NPTEL ONLINE CERTIFICATION COURSES

Course Name: Deep Learning

Faculty Name: Prof. P. K. Biswas

Department : E & ECE, IIT Kharagpur

Topic

Lecture 39: Popular CNN Models III

CONCEPTS COVERED

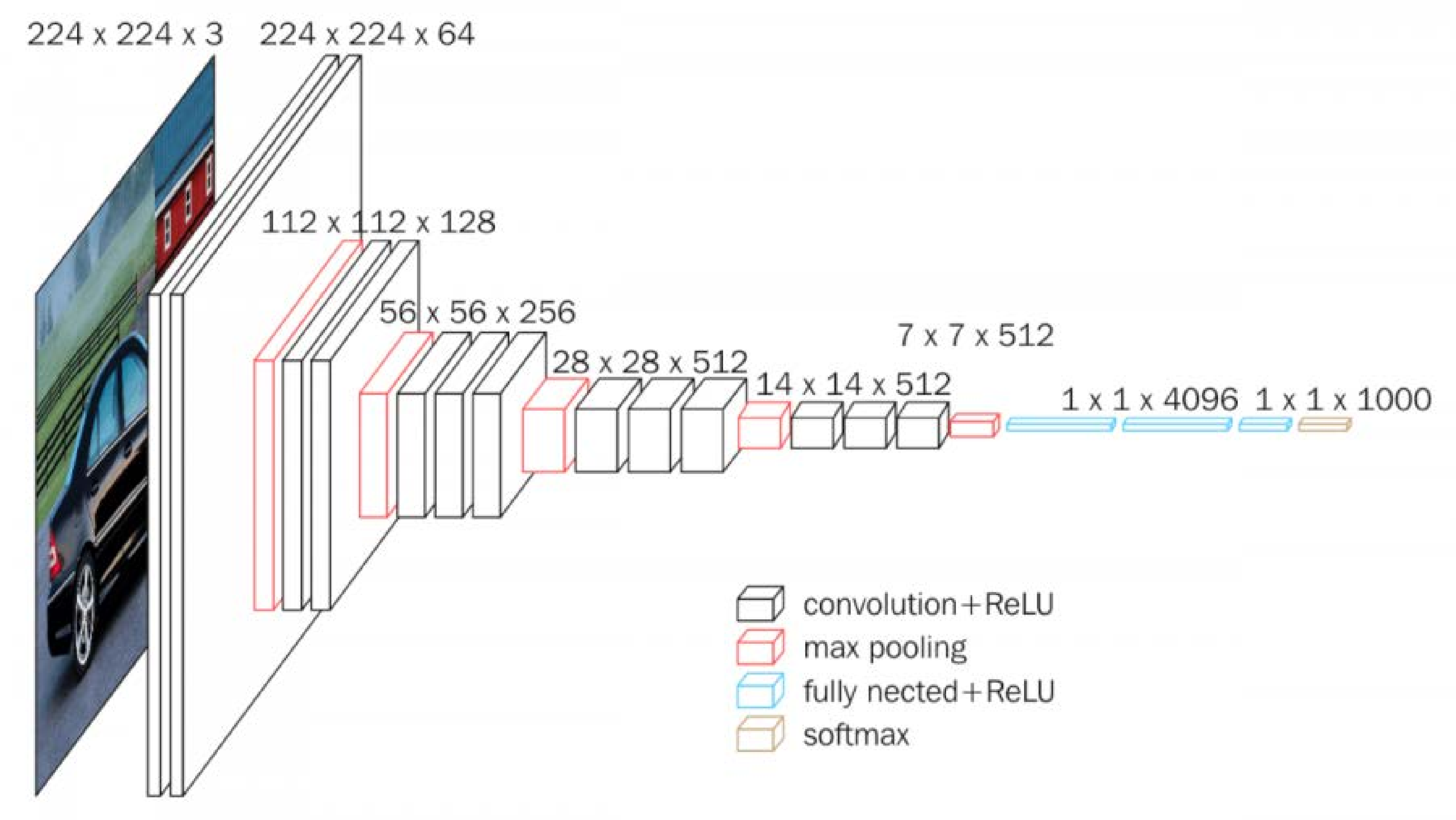
Concepts Covered:

- CNN
    - AlexNet
    - VGG Net
    - Transfer Learning
    - GoogLeNet
    - ResNet
    - etc.

VGG 16
ILSVRC 2014 1st Runner-Up

Visual Geometry Group
Oxford University

VGG 16



Very Deep Convolutional Networks for Large-Scale Image Recognition by Karen Simonyan and Andrew Zisserman

VGG 16

- Input to the architecture are color images of size 224x224.
- The image is passed through a stack of convolutional layers.
- Every convolution filter has very small receptive field: 3×3, Stride 1.
- Uses row and column padding to maintain spatial resolution after convolution.
- There are 13 Convolution Layers.
- There are 5 max-pool layers.
- Max pooling window size 2x2, stride 2.

VGG 16

- Not every convolution layer is followed by max-pool layer.
- 3 Fully connected layers.
- First two FC layers have 4096 channels each.
- Last FC layer has 1000 channels.
- Last layer is a softmax layer with 1000 channels, one for each category of images in ImageNet database.
- Hidden layers have ReLU as activation function.

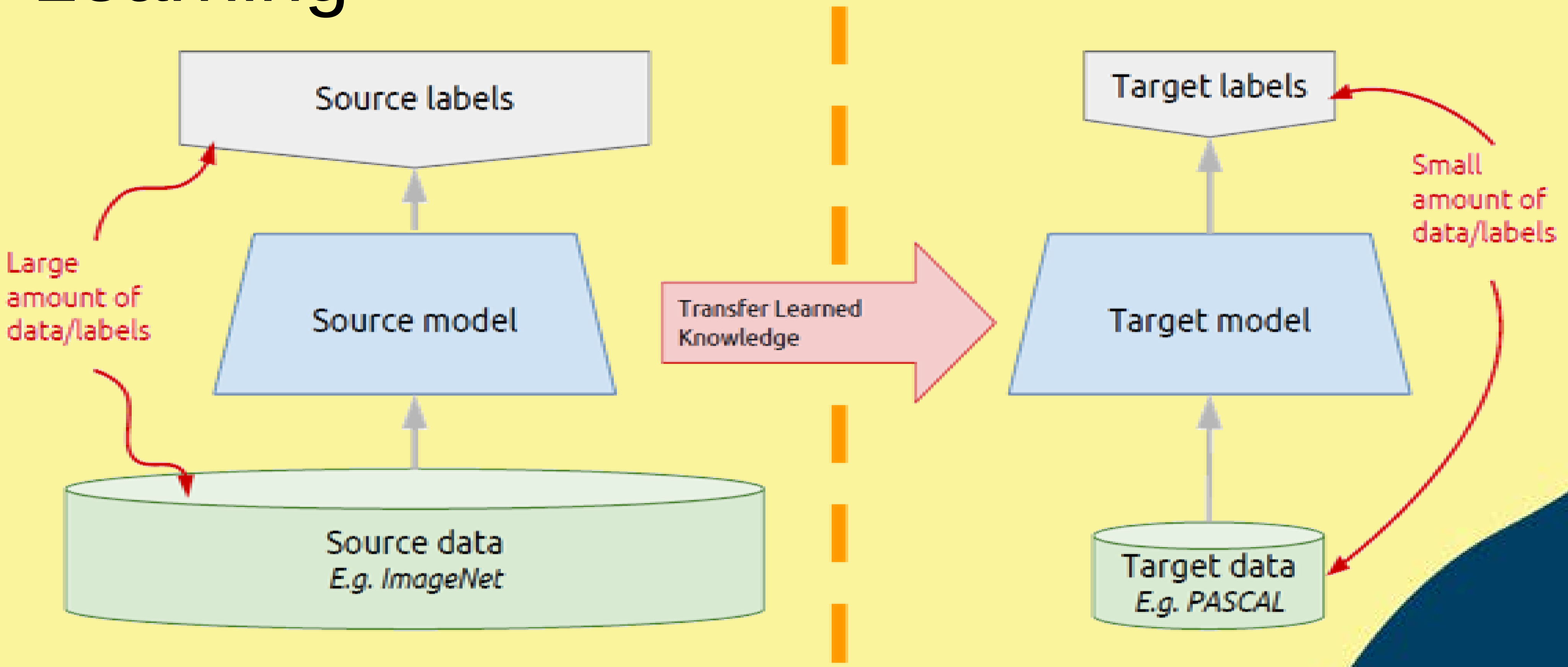
VGG 16

Striking difference from AlexNet

- All convolution kernels are of size 3x3 with stride 1.
- All maxpool kernels are of size 2x2 stride 2
- All variable size kernels as in AlexNet can be realised using multiple 3x3 kernels.
- This realisation is in terms of size of the receptive field covered by the kernels.
- Top-5 error rate ~ 7 %

Transfer Learning

Transfer Learning



Kevin McGuinness
https://www.slideshare.net/xavigiro/transfer-learning-d2l4-insightdcu-machine-learning-workshop-2017

Transfer Learning

CNN as Fixed Feature Extractor:

- Take a pre-trained CNN architecture trained on a large dataset (like ImageNet)
- Remove the last fully connected layer of this pre-trained network
- Remaining CNN acts as a fixed feature extractor for the new dataset

Transfer Learning

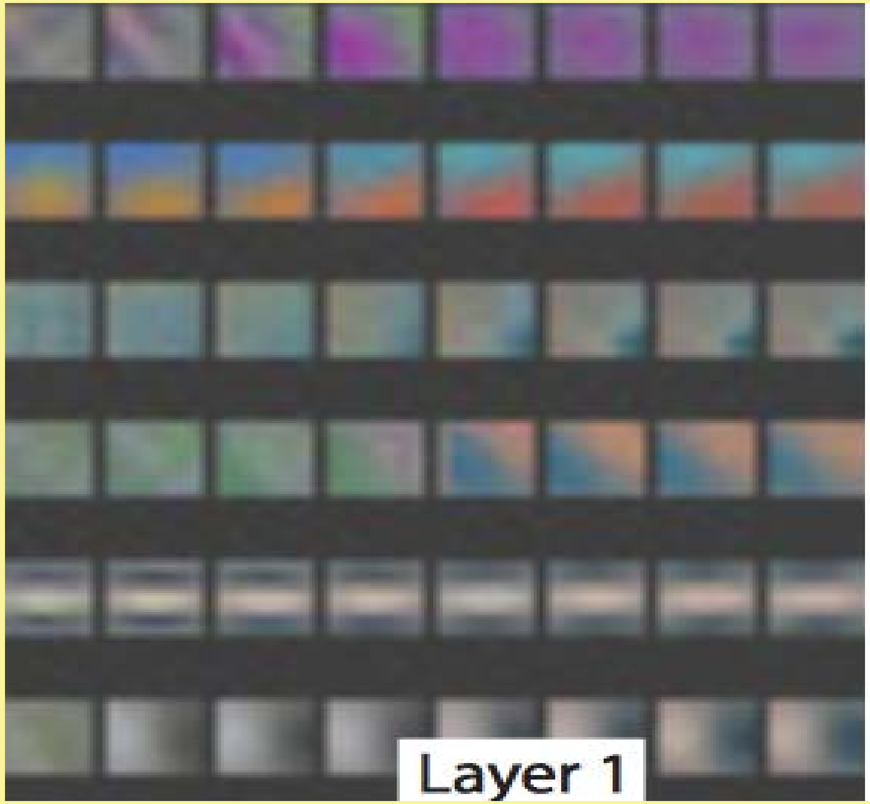


Image Source:-
https://becominghuman.ai/what-exactly-does-cnn-see-4d436d8e6e52

Transfer Learning

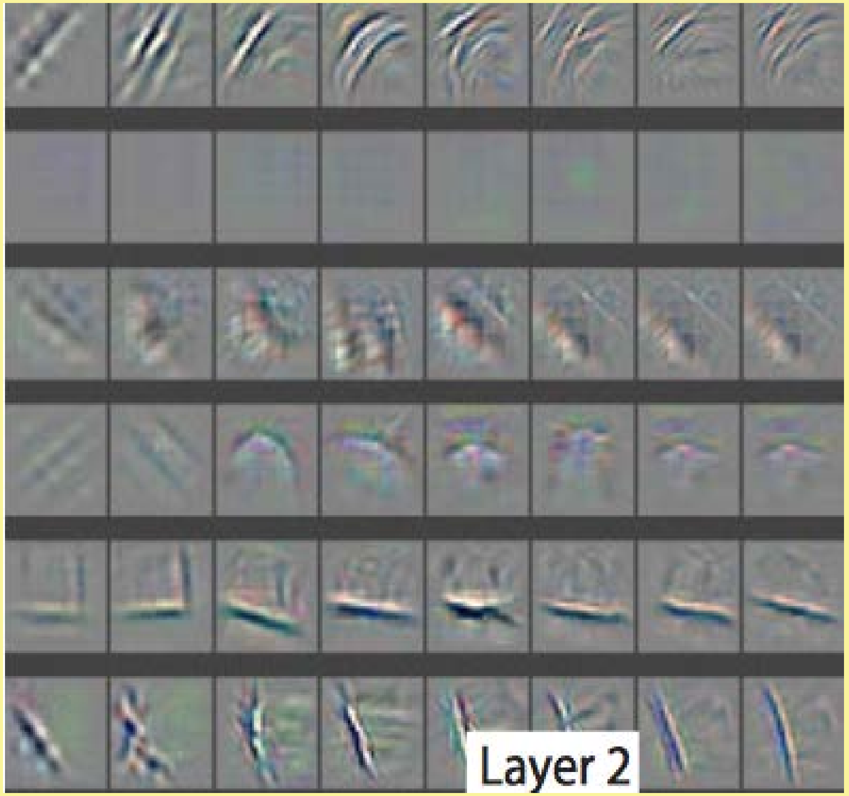


Image Source:-
https://becominghuman.ai/what-exactly-does-cnn-see-4d436d8e6e52

Transfer Learning

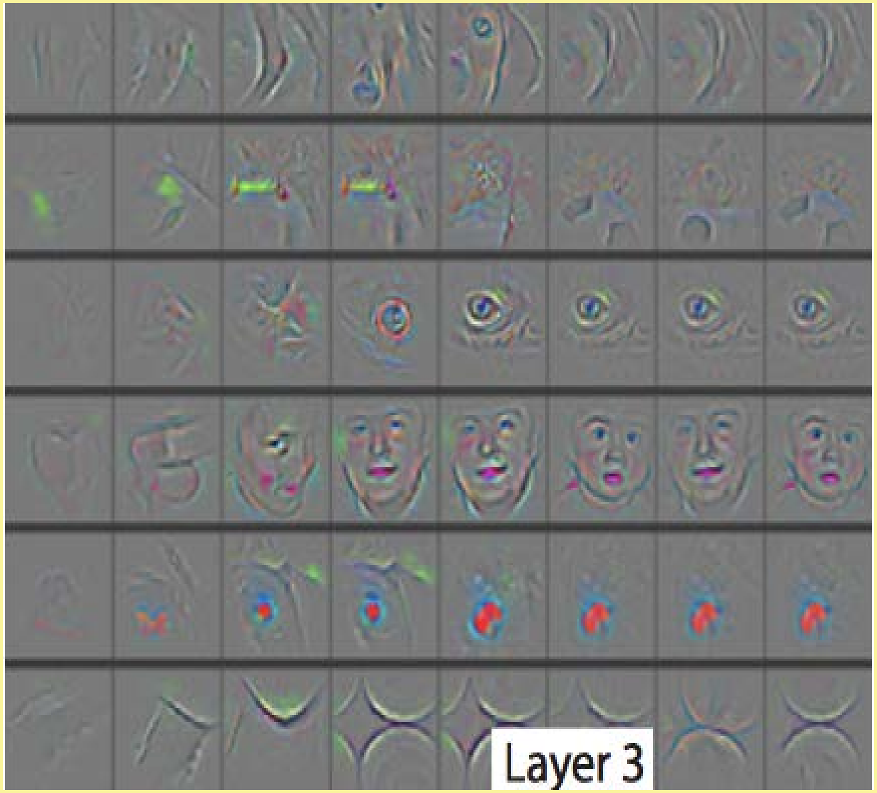
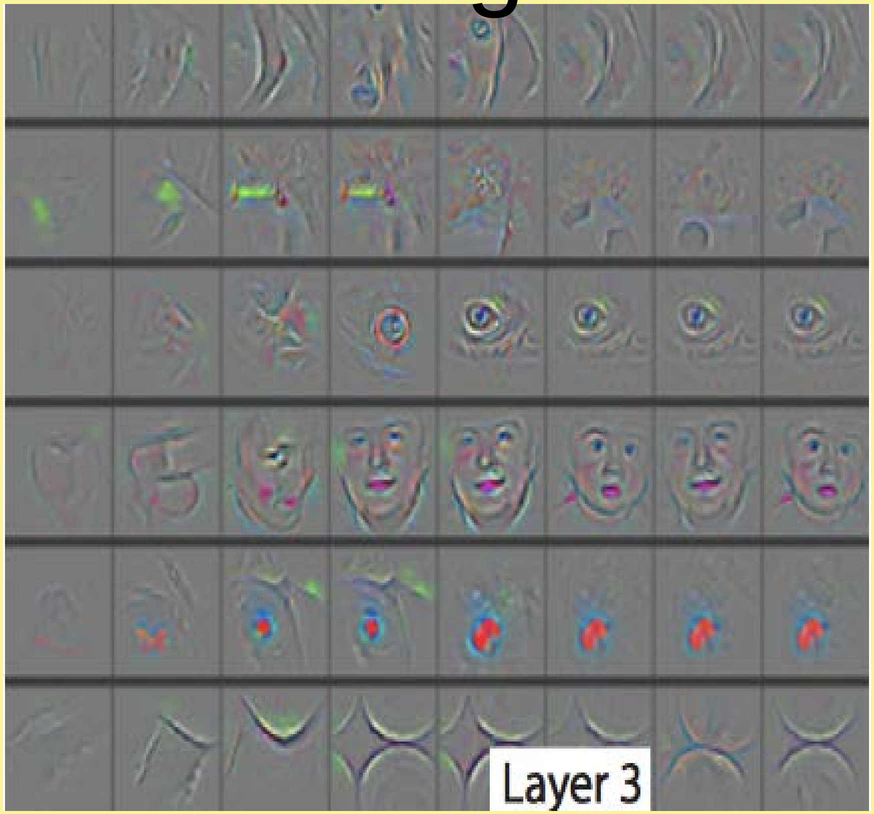


Image Source:-
https://becominghuman.ai/what-exactly-does-cnn-see-4d436d8e6e52

Transfer Learning



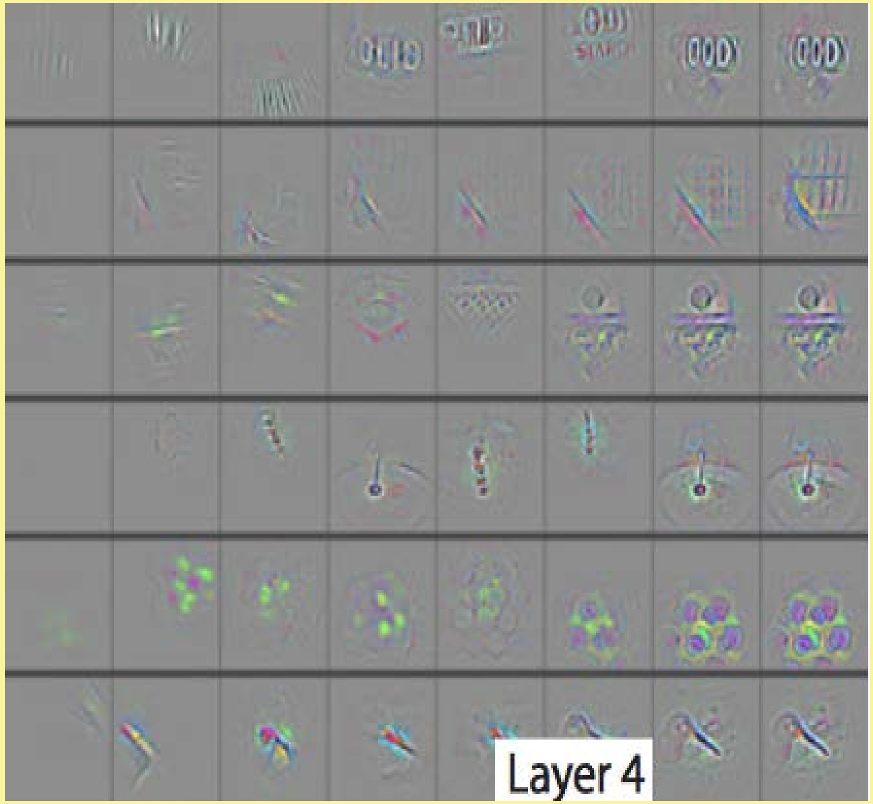
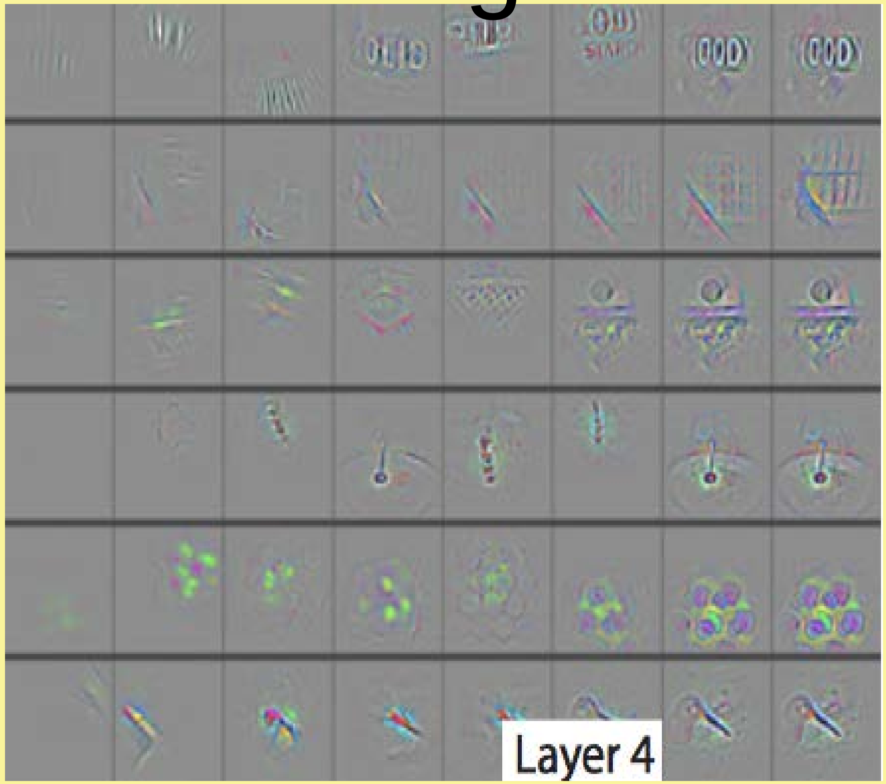


Image Source:- https://becominghuman.ai/what-exactly-does-cnn-see-4d436d8e6e52

Transfer Learning



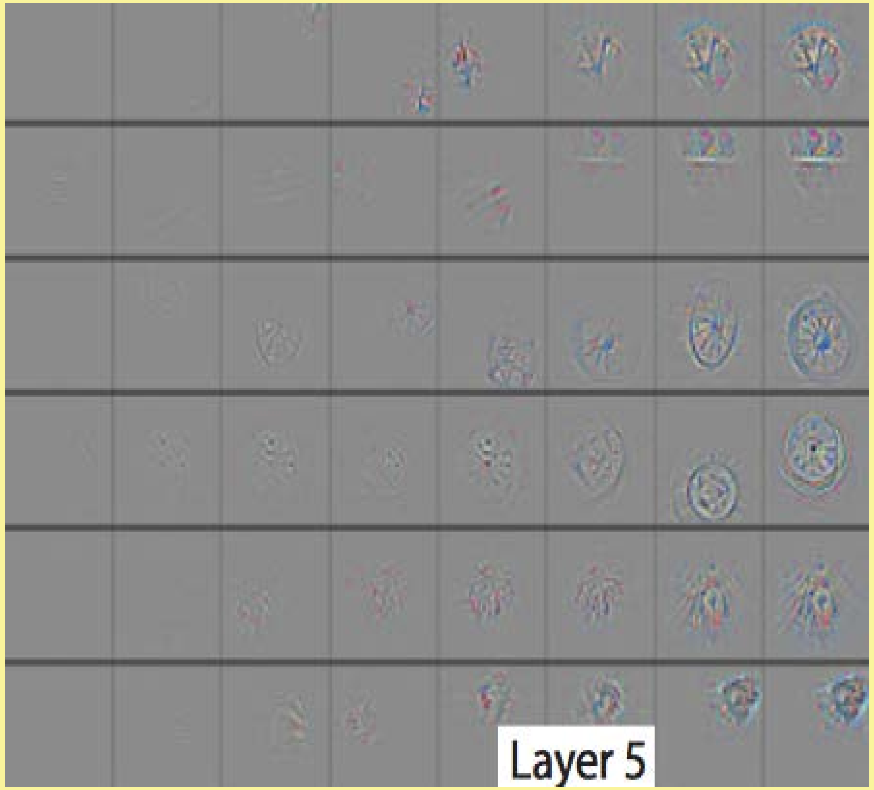


Image Source:- https://becominghuman.ai/what-exactly-does-cnn-see-4d436d8e6e52

Transfer Learning

- Lower layers generate more general features:- knowledge transfers very well to other tasks.
- Higher layers are more task specific.
- Fine-tuning improves generalization when sufficient examples are available.
- Transfer learning and fine tuning often lead to better performance than training from scratch on the target dataset.
- Even features transferred from distant tasks often perform better than random initial weights.

Fine tuning

- Weights of the pre-trained CNN is fine-tuned for the new dataset by continuing the back propagation.
- Fine-tuning can be done for all layers.
- Due to overfitting concern, the earlier layers of the net may be fixed and fine tuning is done only on the higher layers.
- Earlier layers can be fixed as lower layers extract features that are more generic.
- Higher layers on the other hand are task specific.

NPTEL ONLINE CERTIFICATION COURSES

Thank you