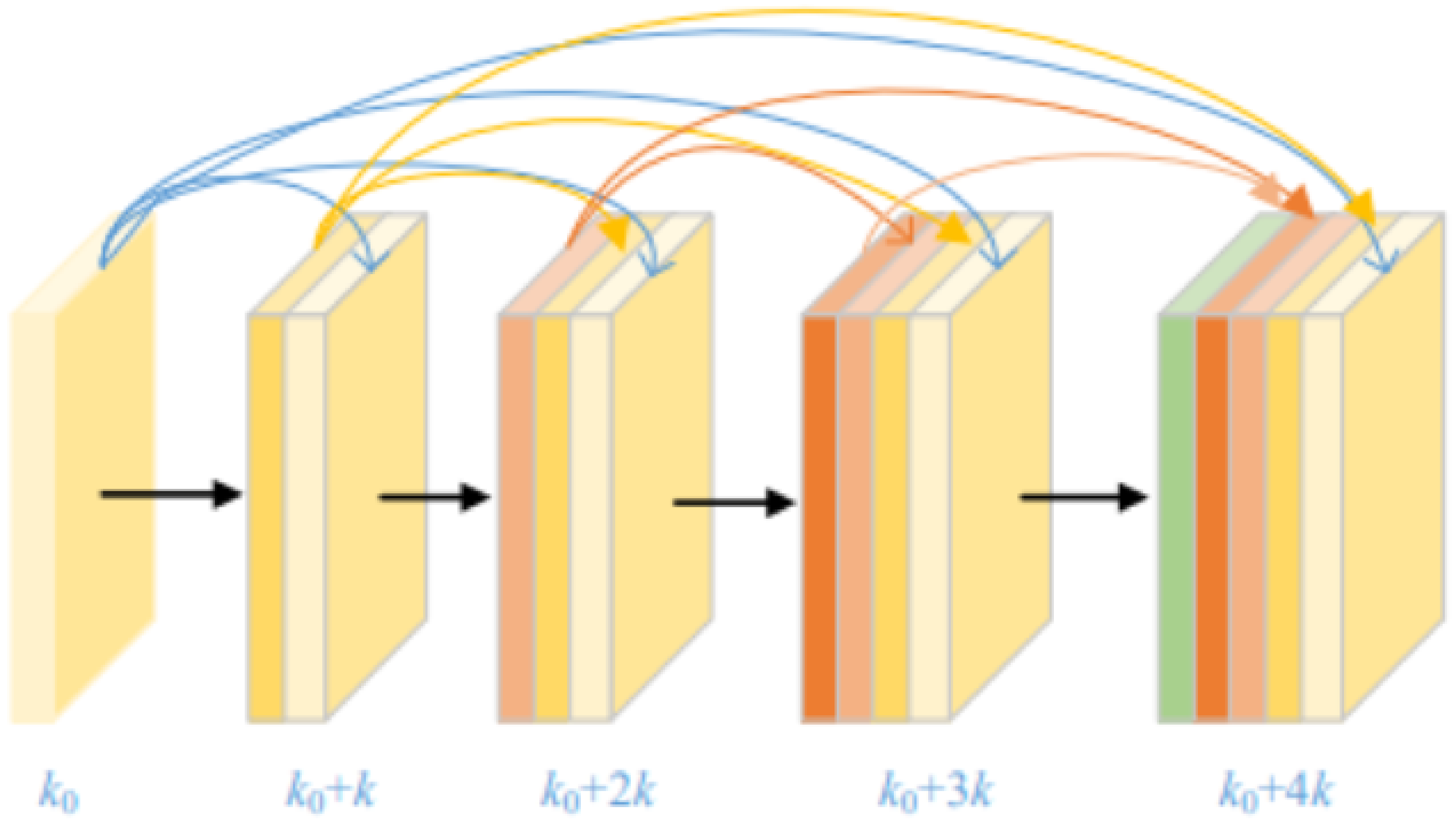
DENSENET

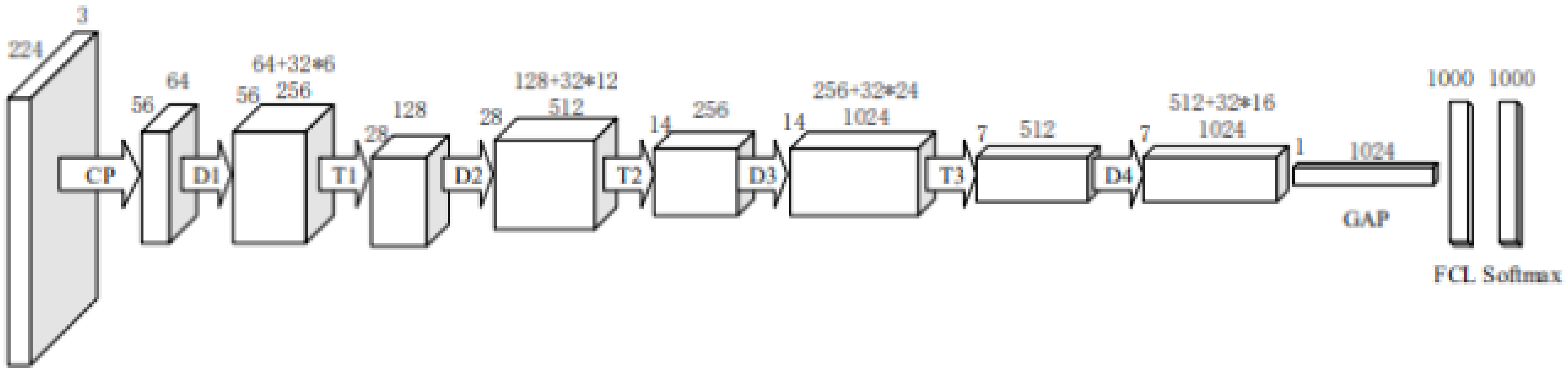


Input
Convolution
Dense Block 1
Transition Layer 1
Dense Block 2
Transition Layer 2
Dense Block 3
GAP+FCL+Softmax
Prediction
Which kind of fruits

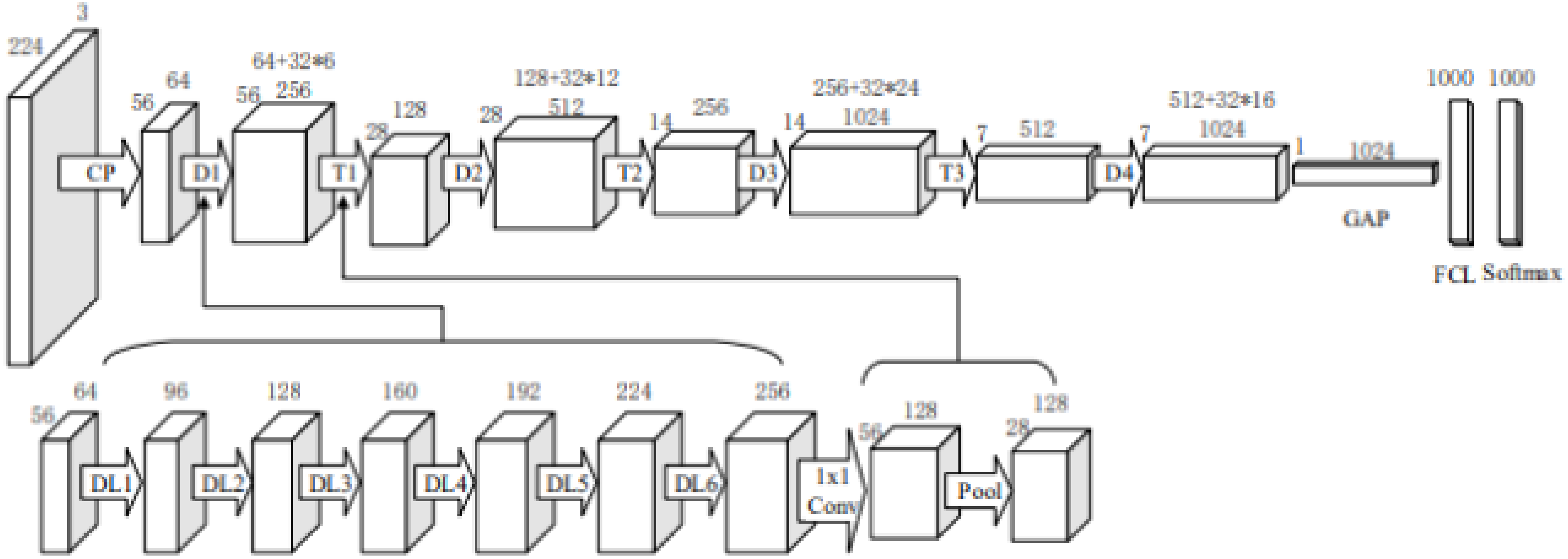
DENSENET

- THE DEFINED FEATURE: THE DENSE BLOCKS.
- CONSISTS OF ALL FEATURE MAPS OF ALL PREVIOUS CONVOLUTIONAL PROCESS
- EXTREME POOLING (1X1) IS NEEDED TO MAKE THE COMPUTATION VIABLE

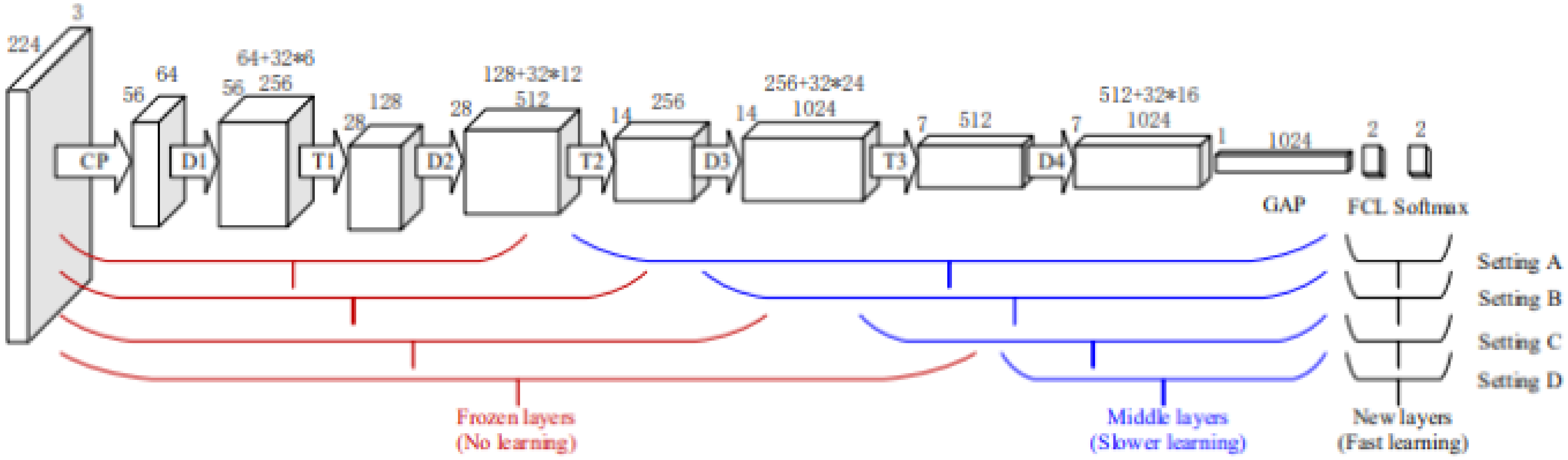




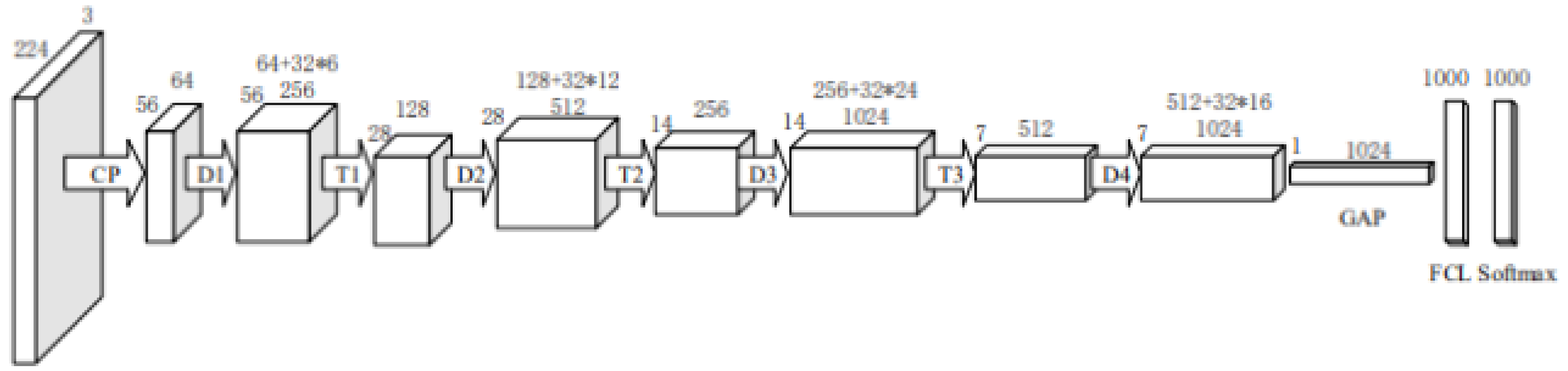
(a) Original pretrained DenseNet-121



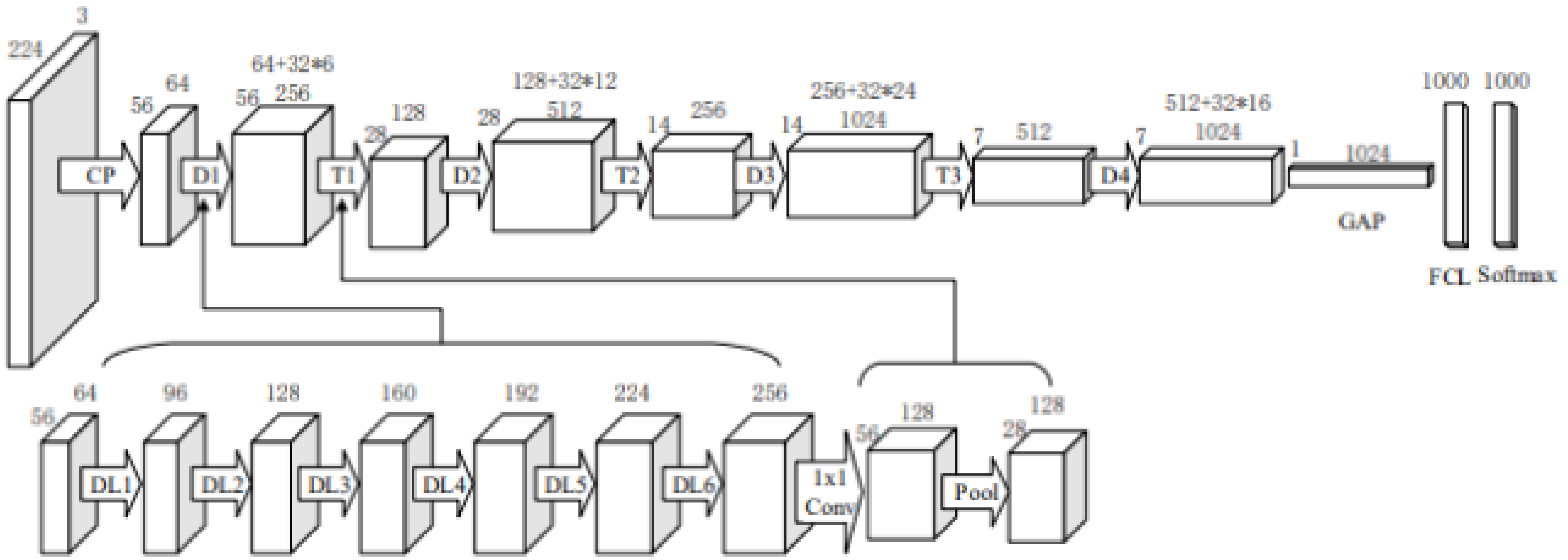
(b) one level deeper view of DenseNet-121



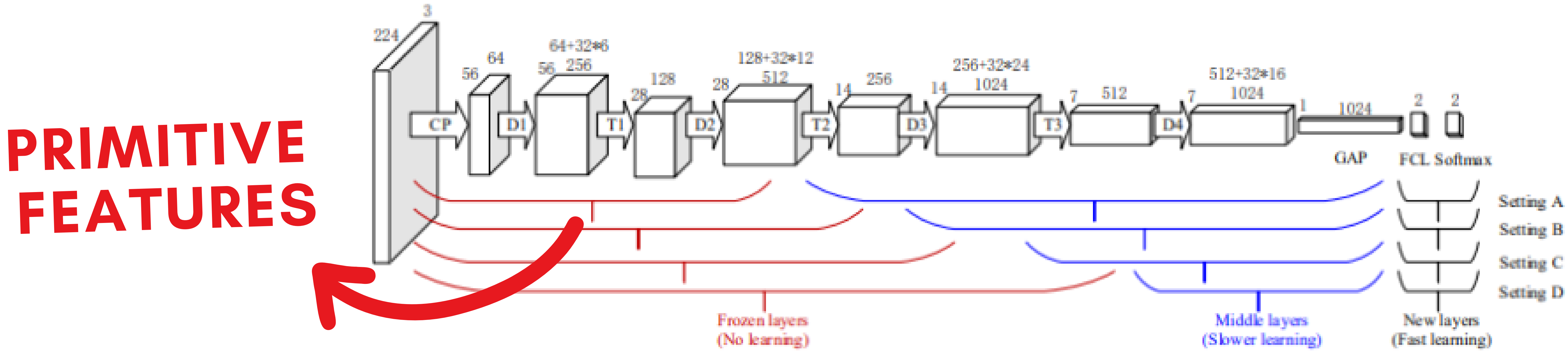
(c) DenseNet-121 transferred to our task



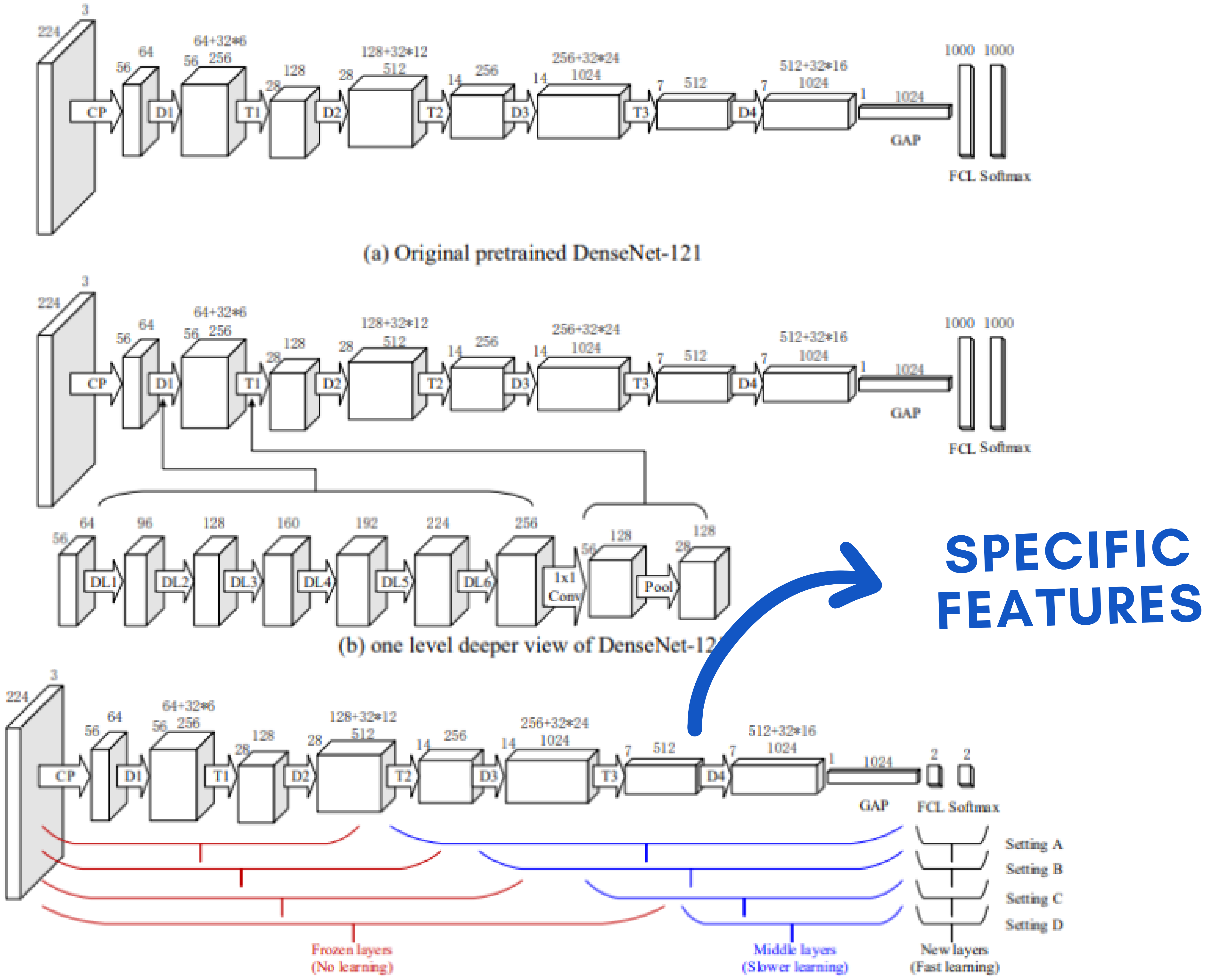
(a) Original pretrained DenseNet-121



(b) one level deeper view of DenseNet-121



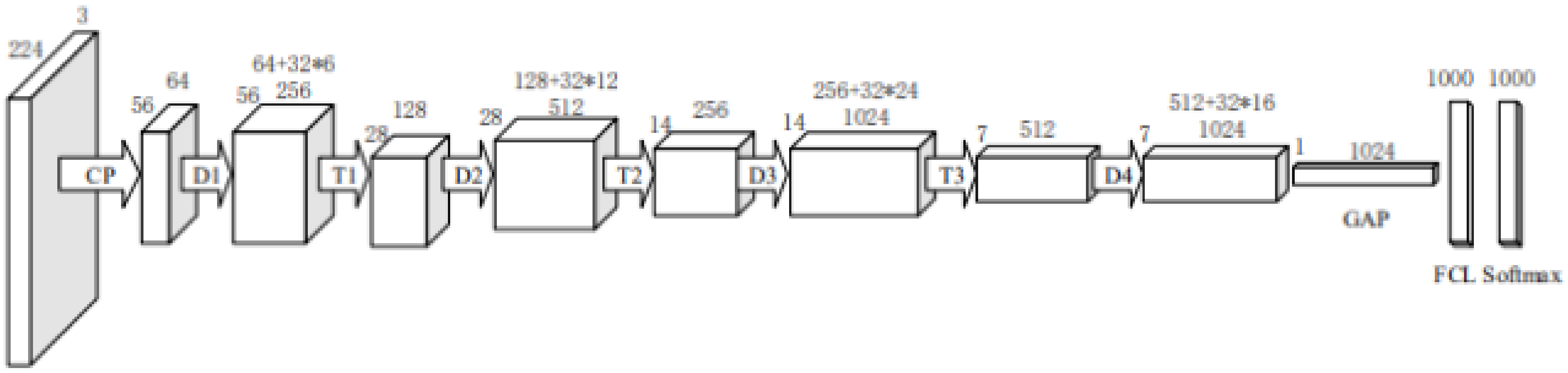
(c) DenseNet-121 transferred to our task



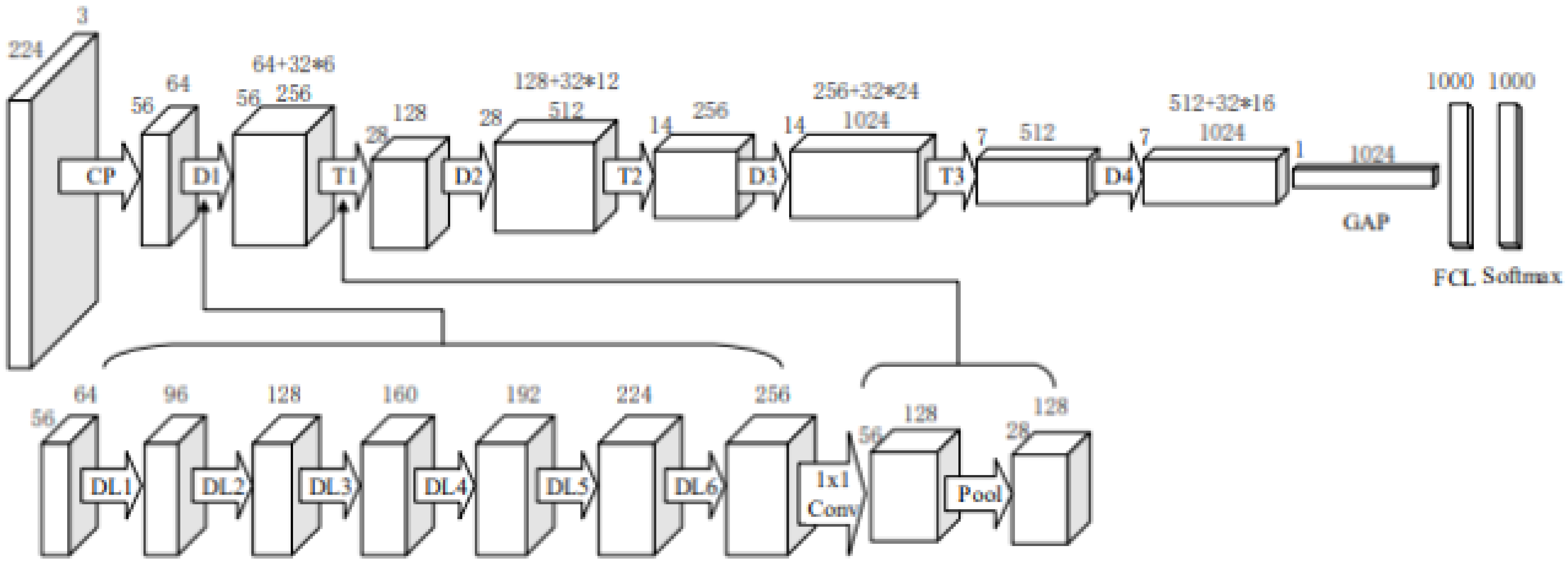
(a) Original pretrained DenseNet-121

(b) one level deeper view of DenseNet-121

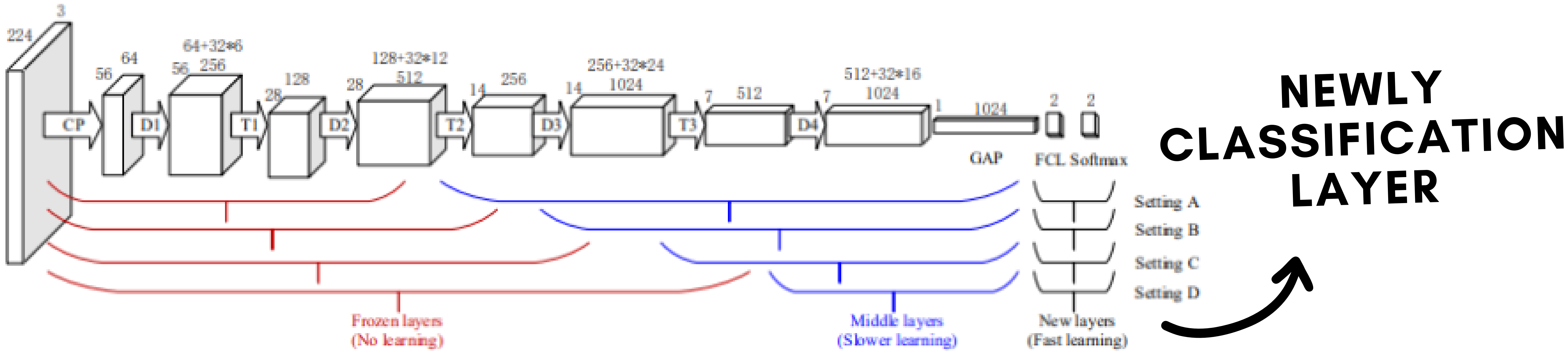
(c) DenseNet-121 transferred to our task



(a) Original pretrained DenseNet-121



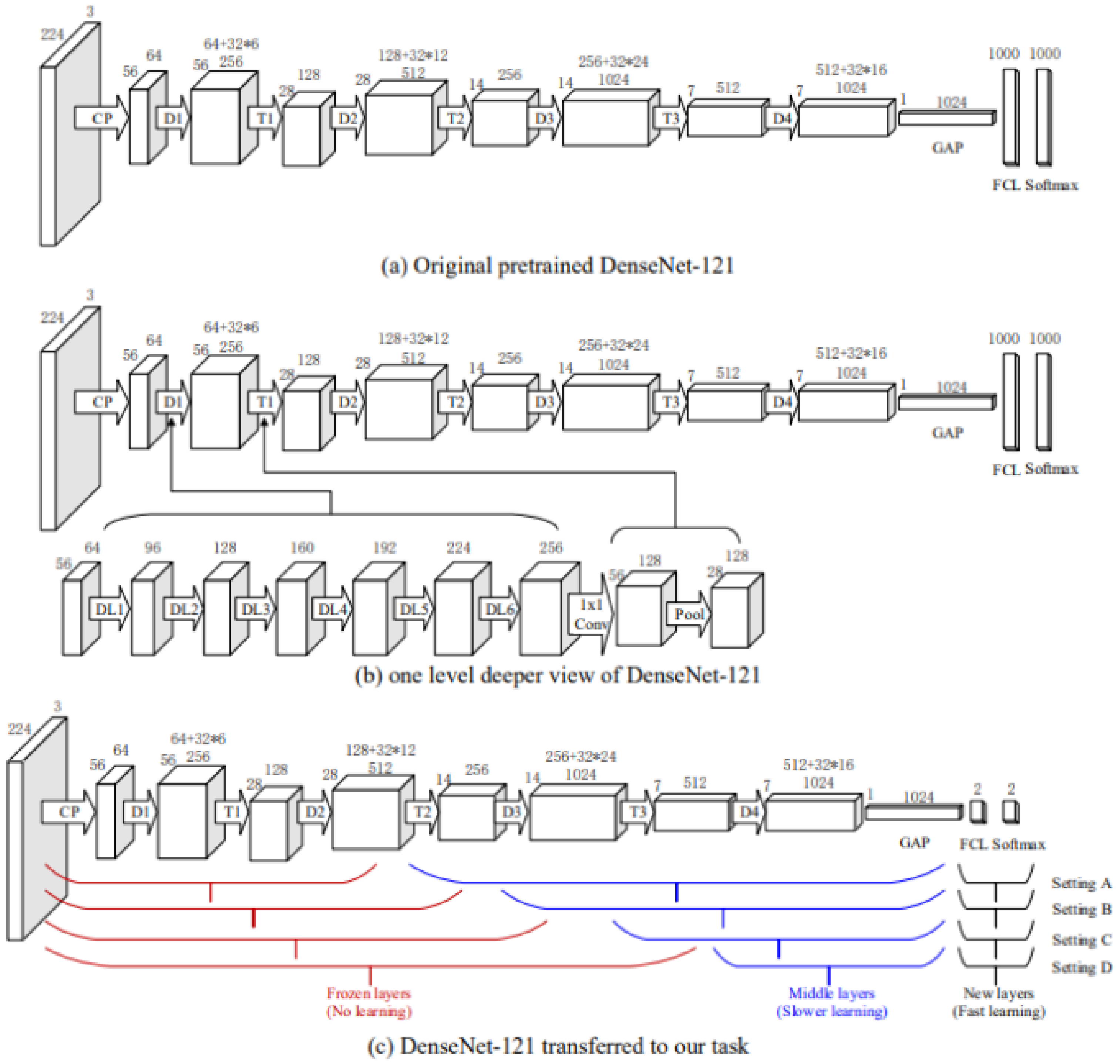
(b) one level deeper view of DenseNet-121



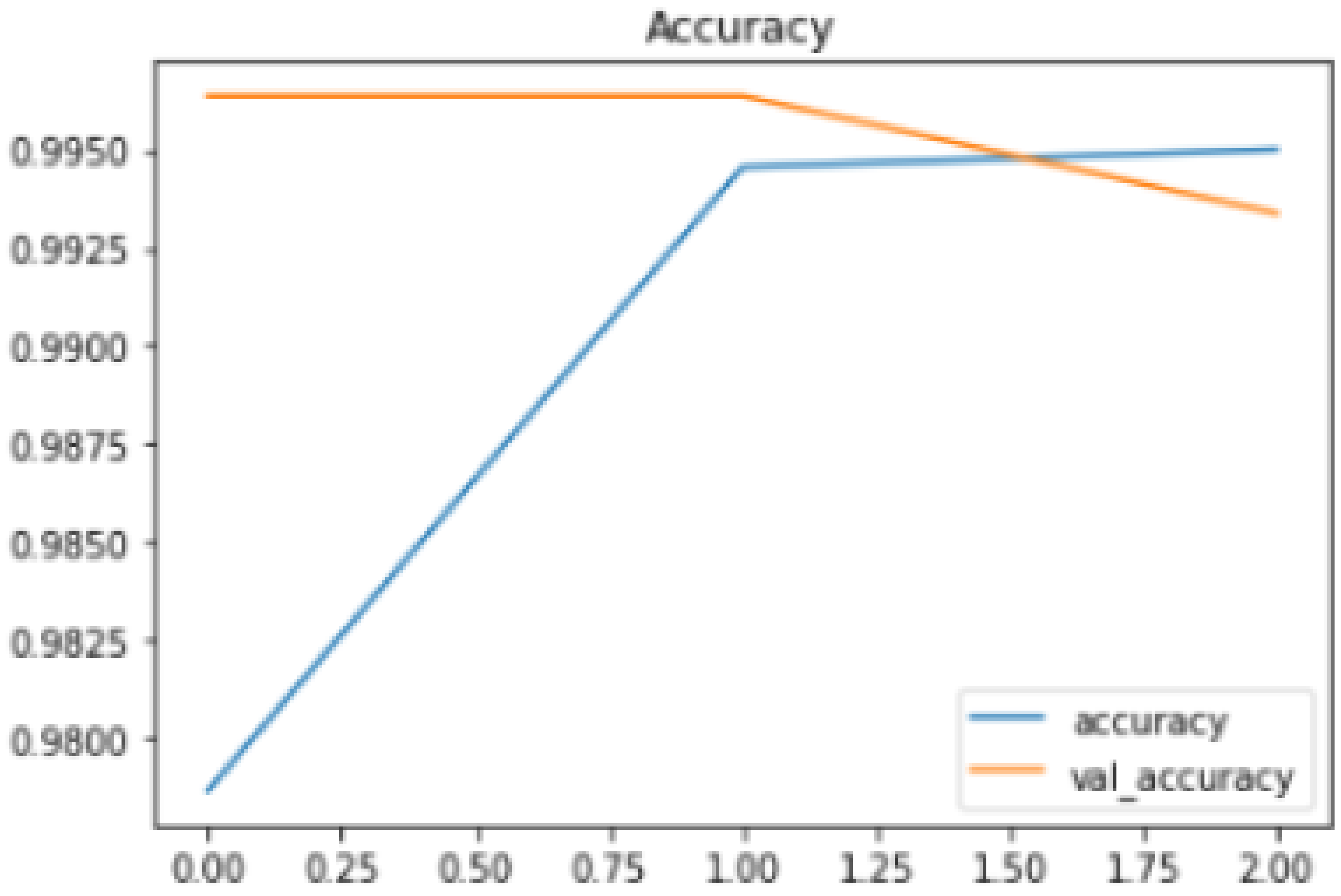
(c) DenseNet-121 transferred to our task

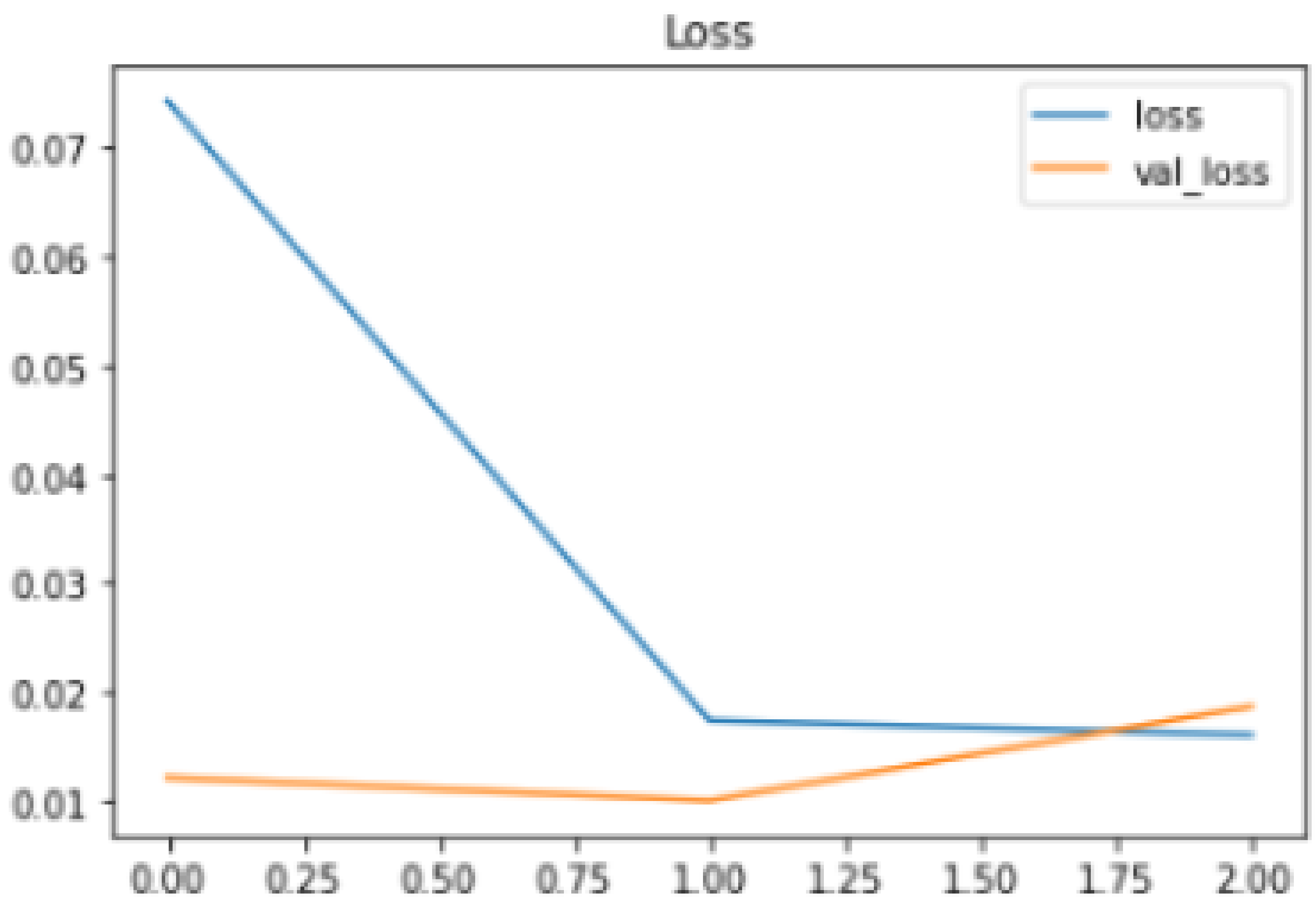
TRANSFER PROCESS

- FROZEN COUPLE FIRST LAYERS WHERE BASIC FEATURE IS TRAINED
- LATER DENSE LAYERS NEED TO BE SLOWLY RETRAINED TO EXTRACT FEATURE MOST REPRESENTITIVE FOR THE TASKS
- COMPLETELY RETRAIN THE CLASSIFICATION AND THE FCL.



RESULT OF DENSENET

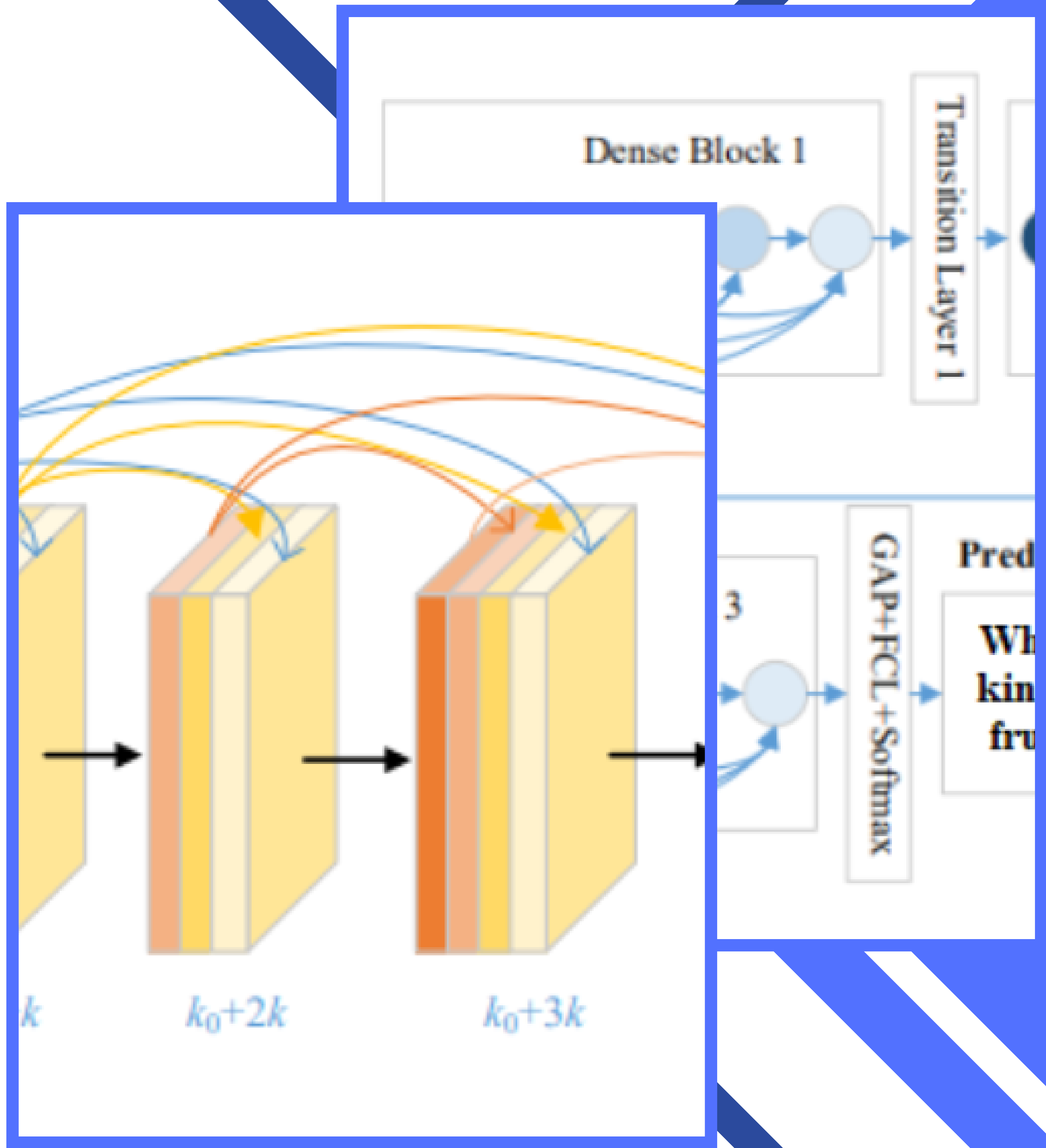





True: Guava
Predicted: Guava
True: Apple
Predicted: Apple
True: Kiwi
Predicted: Kiwi
True: Apple
Predicted: Apple
True: Apple
Predicted: Apple
True: Kiwi
Predicted: Kiwi
True: Tomatoes
Predicted: Tomatoes
True: Guava
Predicted: Guava
True: Pomegranate
Predicted: Pomegranate
True: Kiwi
Predicted: Kiwi
True: Guava
Predicted: Guava
True: Apple
Predicted: Apple
True: Apple
Predicted: Apple
True: Peach
Predicted: Peach
True: Mango
Predicted: Mango
True: Guava
Predicted: Guava
True: Banana
Predicted: Banana
True: Plum
Predicted: Plum
True: Kiwi
Predicted: Kiwi
True: Pear
Predicted: Pear
True: muskmelon
Predicted: muskmelon
True: Apple
Predicted: Apple
True: Guava
Predicted: Guava
True: muskmelon
Predicted: muskmelon

DISCUSION

- IMPRESSIVE RESULT: 99,6 %
- MUCH MORE STABLE IN PREDICTING SIMILAR FRUIT PRODUCT COMPARED TO BASIC CNN



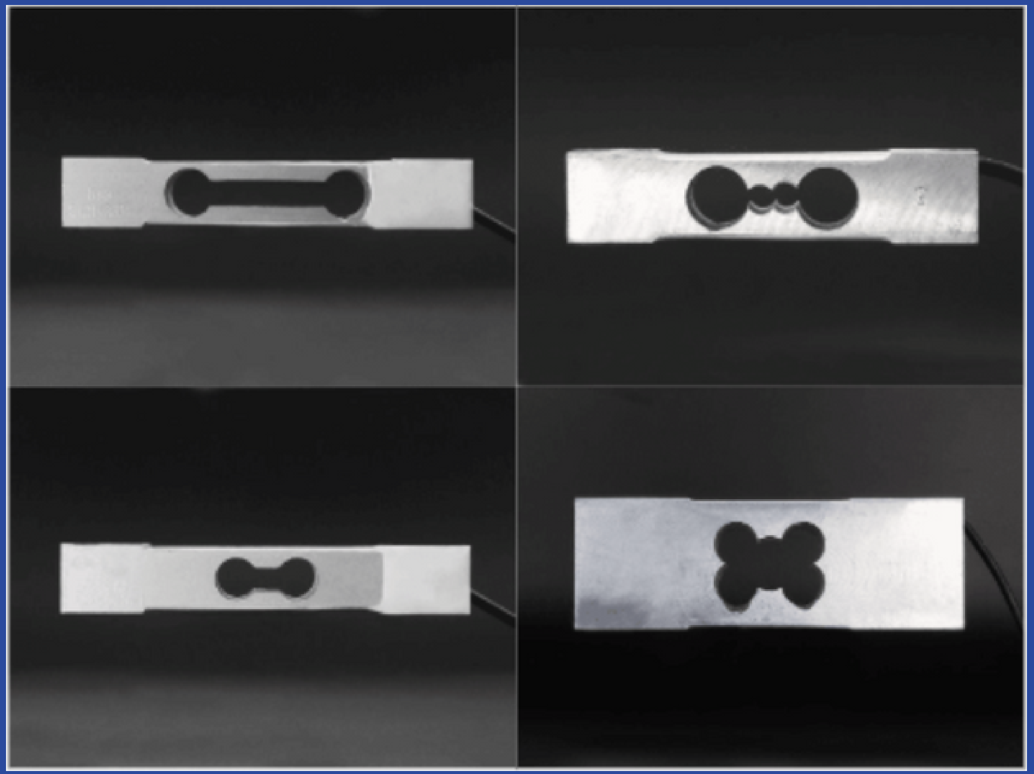
3

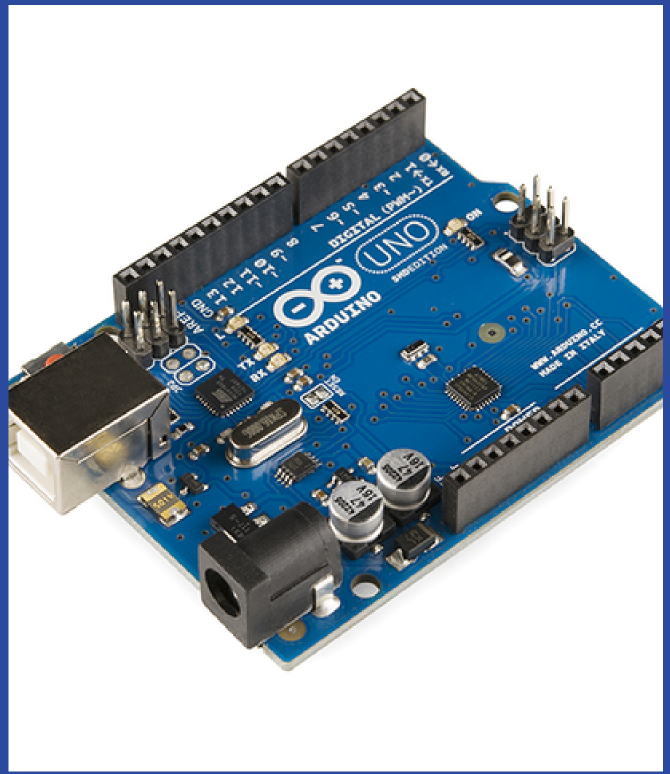
IMPLEMENTATION

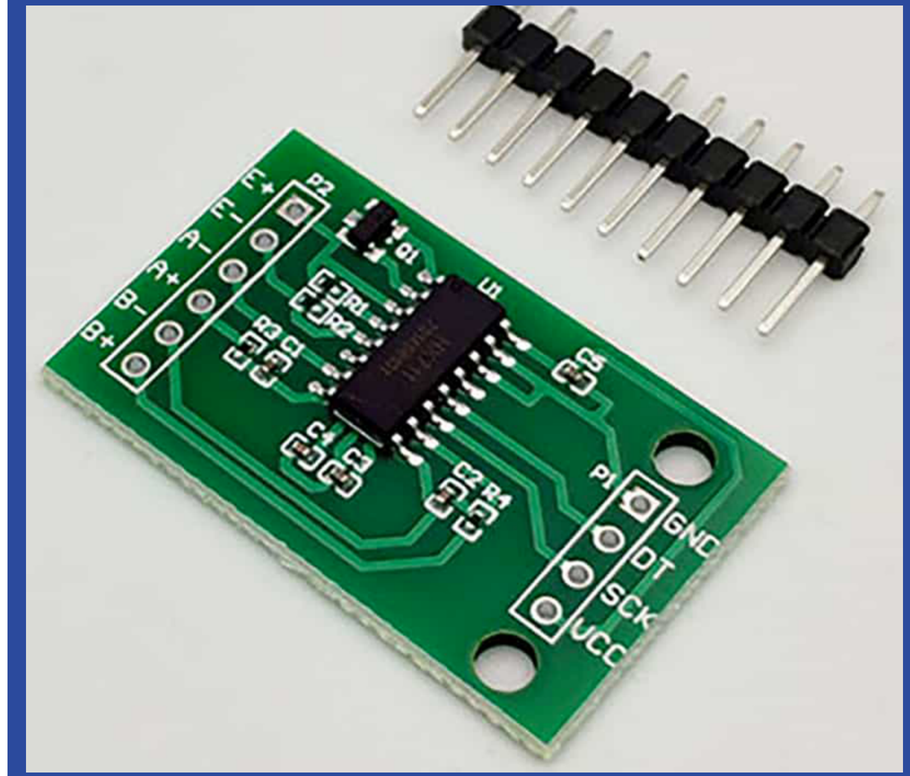
WEIGHT SCALE

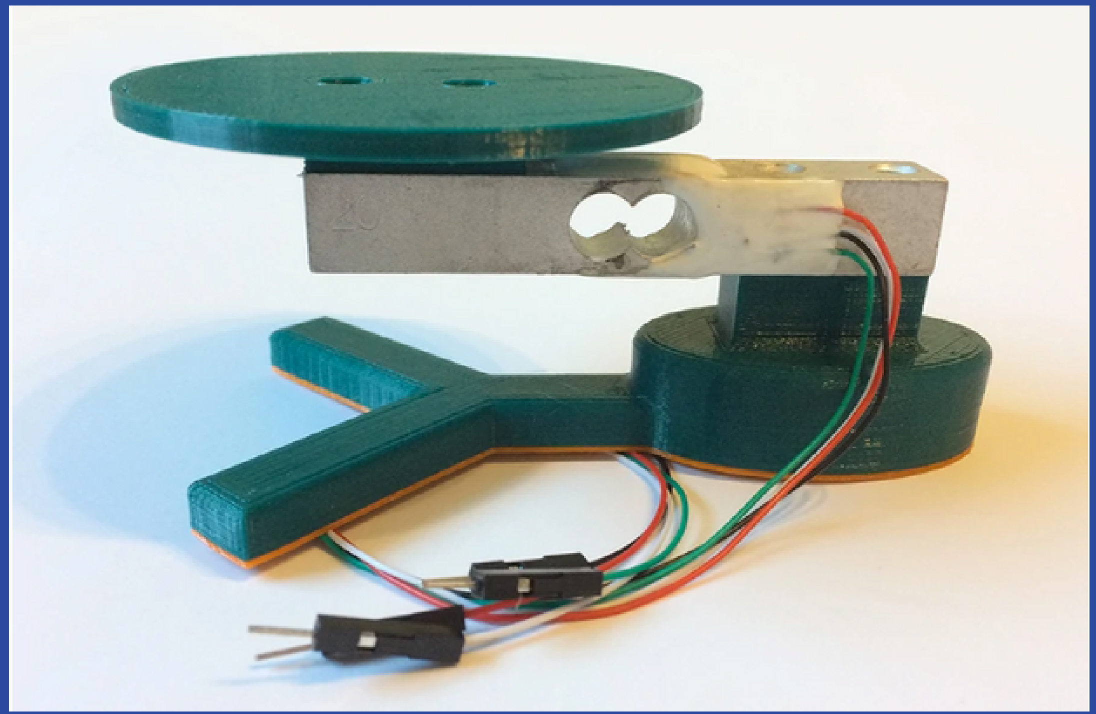
COMPONENTS

- *Load cell 10kg*
- *Arduino kit (Arduino Uno, jumper wires)*
- *HX711 module*
- Scale base



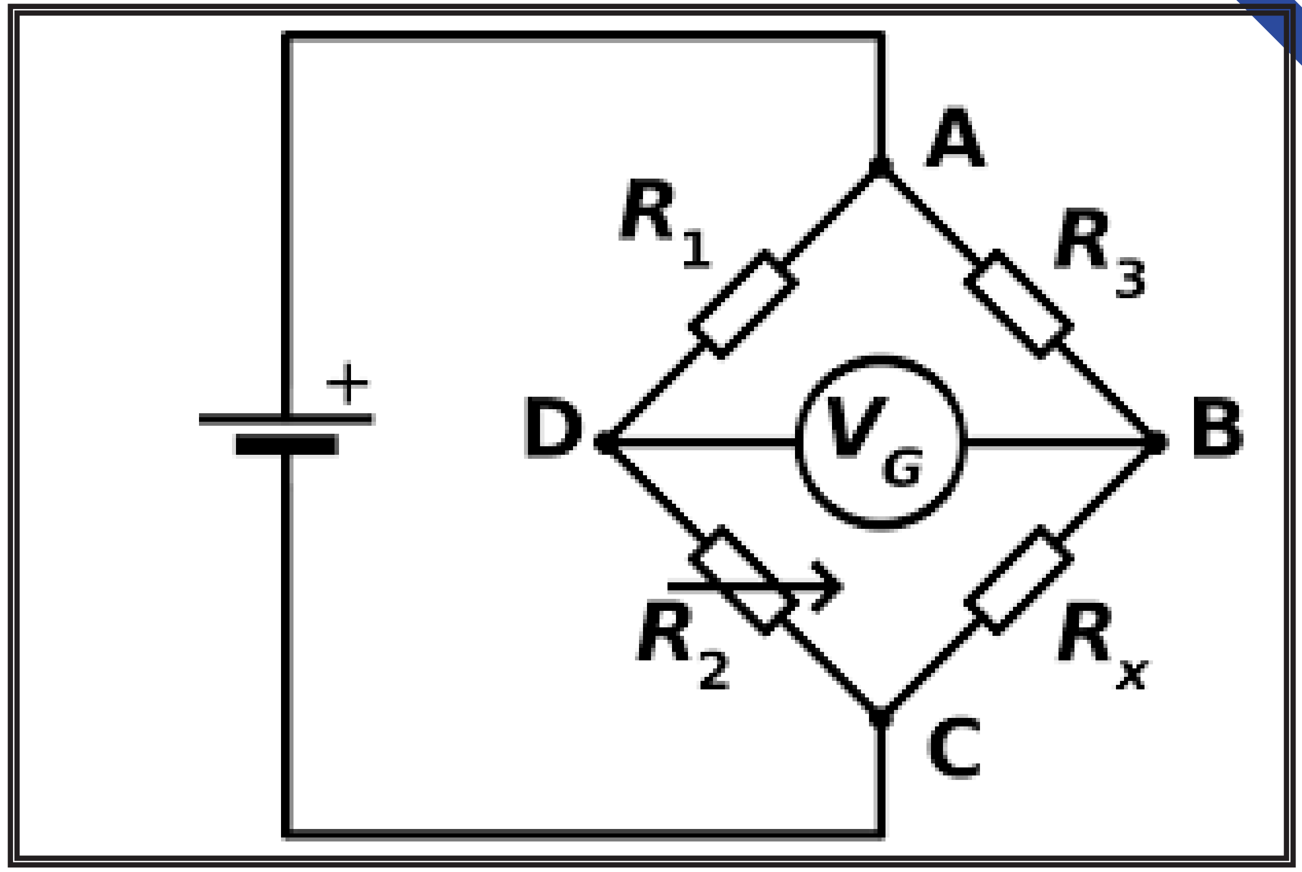


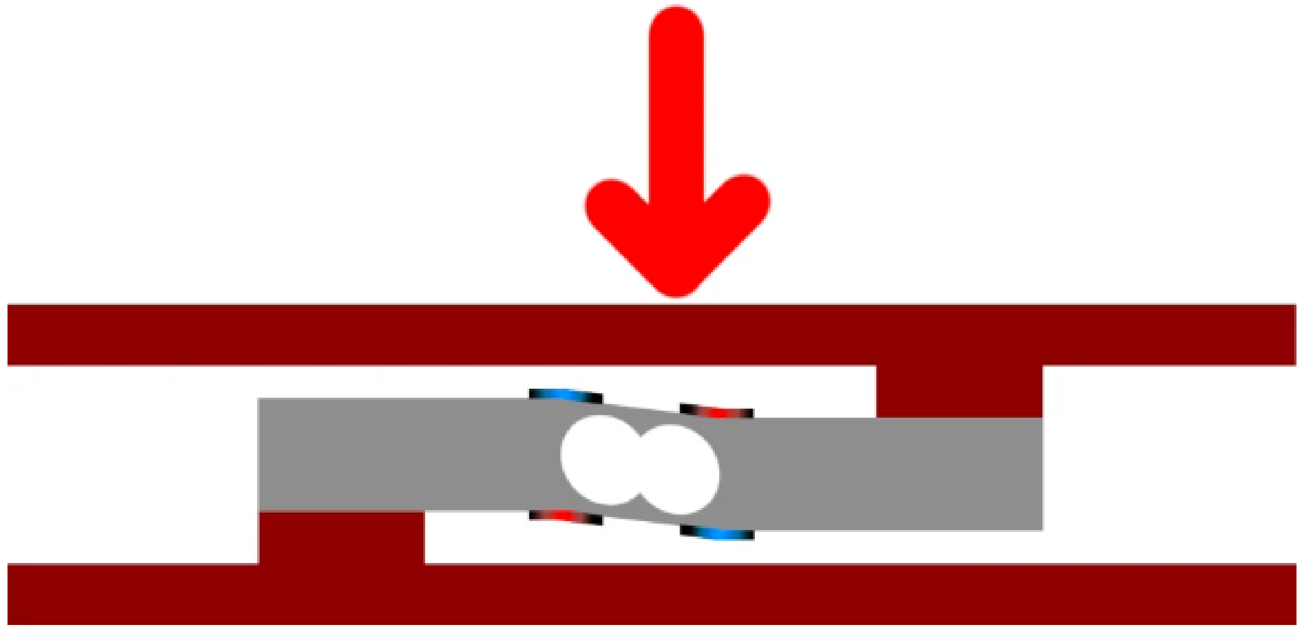




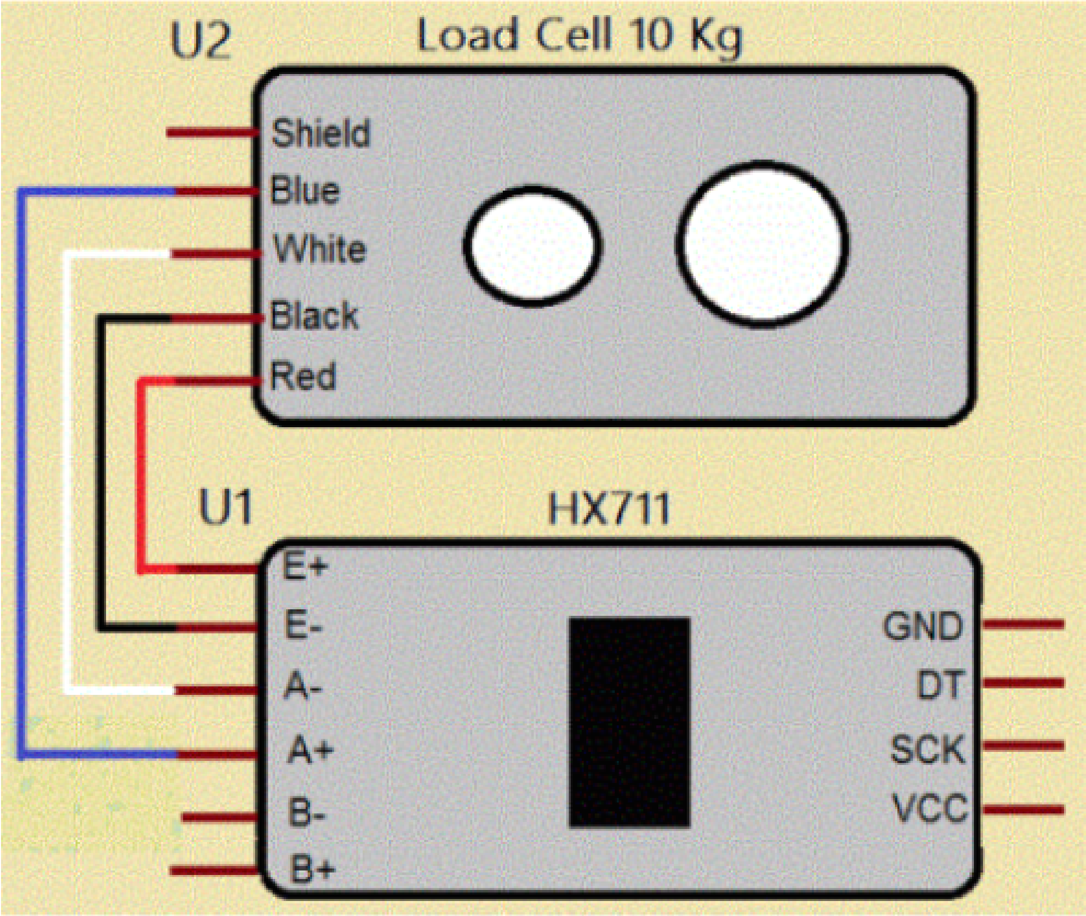
LOAD CELL

- *The load cell has 4 strain gauges that are formed as the Wheatstone bridge formation.*
- *A transducer device for measuring strain and then converting forces into electrical energy that can be measured*
- *When applying force to the load cell, 2 of the 4 strain gauges will compress (marked red) while the rest will stretch (marked blue).*

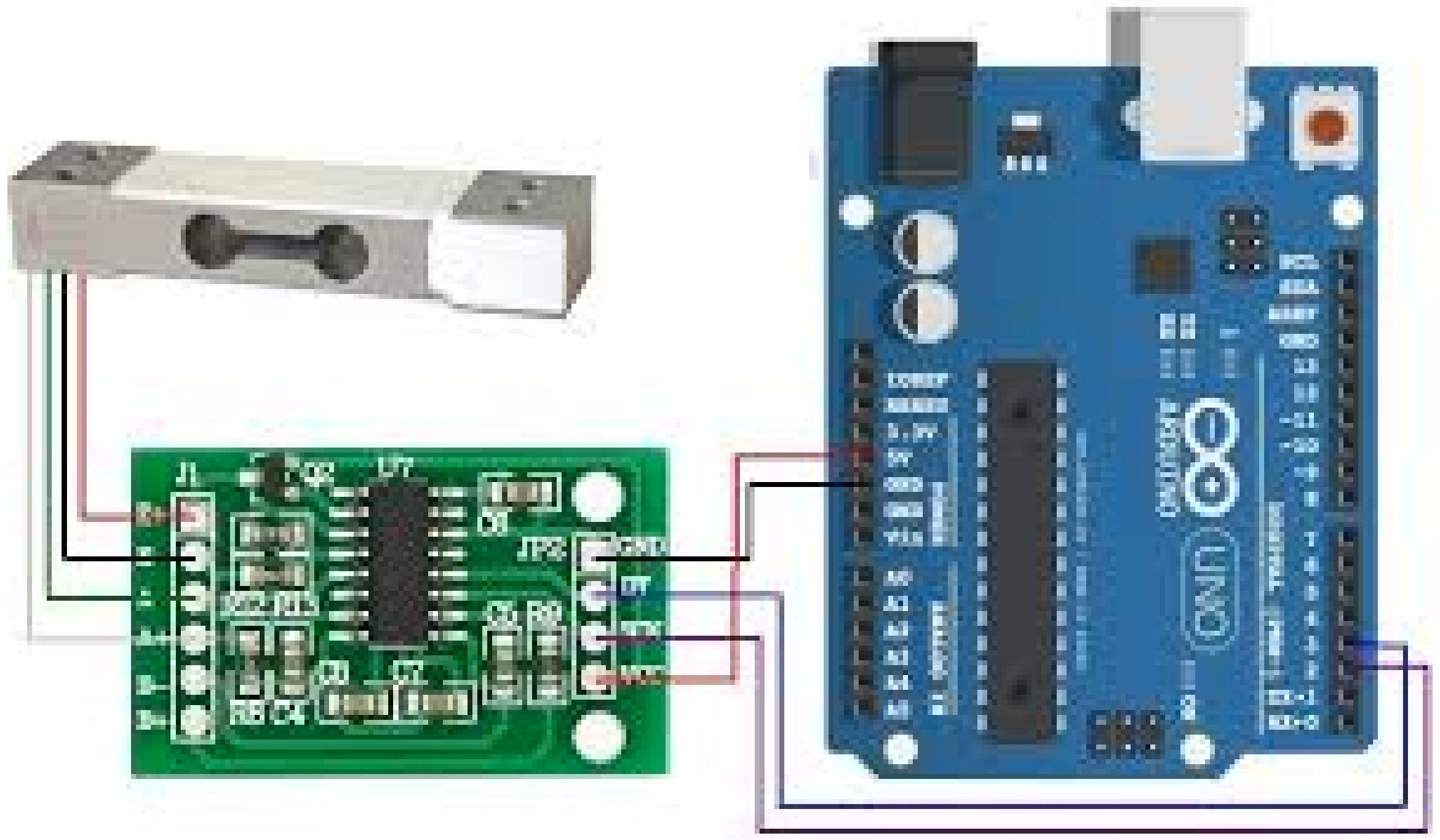




HX711 MODULE



- *Also known as HX711 Load Cell Amplifier*
- *Connect to a microcontroller helps read changes in the resistance of the load cell*
- *Communicates using a two-wire interface, namely Clock and Data*
- HX711_ADC library helps with the calibration: https://github.com/olkal/HX711_ADC

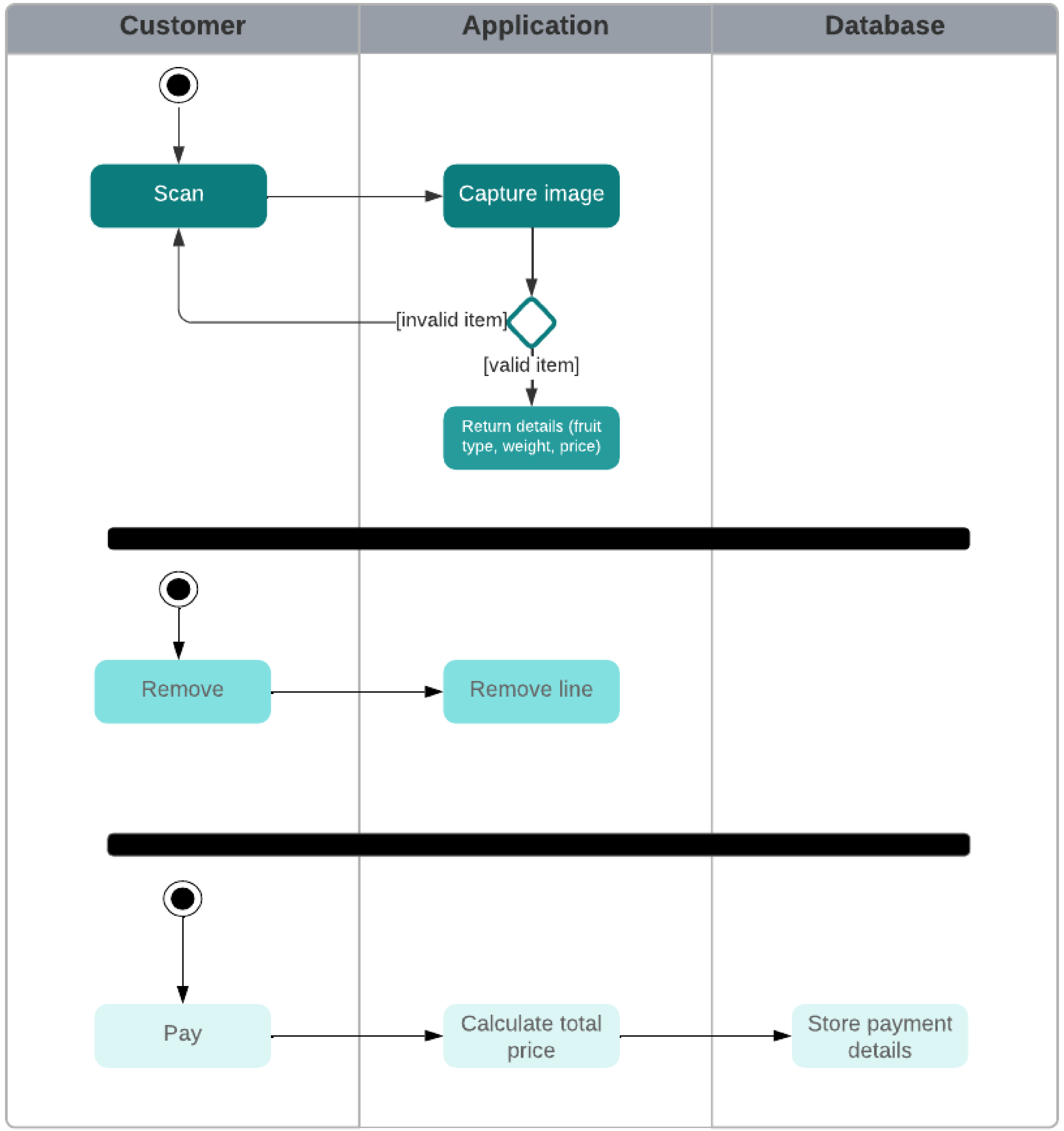


OTHERS

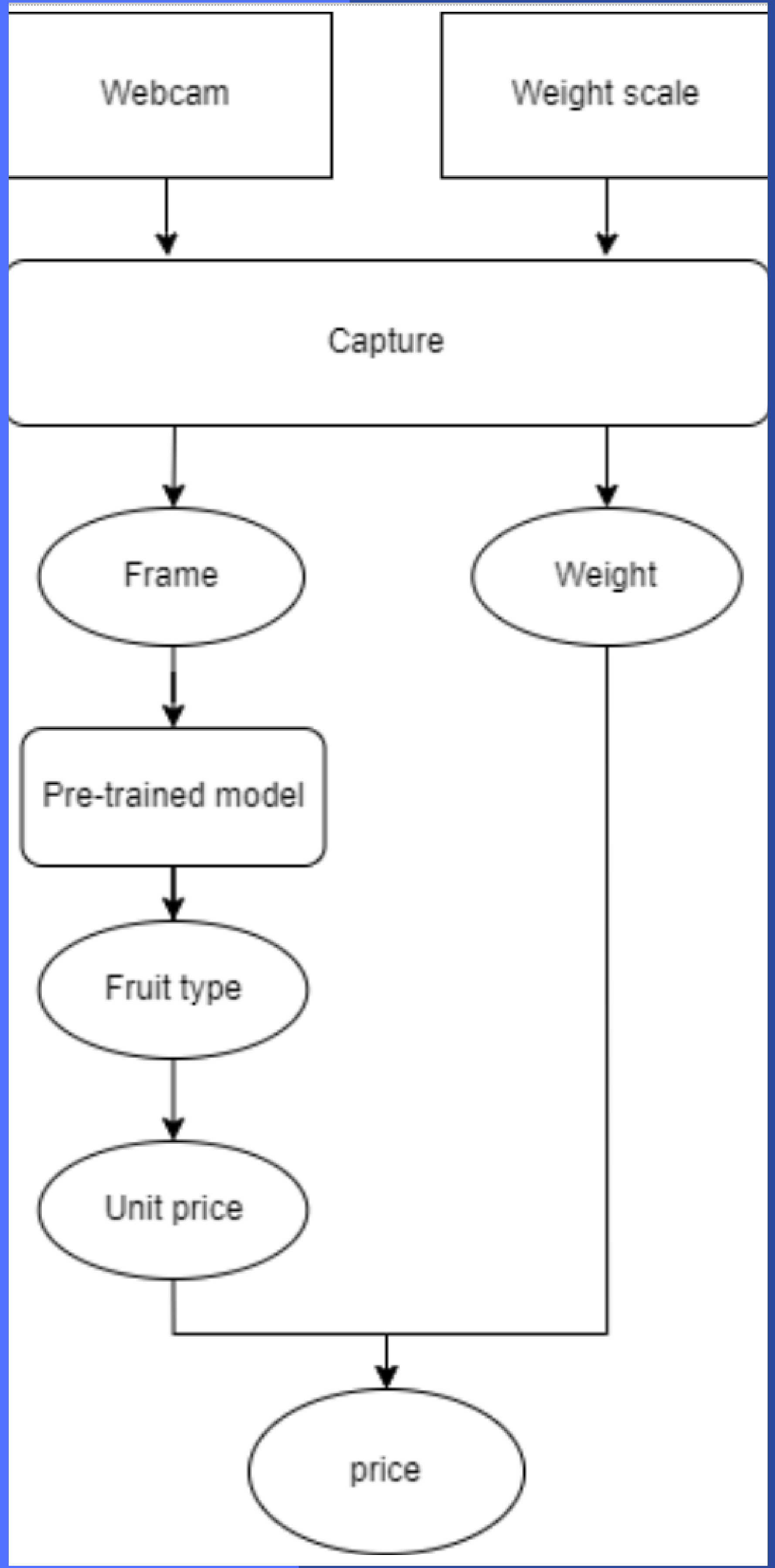
CoolTerm for transmitting weight values

HIKVISION webcam with a resolution of 1920 × 1080 and a USB plug-in

WORKFLOW



FLOWCHART



DEMO

5

CONCLUSION

ACHIEVEMENTS

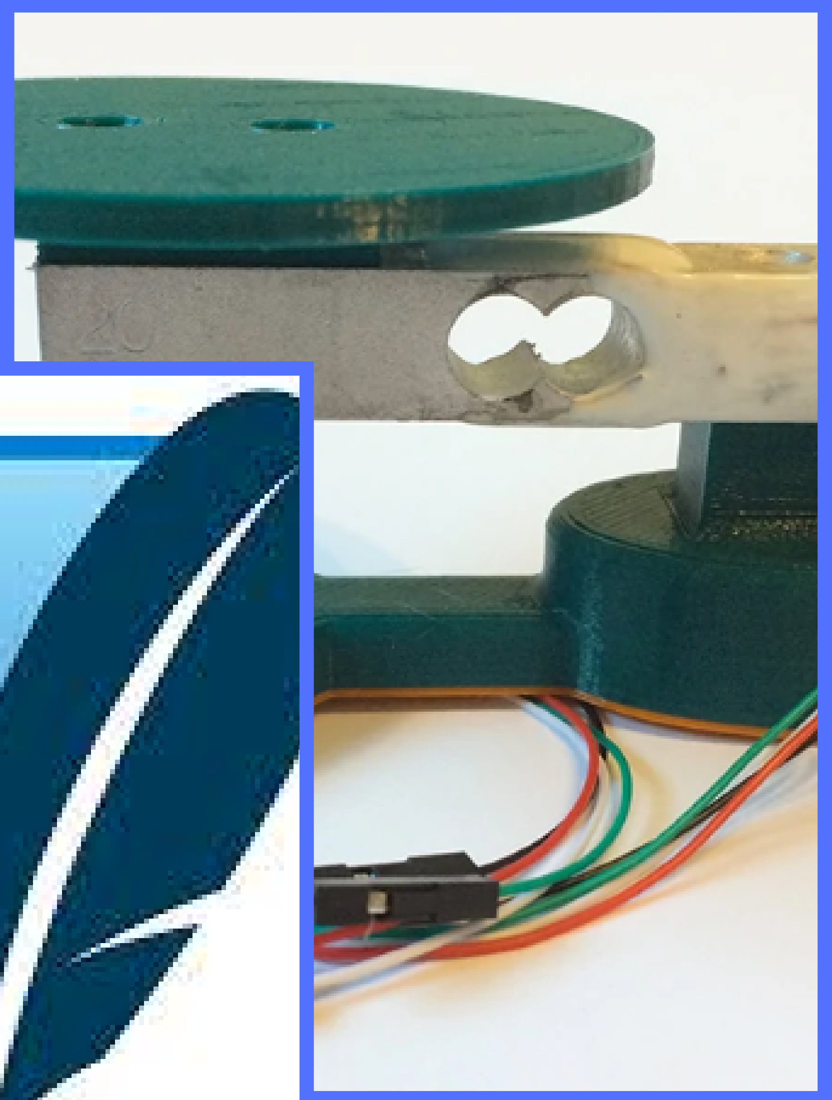
- *A CNN model recognize up to 15 fruit varieties, and detecting anomalies.*

- *A sketch model of a weight scale.*

- *Transmit data value from the weight via a third-party application, which helps to utilize memory capacity.*

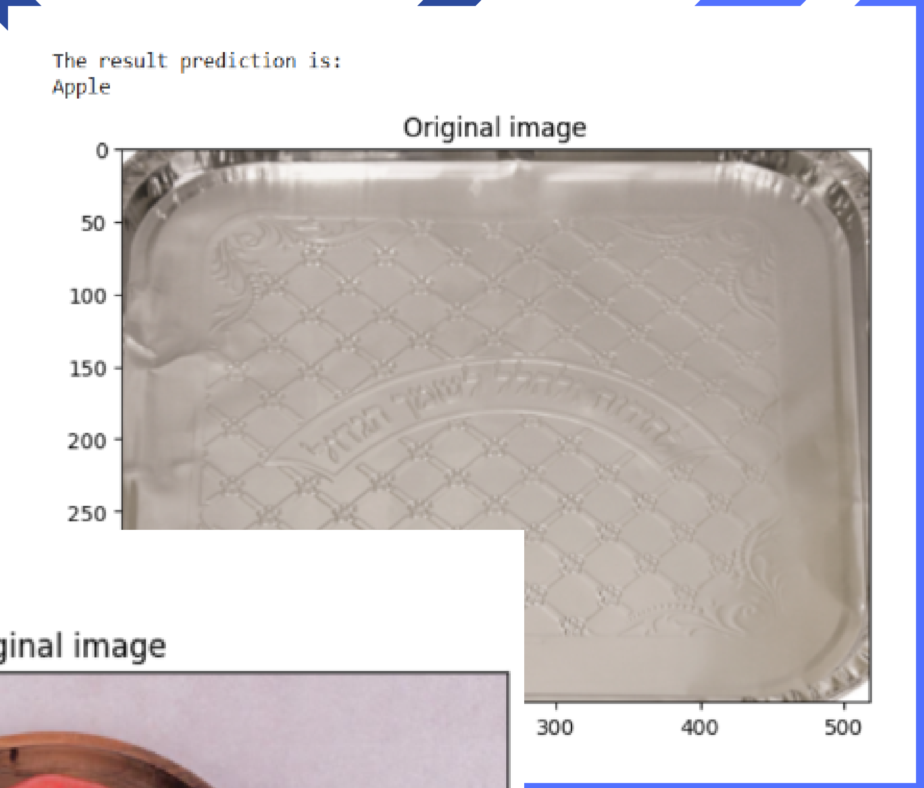
- An application with a user-friendly interface, high stability, and utilized running time that connects all pieces and performs real-time recognition and price result.

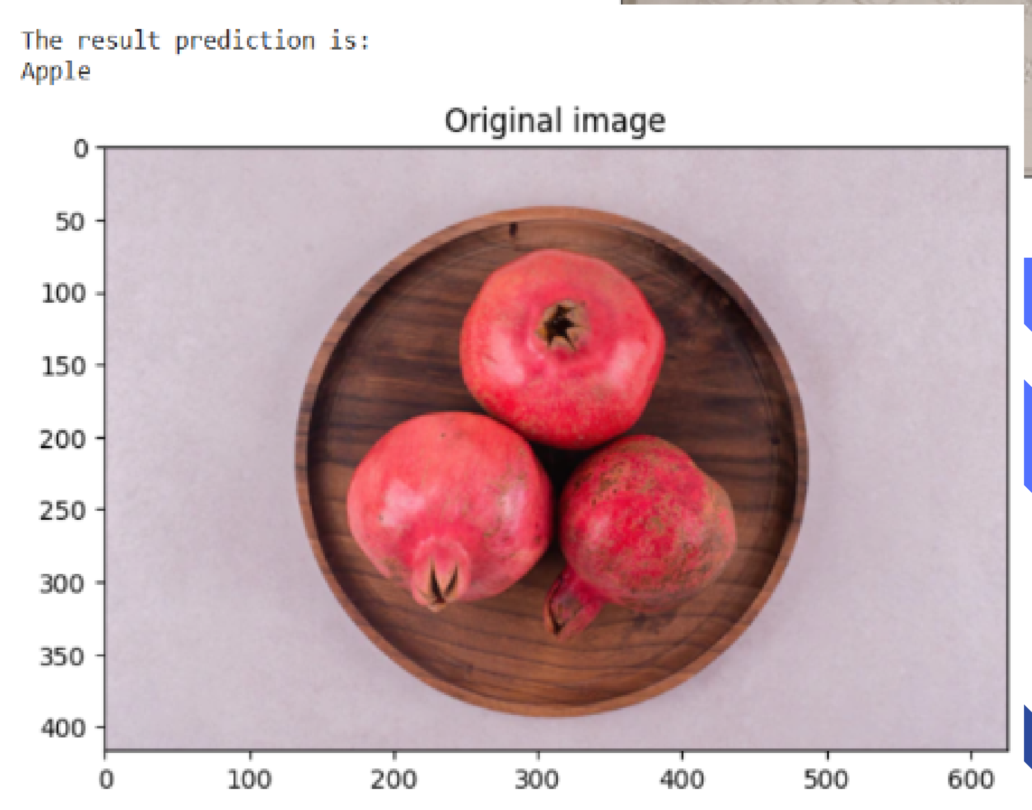
- Established a connection with a database feature for storing payment and sales details.



IMPROVE THE MODEL

- Improve the recognition ability





IMPROVE THE MODEL

- Improve the recognition ability
- Implement segmentation to detect mixed fruits situation
- Widen the range of fruits available to the model

IMPROVE THE MODEL

- Improve the recognition ability
- Implement segmentation to detect mixed fruits situation
- Widen the range of fruits available to the model
- Improve the ability to detect obstacles or non-fruit anomaly

IMPROVE THE MODEL

- Improve the recognition ability
- Implement segmentation to detect mixed fruits situation
- Widen the range of fruits available to the model
- Improve the ability to detect obstacles or non-fruit anomaly
- Upgrade computing power (using GPU) for faster starting, processing time

IMPROVE THE HARDWARE

- The quality and stability of a weighting machine can be improved with more up-to-date and high-end equipment
- Increase the number of sensors, typically load cells, instead of one, thus increase accuracy
- Upgrade to a higher resolution camera for better image quality

IMPROVE THE SOFTWARE

- Adding numerous features like a QR code generator comprised of details such as type, weight, and total price for payment and billing purposes.

- Incorporate with a database feature for storing payments and sales, useful for managing a large supply chain.

FPT Education

FPT UNIVERSITY

THANKS FOR WATCHING