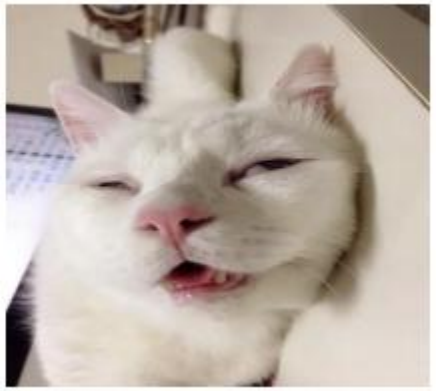
TRANSFER LEARNING

Transfer Learning

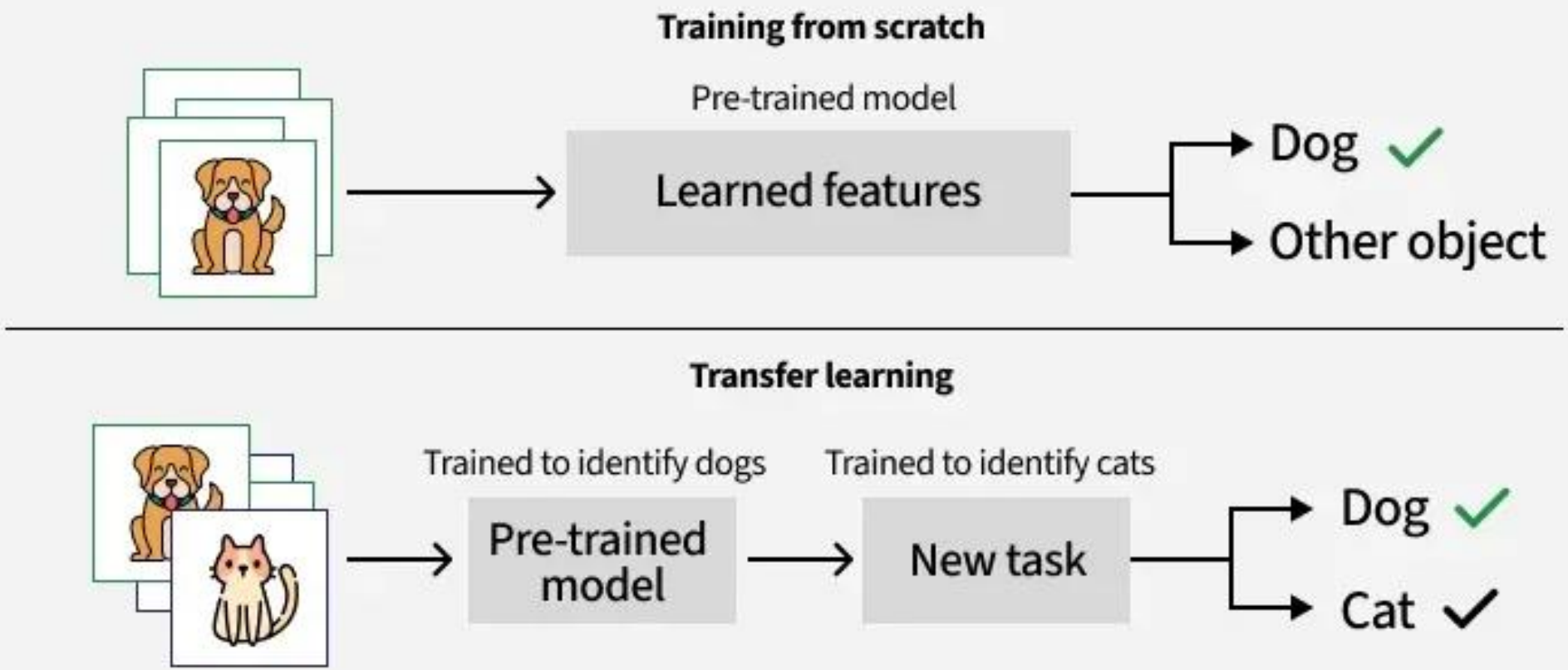


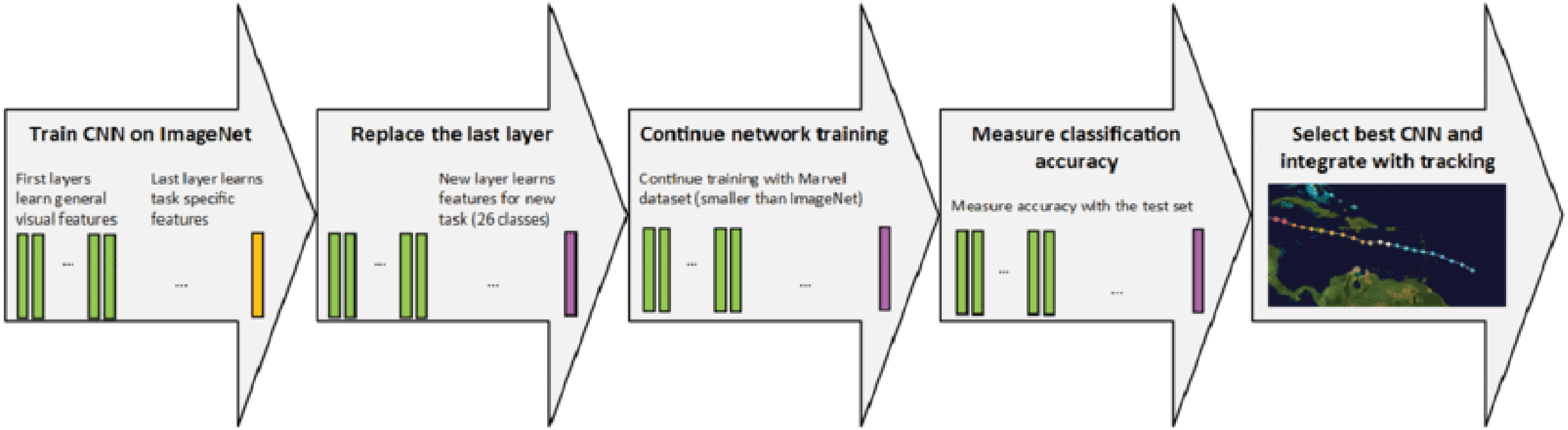
Problem Statement:
Training a Deep Neural Network needs a lot of data Collecting much data is expensive or just not possible

Idea:
– Some problems/ tasks are closely related
– Can we transfer knowledge from one task to another?
– Can we re-use (at least parts of) a pre-trained network for the new task?

How Transfer Learning Works

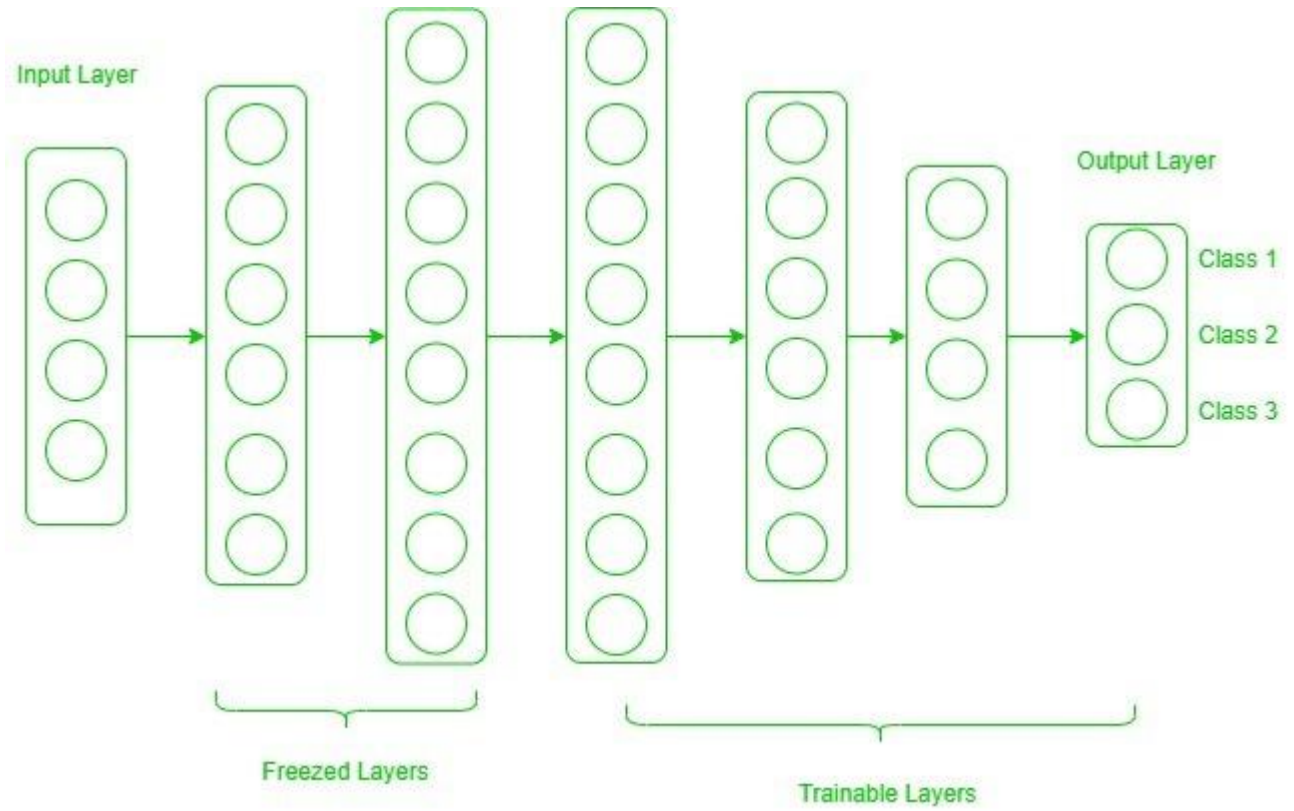


How to implement Transfer Learning?



Obtain the pre-trained model → Create a base model → Freeze layers → Train the new layers on the dataset → Improve the model via fine-tuning

Frozen vs. Trainable Layers in Transfer Learning



How It Works:

1. Take a pre-trained model (e.g., VGG16, ResNet, BERT).
2. Remove the last few layers responsible for classification in the original task.
3. Replace them with new layers specific to the new task (e.g., different output classes).
4. Train only the new layers while freezing the earlier layers to retain the learned representations.

Fine-Tuning

- Fine-Tuning is an extension of Transfer Learning where, after transferring the knowledge from a pre-trained model, some or all of the layers of the pre-trained model are also **unfrozen** and retrained on the new dataset.

How It Works:

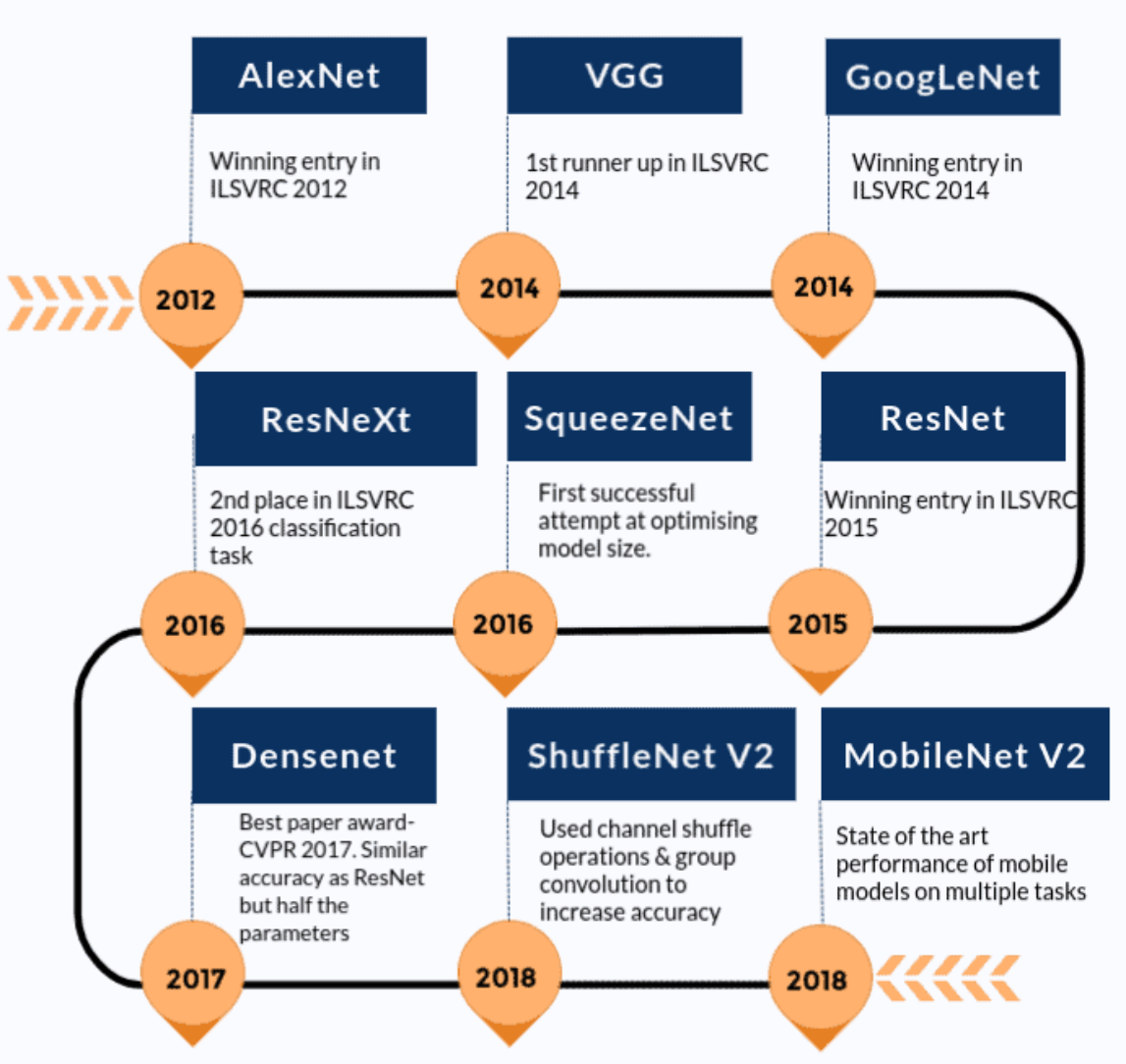
1. Load a pre-trained model.
2. Replace the last few layers with new ones for the specific task.
3. Initially, freeze most layers and train only the new layers (Transfer Learning step).
4. Unfreeze some or all pre-trained layers and retrain them with a **lower learning rate** to fine-tune them for the new dataset.

When to Use Fine-Tuning:

- When the new dataset is **similar but slightly different** from the original dataset.
- When the dataset size is **moderate to large** to avoid overfitting.
- When you need **better feature representations** for your task.

Pre-Trained Models for Image Classification

1. ResNet (Residual Networks) - by Microsoft Research
2. Inception (GoogLeNet) - by Google
3. VGG (Visual Geometry Group) - by the Visual Geometry Group
4. EfficientNet - by Google
5. DenseNet (Dense Convolutional Network) - Cornell University
6. MobileNet - by Google
7. NASNet (Neural Architecture Search Network) - by Google
8. Xception (Extreme Inception) - by Google
9. AlexNet - by Alex Krizhevsky
10. Vision Transformers (ViT) - by Google



Leading deep learning ‘pre-trained models’
(DNN architectures)

Computer Vision

Image classification

- Google InceptionV3
- Microsoft ResNet
- NASNet
- Oxford VGG Model
- MobileNetV2
- etc.

Object detection

- Yolo (You Only Look Once)
- R-CNN
- SPP-net
- Fast R-CNN
- Faster R-CNN
- etc.

Audio and speech

- Wavenet
- espnet
- waveblow
- deepspeech2
- loop
- tacotron
- etc.

NLP (Natural language processing)

Generative Models

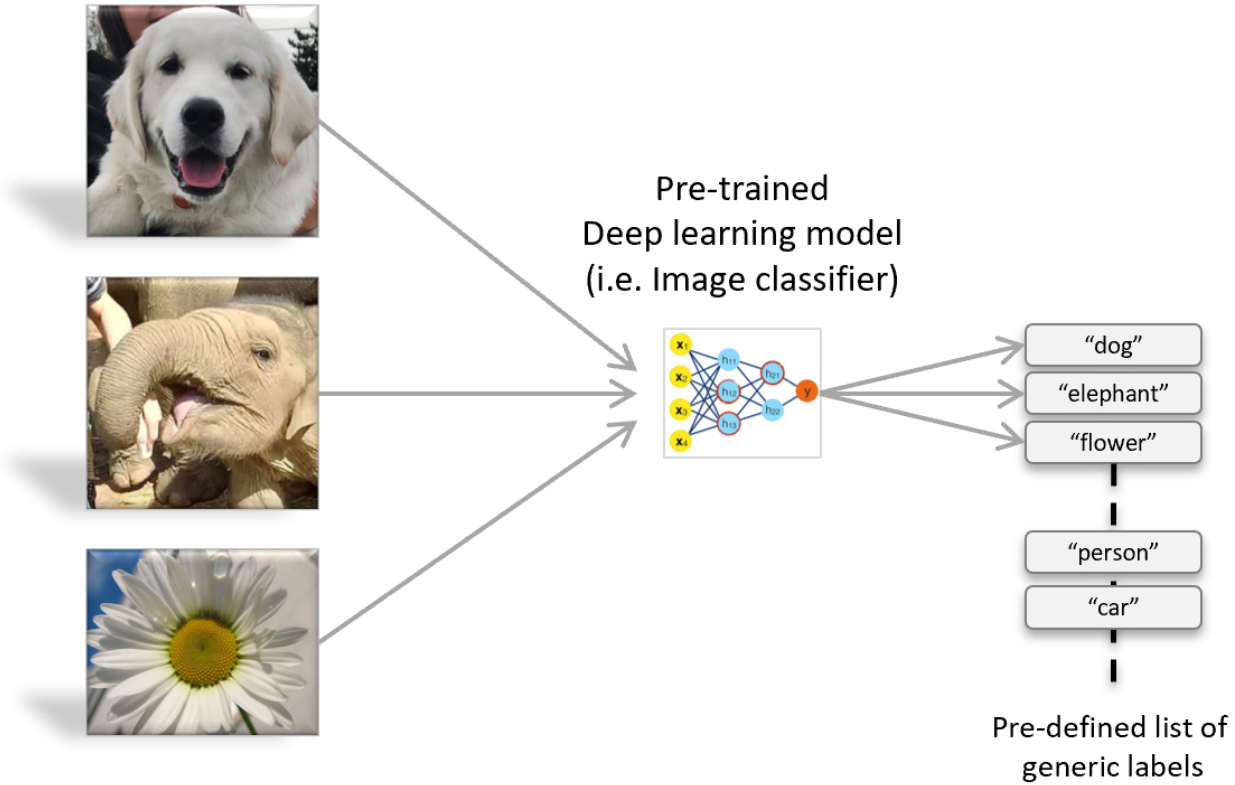
… other domains …

- Huge investment/cost in DNN architecture research plus costly training on large datasets made by organizations such as Google, Microsoft, Facebook, Universities and researchers.
- You can take advantage of it by simply consuming pre-trained models

https://modelzoo.co/categories

Using pre-trained deep learning models with ML.NET

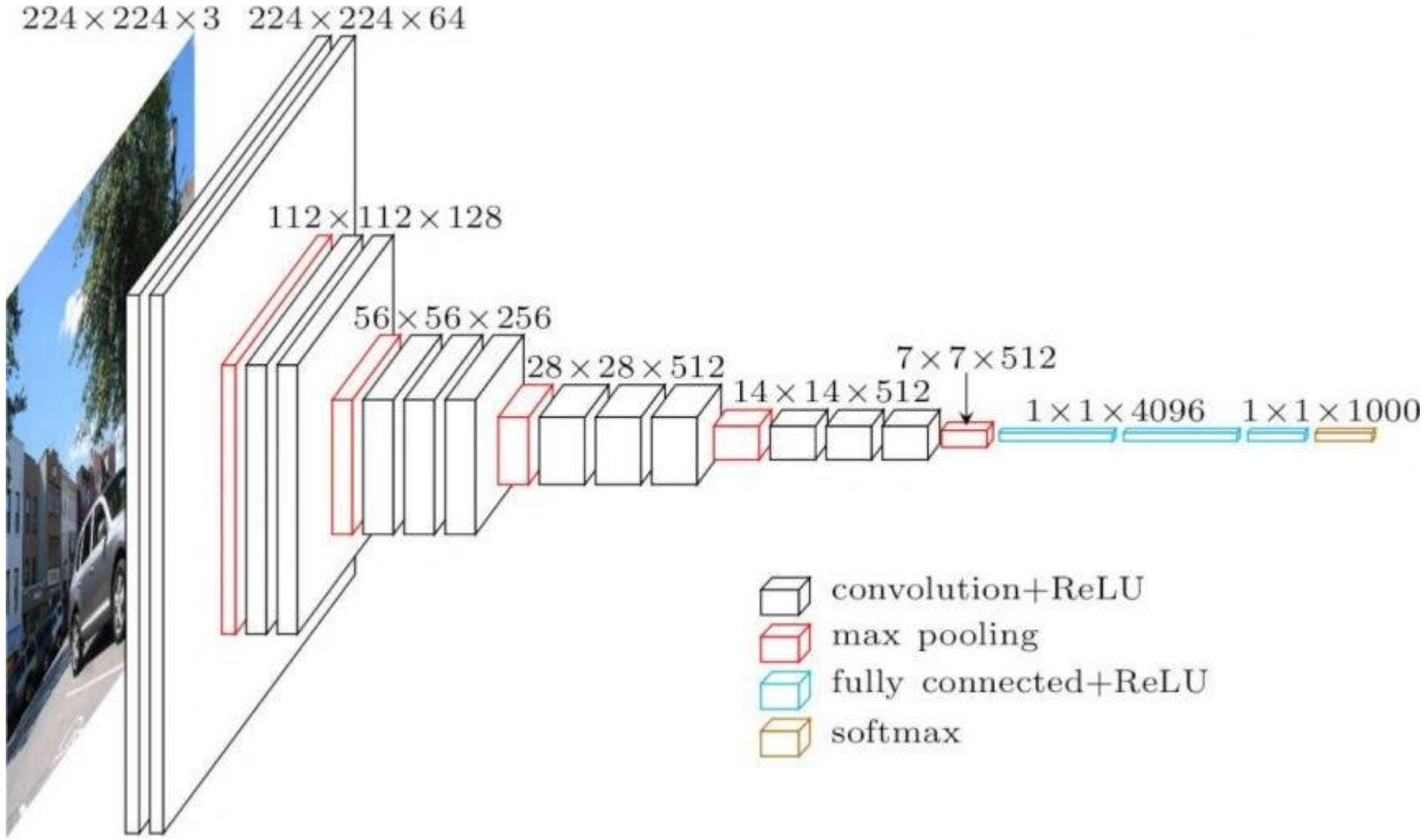
Scenario: **Image classifier** (Consuming the model)

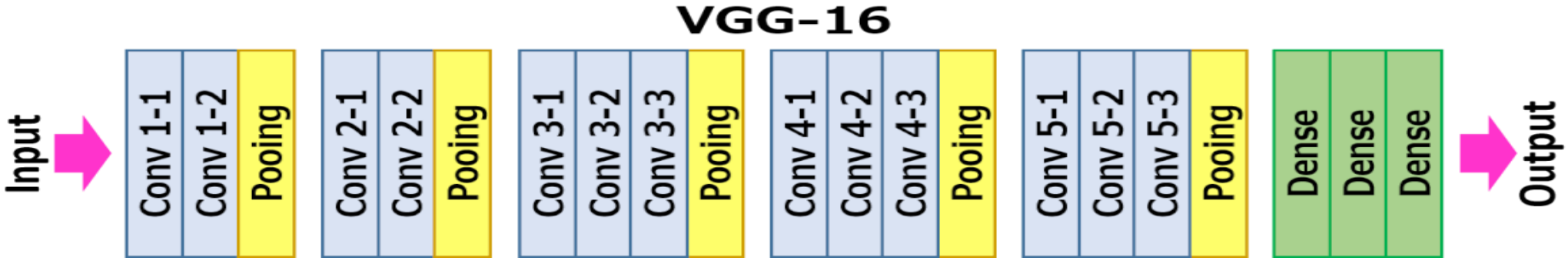


Examples of pre-trained models (Image classifiers):

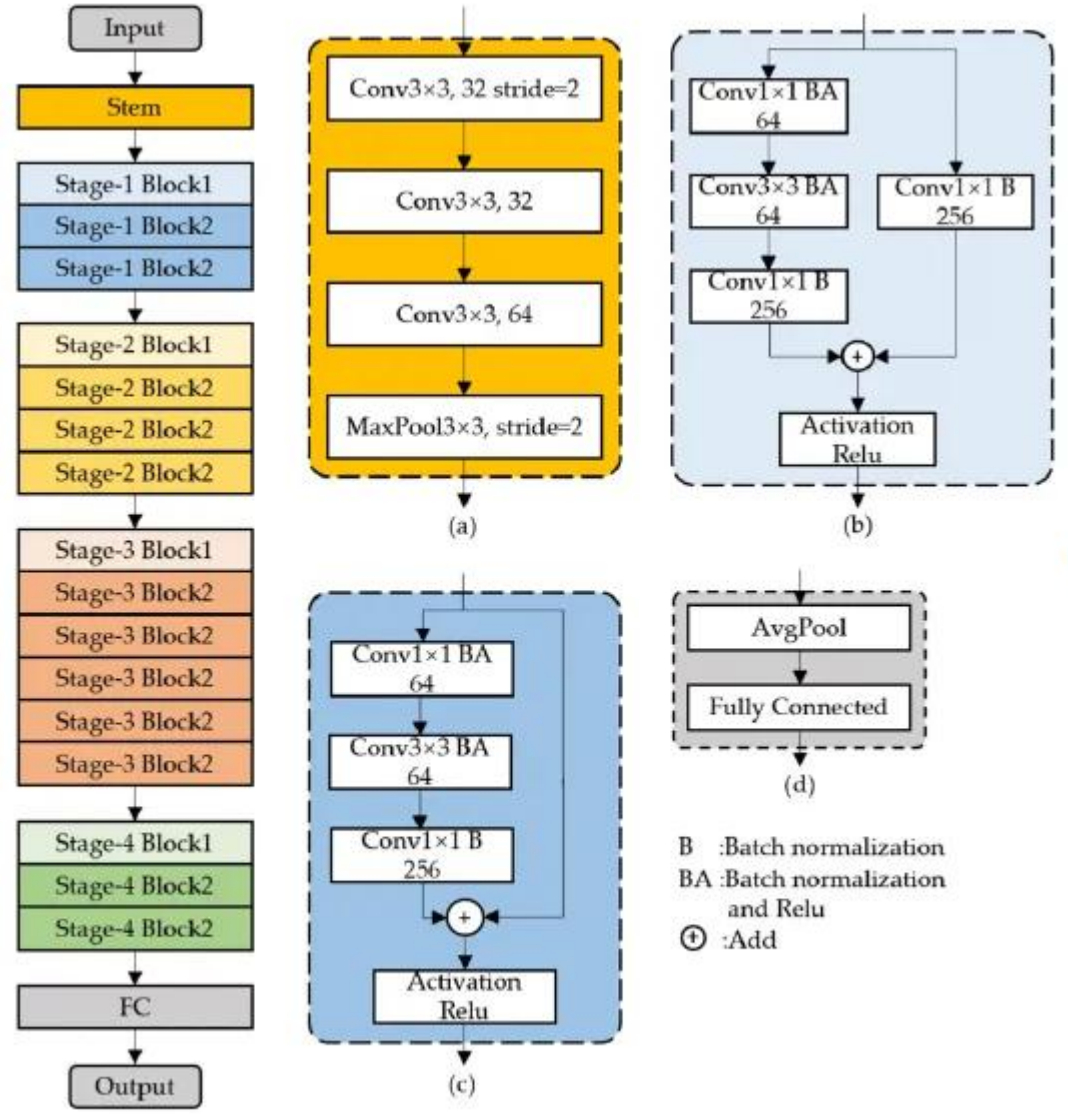
- Google **Inception v3**, **NASNet**
- Microsoft **ResNet**
- Oxford **VGG** Model, etc.

VGG-16


224×224×3
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
softmax

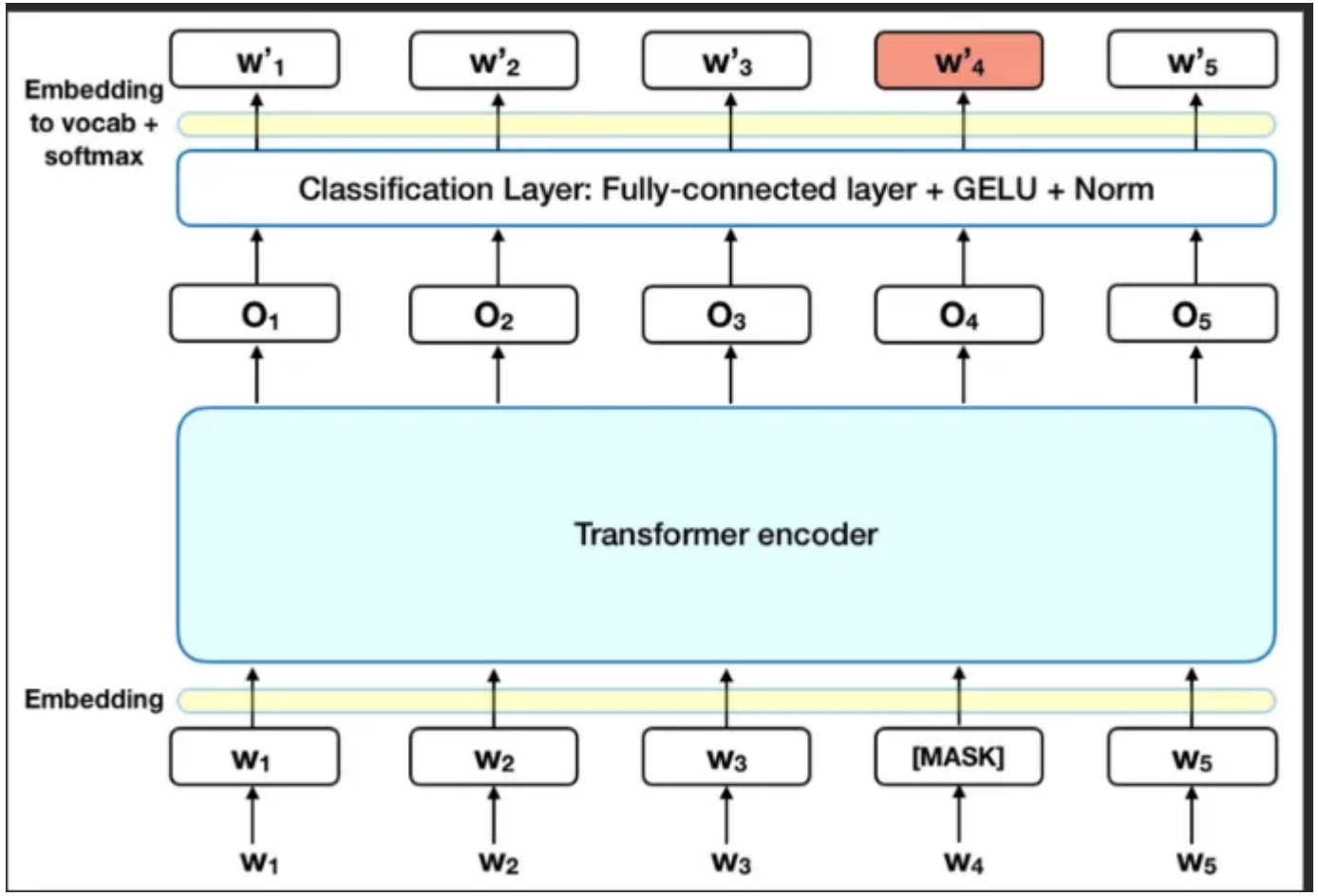

VGG-16
Input
Conv 1-1
Conv 1-2
Pooing
Conv 2-1
Conv 2-2
Pooing
Conv 3-1
Conv 3-2
Conv 3-3
Pooing
Conv 4-1
Conv 4-2
Conv 4-3
Pooing
Conv 5-1
Conv 5-2
Conv 5-3
Pooing
Dense
Dense
Dense
Output

ResNet (Residual Network)



The architecture of ResNet-50-vd. (a) Stem block; (b) Stage1-Block1; (c) Stage1-Block2; (d) FC-Block.

BERT (Bidirectional Encoder Representations from Transformers)



Language model

Popular Image Datasets

- **ImageNet**
- **CIFAR** - https://www.cs.toronto.edu/~kriz/cifar.html
- **MNIST**

Applications of Transfer Learning

- Healthcare: Transfer learning helps develop medical diagnostic tools, leveraging knowledge from general image recognition models to analyze medical images like X-rays or MRIs.

Ex: MegaMolBART — part of the NVIDIA BioNeMo service and framework — can understand the language of chemistry and learn the relationships between atoms in real-world molecules (Drug Industry)

- Finance: Transfer learning in finance assists in fraud detection, risk assessment, and credit scoring by transferring patterns learned from related financial datasets.

- Natural Language Processing (NLP): In NLP, models like BERT, GPT, and ELMo are pre-trained on vast text corpora and later fine-tuned for specific tasks such as sentiment analysis, machine translation, and question-answering.

Ex: Neural Translation model (GNMT) by Google : Cross-lingual translations (Ex: English to Korean), NVIDIA NeMo Megatron

Thank You ☺