

TRANSFER LEARNING

DAO THI TRUC LOAN

Tuesday, November 23th, 2021

Outline

1. What is Transfer Learning?
2. Why is use Transfer Learning?
3. How transferable are features?
4. Approaches to Transfer Learning
5. Experimental
6. Conclusions

Outline

1. What is Transfer Learning?

1. What is Transfer Learning?



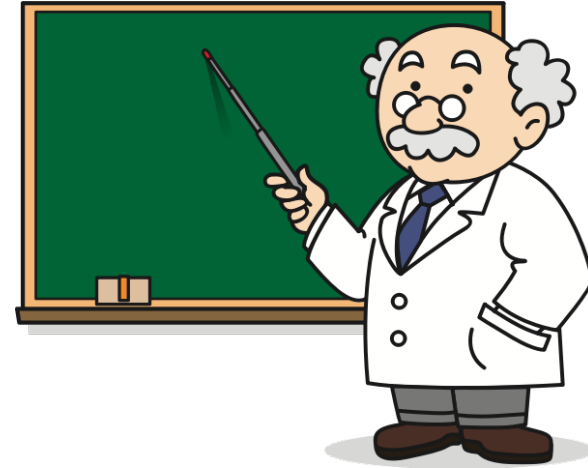
1. What is Transfer Learning?



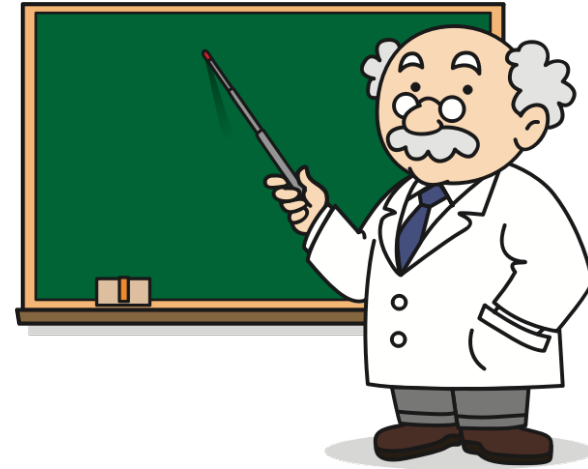
1. What is Transfer Learning?



1. What is Transfer Learning?



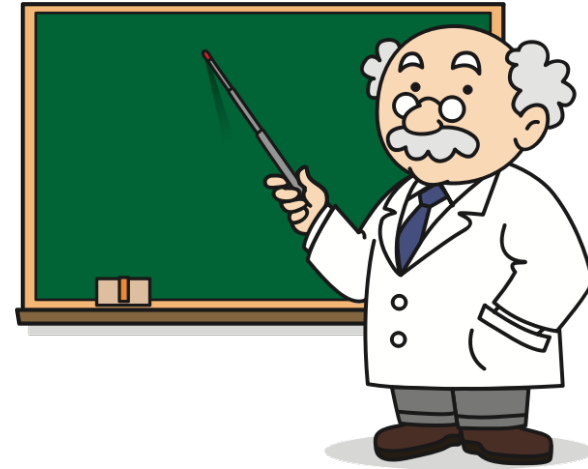
1. What is Transfer Learning?



1. What is Transfer Learning?

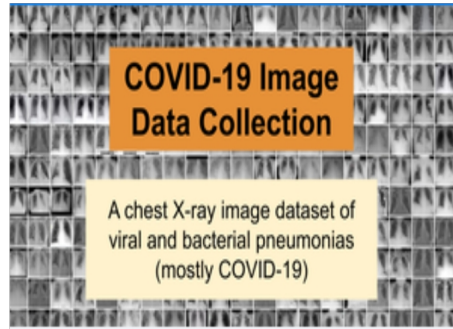


Learning from Scratch

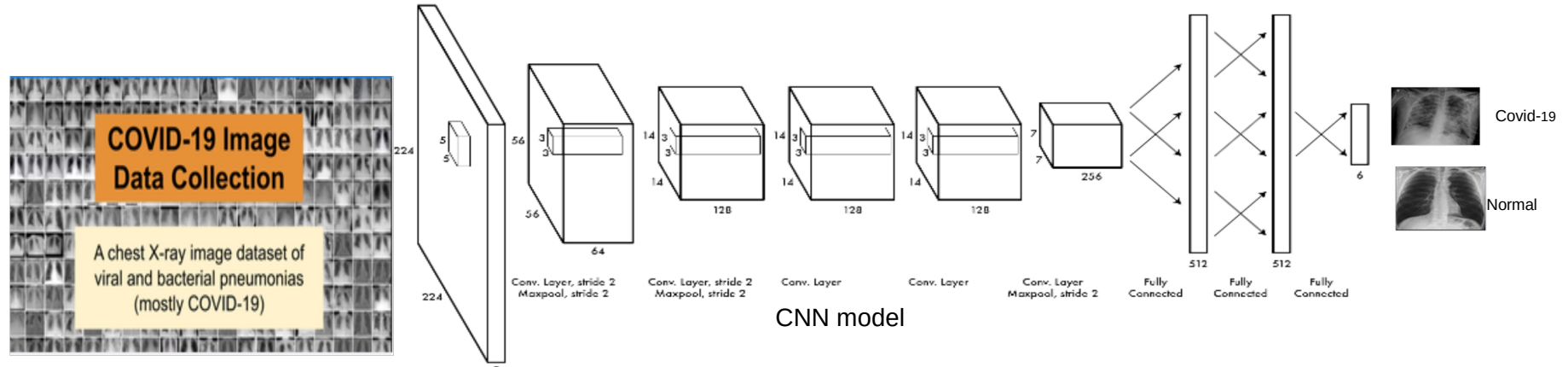


Transfer Learning

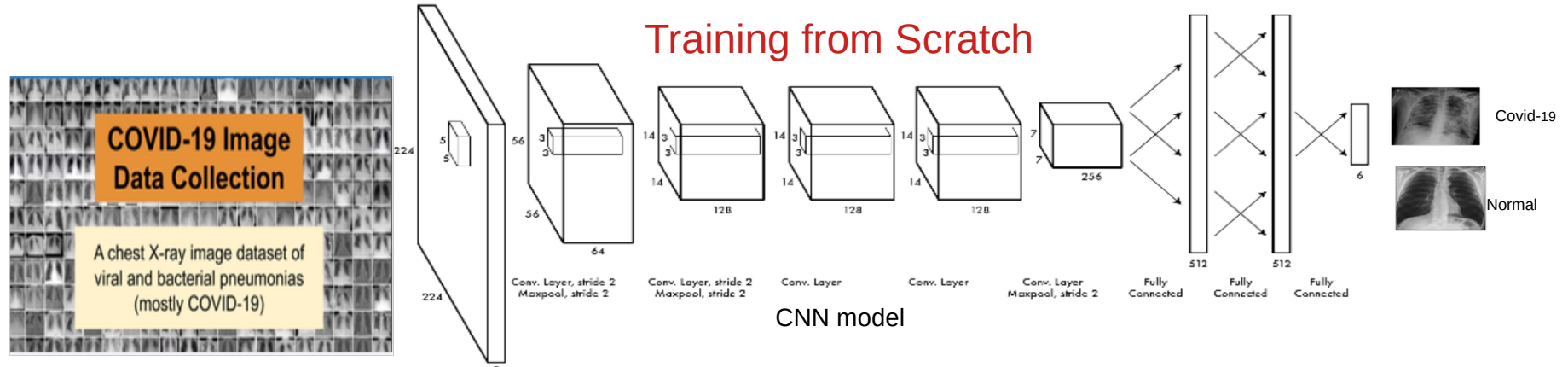
1. What is Transfer Learning?



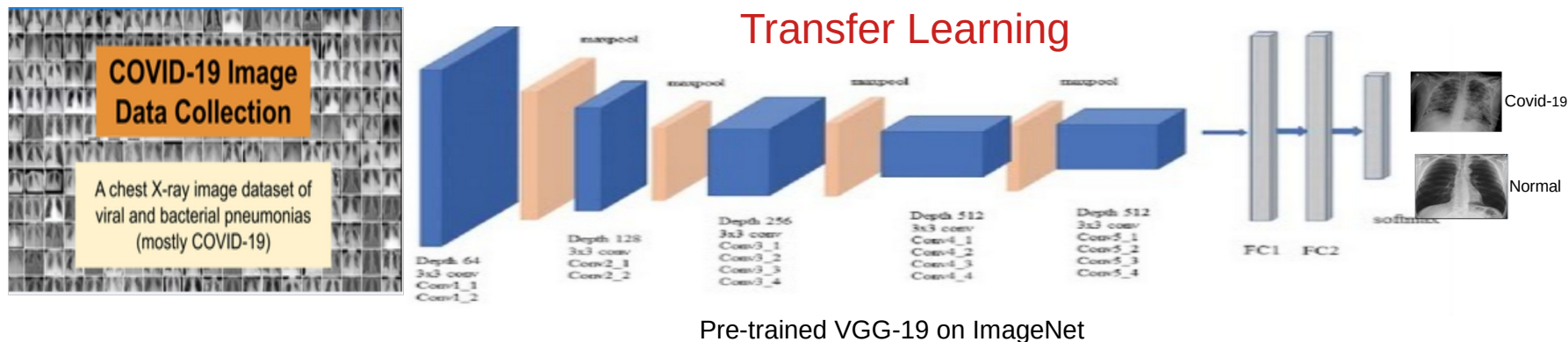
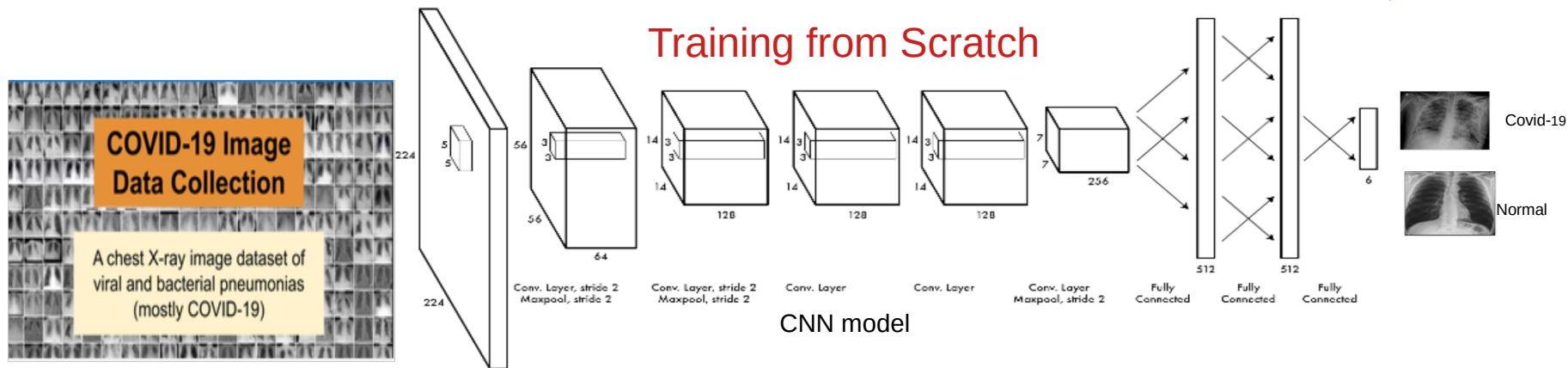
1. What is Transfer Learning?



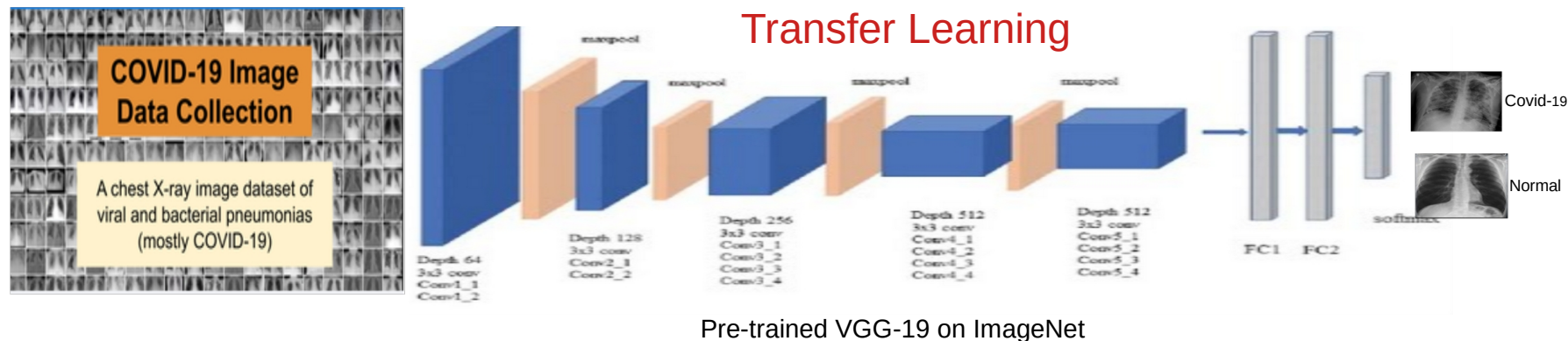
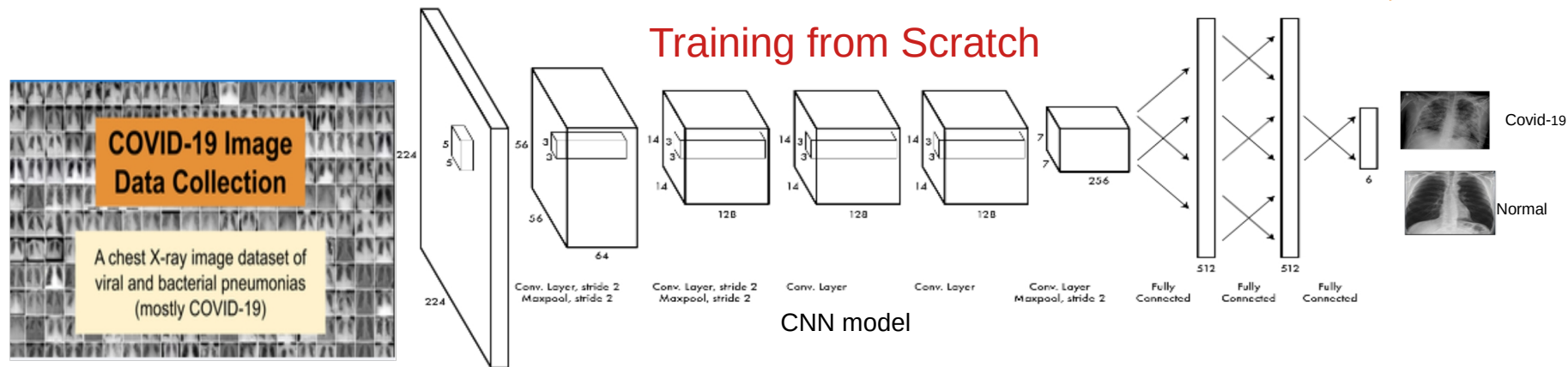
1. What is Transfer Learning?



1. What is Transfer Learning?

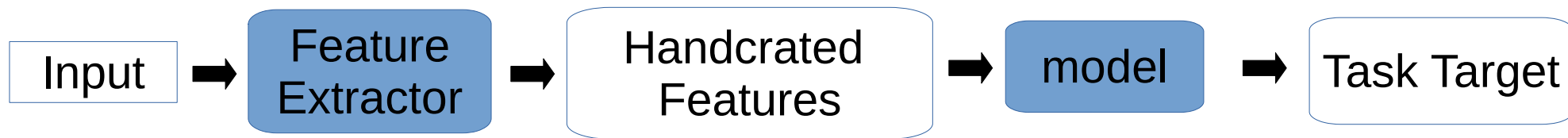


1. What is Transfer Learning?

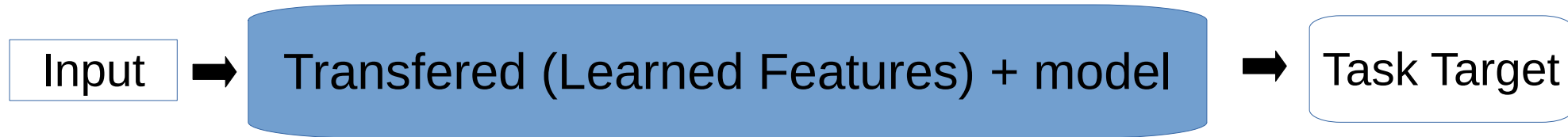


Transfer learning is a method of reusing a model or knowledge for another related task

1. What is Transfer Learning?



Traditional Learning Flow



Transfer Learning Flow

Outline

1. What is Transfer Learning?

2. Why is use Transfer Learning?

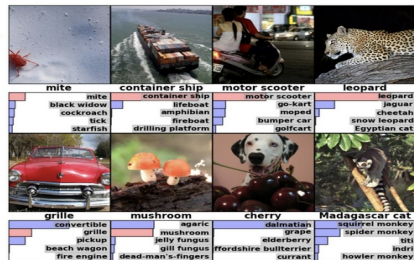
2. Why is use Transfer Learning?

High-quality dataset consisting of **a large** amount of data

ImageNet Challenge

IMAGENET

- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.



2. Why is use Transfer Learning?

High-quality dataset consisting of **a large** amount of data

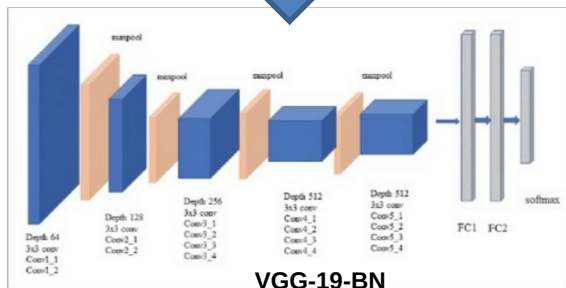
ImageNet Challenge

IMAGENET

- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.



↓ Training



High performance

2. Why is use Transfer Learning?

High-quality dataset consisting of **a large** amount of data

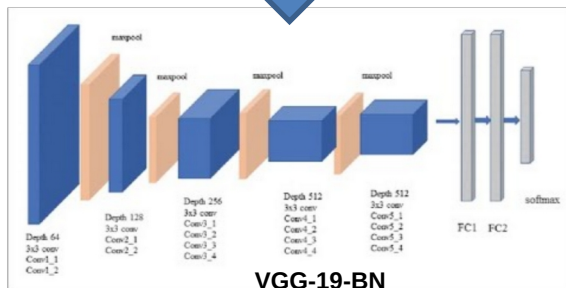
ImageNet Challenge

IMAGENET

- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.

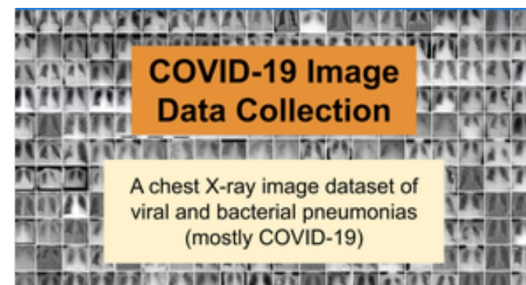


Training

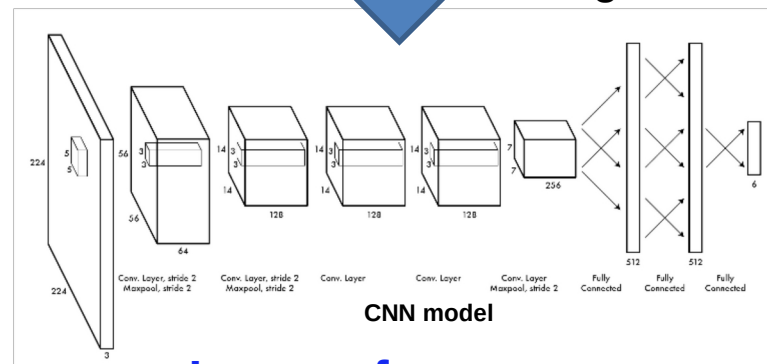


High performance

Dataset consisting of **a small** amount of data



Training



Low performance

2. Why is use Transfer Learning?

High-quality dataset consisting of **a large** amount of data

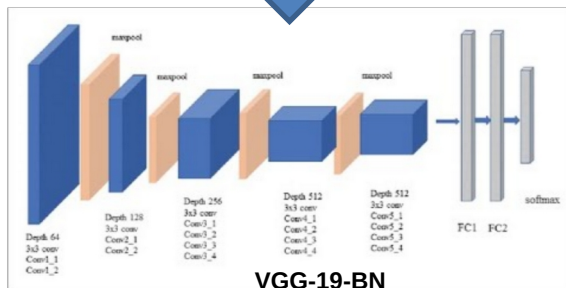
ImageNet Challenge

IMAGENET

- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.

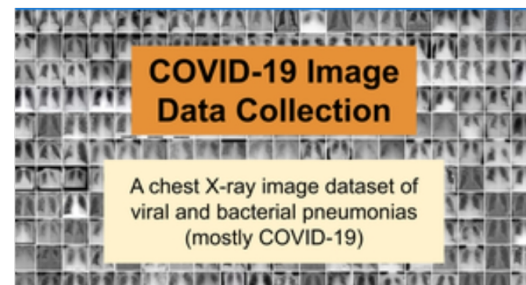


Training

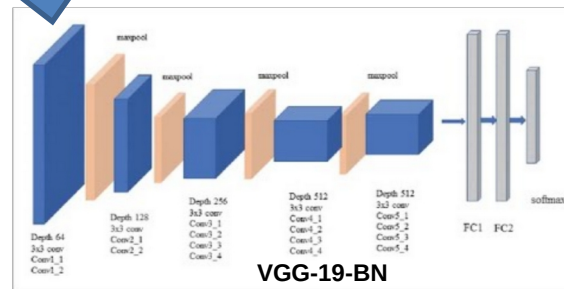


High performance

Dataset consisting of **a small** amount of data



Training



High performance

Reuse

2. Why is use Transfer Learning?

High-quality dataset consisting of **a large** amount of data

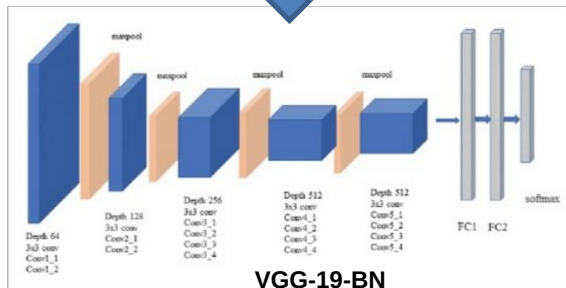
ImageNet Challenge

IMAGENET

- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.

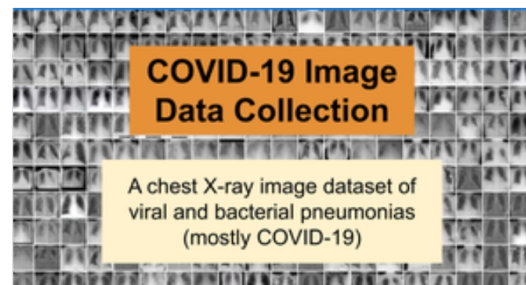


Training

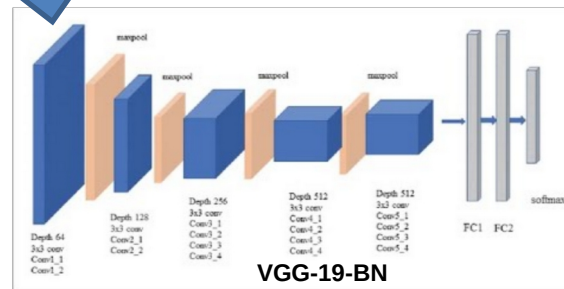


High performance

Dataset consisting of **a small** amount of data



Training



High performance

1. Not needing a lot of data

Reuse

2. Why is use Transfer Learning?

High-quality dataset consisting of **a large** amount of data

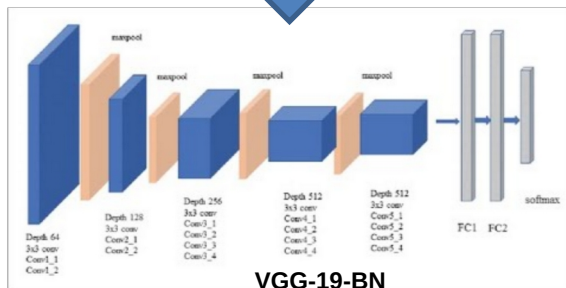
ImageNet Challenge

IMAGENET

- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.

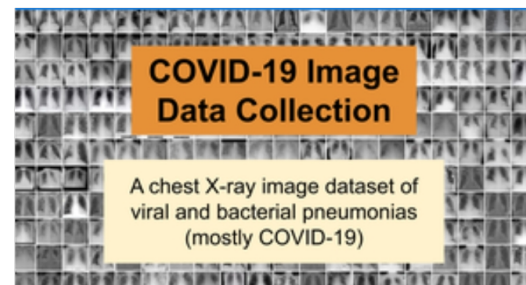


Training

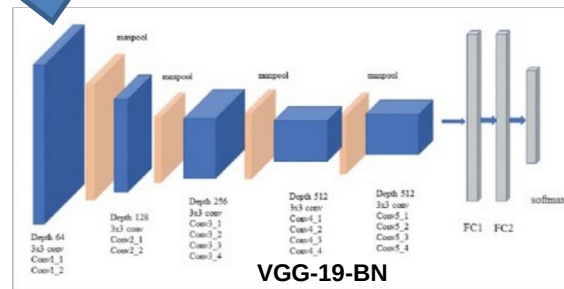


High performance

Dataset consisting of **a small** amount of data



Training



High performance

1. Not needing a lot of data

Reuse
2. Saving training time

2. Why is use Transfer Learning?

High-quality dataset consisting of **a large** amount of data

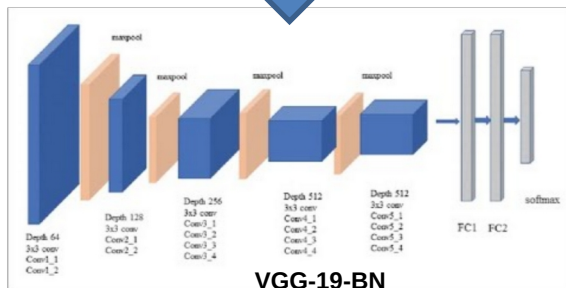
ImageNet Challenge

IMAGENET

- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.

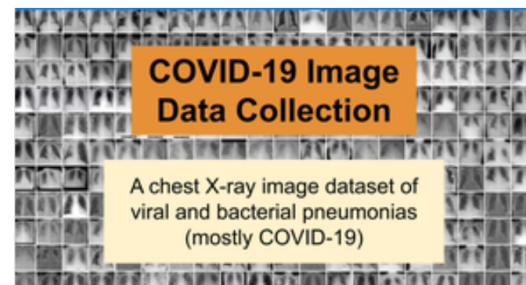


Training

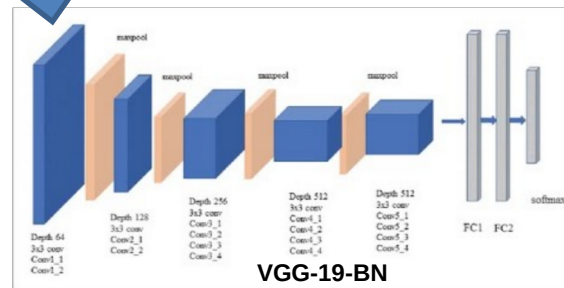


High performance

Dataset consisting of **a small** amount of data



Training



High performance

1. Not needing a lot of data

Reuse
2. Saving training time

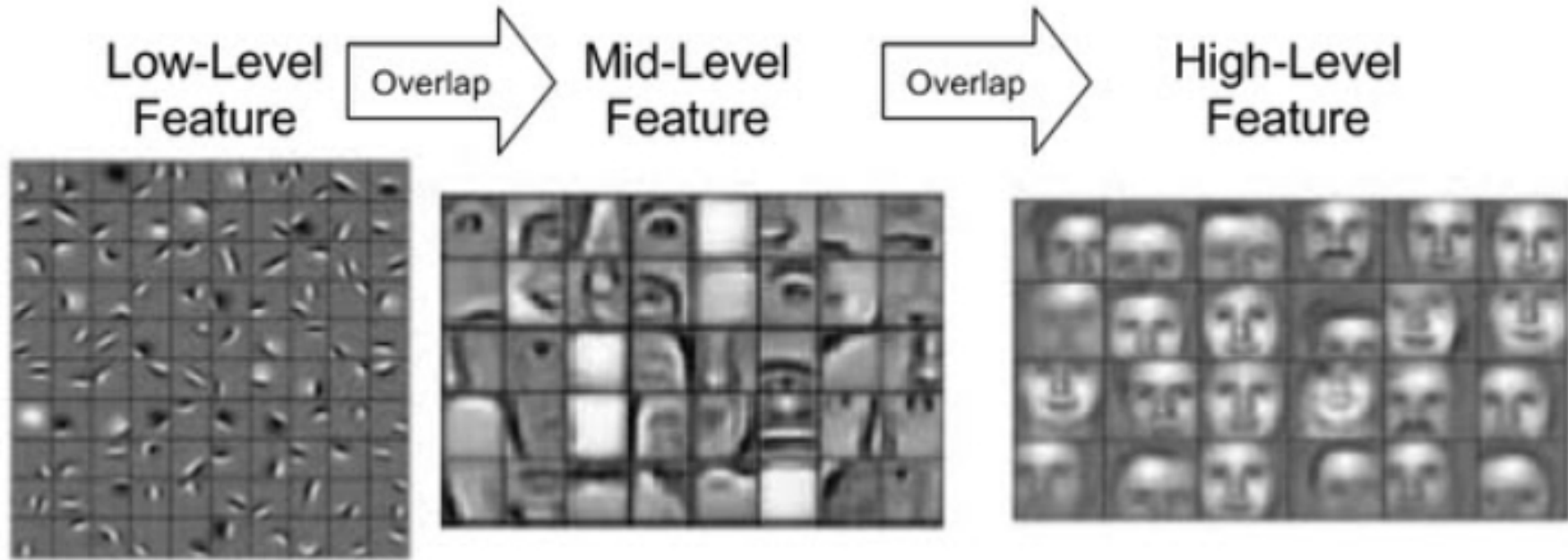
3. Better performance

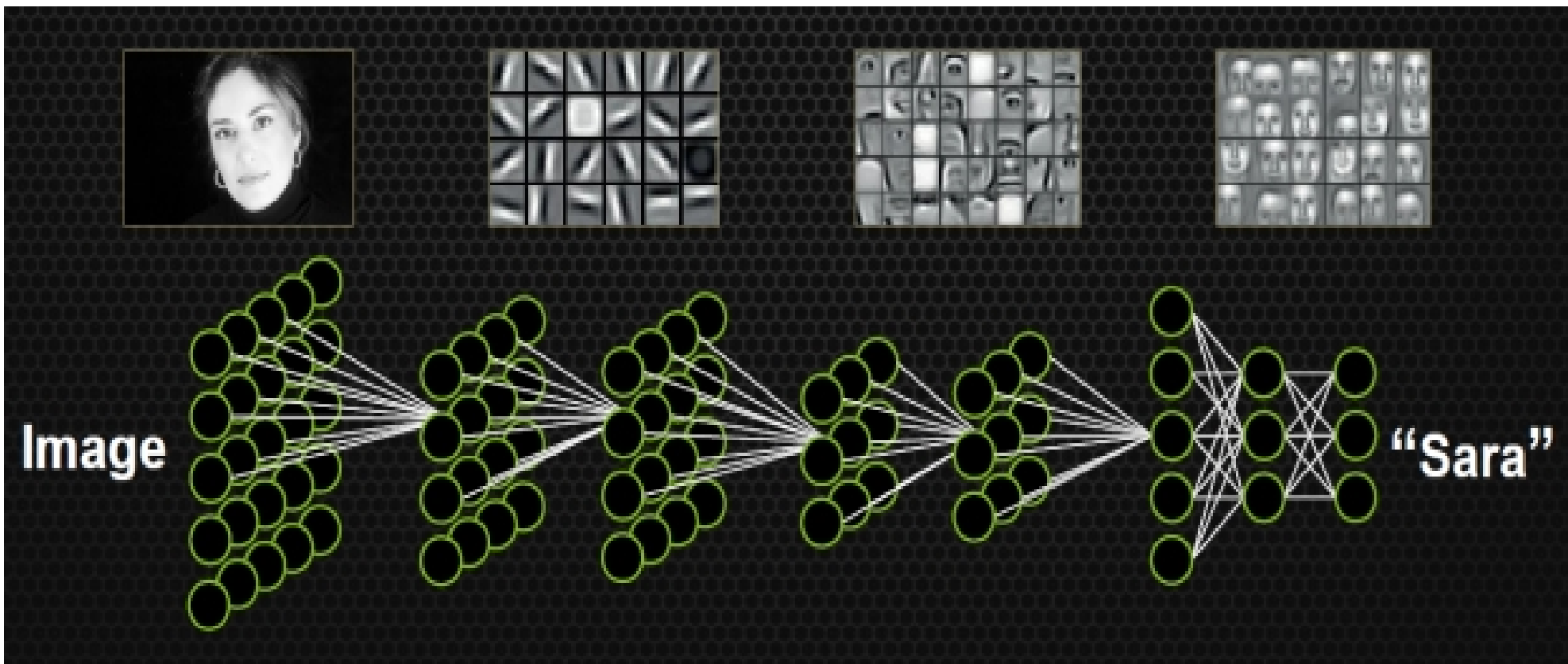
Outline

1. What is Transfer Learning?
2. Why is use Transfer Learning?
3. How transferable are features?

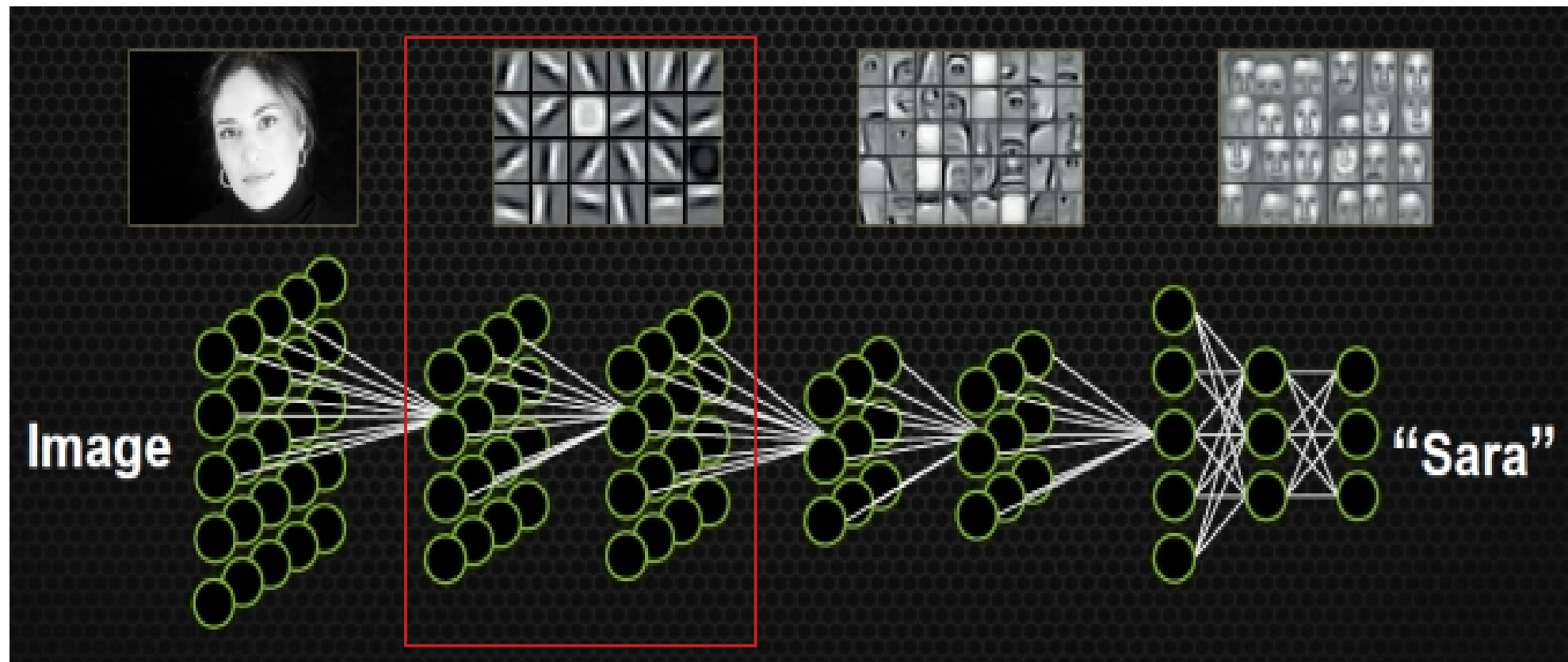
3. How transferable are features?

Feature Map in Convolutional Neural Networks (CNN)

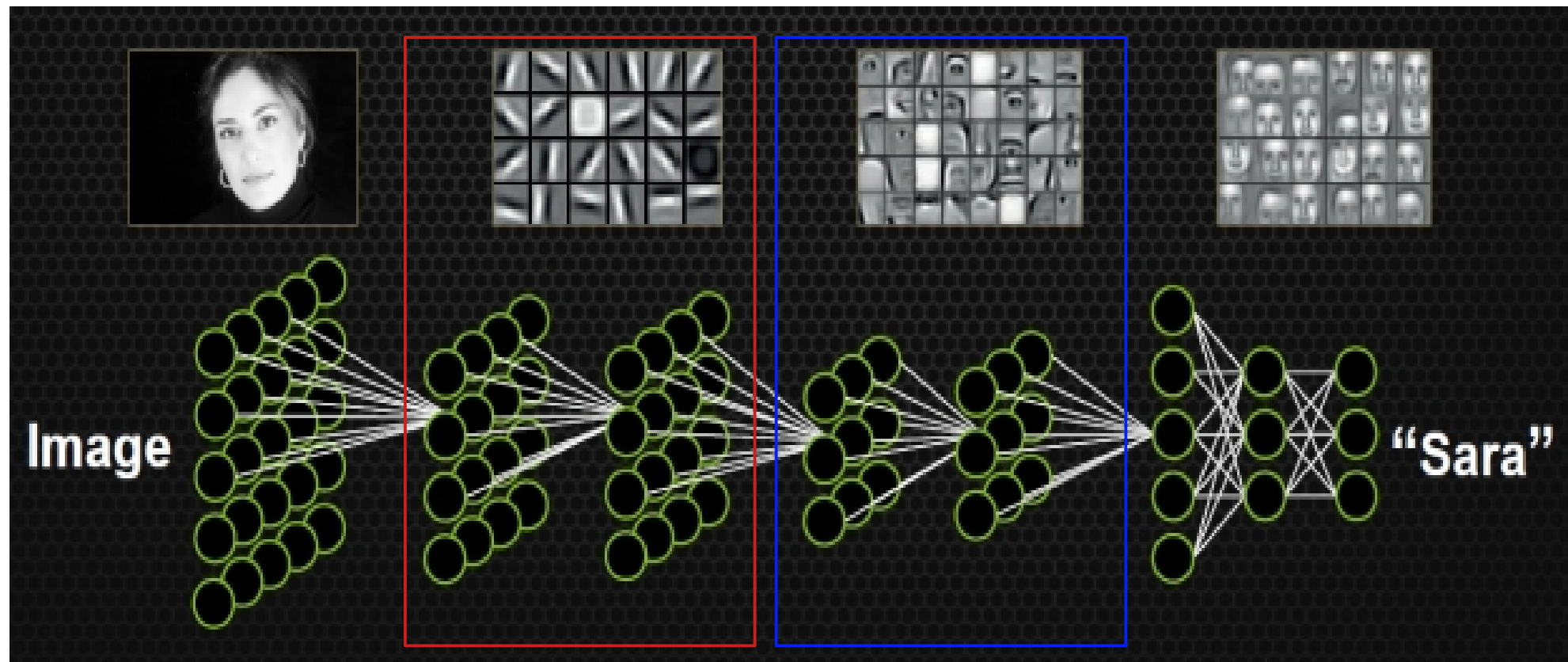




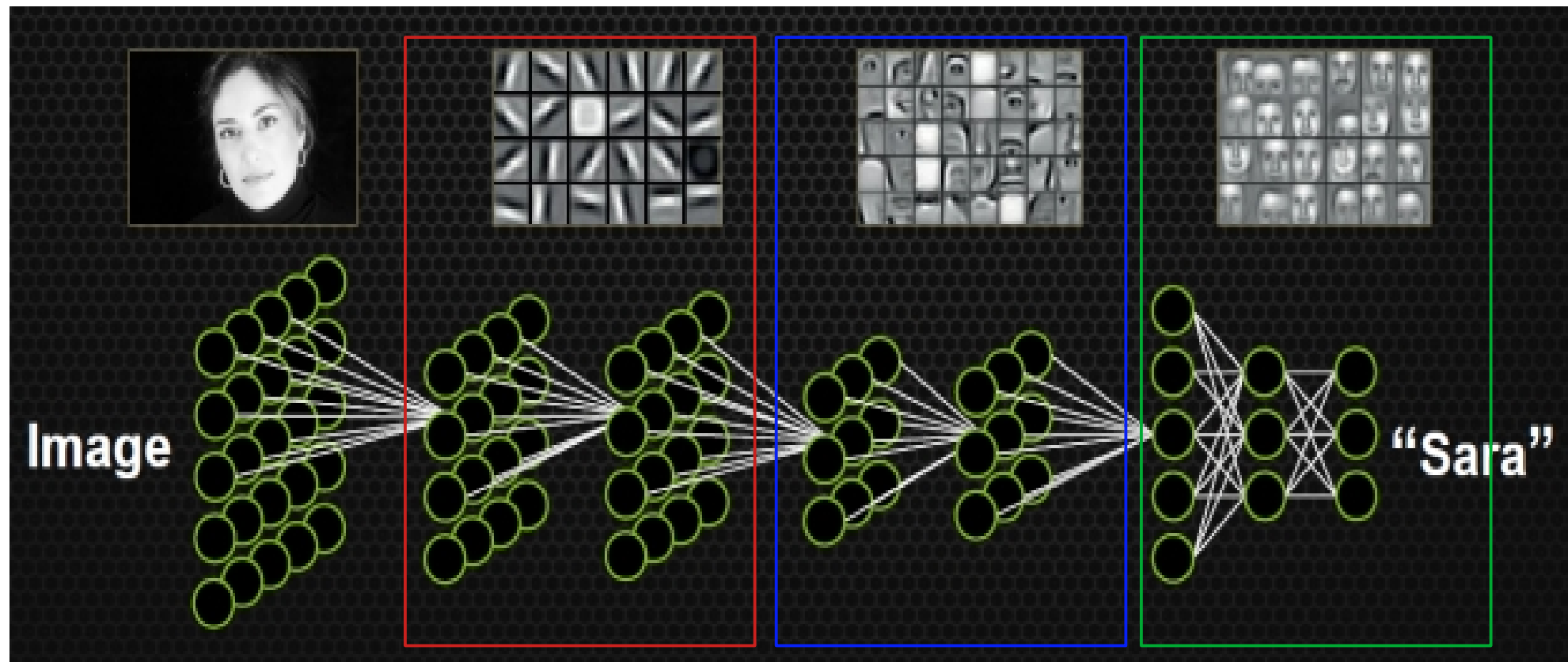
Detect edges in the earlier layers



Detect edges in the earlier layers, shapes in the middle layer



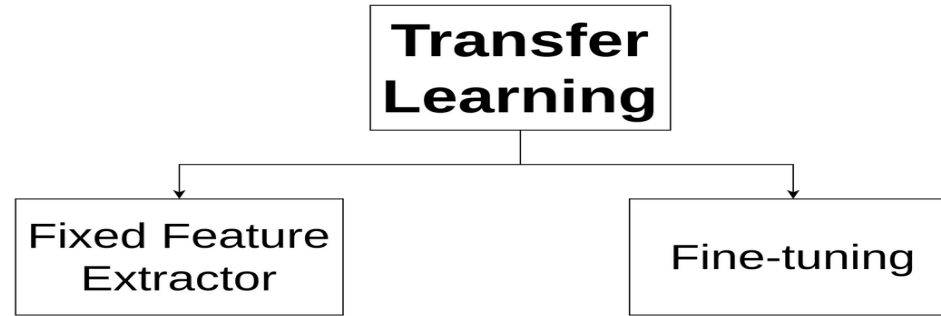
Detect edges in the earlier layers, shapes in the middle layer and some task-specific features in the later layers



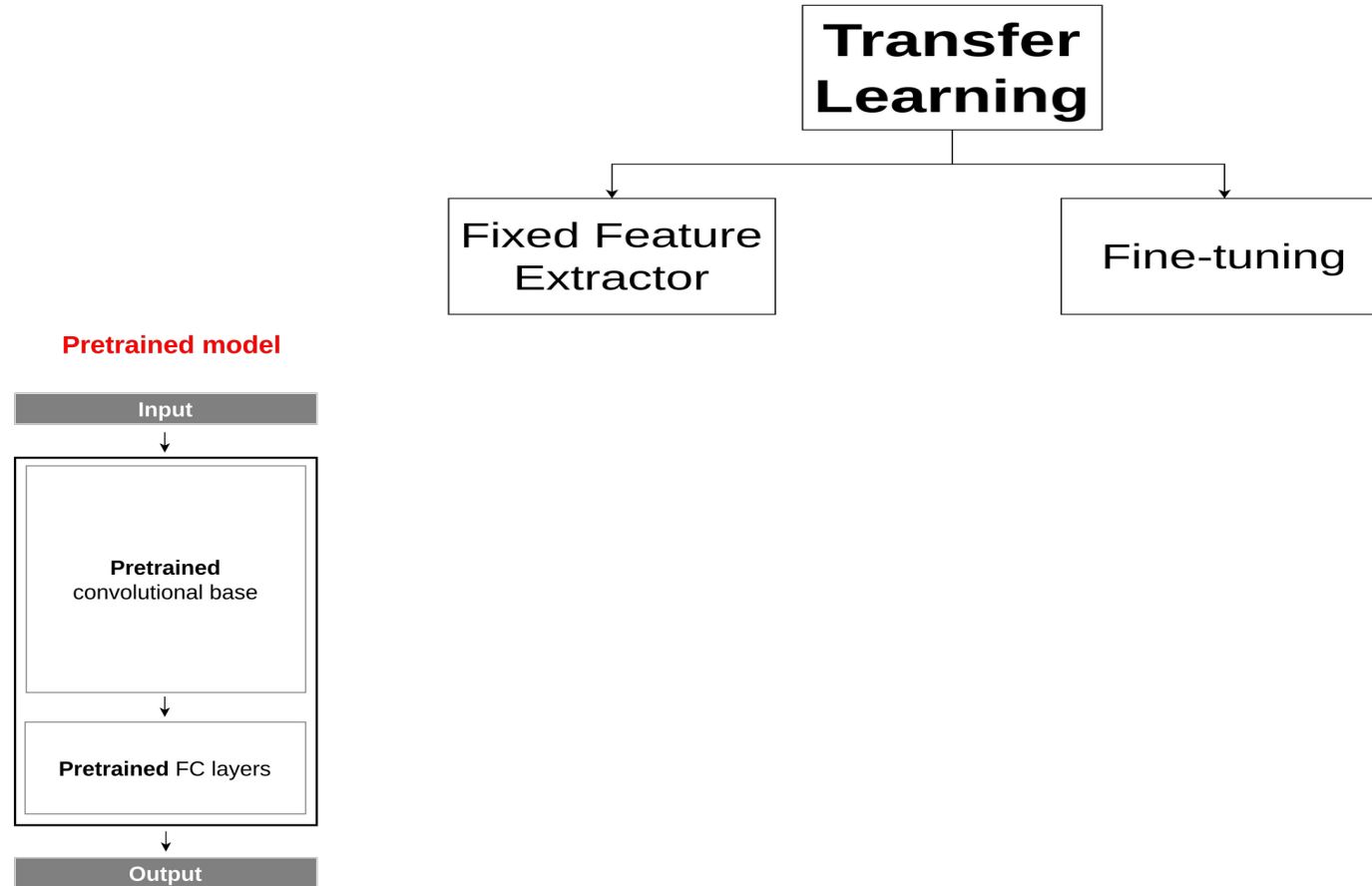
Outline

1. What is Transfer Learning?
2. Why is use Transfer Learning?
3. How transferable are features?
4. Approaches to Transfer Learning

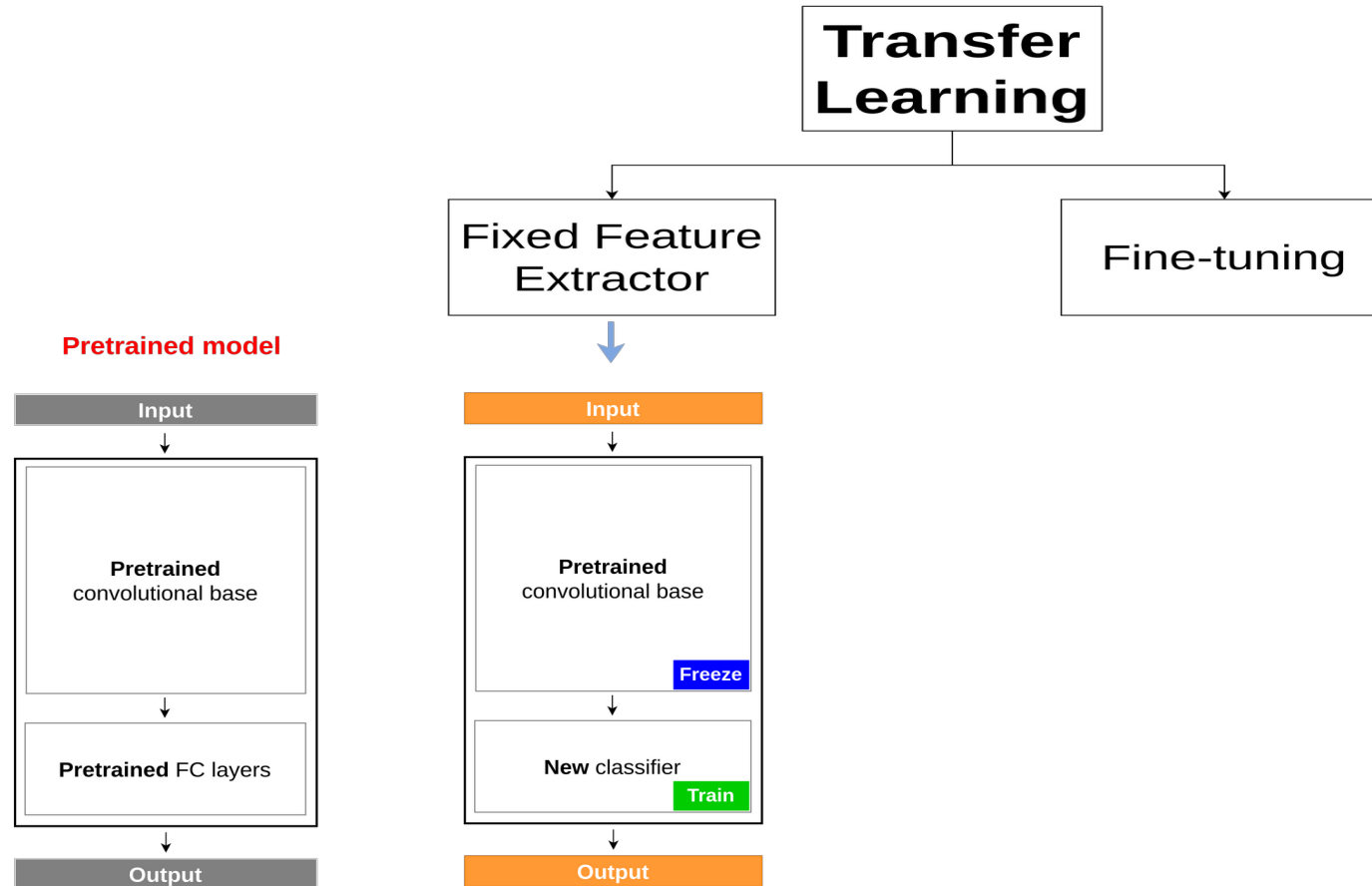
4. Approaches to Transfer Learning



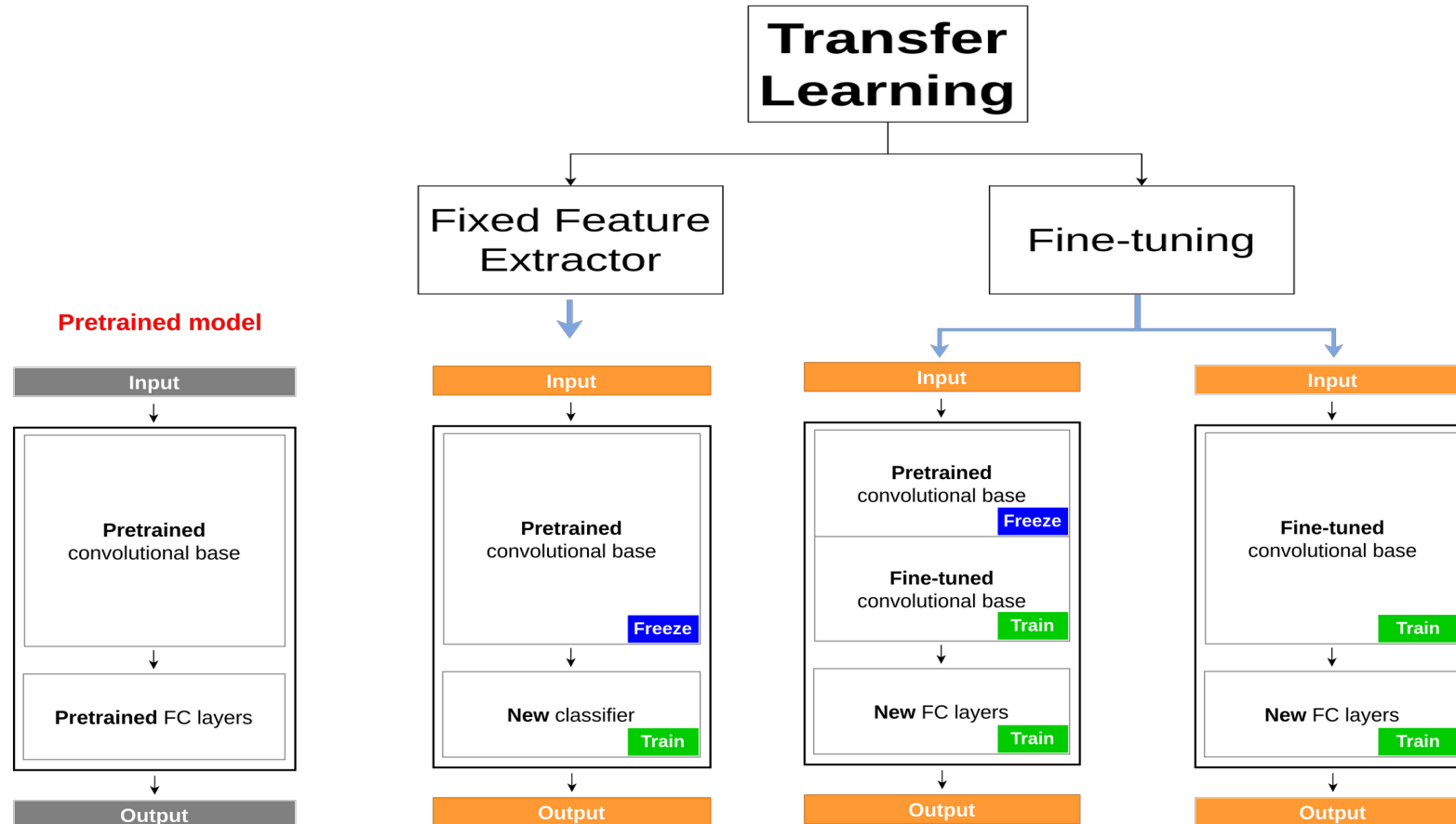
4. Approaches to Transfer Learning



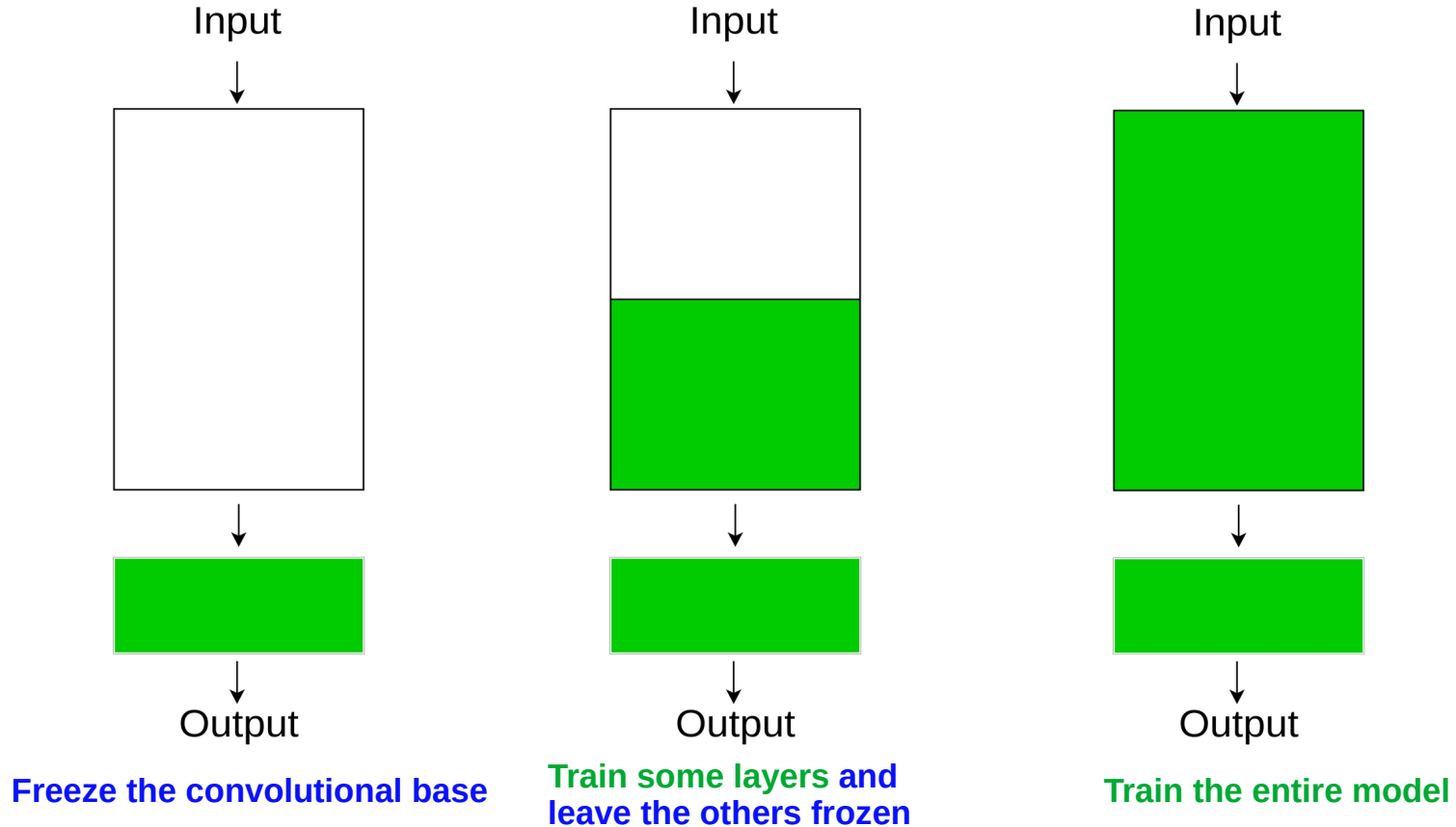
4. Approaches to Transfer Learning



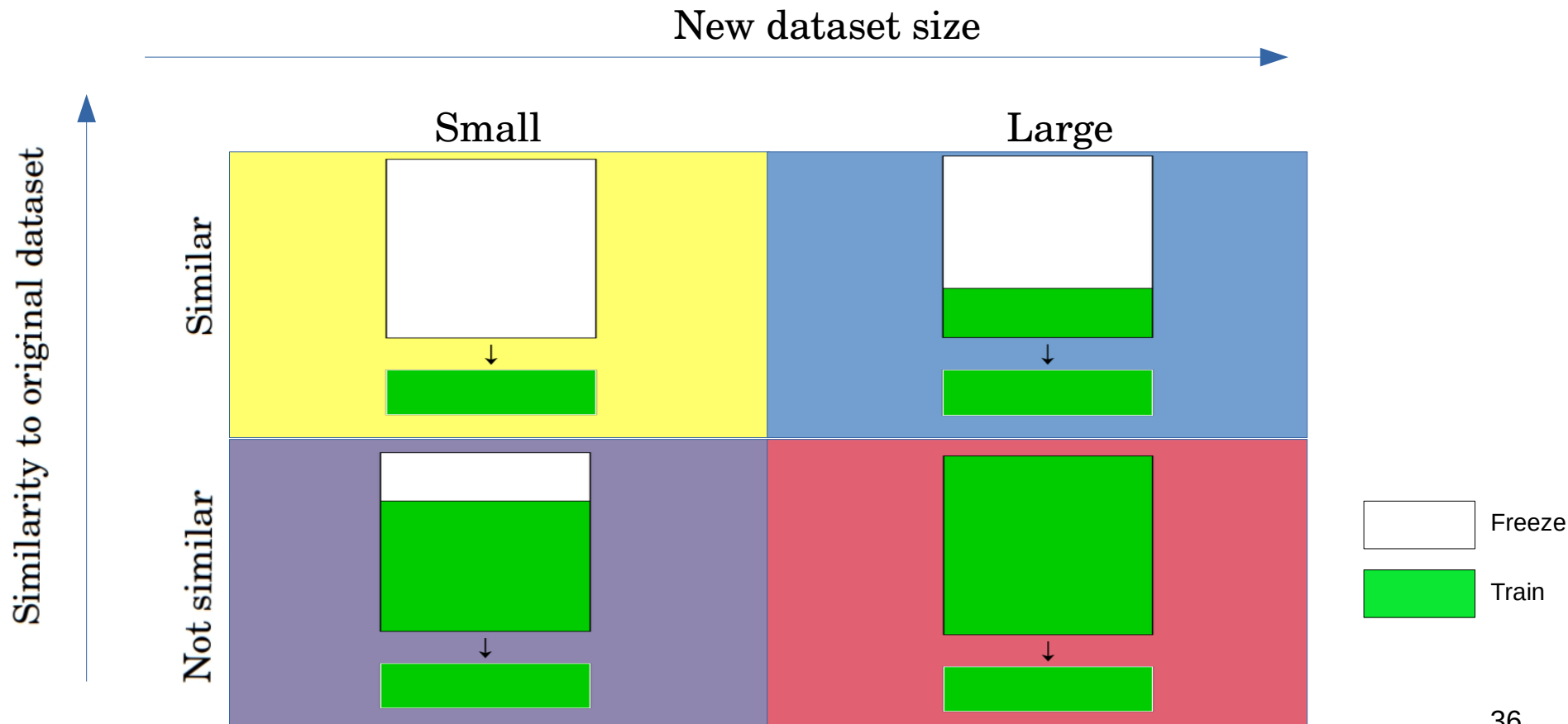
4. Approaches to Transfer Learning



4. Approaches to Transfer Learning



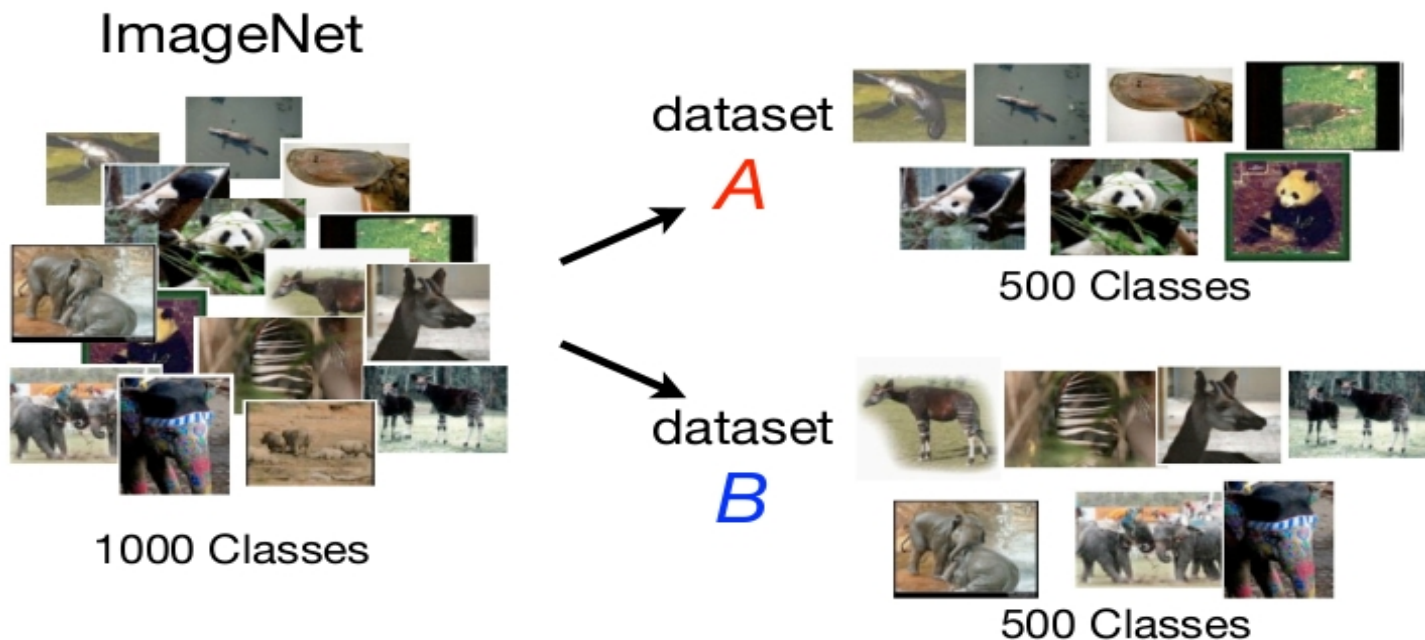
4. Approaches to Transfer Learning



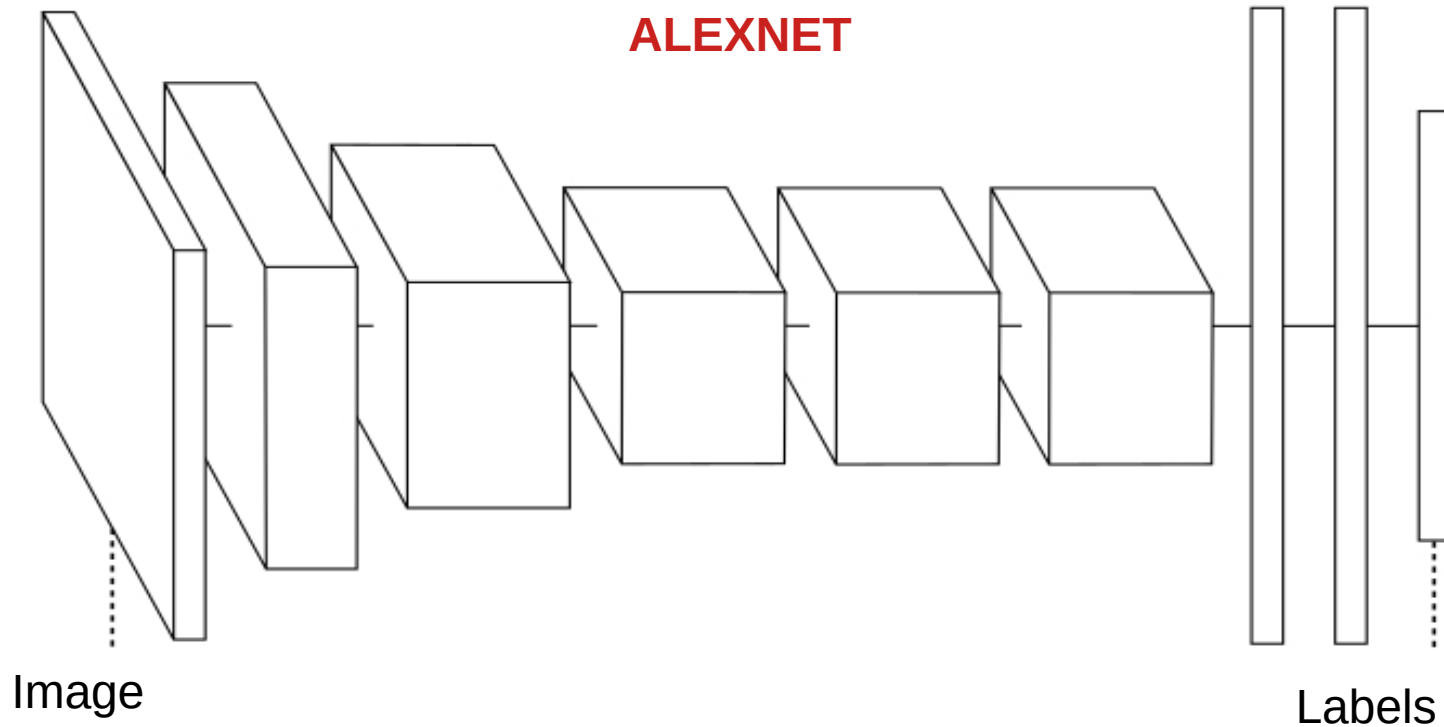
4. Approaches to Transfer Learning



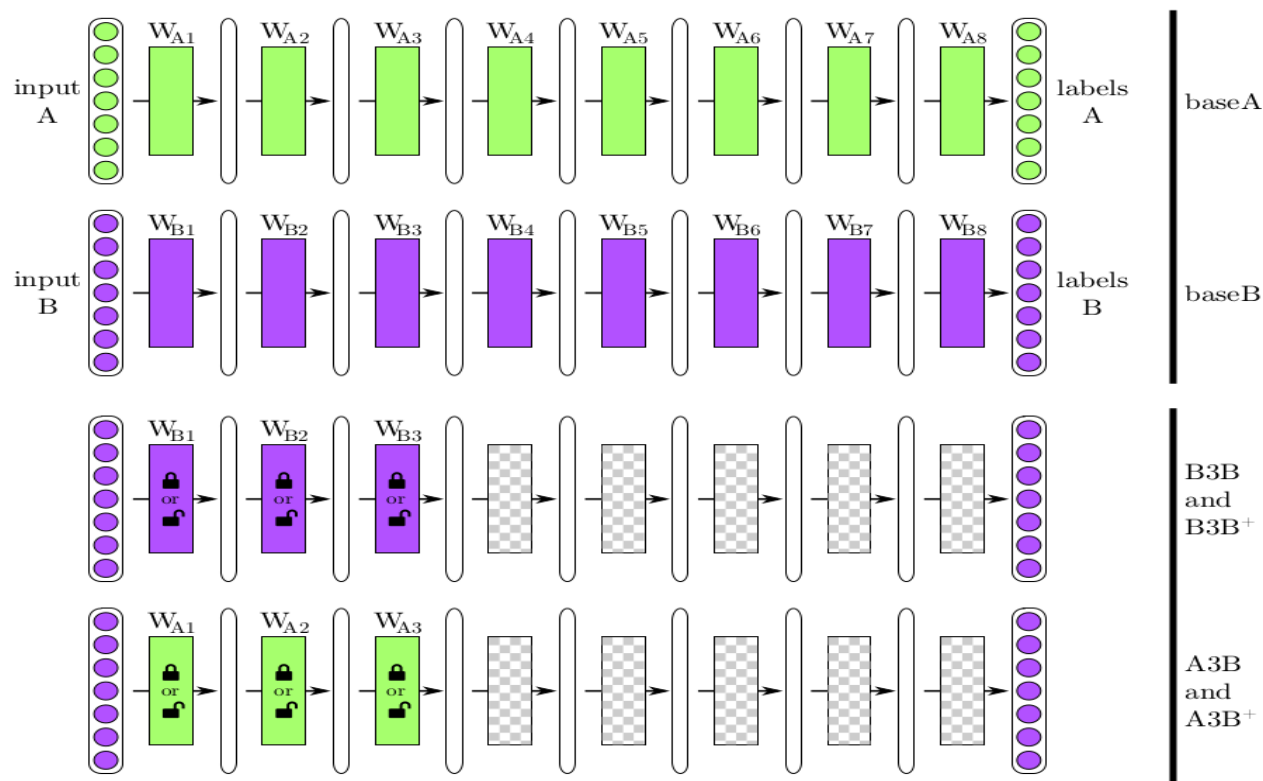
4. Approaches to Transfer Learning



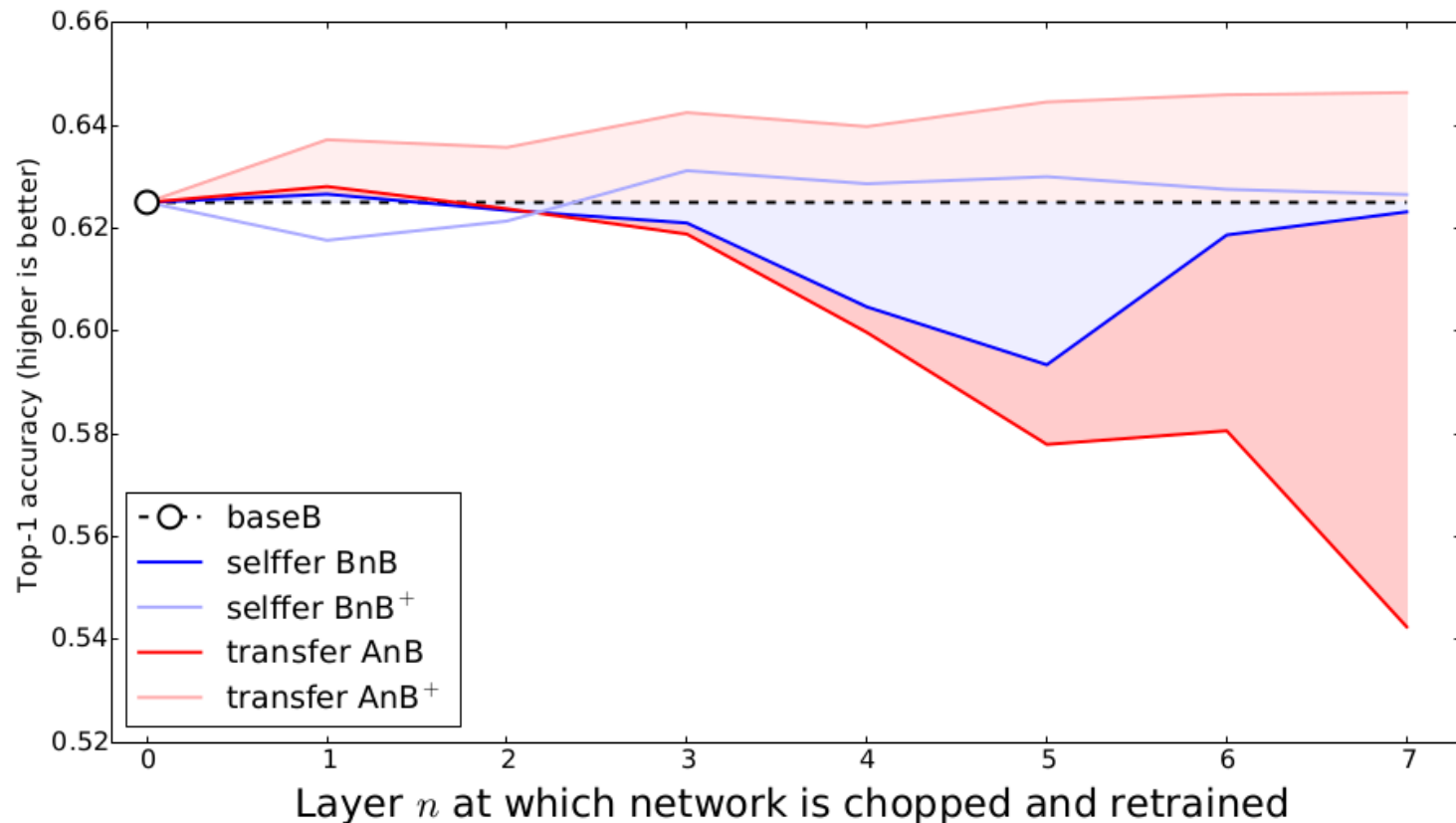
4. Approaches to Transfer Learning



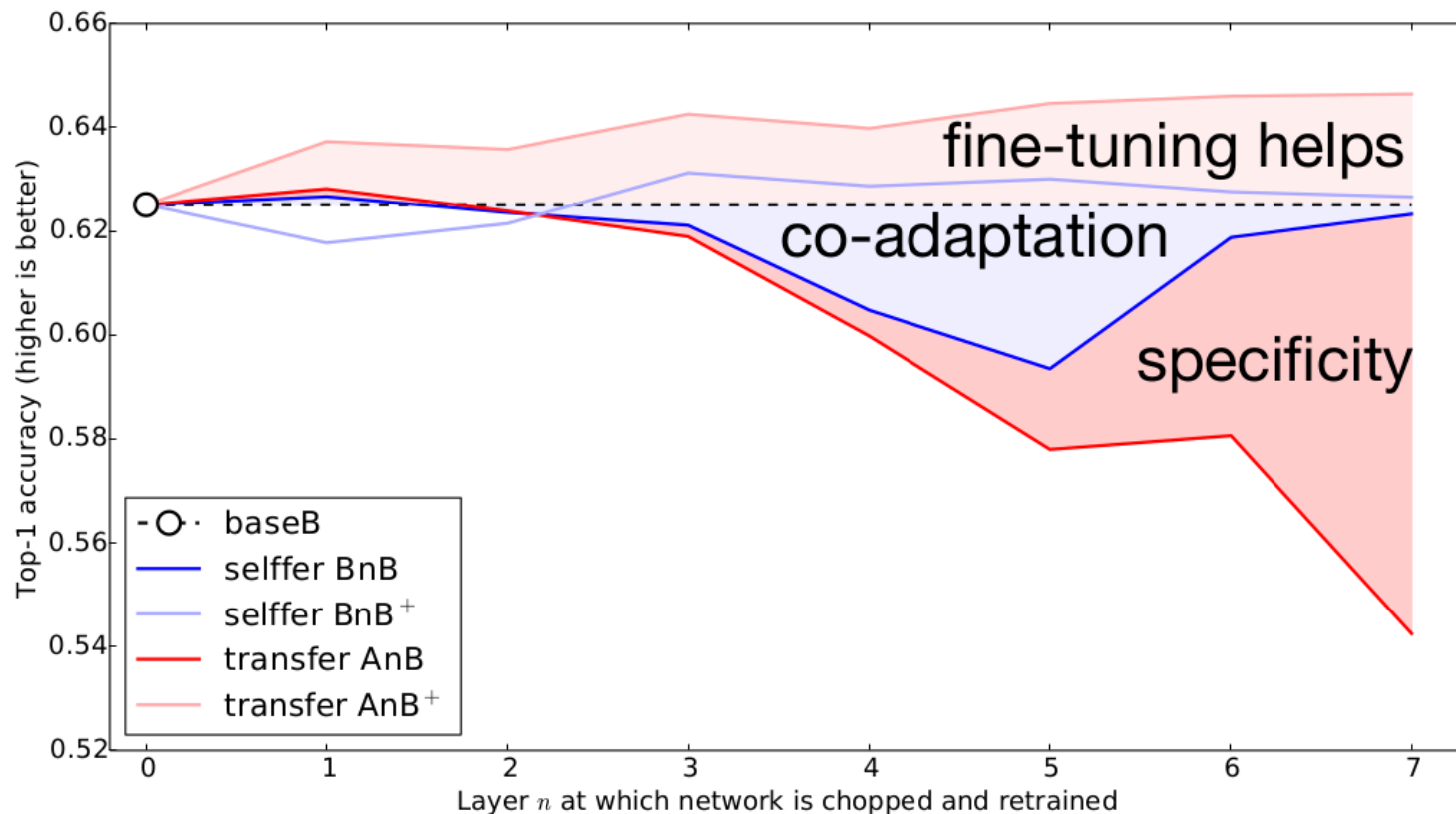
4. Approaches to Transfer Learning



4. Approaches to Transfer Learning



4. Approaches to Transfer Learning



4. Approaches to Transfer Learning

ImageNet has many related categories...

Dataset A: random

gecko

fire truck

baseball

panther

rabbit

gorilla

Dataset B: random

garbage truck

toucan

radiator

binoculars

lion

bookshop

4. Approaches to Transfer Learning

ImageNet has many related categories...

Dataset A: man-made

garbage truck

fire truck

radiator

baseball

binoculars

bookshop

Dataset B: natural

gorilla

gecko

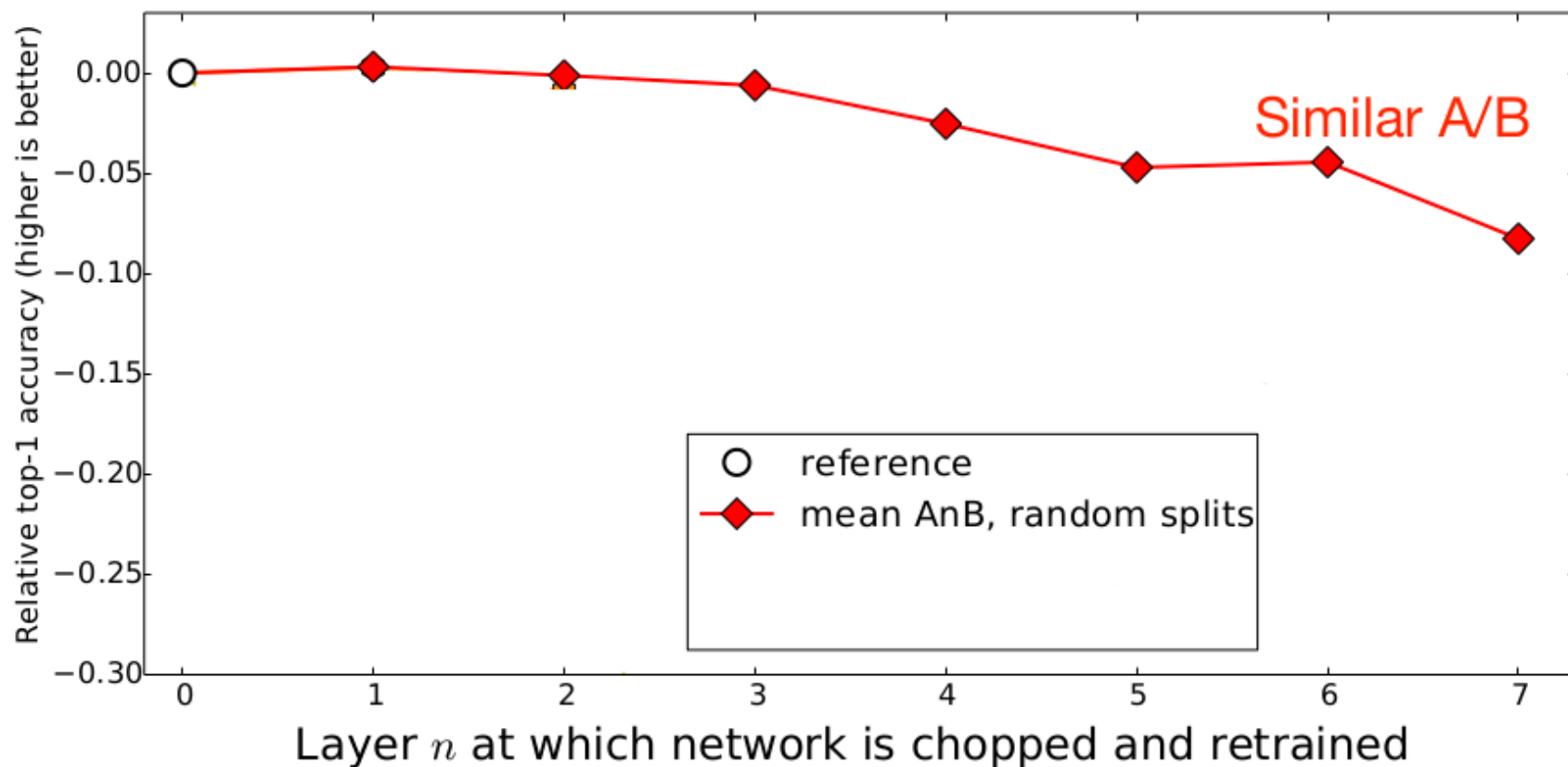
toucan

rabbit

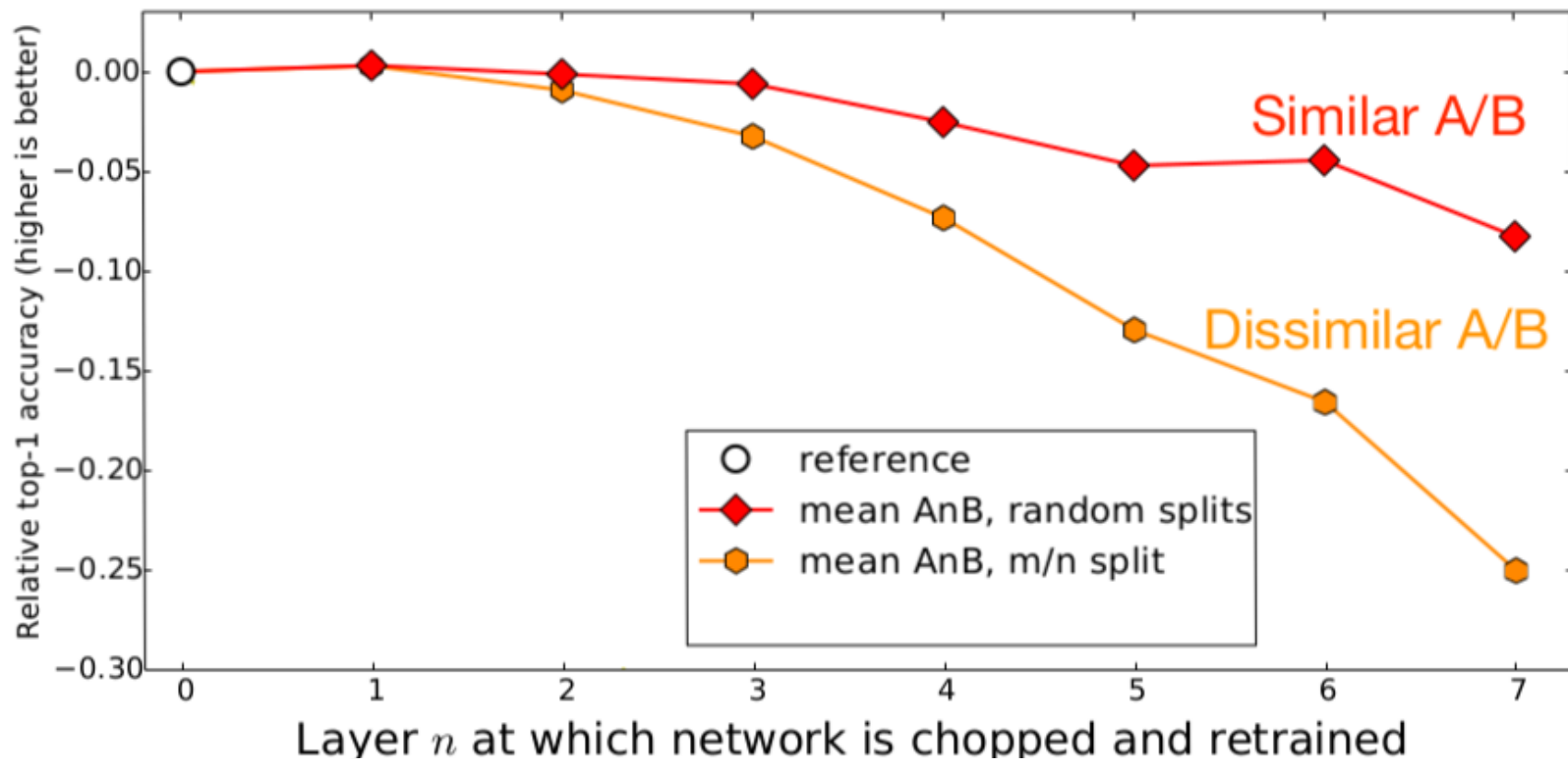
panther

lion

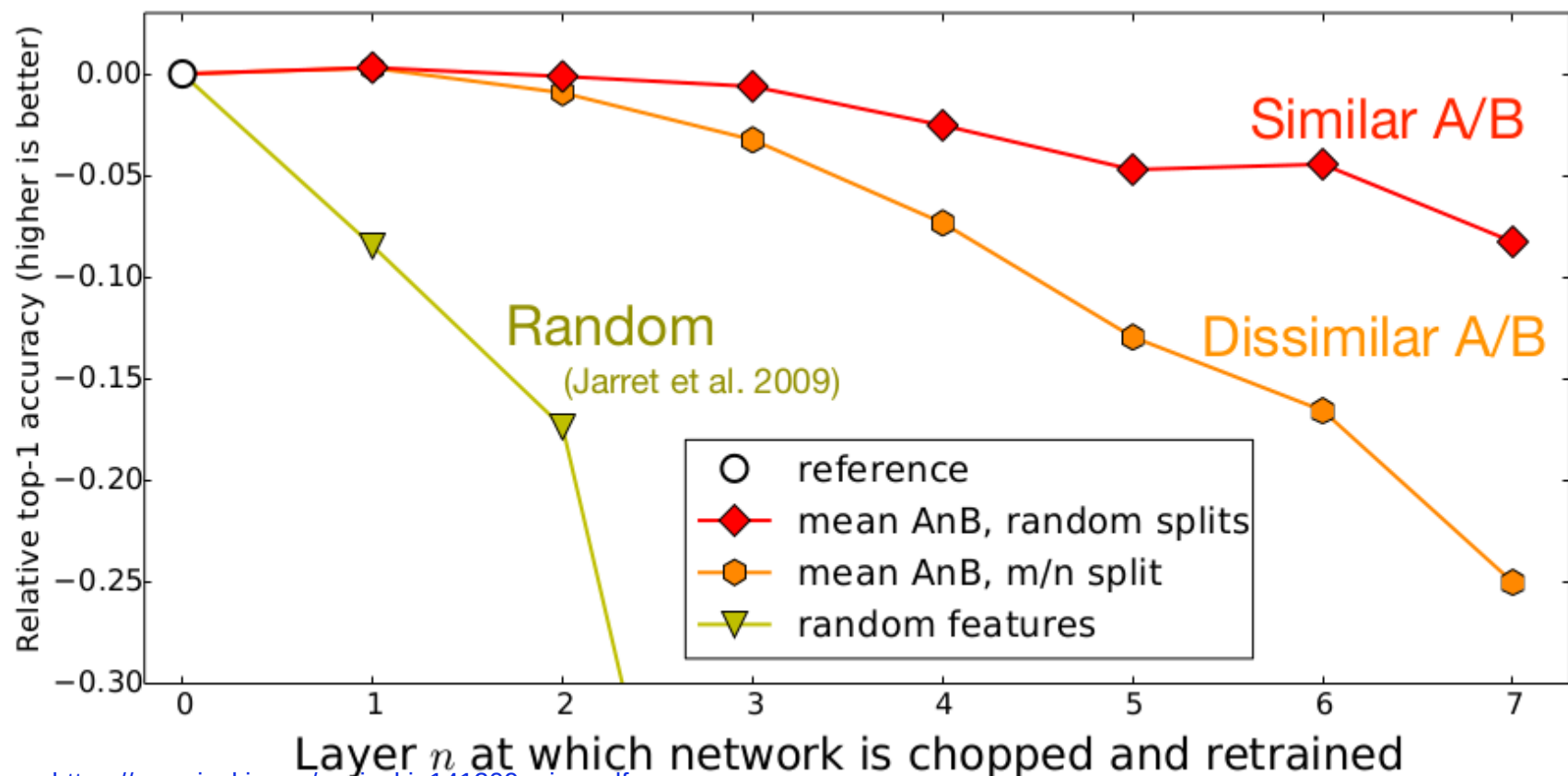
4. Approaches to Transfer Learning



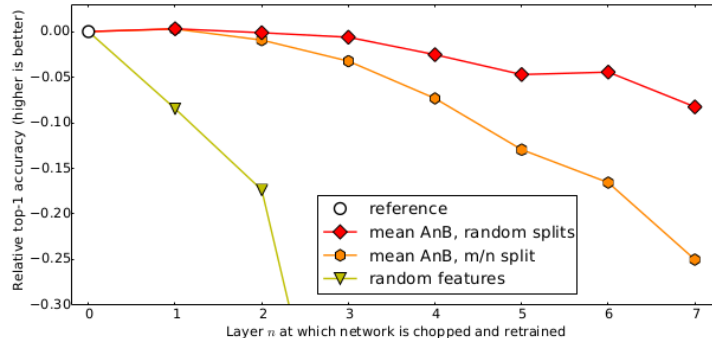
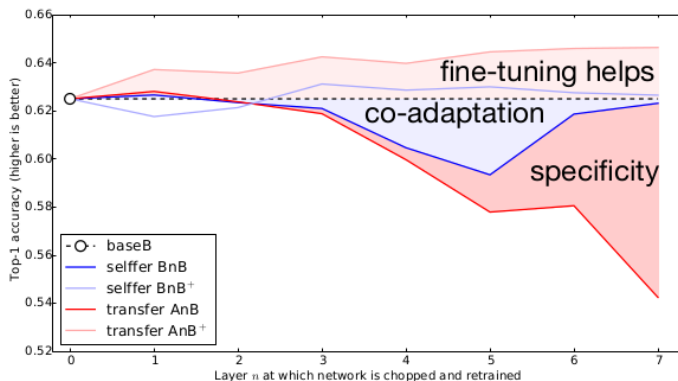
4. Approaches to Transfer Learning



4. Approaches to Transfer Learning



4. Approaches to Transfer Learning



- Measure *general* to *specific* transition layer by layer
- Transferability governed by:
 - lost co-adaptations
 - specificity
 - difference between base and target dataset
- Fine-tuning helps even on large target dataset

Outline

1. What is Transfer Learning?
2. Why is use Transfer Learning?
3. How transferable are features?
4. Approaches to Transfer Learning
5. Experimental

5. Experimental

Dataset: <https://www.kaggle.com/andyczhao/covidx-cxr2?select=train.txt>

Data Structure

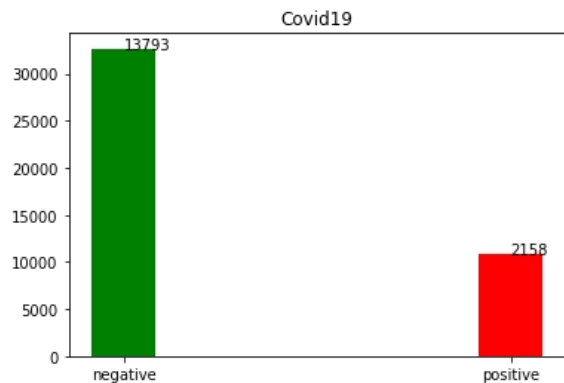
patient id	file_name	class	source
5	ARDSSevere.png	negative	cohen
25	acute-respiratory-distress-syndrome-ards-1.jpg	negative	cohen
26	acute-respiratory-distress-syndrome-ards.jpg	negative	cohen
27	ards-secondary-to-tiger-snake-bite.png	negative	cohen
28	pneumocystis-pneumonia-2-PA.png	negative	cohen

5. Experimental

Dataset: <https://www.kaggle.com/andyczhao/covidx-cxr2?select=train.txt>

Data Structure

patient id	file_name	class	source
5	ARDSSevere.png	negative	cohen
25	acute-respiratory-distress-syndrome-ards-1.jpg	negative	cohen
26	acute-respiratory-distress-syndrome-ards.jpg	negative	cohen
27	ards-secondary-to-tiger-snake-bite.png	negative	cohen
28	pneumocystis-pneumonia-2-PA.png	negative	cohen

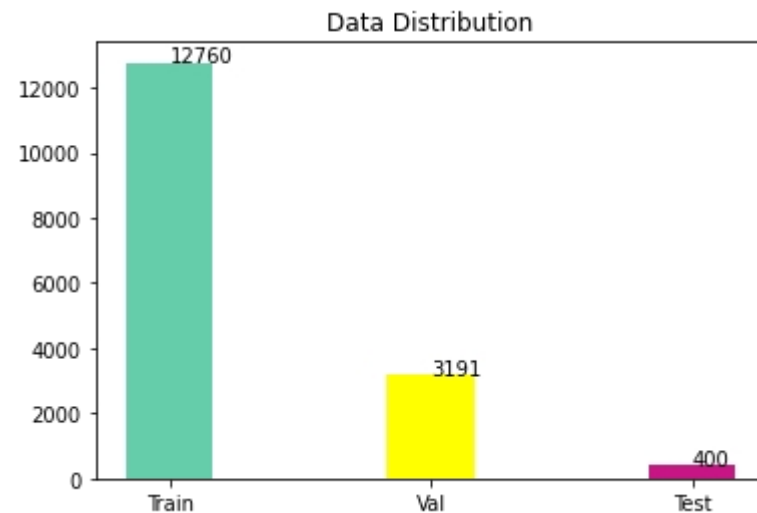
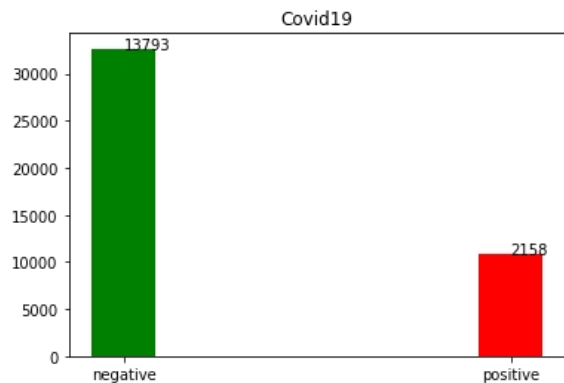


5. Experimental

Dataset: <https://www.kaggle.com/andyczhao/covidx-cxr2?select=train.txt>

Data Structure

patient id	file_name	class	source
5	ARDSSevere.png	negative	cohen
25	acute-respiratory-distress-syndrome-ards-1.jpg	negative	cohen
26	acute-respiratory-distress-syndrome-ards.jpg	negative	cohen
27	ards-secondary-to-tiger-snake-bite.png	negative	cohen
28	pneumocystis-pneumonia-2-PA.png	negative	cohen



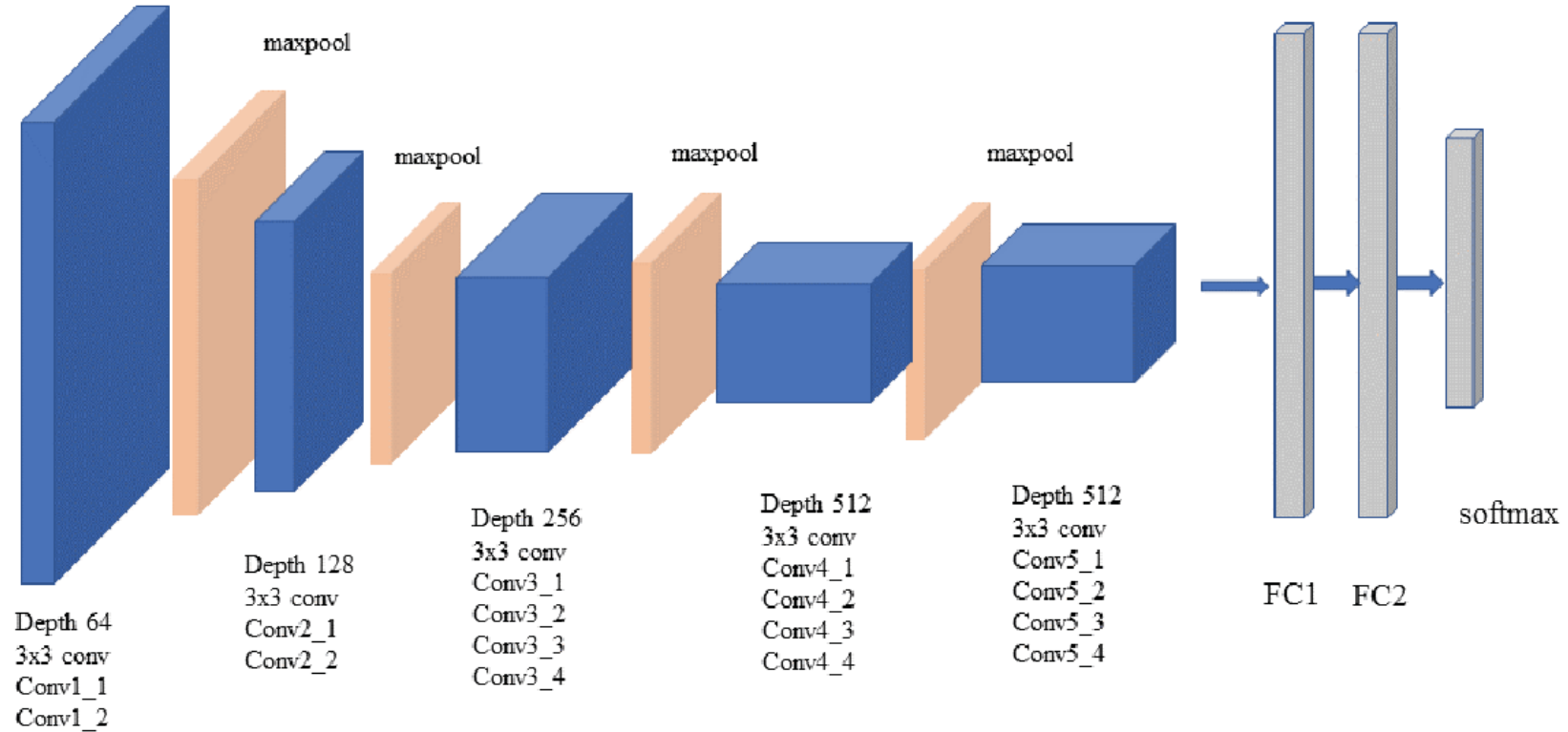
5. Experimental

VGG-NETS

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

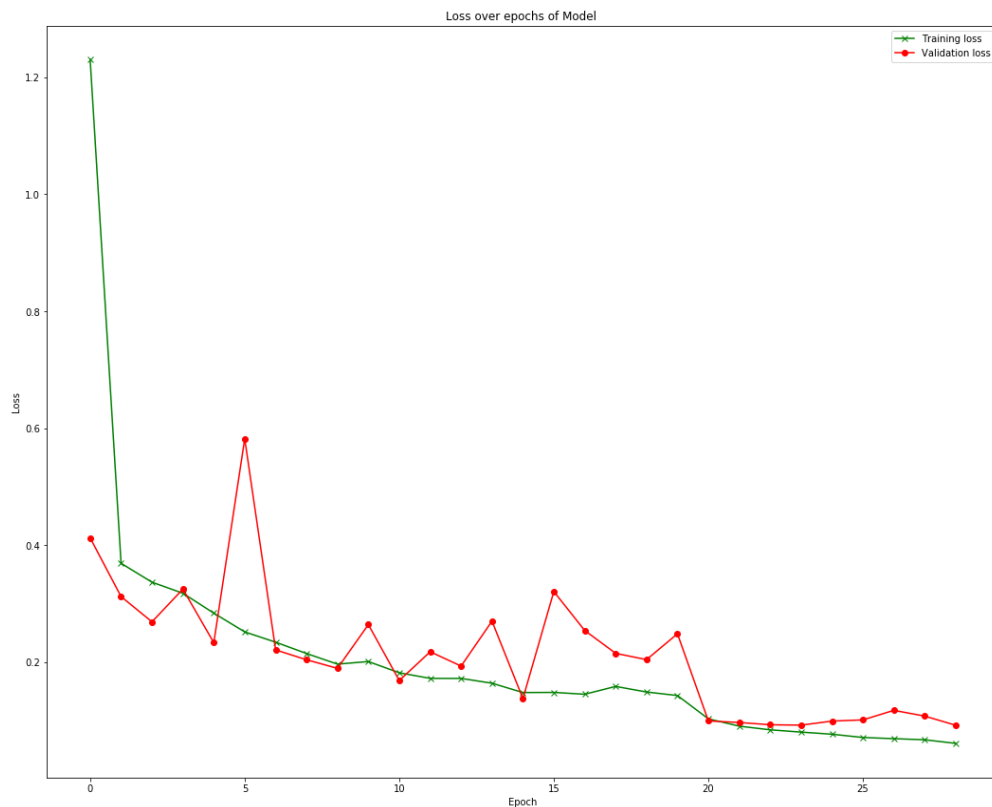
5. Experimental

VGG19

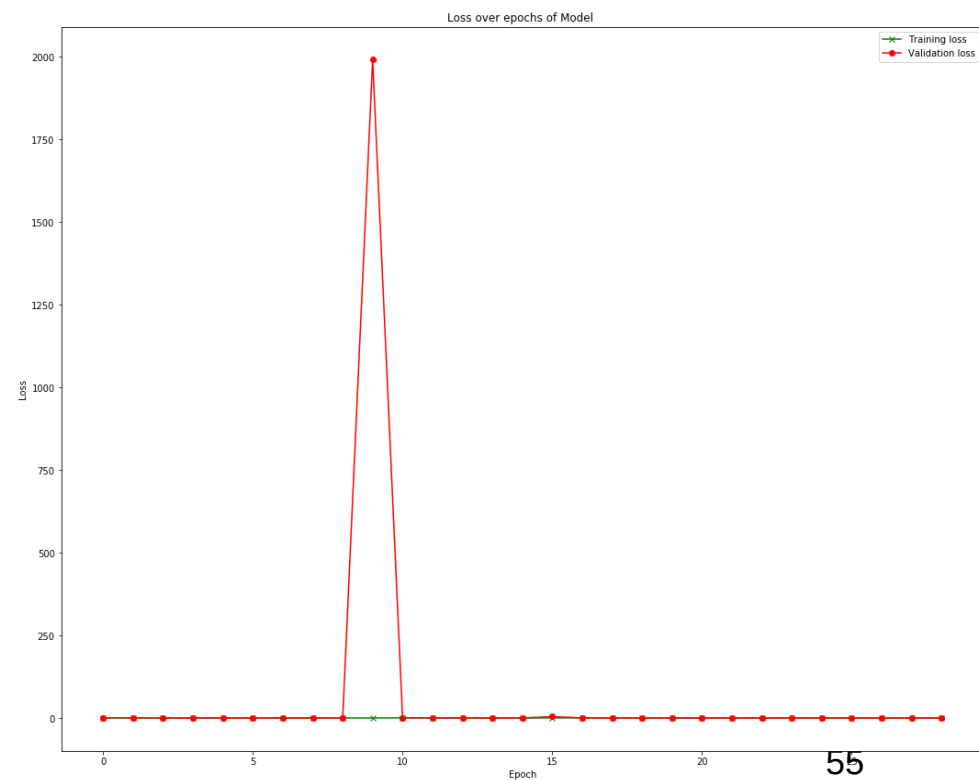


5. Experimental

From the scratch: in 162m 38s

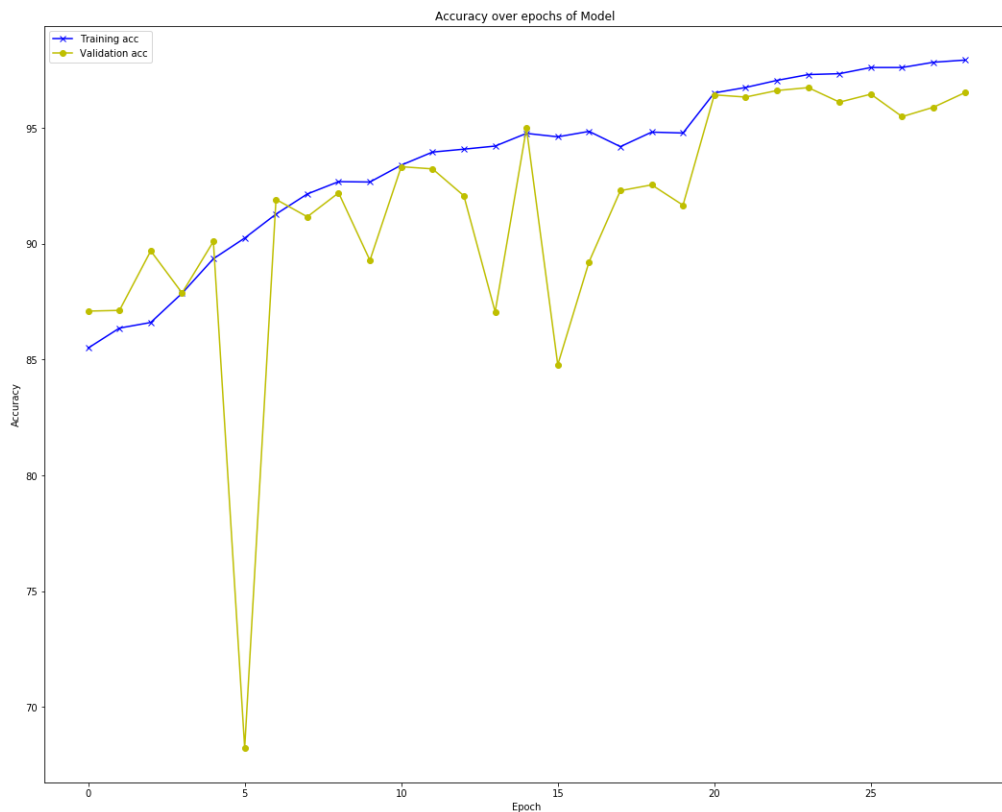


Transfer Learning: in 164m 26s

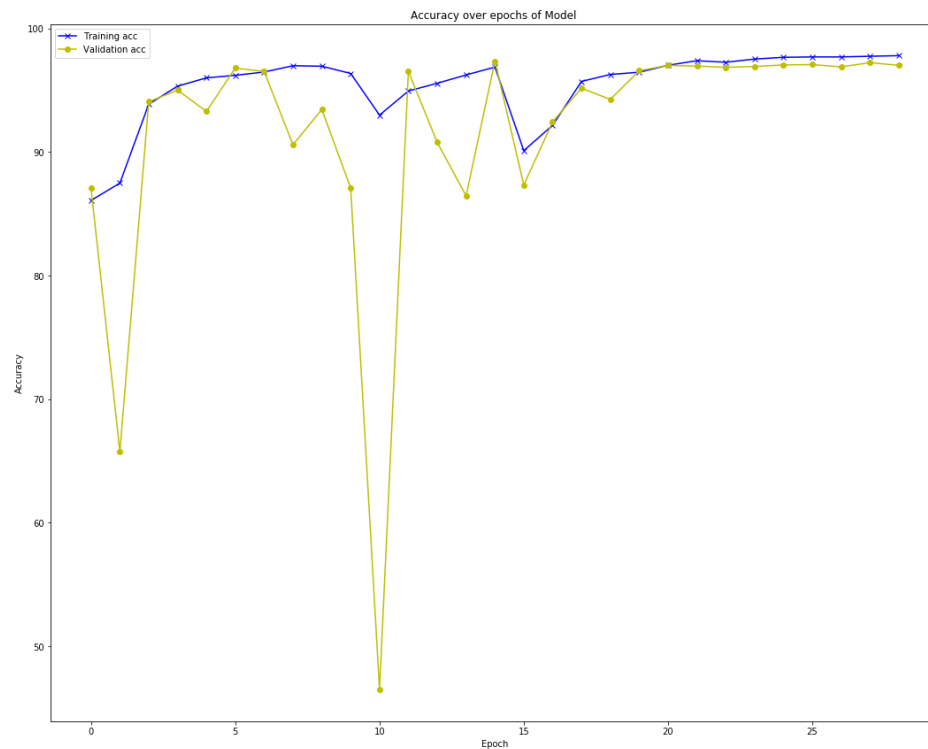


5. Experimental

From the scratch: in 162m 38s

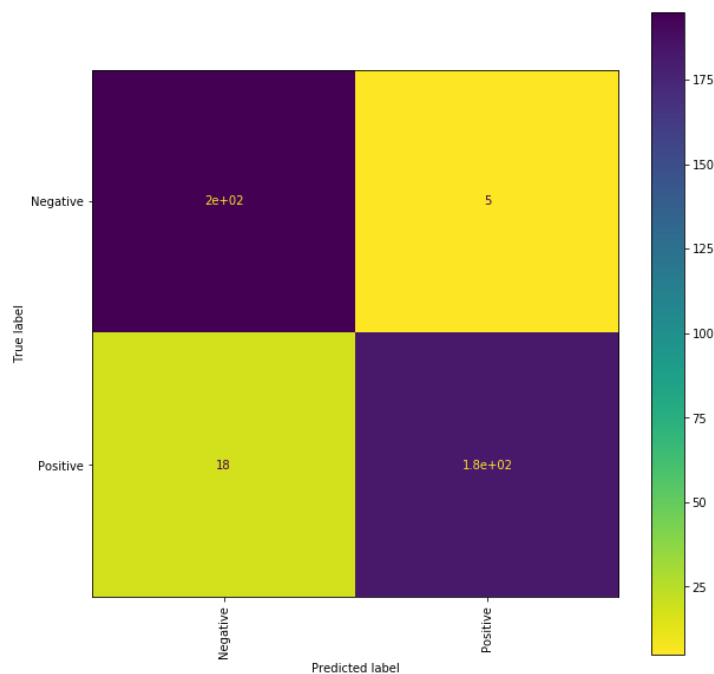


Transfer Learning: in 164m 26s

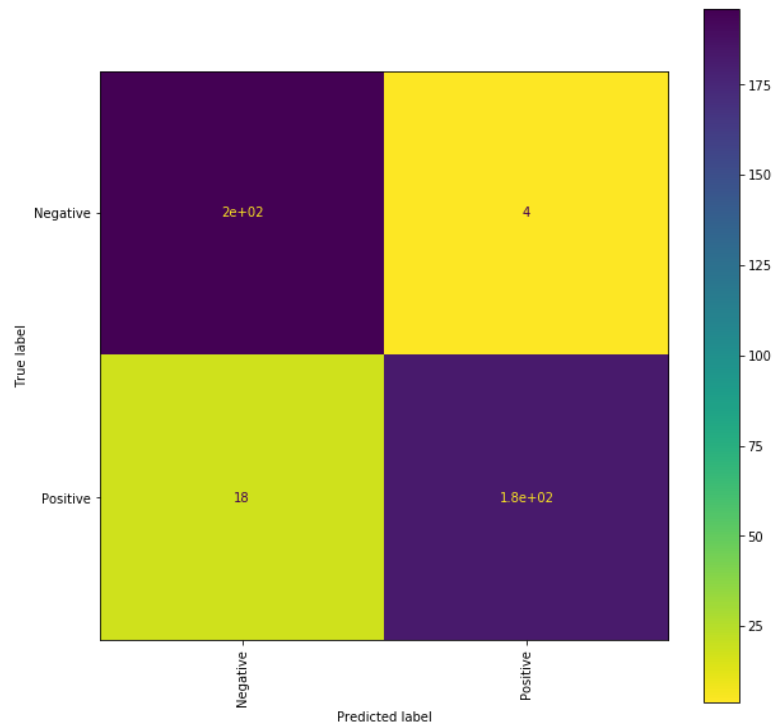


5. Experimental

From the scratch: in 162m 38s



Transfer Learning: in 164m 26s



Outline

1. What is Transfer Learning?
2. Why is use Transfer Learning?
3. How transferable are features?
4. Approaches to Transfer Learning
5. Experimental
6. Conclusions

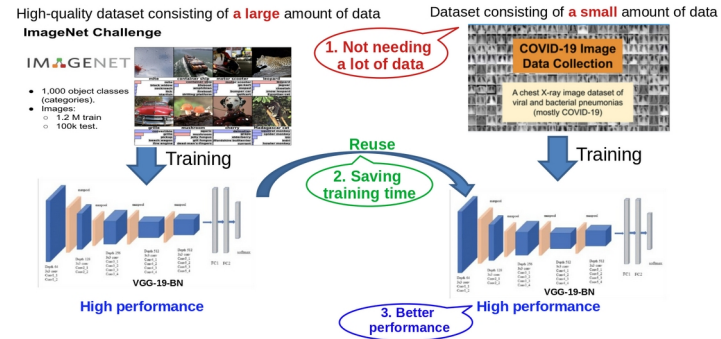
Conclusions

1. What is Transfer Learning?



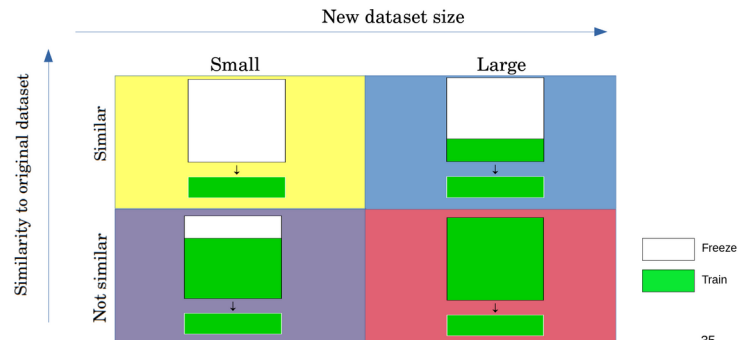
8

2. Why is use Transfer Learning?



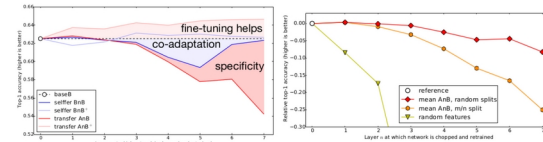
22

4. Approaches to Transfer Learning



35

4. Approaches to Transfer Learning



- Measure *general* to *specific* transition layer by layer
- Transferability governed by:
 - lost co-adaptations
 - specificity
 - difference between base and target dataset
- Fine-tuning helps even on large target dataset

https://s.yosinski.com/yosinski_141209_nips.pdf

47

References

<https://towardsdatascience.com/a-practical-example-in-transfer-learning-with-pytorch-846bb835f2db>
<https://yosinski.com/transfer>
<https://www.analyticsvidhya.com/blog/2019/10/how-to-master-transfer-learning-using-pytorch/>
https://pytorch.org/tutorials/beginner/transfer_learning_tutorial.html
<https://hackernoon.com/transfer-learning-approaches-and-empirical-observations-efeff9dfeca6>
<https://towardsdatascience.com/transfer-learning-from-pre-trained-models-f2393f124751>
https://pytorch.org/tutorials/beginner/finetuning_torchvision_models_tutorial.html



**Thank you for
Watching**