machinelearningmastery.com

A Gentle Introduction to Transfer Learning for Deep Learning - Machine Learning Mastery

Jason Brownlee

11-14 minutes

Last Updated on September 16, 2019

Transfer learning is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task.

It is a popular approach in deep learning where pre-trained models are used as the starting point on computer vision and natural language processing tasks given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in skill that they provide on related problems.

In this post, you will discover how you can use transfer learning to speed up training and improve the performance of your deep learning model.

After reading this post, you will know:

- What transfer learning is and how to use it.
- Common examples of transfer learning in deep learning.

- When to use transfer learning on your own predictive modeling problems.

Kick-start your project with my new book Deep Learning for Computer Vision, including *step-by-step tutorials* and the *Python source code* files for all examples.

Let’s get started.

For an example of how to use transfer learning in computer vision, see the post:

- Transfer Learning in Keras with Computer Vision Models

A Gentle Introduction to Transfer Learning with Deep Learning
Photo by Mike's Birds, some rights reserved.

What is Transfer Learning?

Transfer learning is a machine learning technique where a model trained on one task is re-purposed on a second related task.

> Transfer learning and domain adaptation refer to the situation where what has been learned in one setting … is exploited to improve generalization in another setting

— Page 526, Deep Learning, 2016.

Transfer learning is an optimization that allows rapid progress or improved performance when modeling the second task.

> Transfer learning is the improvement of learning in a new task through the transfer of knowledge from a related task that has already been learned.

— Chapter 11: Transfer Learning, Handbook of Research on Machine Learning Applications, 2009.

Transfer learning is related to problems such as multi-task learning and concept drift and is not exclusively an area of study for deep learning.

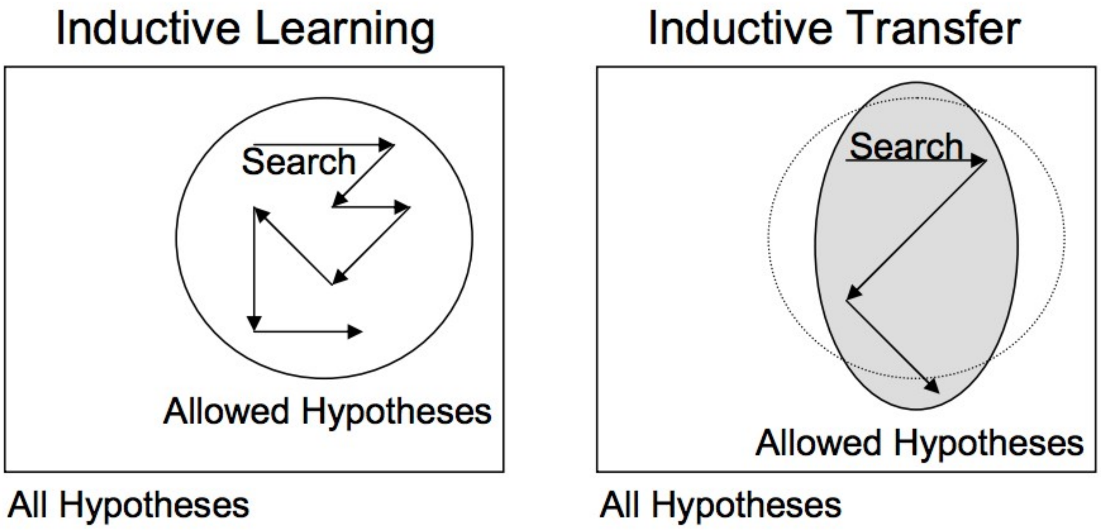
Nevertheless, transfer learning is popular in deep learning given the enormous resources required to train deep learning models or the large and challenging datasets on which deep learning models are trained.

Transfer learning only works in deep learning if the model features learned from the first task are general.

> In transfer learning, we first train a base network on a base dataset and task, and then we repurpose the learned features, or transfer them, to a second target network to be trained on a target dataset and task. This process will tend to work if the features are general, meaning suitable to both base and target tasks, instead of specific to the base task.

— How transferable are features in deep neural networks?

This form of transfer learning used in deep learning is called inductive transfer. This is where the scope of possible models (model bias) is narrowed in a beneficial way by using a model fit on a different but related task.



Depiction of Inductive Transfer
Taken from “Transfer Learning”

How to Use Transfer Learning?

You can use transfer learning on your own predictive modeling problems.

Two common approaches are as follows:

1. Develop Model Approach
2. Pre-trained Model Approach

Develop Model Approach

1. **Select Source Task**. You must select a related predictive modeling problem with an abundance of data where there is some relationship in the input data, output data, and/or concepts learned during the mapping from input to output data.
2. **Develop Source Model**. Next, you must develop a skillful model for this first task. The model must be better than a naive model to ensure that some feature learning has been performed.
3. **Reuse Model**. The model fit on the source task can then be used as the starting point for a model on the second task of interest. This may involve using all or parts of the model, depending on the modeling technique used.
4. **Tune Model**. Optionally, the model may need to be adapted or refined on the input-output pair data available for the task of interest.

Pre-trained Model Approach

1. **Select Source Model**. A pre-trained source model is chosen from

available models. Many research institutions release models on large and challenging datasets that may be included in the pool of candidate models from which to choose from.

2. **Reuse Model**. The model pre-trained model can then be used as the starting point for a model on the second task of interest. This may involve using all or parts of the model, depending on the modeling technique used.
3. **Tune Model**. Optionally, the model may need to be adapted or refined on the input-output pair data available for the task of interest.

This second type of transfer learning is common in the field of deep learning.

Examples of Transfer Learning with Deep Learning

Let's make this concrete with two common examples of transfer learning with deep learning models.

Transfer Learning with Image Data

It is common to perform transfer learning with predictive modeling problems that use image data as input.

This may be a prediction task that takes photographs or video data as input.

For these types of problems, it is common to use a deep learning model pre-trained for a large and challenging image classification task such as the ImageNet 1000-class photograph classification competition.

The research organizations that develop models for this competition and do well often release their final model under a permissive license for reuse. These models can take days or weeks to train on modern hardware.

These models can be downloaded and incorporated directly into new models that expect image data as input.

Three examples of models of this type include:

- Oxford VGG Model
- Google Inception Model
- Microsoft ResNet Model

For more examples, see the Caffe Model Zoo where more pre-trained models are shared.

This approach is effective because the images were trained on a large corpus of photographs and require the model to make predictions on a relatively large number of classes, in turn, requiring that the model efficiently learn to extract features from photographs in order to perform well on the problem.

In their Stanford course on Convolutional Neural Networks for Visual Recognition, the authors caution to carefully choose how much of the pre-trained model to use in your new model.

> [Convolutional Neural Networks] features are more generic in early layers and more original-dataset-specific in later layers

— Transfer Learning, CS231n Convolutional Neural Networks for Visual Recognition

Transfer Learning with Language Data

It is common to perform transfer learning with natural language processing problems that use text as input or output.

For these types of problems, a word embedding is used that is a mapping of words to a high-dimensional continuous vector space where different words with a similar meaning have a similar vector representation.

Efficient algorithms exist to learn these distributed word representations and it is common for research organizations to release pre-trained models trained on very large corpa of text documents under a permissive license.

Two examples of models of this type include:

- Google's word2vec Model
- Stanford's GloVe Model

These distributed word representation models can be downloaded and incorporated into deep learning language models in either the interpretation of words as input or the generation of words as output from the model.

In his book on Deep Learning for Natural Language Processing, Yoav Goldberg cautions:

> … one can download pre-trained word vectors that were trained on very large quantities of text […] differences in training regimes and underlying corpora have a strong influence on the resulting representations, and that the available pre-trained representations may not be the best choice for [your] particular use case.

— Page 135, Neural Network Methods in Natural Language Processing, 2017.

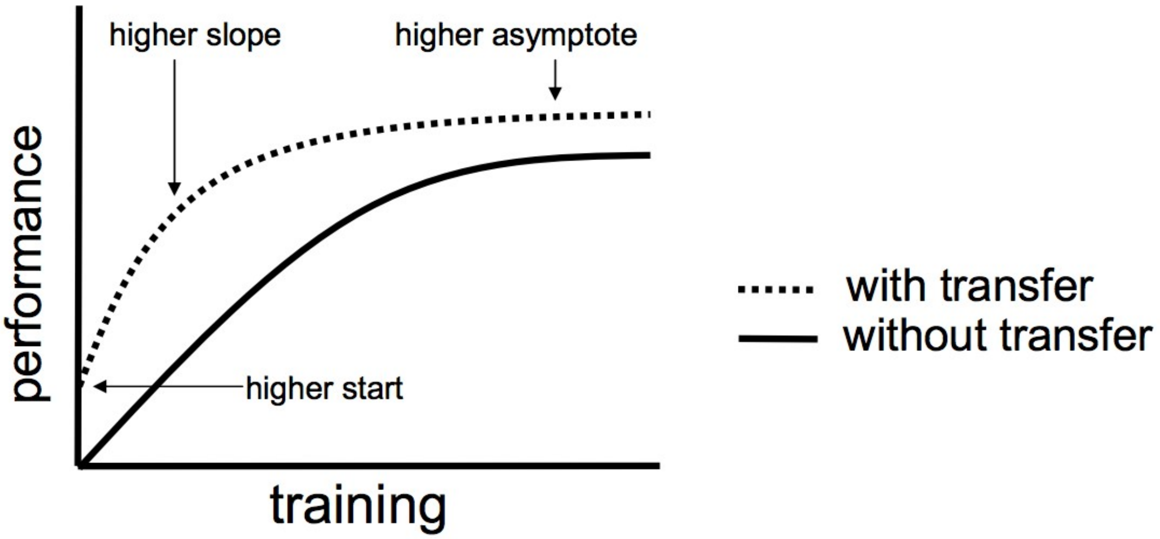
When to Use Transfer Learning?

Transfer learning is an optimization, a shortcut to saving time or getting better performance.

In general, it is not obvious that there will be a benefit to using transfer learning in the domain until after the model has been developed and evaluated.

Lisa Torrey and Jude Shavlik in their chapter on transfer learning describe three possible benefits to look for when using transfer learning:

1. **Higher start**. The initial skill (before refining the model) on the source model is higher than it otherwise would be.
2. **Higher slope**. The rate of improvement of skill during training of the source model is steeper than it otherwise would be.
3. **Higher asymptote**. The converged skill of the trained model is better than it otherwise would be.



Three ways in which transfer might improve learning.
Taken from "Transfer Learning".

Ideally, you would see all three benefits from a successful application of transfer learning.

It is an approach to try if you can identify a related task with abundant data and you have the resources to develop a model for that task and reuse it on your own problem, or there is a pre-trained model available that you can use as a starting point for your own model.

On some problems where you may not have very much data, transfer learning can enable you to develop skillful models that you simply could not develop in the absence of transfer learning.

The choice of source data or source model is an open problem and may require domain expertise and/or intuition developed via experience.

Further Reading

This section provides more resources on the topic if you are looking to go deeper.

Tutorials

- How to Use Transfer Learning when Developing Convolutional Neural Network Models
- How to Use The Pre-Trained VGG Model to Classify Objects in Photographs
- How to Improve Performance With Transfer Learning for Deep Learning Neural Networks

Books

- Deep Learning, 2016.
- Neural Network Methods in Natural Language Processing, 2017.

Papers

- A survey on transfer learning, 2010.
- Chapter 11: Transfer Learning, Handbook of Research on Machine Learning Applications, 2009.
- How transferable are features in deep neural networks?

Pre-trained Models

- Oxford VGG Model
- Google Inception Model
- Microsoft ResNet Model
- Google's word2vec Model
- Stanford's GloVe Model
- Caffe Model Zoo

Articles

- Transfer learning on Wikipedia
- Transfer Learning – Machine Learning's Next Frontier, 2017.
- Transfer Learning, CS231n Convolutional Neural Networks for Visual Recognition
- How does transfer learning work? on Quora

Summary

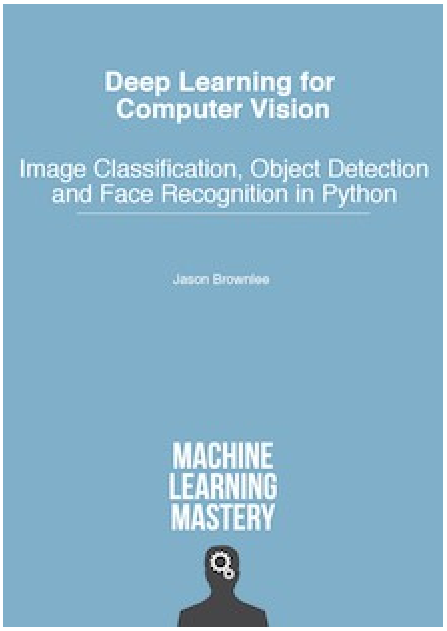
In this post, you discovered how you can use transfer learning to speed up training and improve the performance of your deep learning model.

Specifically, you learned:

- What transfer learning is and how it is used in deep learning.
- When to use transfer learning.
- Examples of transfer learning used on computer vision and natural language processing tasks.

Do you have any questions?
Ask your questions in the comments below and I will do my best to answer.

Develop Deep Learning Models for Vision Today!



Deep Learning for Computer Vision

Image Classification, Object Detection and Face Recognition in Python

Jason Brownlee

MACHINE LEARNING MASTERY

Develop Your Own Vision Models in Minutes

...with just a few lines of python code

Discover how in my new Ebook:
Deep Learning for Computer Vision

It provides **self-study tutorials** on topics like:
classification, *object detection (yolo and rcnn)*, *face recognition*