THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學

COMP4434 Big Data Analytics

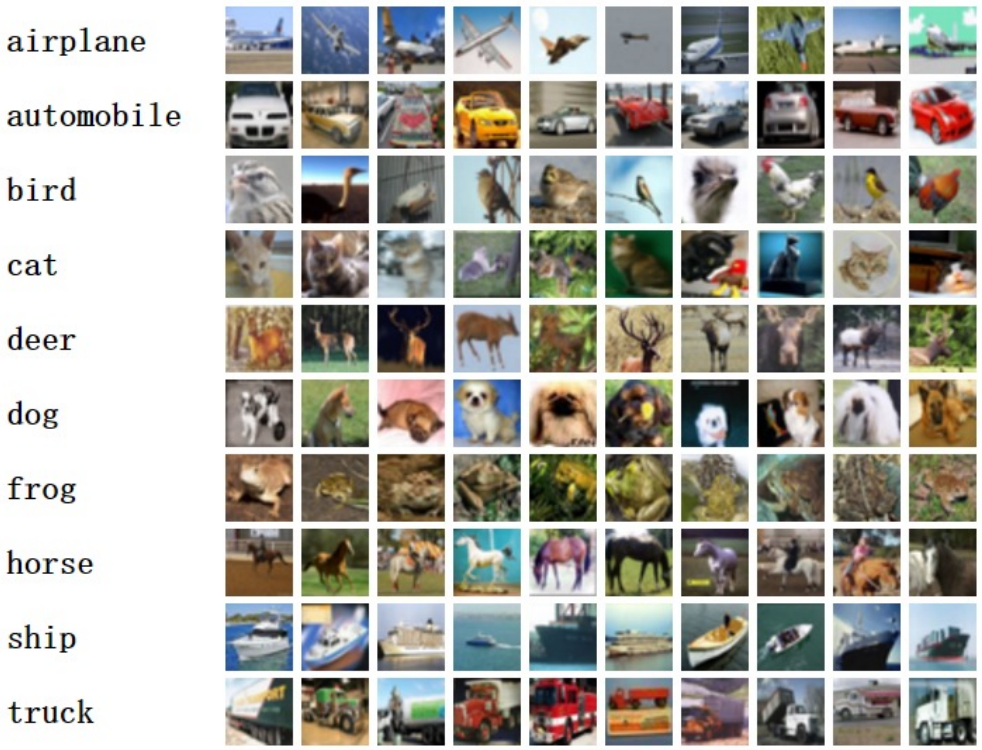
Lab 7
Convolutional Neural Networks

HUANG Xiao
xiaohuang@comp.polyu.edu.hk

CNN-based Classification

- A multi-class image dataset are considered for classification.
  - Contain training and test set.
- The CNN classifier is trained on training set to classify the images on test set with more propability.
- Typical CNN image generally contains:
  - Input layer (after preprocessing)
  - Convolutional layer
  - Pooling layer
  - Fully connected layer
  - Output layer

CIFAR10 dataset and state-of-the-art

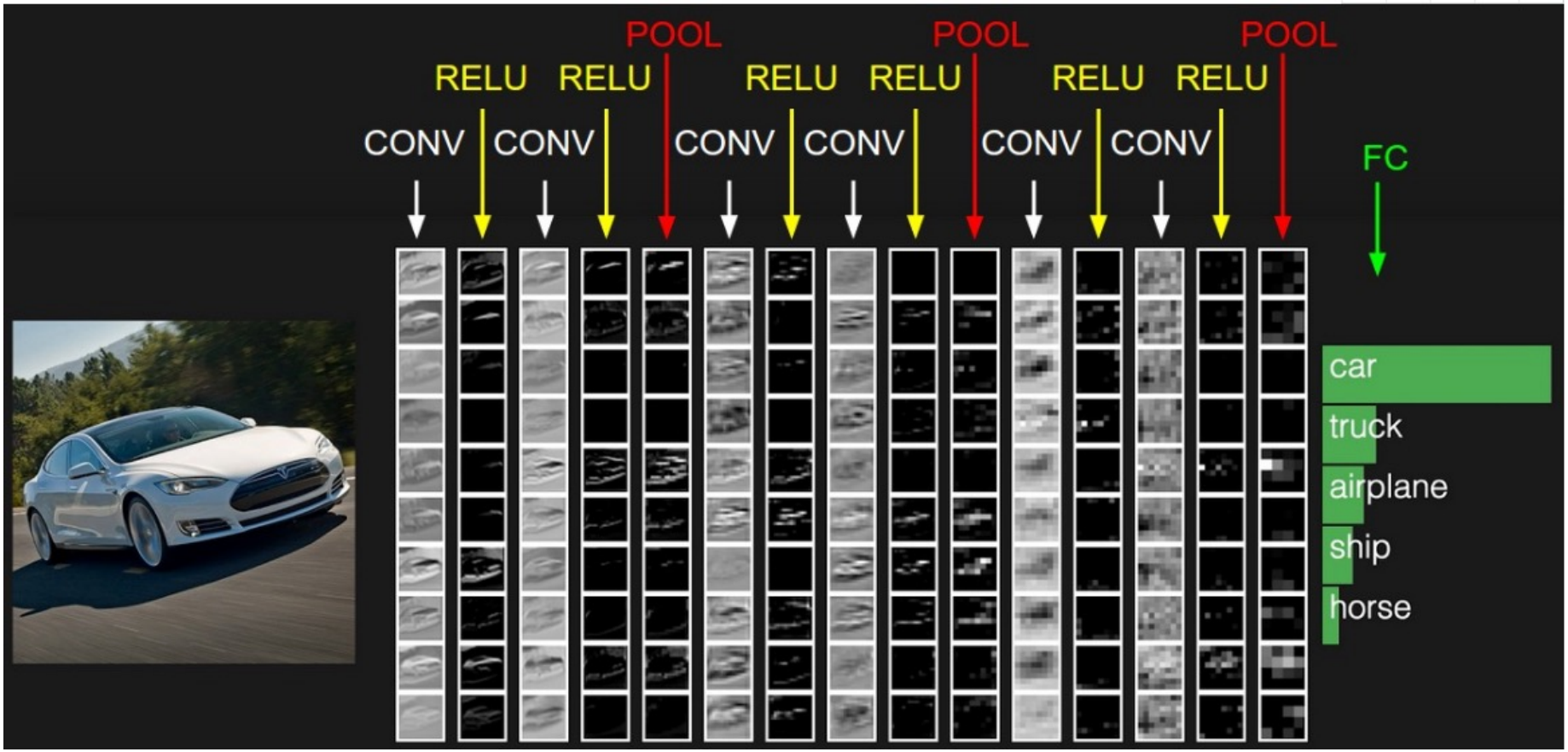


Accuracy

Model	Acc.
VGG16	92.64%
ResNet18	93.02%
ResNet50	93.62%
ResNet101	93.75%
ResNeXt29(32x4d)	94.73%
ResNeXt29(2x64d)	94.82%
DenseNet121	95.04%
PreActResNet18	95.11%
DPN92	95.16%

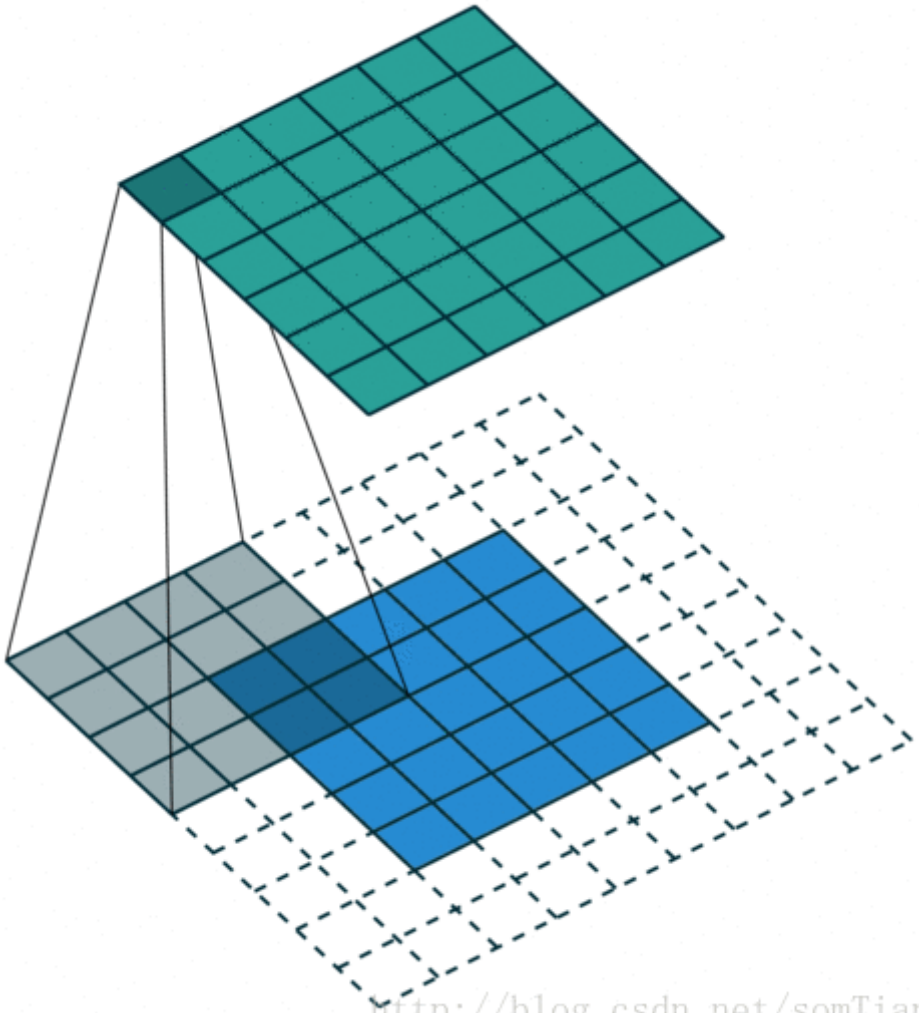
- The CIFAR-10 dataset consists of 60000 32x32x3 color images in 10 classes
- with 6000 images per class. There are 50000 training images and 10000 test images

A Simple CNN Structure Overview



- CONV: Convolutional layer
- RELU: Activation function
- POOL: Pooling layer
- FC: Fully connection layer

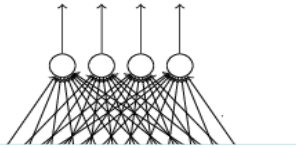
Recap: Convolutional kernel



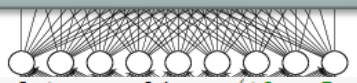
- A convolutional layer has a number of filters that does convolutional operation
- This image show the convolutional operation for one filter
- Each filter detects a small pattern and learns its parameter

CNN-based Classifier

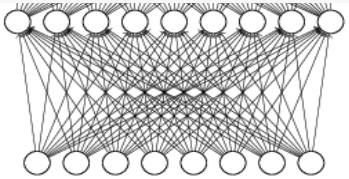
Output



Fully connected feedforward network



```
self.fc1 = nn.Linear(16 * 5 * 5, 120)
self.fc2 = nn.Linear(120, 84)
self.fc3 = nn.Linear(84, 10)
```



```
self.conv1 = nn.Conv2d(3, 6, 5)
```

Convolution

```
self.pool = nn.MaxPool2d(2, 2)
```

Max Pooling

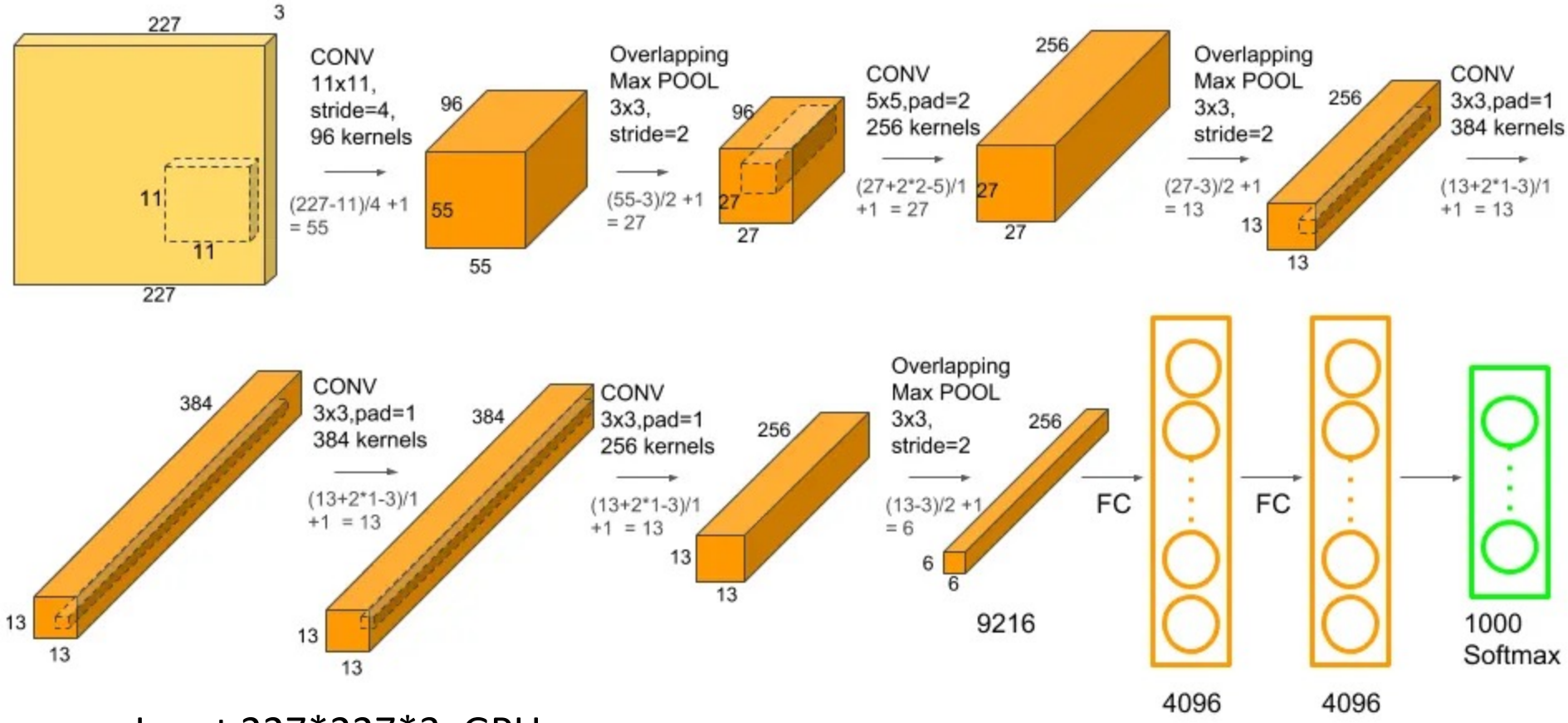
```
self.conv2 = nn.Conv2d(6, 16, 5)
```

Convolution

Flattened

```
x = torch.flatten(x, 1)
```

Convolutional Neural Networks in 2012



- Input 227*227*3. GPU.
- AlexNet: a layered model composed of convolution, subsampling, and further operations followed by a holistic representation and all-in-all a landmark classifier on ImageNet Large Scale Visual Recognition Challenge 2012
- + data; + gpu; + non-saturating nonlinearity; + regularization