Machine Learning (ML)

Chapter 10:

Transfer Learning (TL) in Machine Learning Models

Saeed Saeedvand, Ph.D.

Outline

In this Chapter:

- ✓ Define the Transfer Learning
- ✓ Why and when to Use Transfer Learning
- ✓ Transfer Learning Types
  - Inductive Transfer Learning:
  - Transductive Transfer Learning:
  - Unsupervised Transfer Learning:
- ✓ Relationship between DL and TL?
- ✓ Visual Geometry Group (VGG) and ResNet structures
- ✓ Self-Supervised Learning
- ✓ Recent Advances in Transfer Learning

Aim of this chapter:

- ✓ Understanding the Transfer Learning and see its different approaches, talk about its applications and understand some examples.

Transfer Learning (TL)

What is the Transfer Learning?

- ✓ Training a DL model from scratch often requires a large amount of labeled data and computational resources.
- ✓ Transfer Learning (TL) a type of learning strategy, allows to transfer the knowledge learned in one ML problem to another problem.
- ✓ So a model trained on one task is reused as the starting point for a model on a second task.

Transfer Learning (TL)

What is the Transfer Learning?

- ✓ The main goal of transfer learning is to use the knowledge (features, patterns, models etc.) first.
- ✓ Then use them to improve the performance **or** speed up learning in another related task (the target task).

Why and when to Use Transfer Learning

When to Use Transfer Learning?

- **Similar Tasks:**
  - ✓ When your target task is similar to some tasks that a model has already learned.
  - ✓ Transfer Learning allows you to influence that knowledge in this case.

- **Insufficient Resources:**
  - ✓ When computational resources or time is limited
  - ✓ Transfer Learning can reduce the need for training from scratch.

Why and when to Use Transfer Learning

Why to Use Transfer Learning?

- **Improved Performance:**
  - ✓ Transfer Learning often leads to better model performance, especially in the cases we have limited data for the new task.

- **Efficiency:**
  - ✓ It can save computational resources and time, because we reusing a pre-trained model.

- **Avoid Overfitting:**
  - ✓ TL can help to avoid overfitting, **when** the amount of data for the target task is small (if we only train for new task, it may overfit easily and not generalized).

Transfer Learning (TL)

Transfer Learning Types

✓ TL usually categorized based on the relationship between the **source** and **target** tasks.

Inductive Transfer Learning:

✓ In Inductive Transfer Learning, the source and target tasks are different but they share the same input space or domain.

- E.g. both classification and regression tasks, but the visual image are the same.

Transductive Transfer Learning:

✓ Both source and target tasks are the same, but the domains are different.

- E.g. both classification tasks, but the visual image different.

Unsupervised Transfer Learning:

✓ Both the source and target tasks are unsupervised learning tasks.

Transfer Learning (TL) Types

Inductive Transfer Learning:

✓ Learn a source task and use the learned knowledge to improve the learning of the target task.

Example

Source task

✓ Train a network model to recognize various types of animals in images.

- Our model will learn to identify features and patterns that are important for recognizing animals:
- For example: colors, shapes, textures, etc.

Target task

✓ Train a new model to recognize different breeds of dogs:

- Instead of training a new model from scratch, you can use the previous trained model as a starting point.
- **This model** already has useful information about recognizing animals.

Deep Learning and Transfer Learning

Relationship between DL and TL?

Synergy:

- ✓ DL and TL work well together.
- ✓ DL models can learn hierarchical, abstract representations of data, which are often useful for many tasks.

Pre-trained Models:

- ✓ DL models (like VGG for images or BERT for text) trained on large datasets that already captured useful features and patterns.
- ✓ These models can be used as starting points for other tasks via Transfer Learning.

Deep Learning and Transfer Learning

Relationship between DL and TL?

Feature Extraction:

- ✓ Higher-level representation in DL can be used as pre-trained model to transform the input data into new model (for a specific task).

Fine-Tuning:

- ✓ In this method, continue training (i.e., updating the weights) on the same or new task.

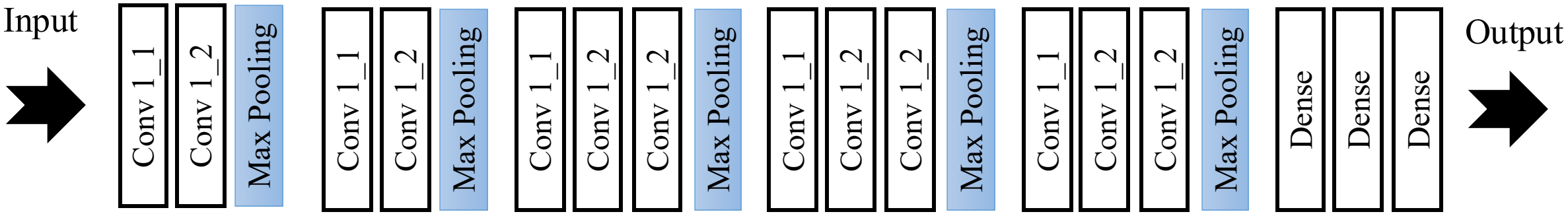
Image Classification with Transfer Learning

Visual Geometry Group (VGG)

- ✓ From University of Oxford for image classification
- ✓ It has simple and uniform architecture.
- ✓ The main contribution of the VGG networks:
  - Showing that the depth of the network is a critical factor for achieving good performance.

VGG16

Has different version:
VGG16 and VGG19



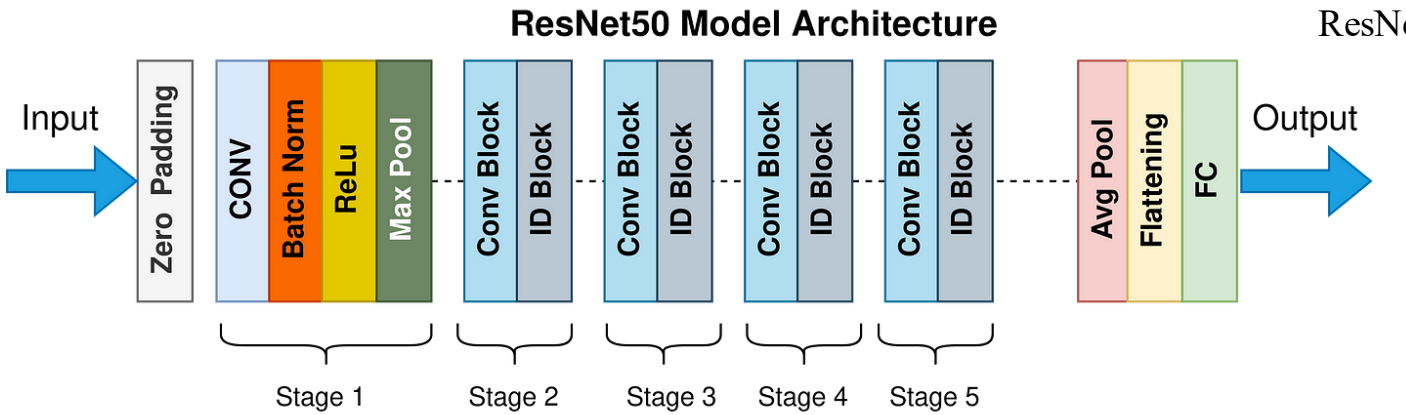

```
vgg_model = models.vgg16()
```

Image Classification with Transfer Learning

Residual Neural Network (ResNet)

- ✓ It was highly influential in the field of DL.
- ✓ The winners of the ImageNet competition in **2015,** on more than 1.2 million training images.
- ✓ Its key innovation is the introduction of skip connections.

Has different version:
ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152



- ✓ **Identity blocks:** shortcuts connection that allows the network to learn residual mappings effectively.
- ✓ Identity blocks in ResNet helps dealing the of gradients vanishing problem.

Image Classification with Transfer Learning

Image Classification with TL example

- ✓ We can use a pre-trained model like **VGG**, **ResNet**, etc., which trained on a large-scale image dataset (like ImageNet), as a feature extractor.
- ✓ The extracted features can then be used to train a classifier for your specific task.
- ✓ This often leads to better performance than training a model from scratch, especially when you have limited data for your task.

```
vgg_model = models.vgg16(pretrained=True)
```

Challenges and Limitations in Transfer Learning

Challenges

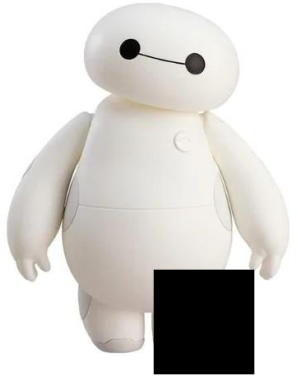
- ✓ Transfer Learning is not always beneficial.
- ✓ For instance, when the source and target tasks are not related (called Negative Transfer).
- ✓ Negative Transfer can degrade the performance.
- ✓ Some times differences between the source and target domains can lead to have more challenges in transferring knowledge.

Self-Supervised Learning

What is the Self-Supervised Learning

- ✓ Self-Supervised Learning subfield of ML where the model learns to predict the part of the missing data.
- ✓ Term self-supervised idea comes from the design of a task where the labels or answers come from the input data.
  - For instance, we remove part of text or image, or … and our model learns to guess what is there

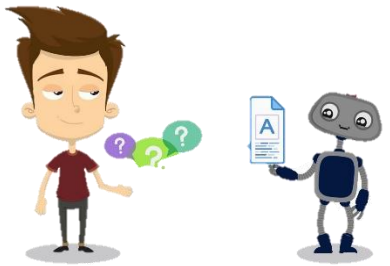
Image Inpainting



Recent Advances in Transfer Learning

In Natural Language Processing (NLP)

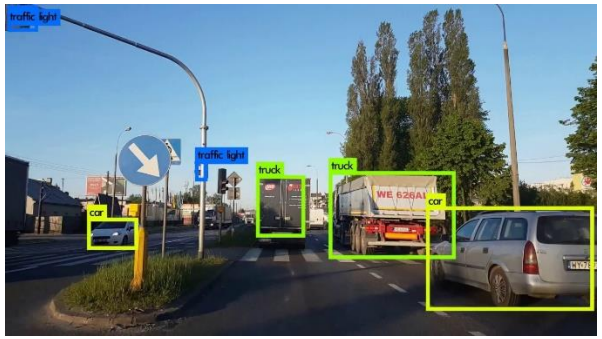
- ✓ Transfer learning has revolutionized NLP, which usually uses it **in combination with** Self-Supervised Learning (e.g. in GPT-3, GPT-4 and RoBERTa).
- ✓ These models first are pre-trained on a massive corpus (larger collections of text), learning the general structure and semantics of the language.
- ✓ Then, they are fine-tuned on specific tasks like question answering, sentiment analysis, or translation, … and they often achieve state-of-the-art performance.

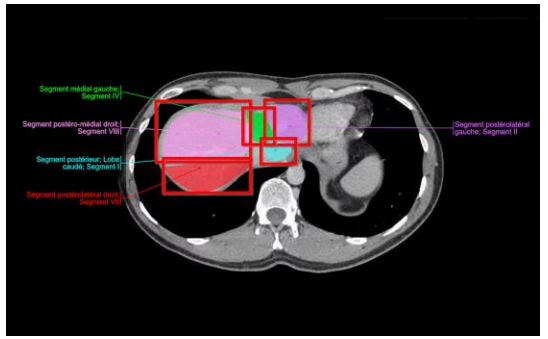


Recent Advances in Transfer Learning

In Computer Vision:

- ✓ Models such as ResNet, VGG, … initially trained on large image datasets like ImageNet.
- ✓ They are often fine-tuned for specific tasks like object detection, image segmentation, or medical image analysis.





Recent Advances in Transfer Learning

Cross-Modal Transfer

- ✓ Cross-Modal transfer involves transferring knowledge between different types of data.

For example:

Computer Vision:

- ✓ Pre-trained models from computer vision can be used for generate a textual description of an image (image captioning).

Audio-visual speech recognition task (recognize spoken words):

- ✓ Pre-trained models Audio Speech Recognition (ASR) to Visual Lip Movement Recognition.

Recent Advances in Transfer Learning

Deep Reinforcement Learning:

- ✓ In DRL, transfer learning can help to transfer information from one environment to another.
- ✓ This speeds up the training process and leads to better overall performance.

The items we can transfer in DRL:

Parameter Transfer:

- Transferer parameters to have a good initialization (network parameters).

Policy Transfer:

- We can set the policy as a starting point for training in the target task (agent acts based on previous network first).
- E.g. drone to other mobile robots, …

Recent Advances in Transfer Learning

✓ **Transfer in DRL (continue):**

Knowledge Distillation:

- Transferring knowledge from a complex RL agent to a simpler agent.
- For example: a student model can learn to mimic the behavior or predictions of the teacher model.
- This can be done by moving experience replay, value function, heuristics or rules, ….

Multi-Task Learning:

- Multiple related tasks are learned simultaneously.
- Share representations or policies among the tasks.

Transfer Learning

Can we apply transfer learing if network structures are different?

✓ Yes, transfer learning is still possible even if the network structures are different.

Approaches:

Partial Transfer:

✓ If the early layers or a some of layers in the source network are similar to the target network.

Feature Extraction:

✓ we can use the pre-trained network with a different structure as a feature extractor.

Model Conversion or Adaptation:

✓ In some cases, you may be able to convert or adapt the weights of the source network to match the structure of the target network.

Transfer Learning

How to save network Weights?

- ✓ Simplest approach is to save our models weights and load on new network to learn new task as initial weights.

```
torch.save(model.state_dict(), 'weights.pth')

model.load_state_dict(torch.load('weights.pth'))
```

Practice

Continue given example code for DNN for image classification by: 1) apply transfer learning and demonstrate its effects on two problems (use cross-validation).

Summery

- ✓ We understood the Transfer Learning and why we should use it
- ✓ We discussed different types of transfer learning
    - Inductive Transfer Learning:
    - Transductive Transfer Learning:
    - Unsupervised Transfer Learning:
- ✓ We saw Deep model architectures like VGG and ResNet
- ✓ We introduced Self-Supervised Learning
- ✓ We discussed the recent advances in TL