INT 307
Multimedia Security System

Neural Network and Adversarial Attack I

Sichen.Liu@xjtlu.edu.cn

XJTLU | SCHOOL OF FILM AND TV ARTS

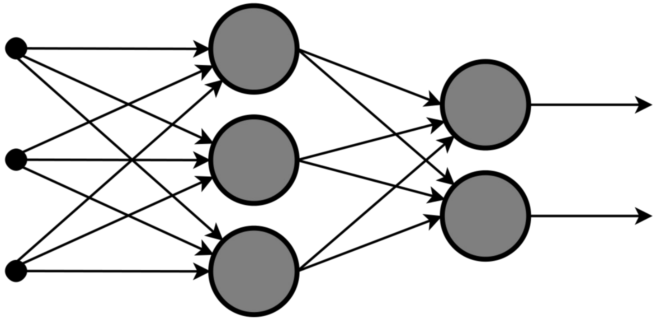
Xi'an Jiaotong-Liverpool University
西交利物浦大学

Aims

- Master the working principle of deep learning
- Understand basic knowledge related to deep learning

Recall INT104

- The boundaries between classes are not necessary linear but can be approximate as a combination of single layers.



- Could be single layer or multiple layer
- There is a threshold process after the output of each neuron, which is named as activation function

Artificial Neural Networks

- Data
- Input Layer
- Hidden Layer
- Output Layer

- Feature Extraction
- Classification

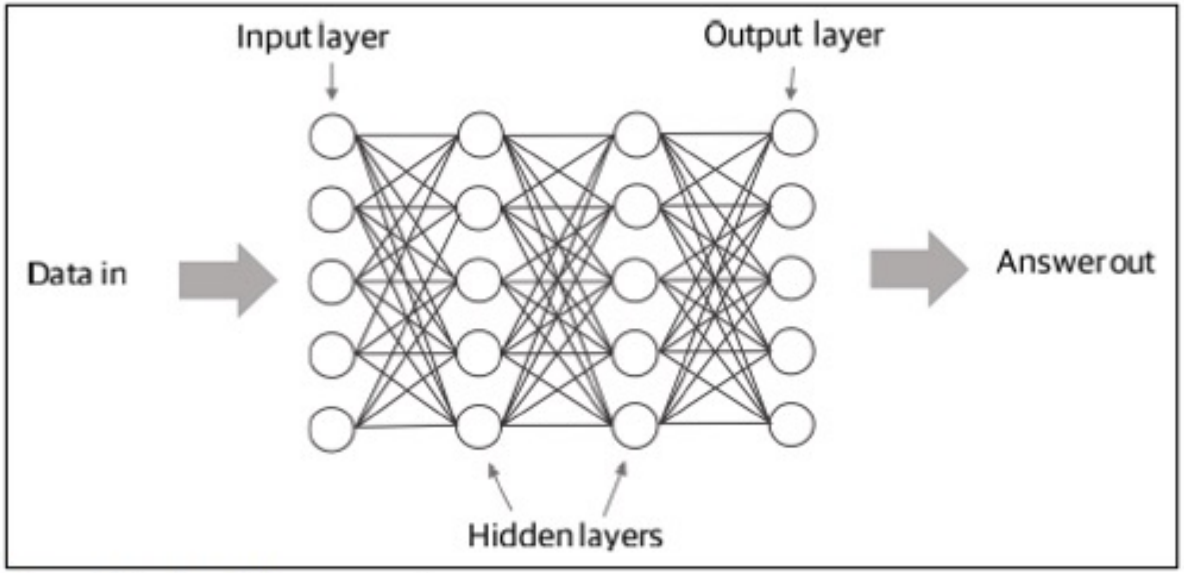


Figure 3-1. A multilayer perceptron

Forward Propagation

- Neurons effectively represent a mapping between feature spaces
- In neural networks, the mapping is represented as weighted sums with activation functions

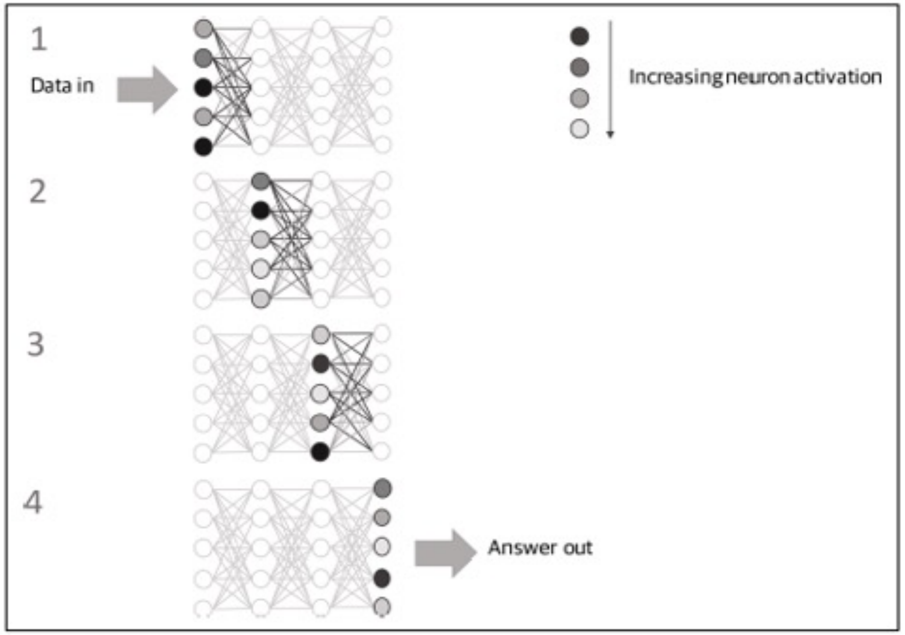


Figure 3-2. Incremental calculation steps in a multilayer perceptron

Forward Propagation

- Diagram 28 × 28
- 784 input neurons
- Two hidden layers with 56 neurons each
- RELU as activation functions

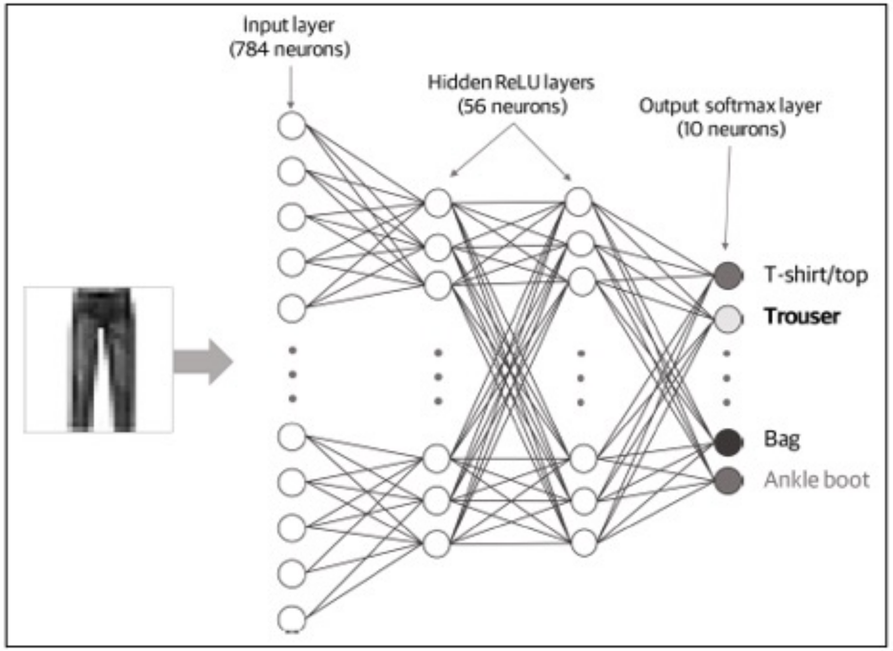
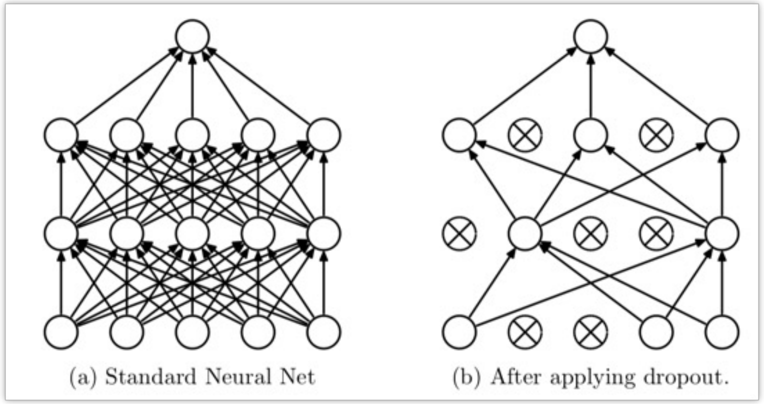


Figure 3-7. DNN architecture for image classification using Fashion-MNIST data

Common Tricks

- Dropout



- Normalization

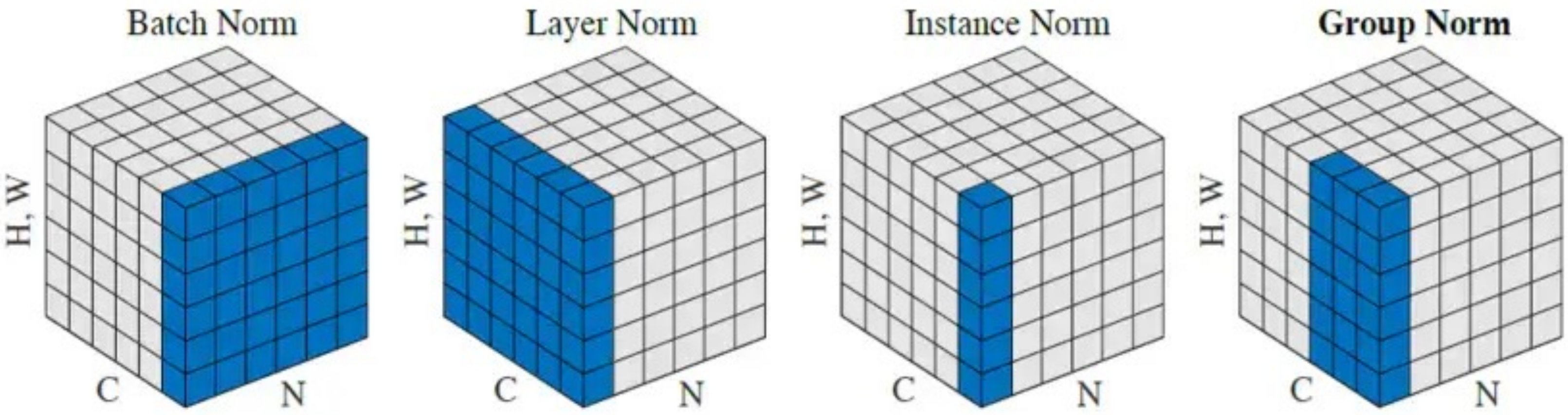
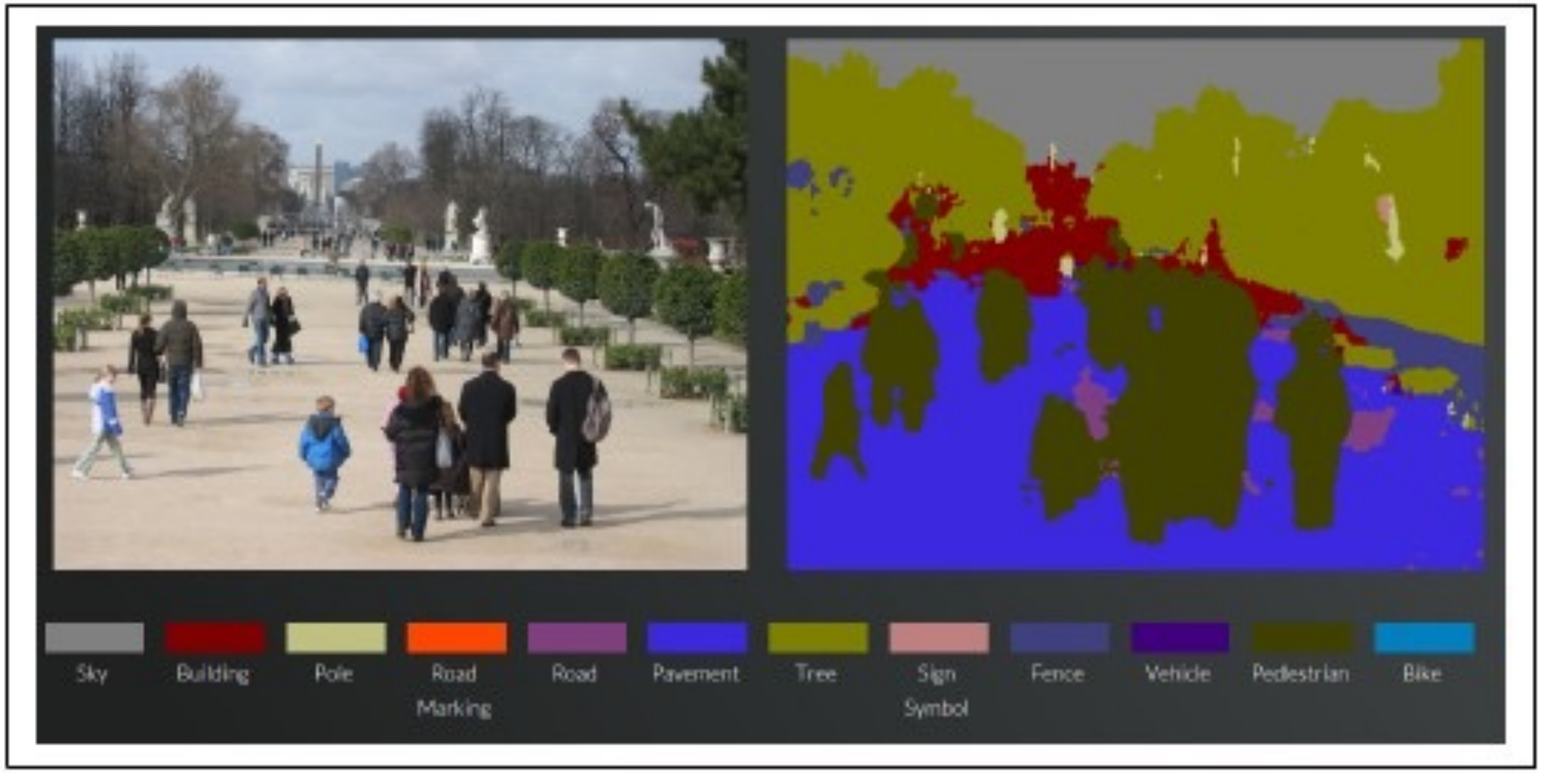
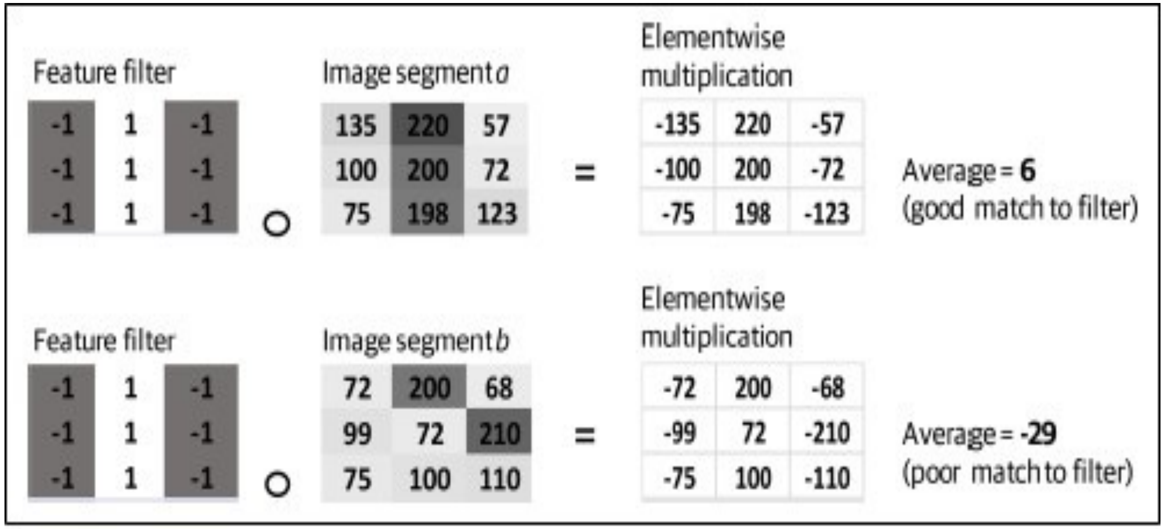


Image Processing with Deep Learning

- Scene classification
- Object detection and localisation
- Semantic segmentation
- Facial recognition



Filter and Convolution



Feature filter

-1	1	-1
-1	1	-1
-1	1	-1

○

Image segment *a*

135	220	57
100	200	72
75	198	123

=

Elementwise multiplication

-135	220	-57
-100	200	-72
-75	198	-123

Average = **6**
(good match to filter)

Feature filter

-1	1	-1
-1	1	-1
-1	1	-1

○

Image segment *b*

72	200	68
99	72	210
75	100	110

=

Elementwise multiplication

-72	200	-68
-99	72	-210
-75	100	-110

Average = **-29**
(poor match to filter)

Figure 4-4. Application of a simple 3 x 3 filter to two different image segments

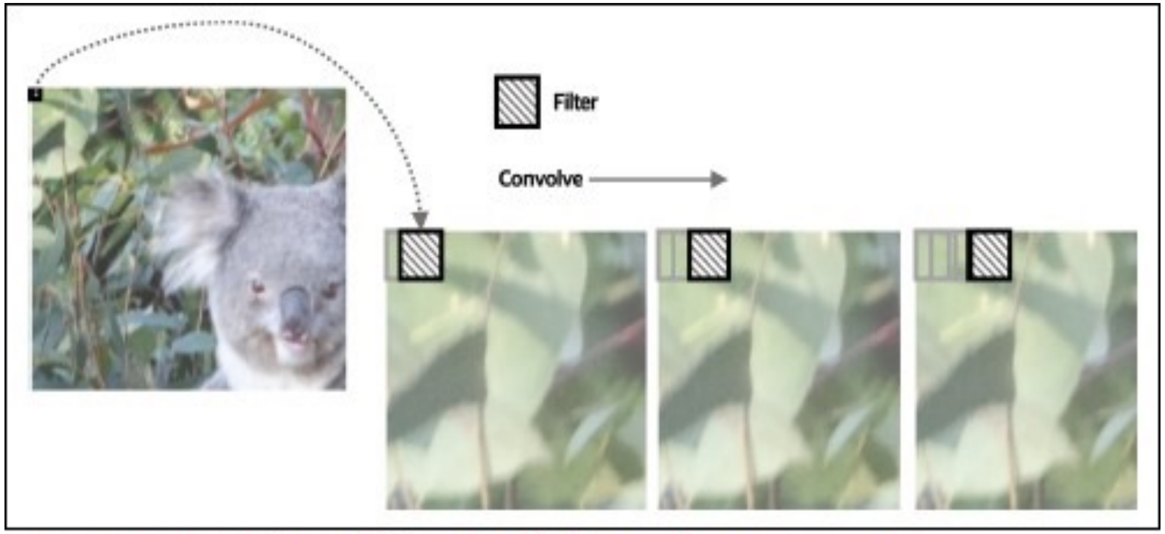


Figure 4-3. A convolutional filter is applied iteratively across an image

Convolutional Layers

- Kernel
  - Size
  - Padding
  - Stride
- Feature Maps

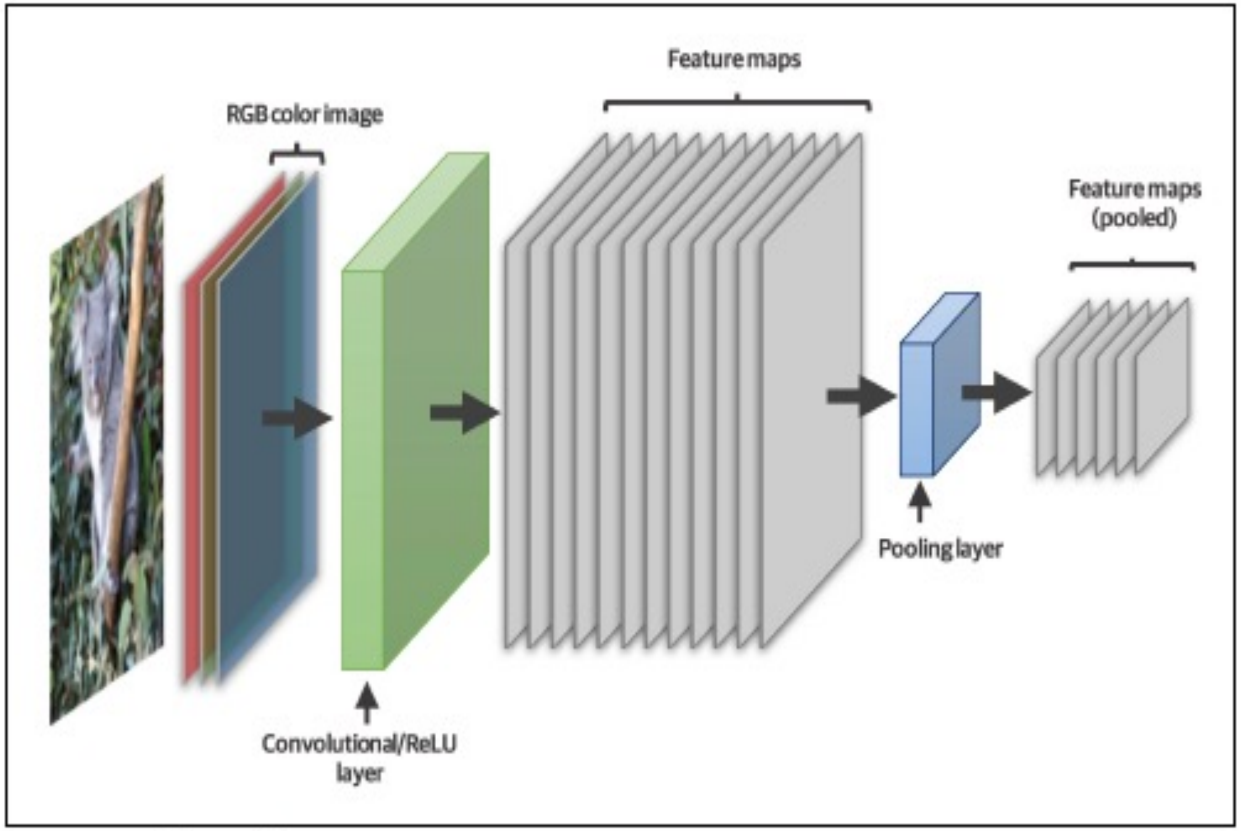


Figure 4-5. Typical layering pattern in a CNN

Convolutional Neural Network

- Convolutional Layers
- Pooling Layers
- Fully Connected Layers
- Classifier
- VGG
  - VGG-16
  - VGG-19

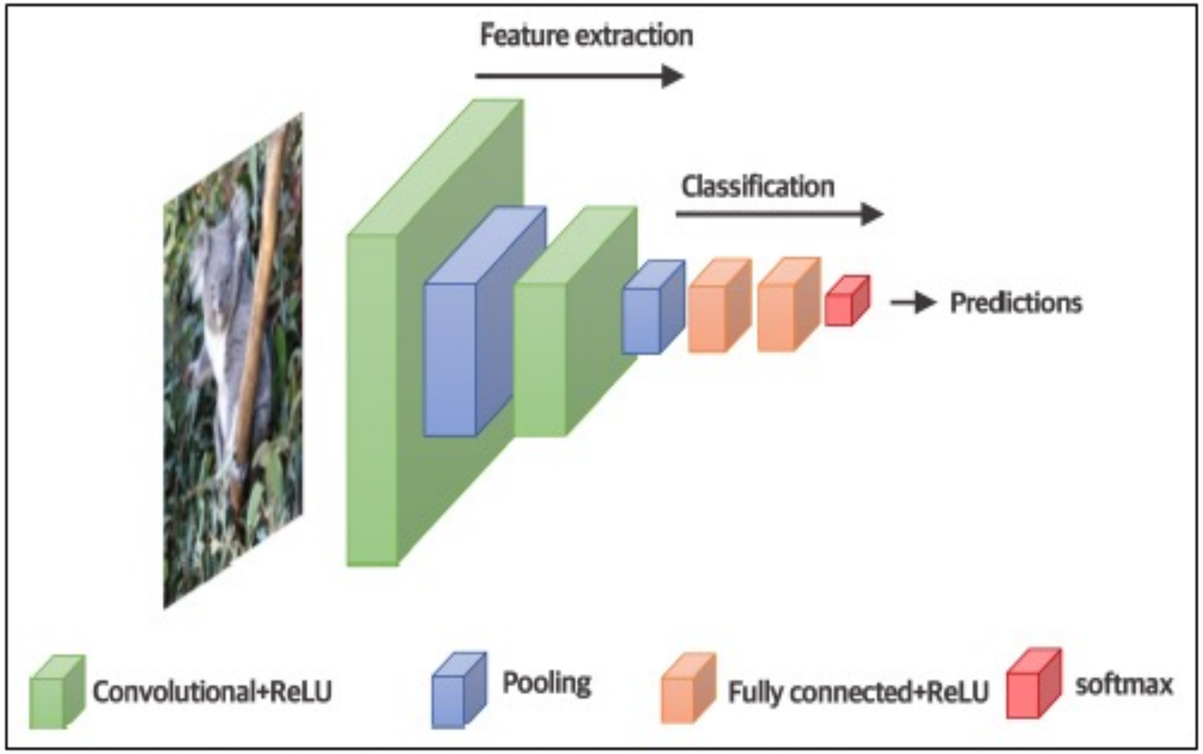
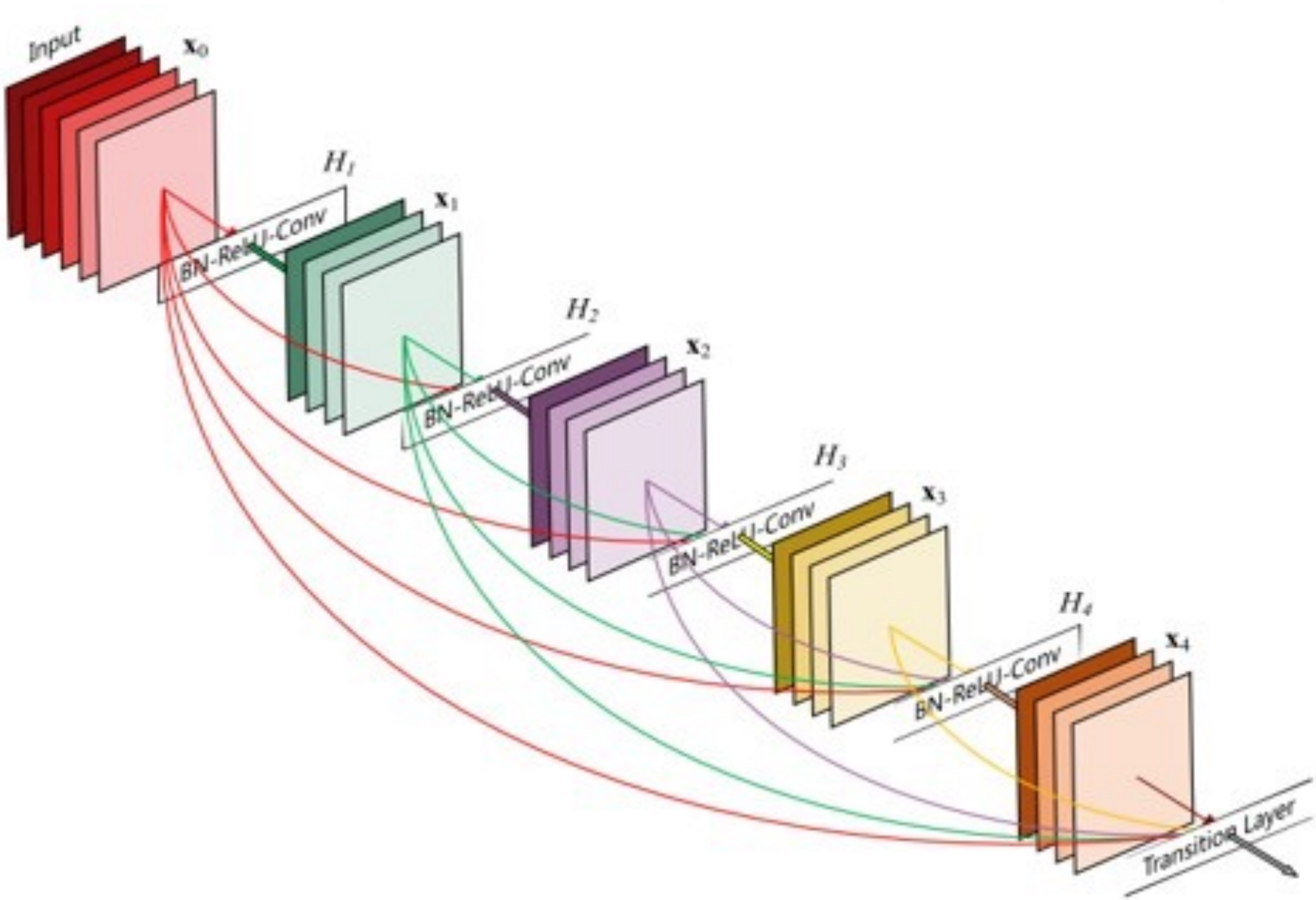


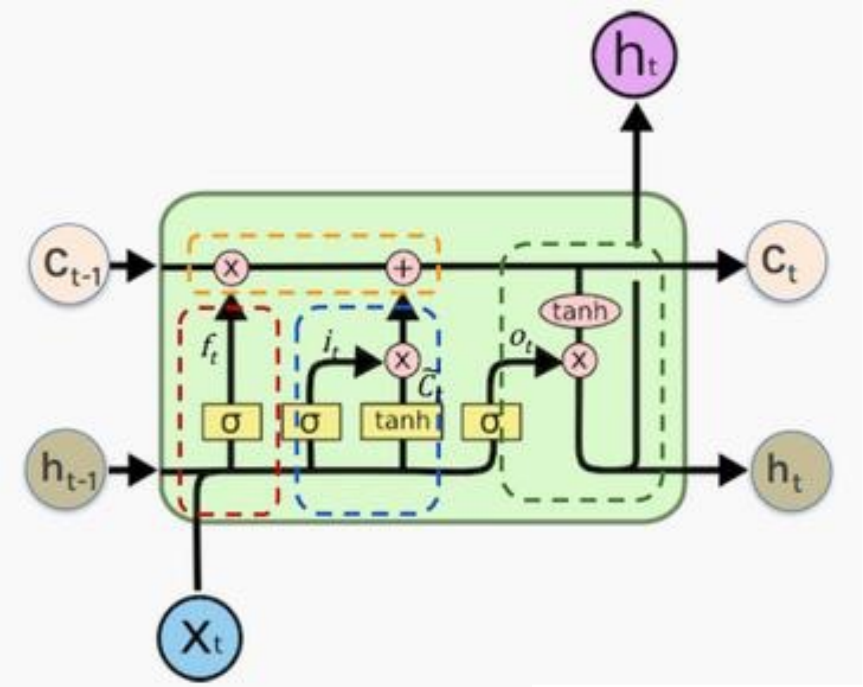
Figure 4-6. An example CNN image classification architecture

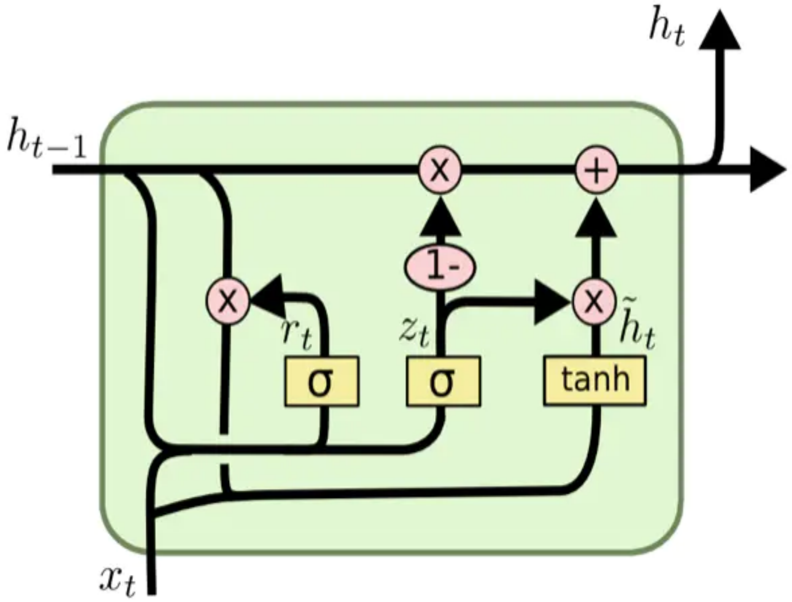
Residue Network



Recurrent Neural Network

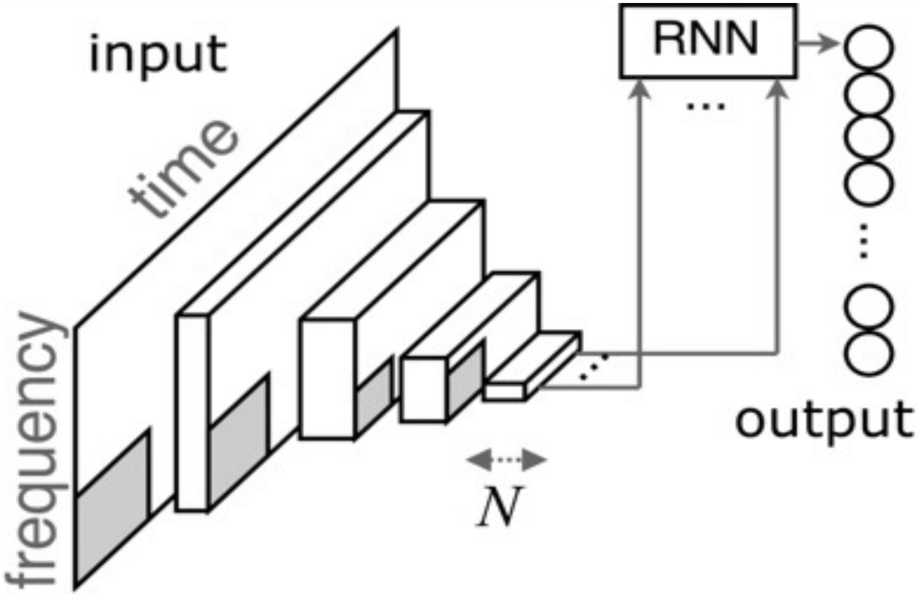
- Recurrent Neural Network is commonly used to process sequential media
- Commonly used transforms are:
  - LSTM (Long Short Time Memory)
  - GRU (Gated Recurrent Unit)



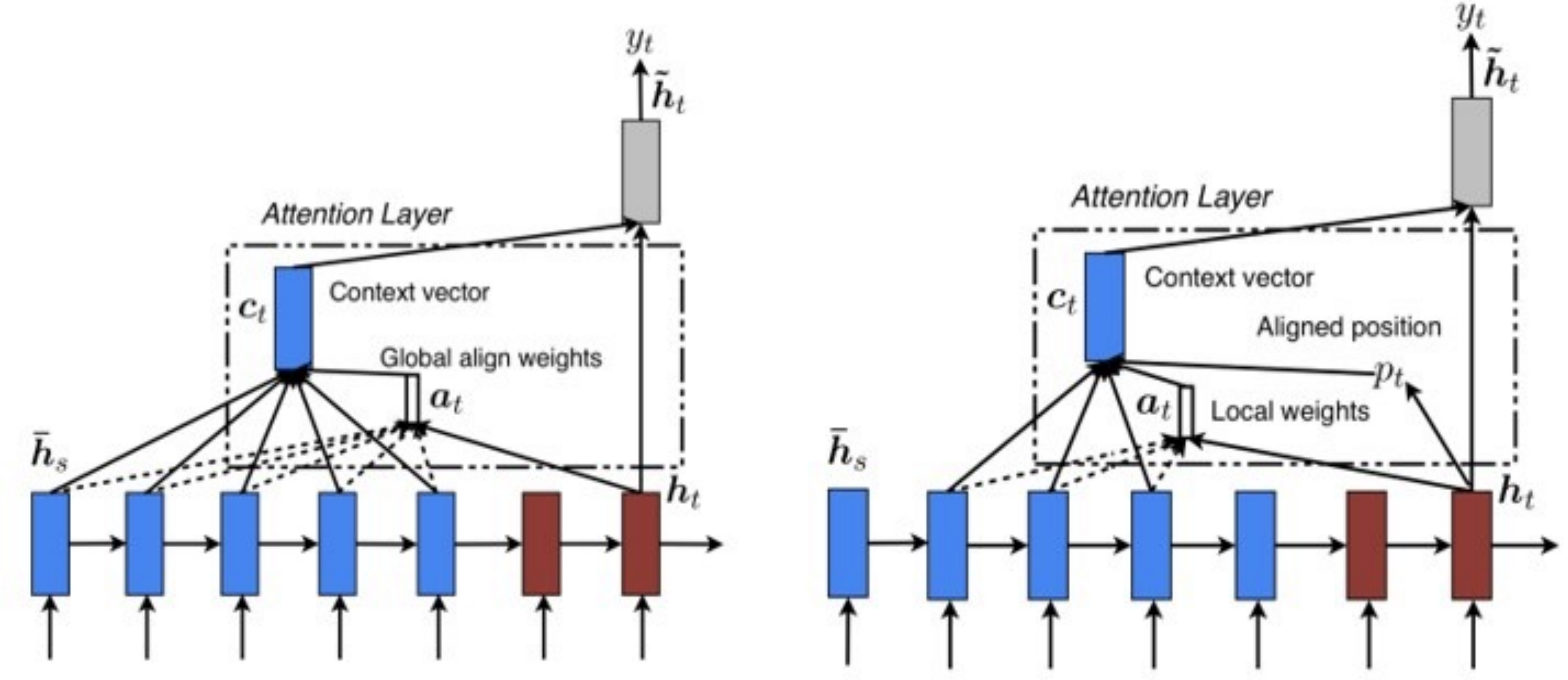


Complex Networks

- A deep learning neural network can combine multiple types of structures
    - CNN = CNN + DNN
    - CRNN = CNN + RNN + DNN
- Discussion: Why CRNN can be considered as a way to analysis signal in multi-scale?



Attention



Global Attention Model

Local Attention Model

Transformer

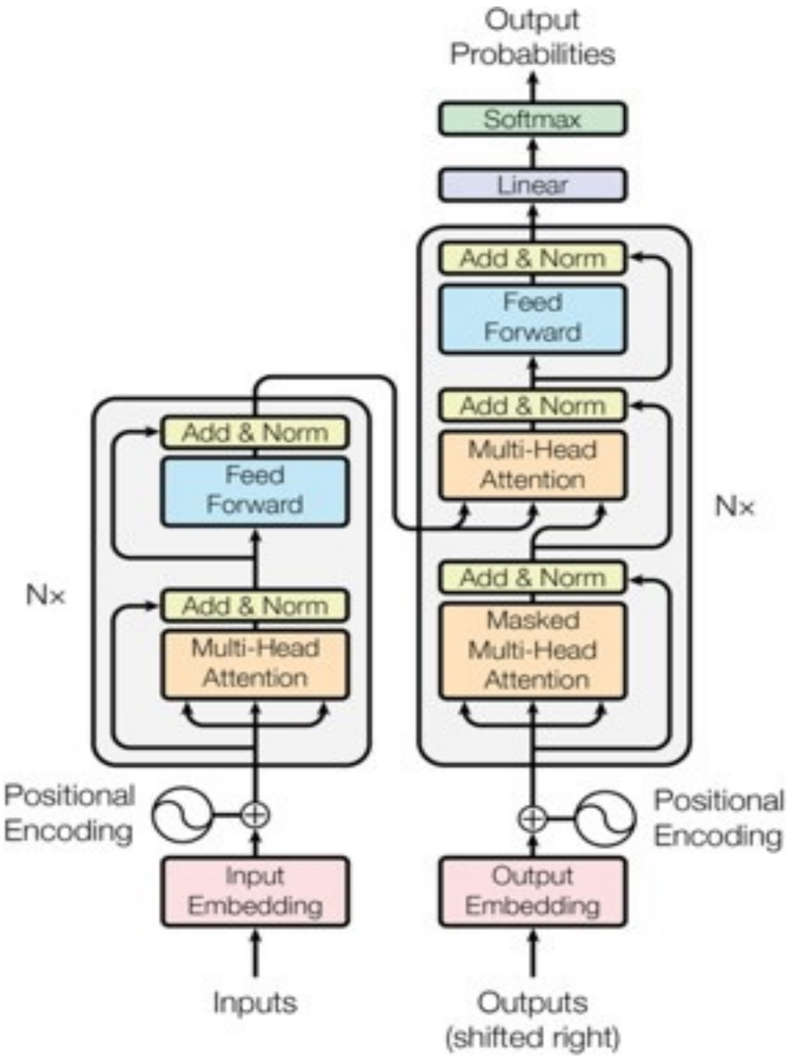
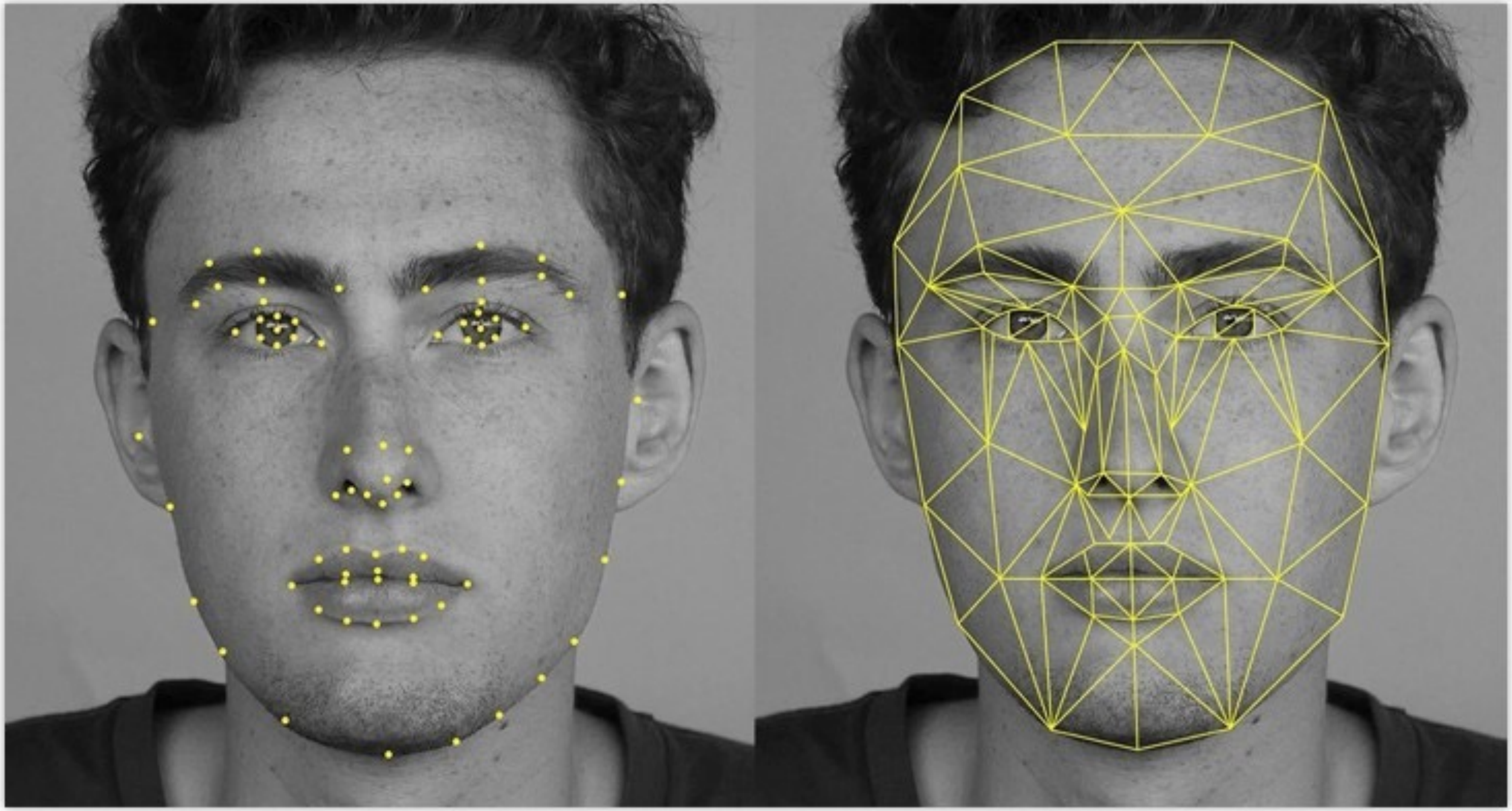
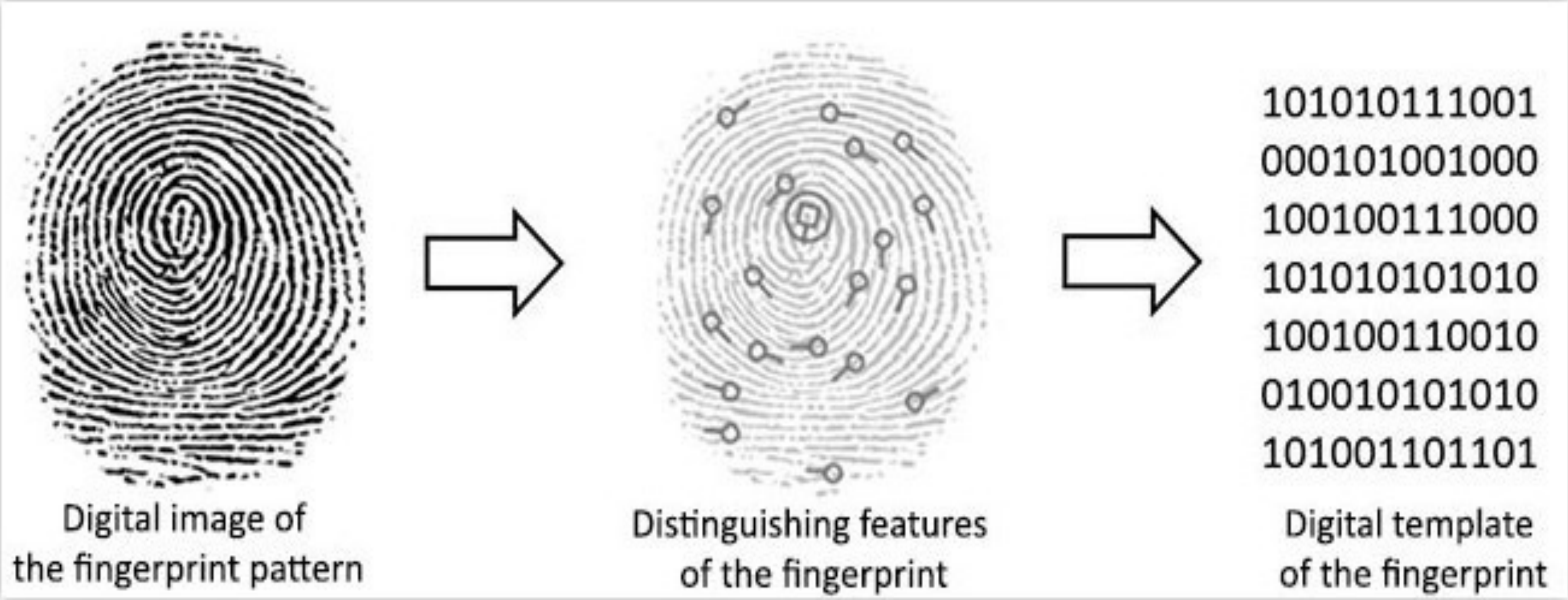


Figure 1: The Transformer - model architecture.

Face Recognition



Fingerprint Recognition

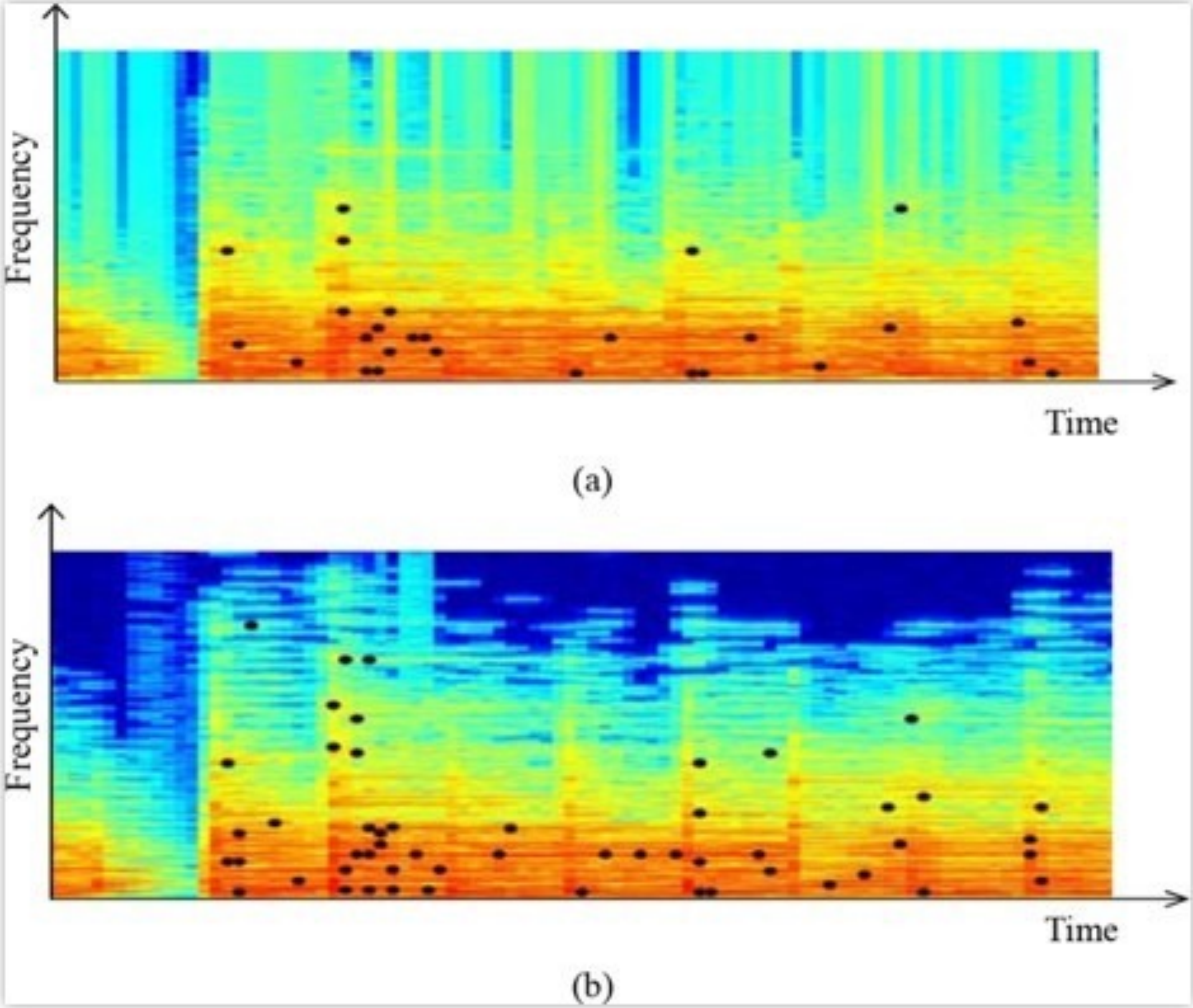


Digital image of the fingerprint pattern

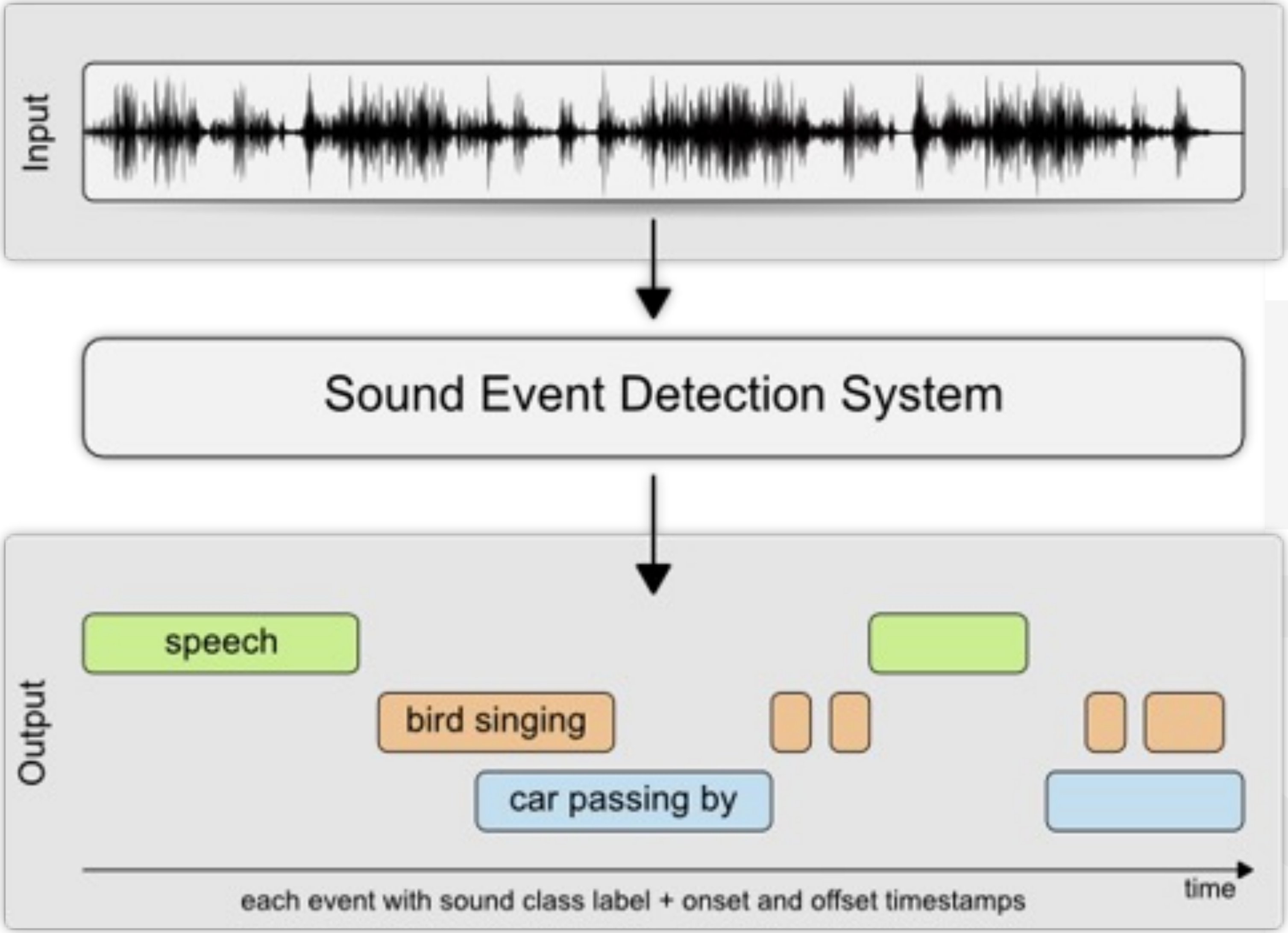
Distinguishing features of the fingerprint

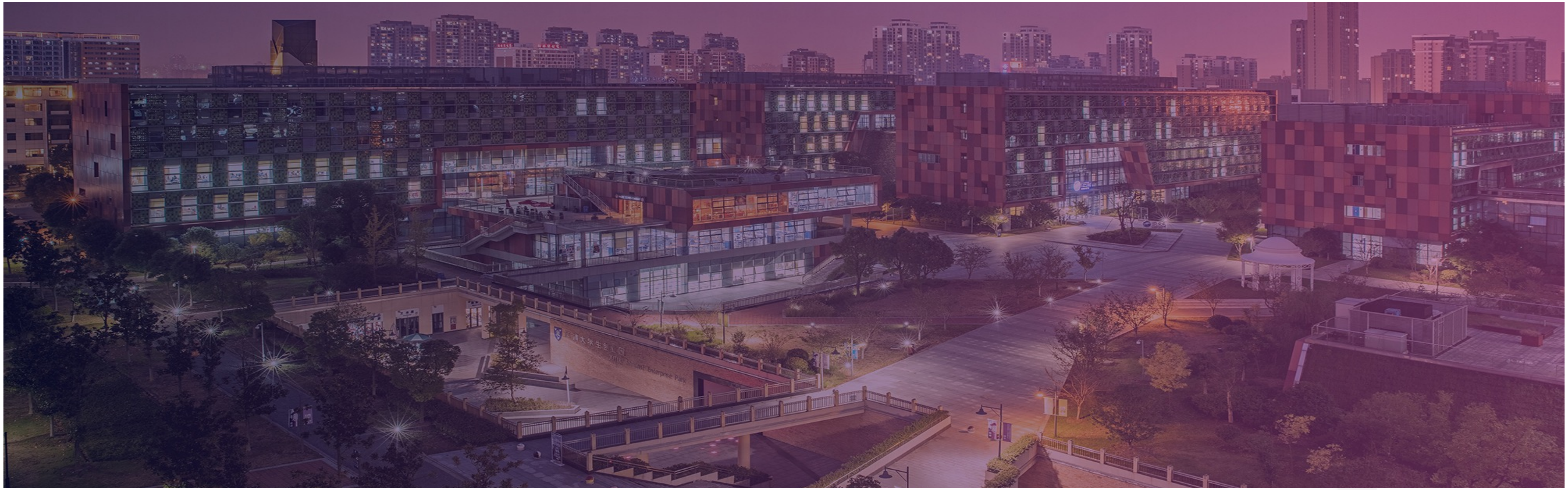
Digital template of the fingerprint

Audio Fingerprint



Audio Event Detection


Input
Sound Event Detection System
Output
speech
bird singing
car passing by
time
each event with sound class label + onset and offset timestamps



THANK YOU

VISIT US

WWW.XJTLU.EDU.CN

FOLLOW US

@XJTLU

Xi'an Jiaotong-Liverpool University
西交利物浦大学

XJTLU | SCHOOL OF FILM AND TV ARTS