# Intro

- In this video we will talk about tricks that will make training of new neural networks much faster!

Transfer learning

- Deep networks learn complex features extractor, but we need lots of data to train it from scratch!

- What if we can reuse an existing features extractor for a new task?

Features extractor

ImageNet
1000 classes

New task

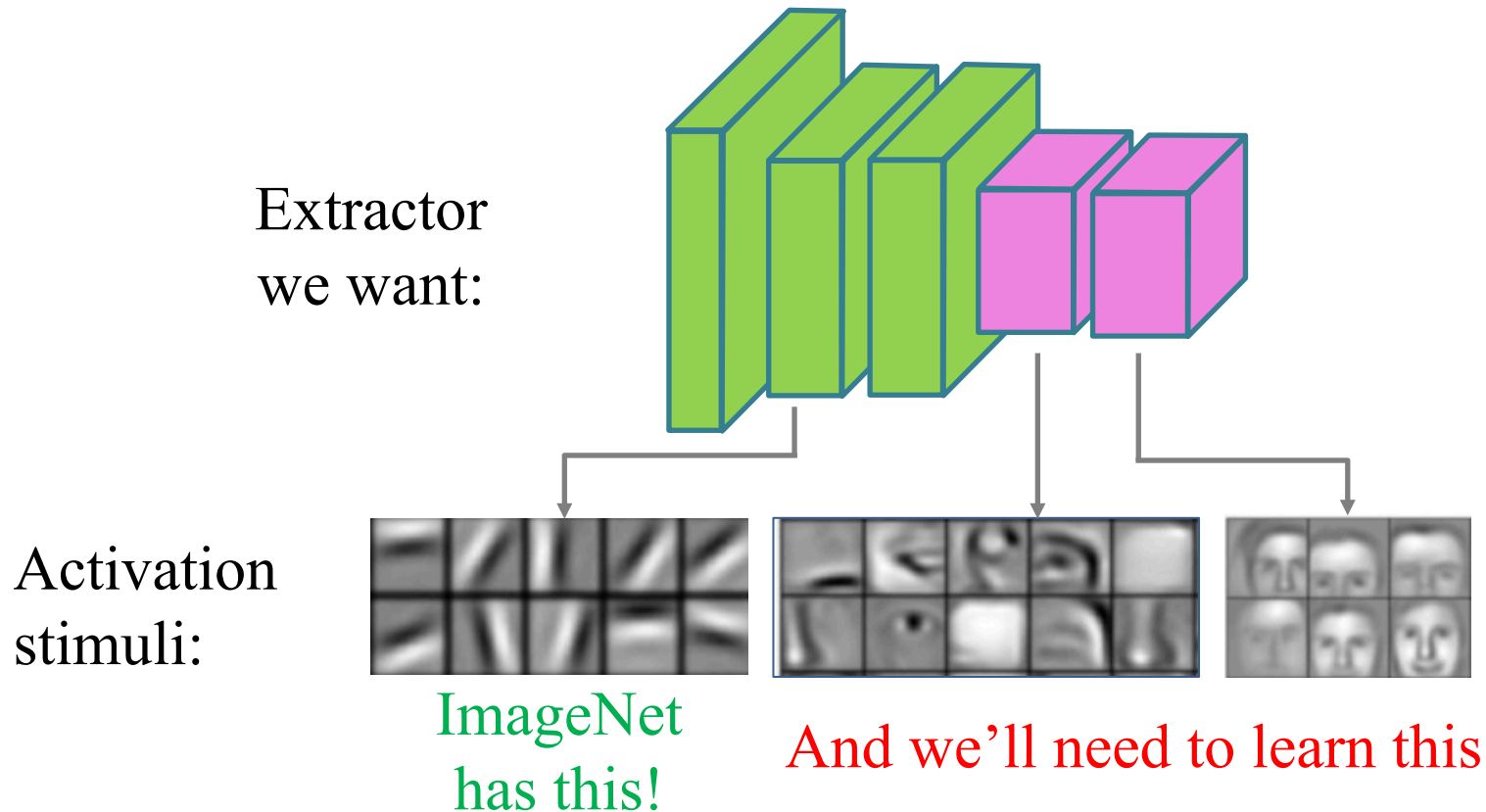Trained on ImageNet

All we need to train
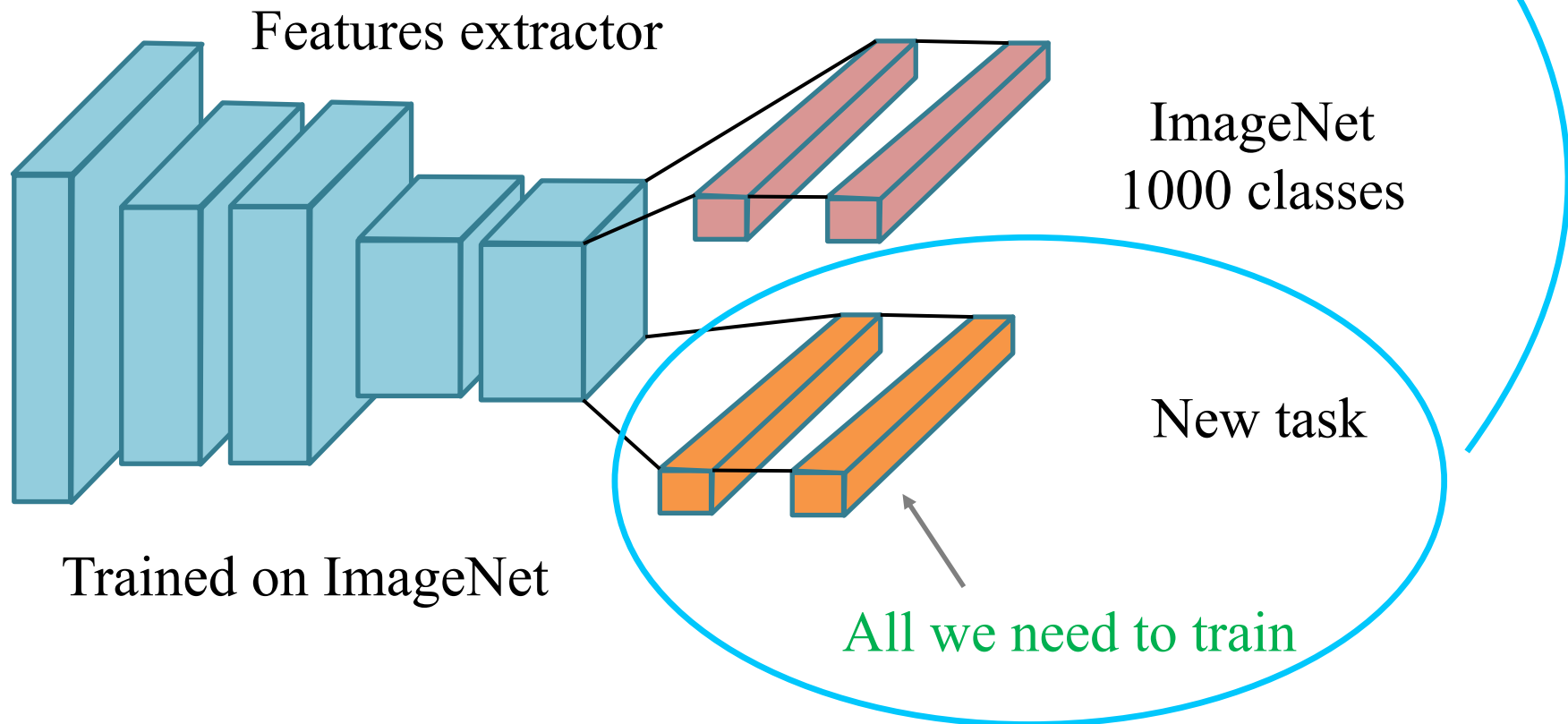
Transfer learning

- You need less data to train (for training only final MLP)

- It works if a domain of a new task is similar to ImageNet's

- Won't work for human emotions classification, ImageNet doesn't have people faces in the dataset!

Transfer learning

- But what if we need to classify human emotions?

- Maybe we can partially reuse ImageNet features extractor?

Extractor
we want:

Activation
stimuli:

ImageNet
has this!

And we'll need to learn this

Honglak Lee, http://web.eecs.umich.edu/~honglak/icml09-ConvolutionalDeepBeliefNetworks.pdf

We use the first few convolution +pool layers. These ones have low level activation stimuli.
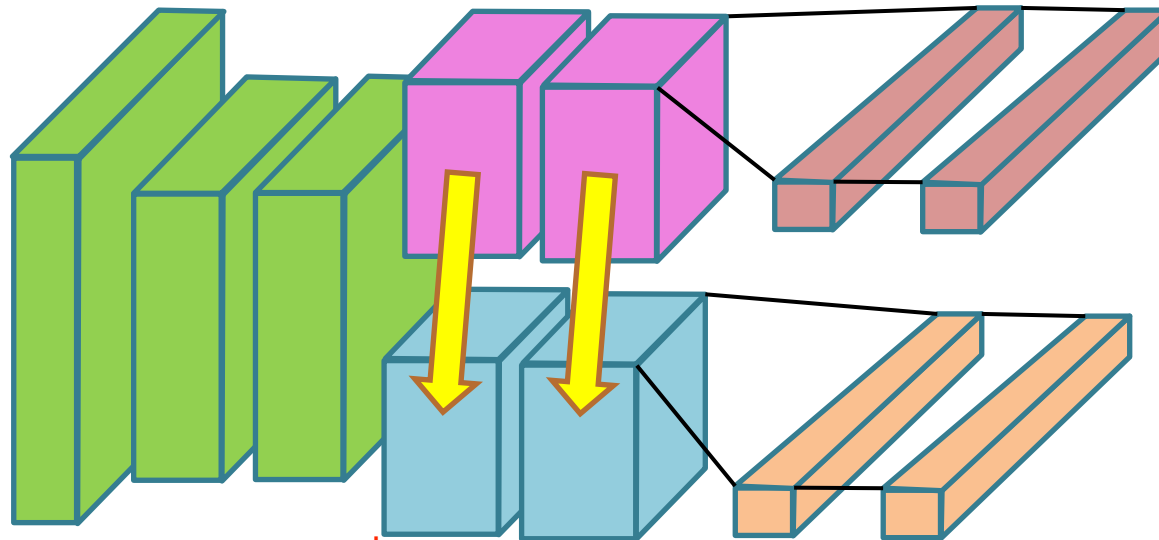
Fine-tuning

The intuition for fine tuning is that we start with features useful for some other task.
They are not perfect for our task, but much better than random.

ImageNet features extractor

We train the deeper layers because these will detect higher-level human face features!

ImageNet
1000 classes

New task

^ all we need to train

- You can initialize deeper layers with values from ImageNet.

- This is called **fine-tuning**, because you don't start with a random initialization.

- Propagate all gradients with smaller learning rate.

Smaller so that we don't lose the initialization weights very quickly!

Fine-tuning

- Very frequently used thanks to wide spectrum of ImageNet classes

- Keras has the weights of pre-trained VGG, Inception, ResNet architectures

- You can fine-tune a bunch of different architectures and make an ensemble out of them!

Takeaways

OK. So basically, use transfer learning with pre-trained models if our domain is similar to that model.

	ImageNet domain	Not similar to ImageNet
Small dataset	Train last MLP layers	Collect more data
Big dataset	Fine-tuning of deeper layers	Train from scratch

Summary

- In the next video we will take a look at other computer vision problems