Transfer learning is a technique which is used when you want to create a neural network but don't have enough labelled data.

In that case we can use an already existing training model which has a different but similar dataset and train it. The resultant neural network can be used with whatever available data we have.

for eg. we have to recognize handwritten digits 1 - 9, but don't have enough data. On the other hand you have pictures of a million cats, dogs, cars, people

You can use those pictures of cats, dogs, cars, etc. to train a different neural network.
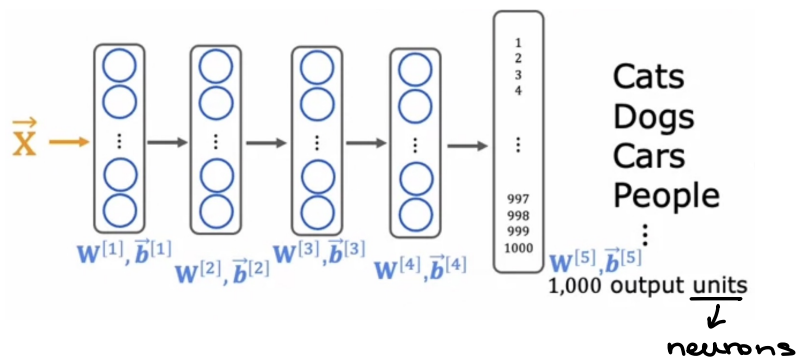
You will copy the above trained neural network, but replace it with a much smaller output layer, keeping the parameters same for all layers except for output layer.
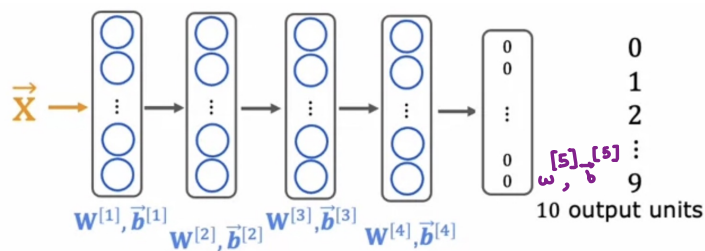

We've two options to train this neural network:-

Option 1: only train output layers parameters. → better if very small training set
Option 2: train all parameters. → if training set is a bit larger.

The reasoning behind this is that by 1-9 digits are basically 'images' but we don't have enough of them. Instead we have 'images' of cats, dogs, people, cars, etc. By training on them we obtain a model which can recognize most images. It can distinguish basic pixels of lines, edges, etc. (SUPERVISED PRE TRAINING)

The only input it hasn't seen is our handwritten digits, but even if we feed it our small dataset, we will get a good model because by now it can recognize images easily. (FINE TUNING)
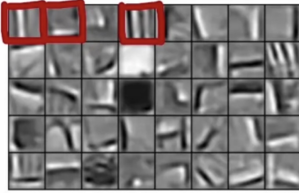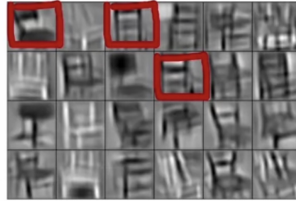
$\vec{x} \rightarrow$
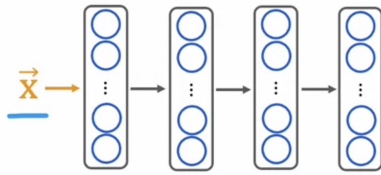
detects edges   detects corners   detects curves/basic shapes
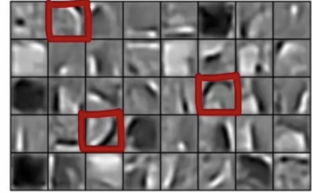
use the same input type

] → learns to detect some generic things needed to identify other images

Edges   Corners   Curves / basic shapes

\# One thing to keep in mind is that to train an image recognition algorithm, supervised training must be done using images. Similarly, for training audio, you need pretrained audio models.

It's not like you train on images and transfer to learn audio.

Transfer learning summary

1. Download neural network parameters pretrained on a large dataset with same input type (e.g., images, audio, text) as your application (or train your own). *1 million images*

2. Further train (fine tune) the network on your own data.

1000 images

50 images