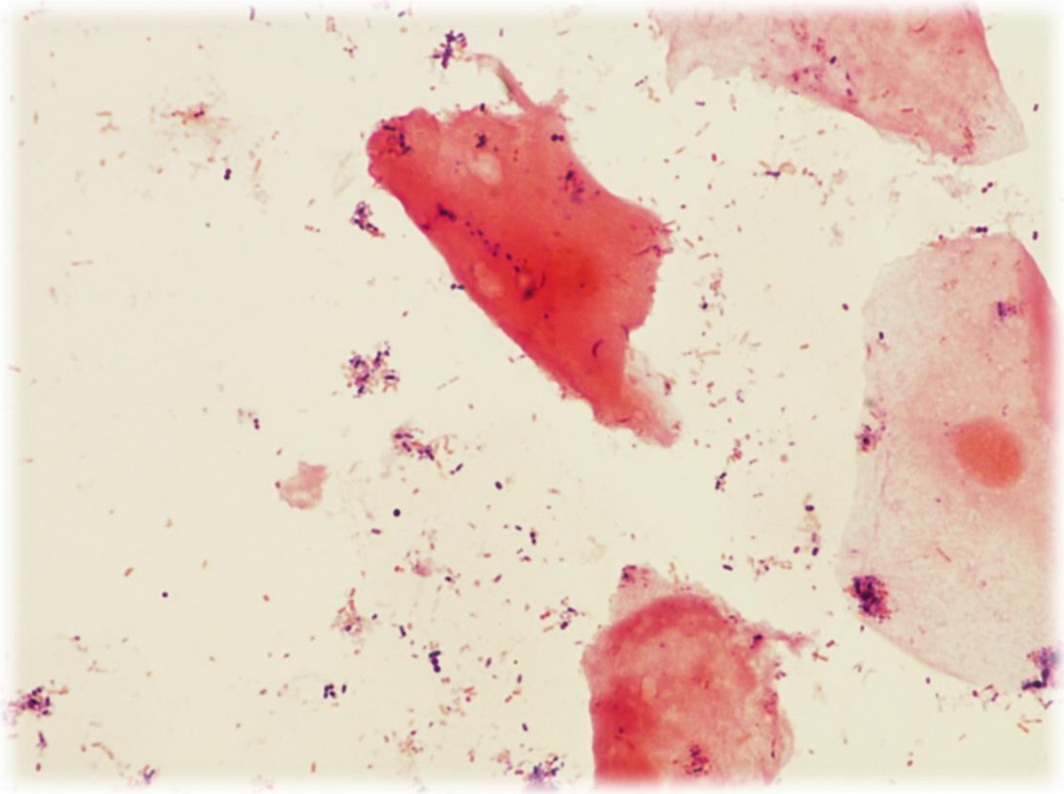
Michael Chi Ian Tang
3035209241

Classification for pathological images using machine learning

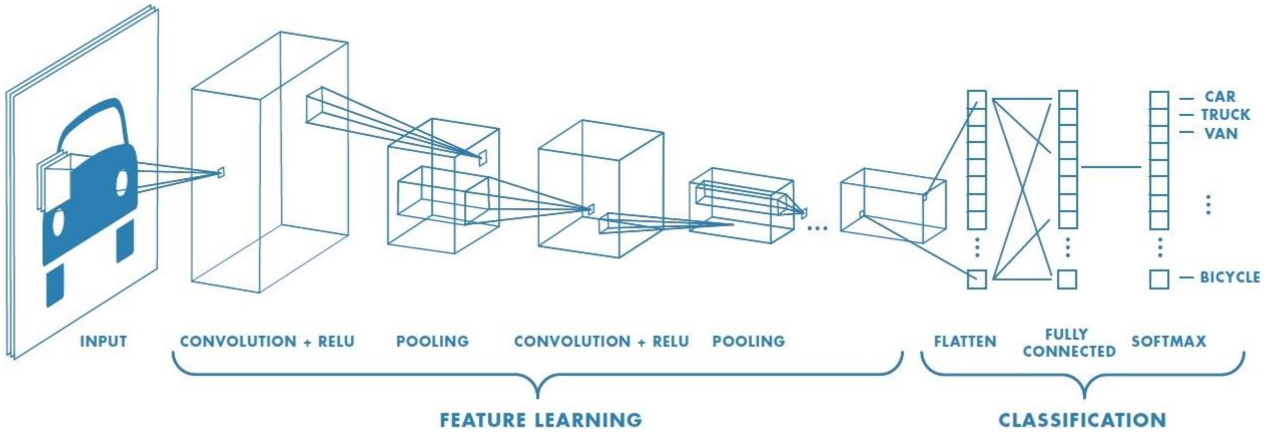
Outline

- **Digital Pathology**
- Convolutional neural networks (CNN)
- From CNN to Residual Network
- First prototype
- Second prototype
- Future direction



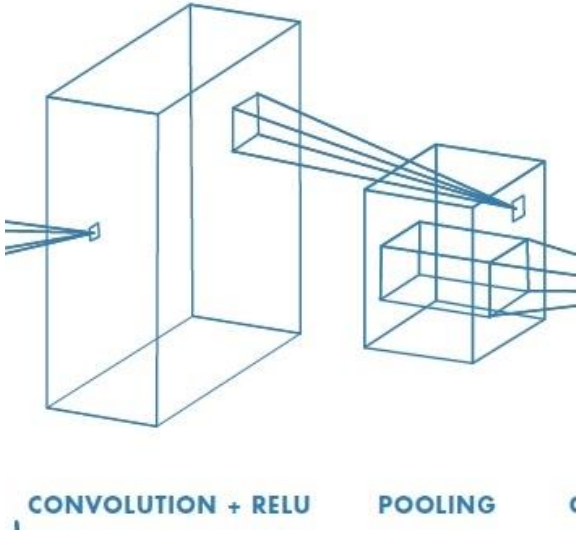
Outline

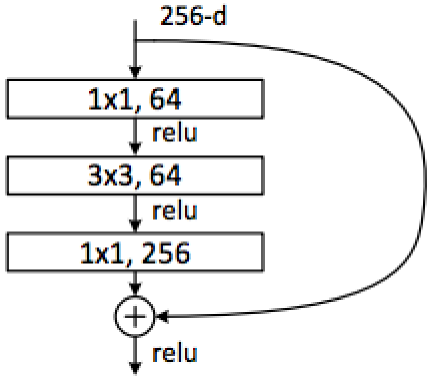
- Digital Pathology
- **Convolutional neural networks (CNN)**
- From CNN to Residual Network
- First prototype
- Second prototype
- Future direction



Outline

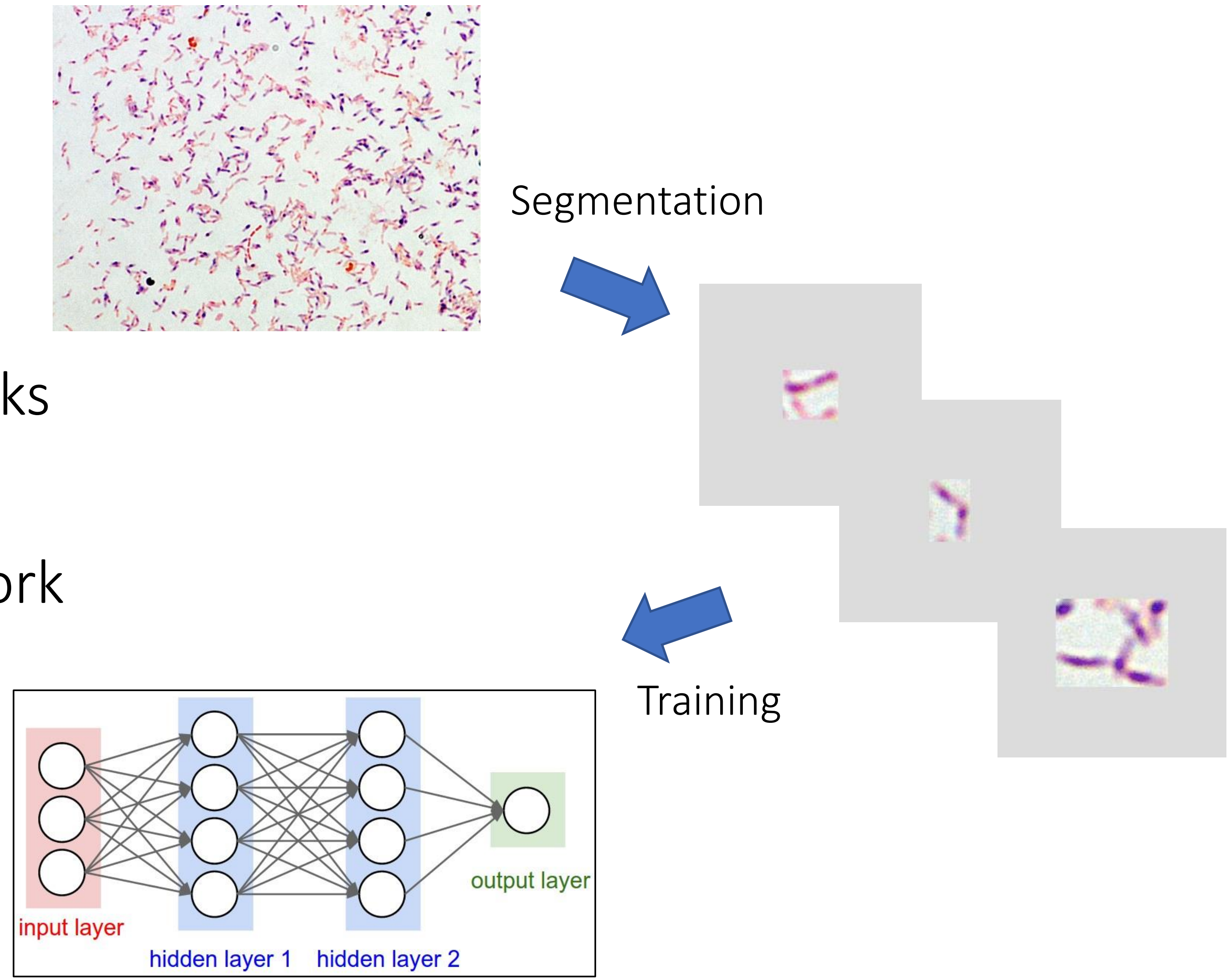
- Digital Pathology
- Convolutional neural networks (CNN)
- **From CNN to Residual Network**
- First prototype
- Second prototype
- Future direction





Outline

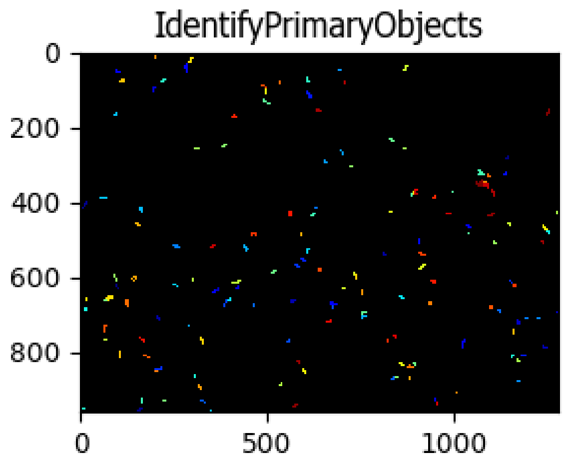
- Digital Pathology
- Convolutional neural networks (CNN)
- From CNN to Residual Network
- **First prototype**
- Second prototype
- Future direction

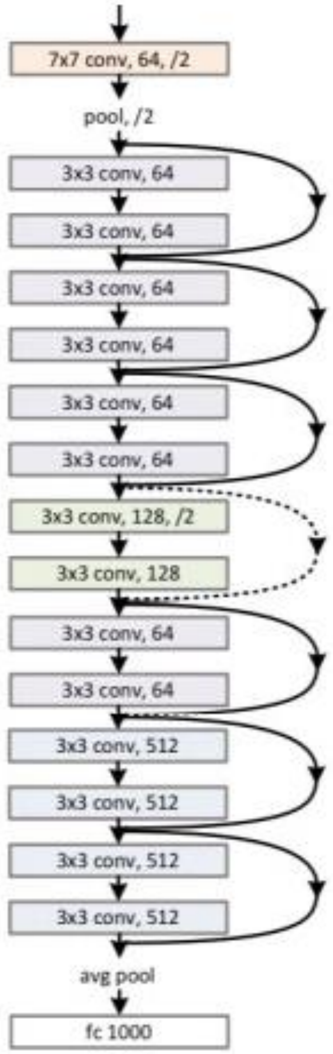


Machine Learning Model

Outline

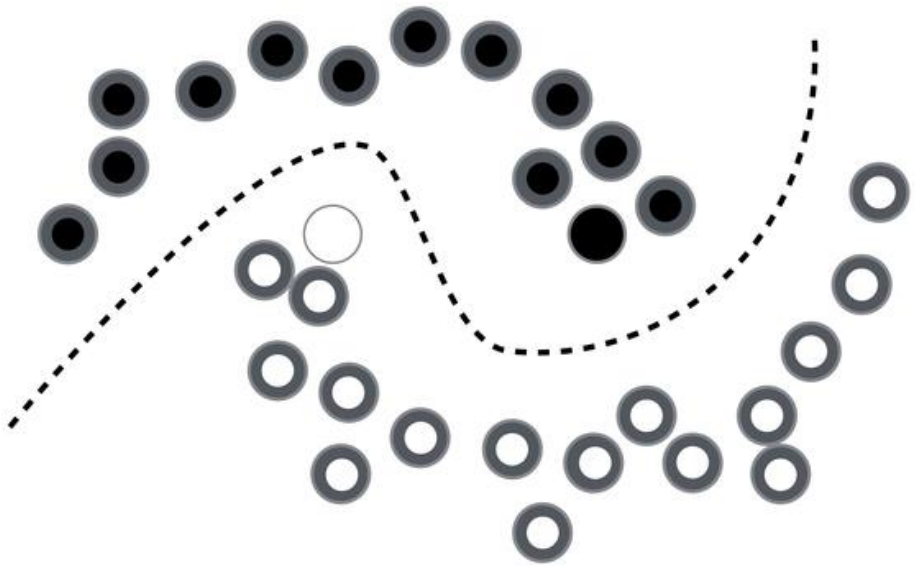
- Digital Pathology
- Convolutional neural networks (CNN)
- From CNN to Residual Network
- First prototype
- **Second prototype**
- Future direction





Outline

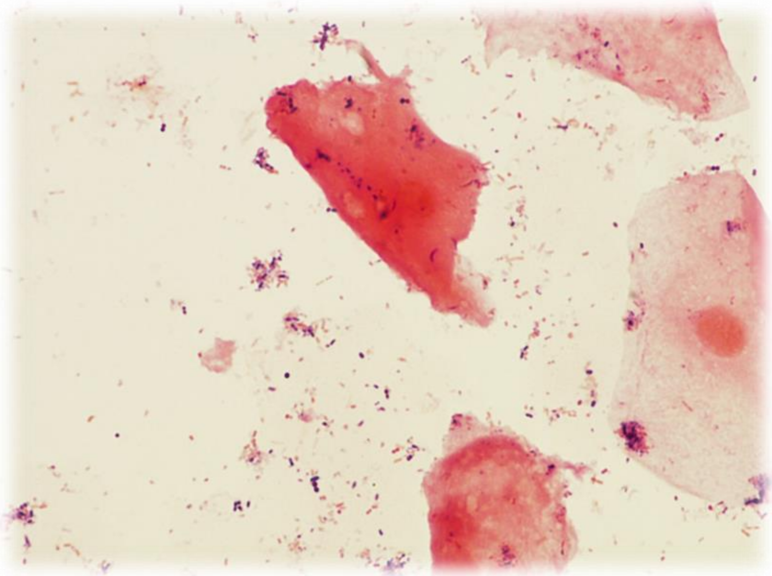
- Digital Pathology
- Convolutional neural networks (CNN)
- From CNN to Residual Network
- First prototype
- Second prototype
- **Future direction**



Digital Pathology

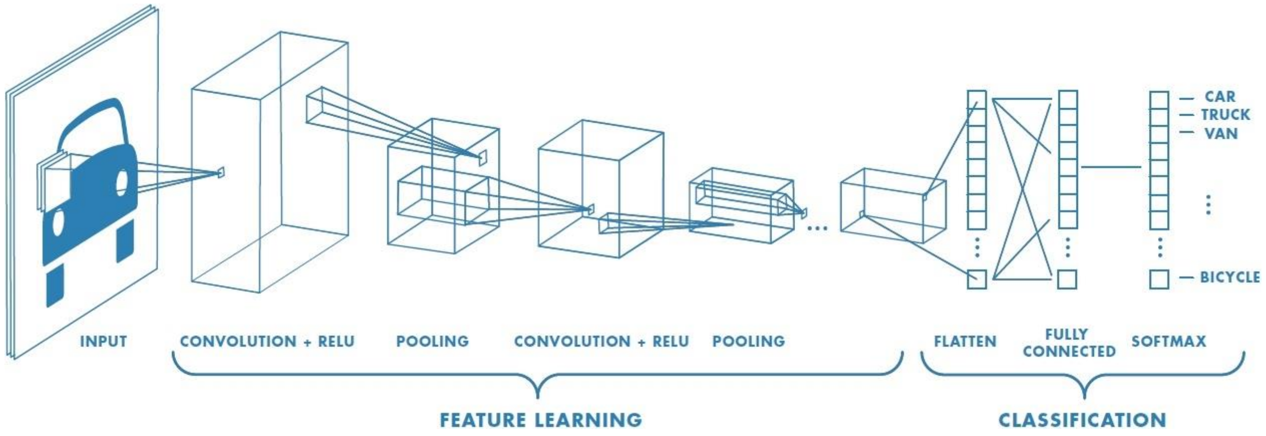
Highly accurate

Time-consuming, Expensive, Dependent on experience



Outline

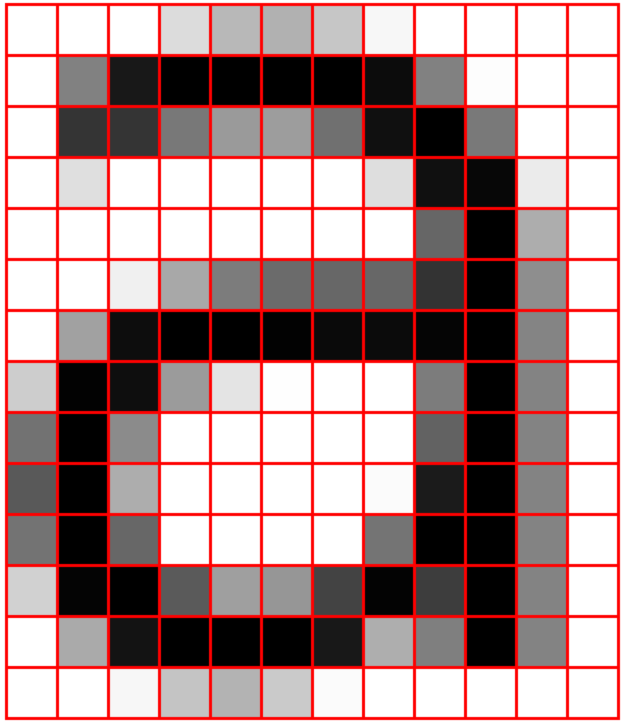
- Digital Pathology
- **Convolutional neural networks (CNN)**
- From CNN to Residual Network
- First prototype
- Second prototype
- Future direction



Convolutional Neural Networks

An image is just a collection of numbers stored in the memory

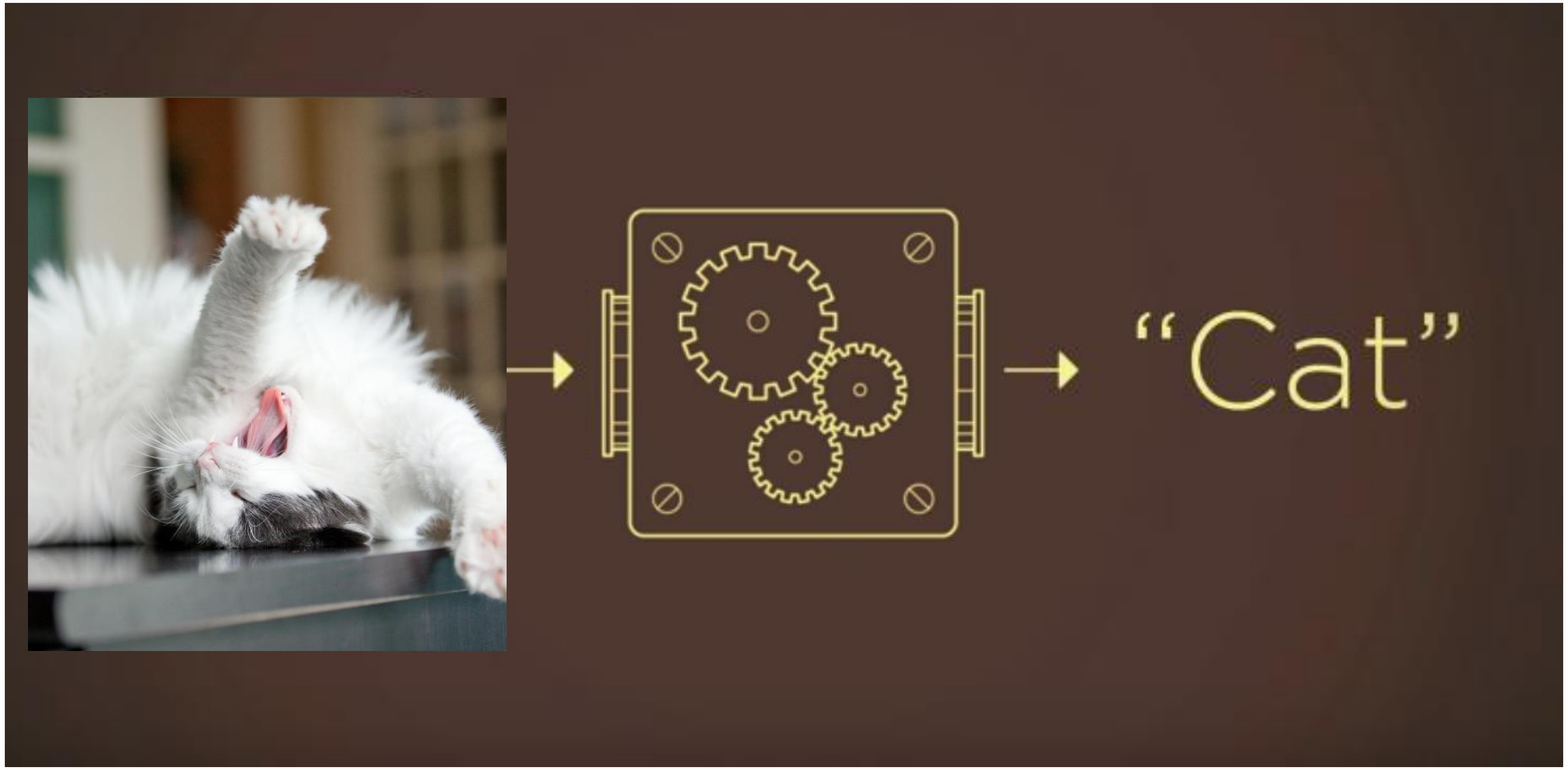
a



1.0	1.0	1.0	0.9	0.6	0.6	0.6	1.0	1.0	1.0	1.0	1.0
1.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0	1.0	1.0
1.0	0.2	0.2	0.5	0.6	0.6	0.5	0.0	0.0	0.5	1.0	1.0
1.0	0.9	1.0	1.0	1.0	1.0	1.0	0.9	0.0	0.0	0.9	1.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
1.0	1.0	1.0	0.5	0.5	0.5	0.5	0.5	0.4	0.0	0.5	1.0
1.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0
0.9	0.0	0.0	0.6	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
0.5	0.0	0.6	1.0	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
0.5	0.0	0.7	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.5	1.0
0.6	0.0	0.6	1.0	1.0	1.0	1.0	0.5	0.0	0.0	0.5	1.0
0.9	0.1	0.0	0.6	0.7	0.7	0.5	0.0	0.5	0.0	0.5	1.0
1.0	0.7	0.1	0.0	0.0	0.0	0.1	0.9	0.8	0.0	0.5	1.0
1.0	1.0	1.0	0.8	0.8	0.9	1.0	1.0	1.0	1.0	1.0	1.0

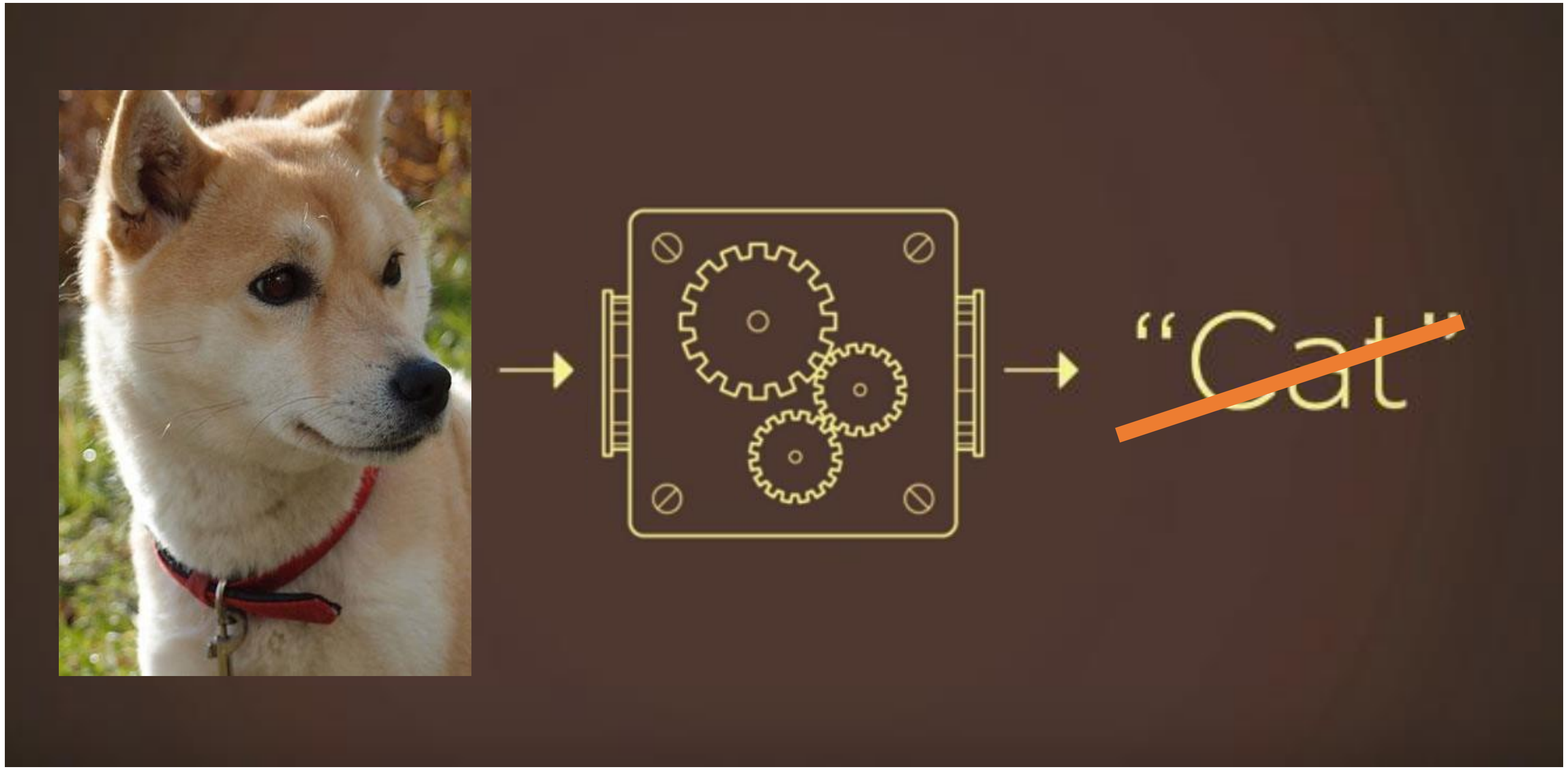
Convolutional Neural Networks

How to build a machine that recognizes objects?



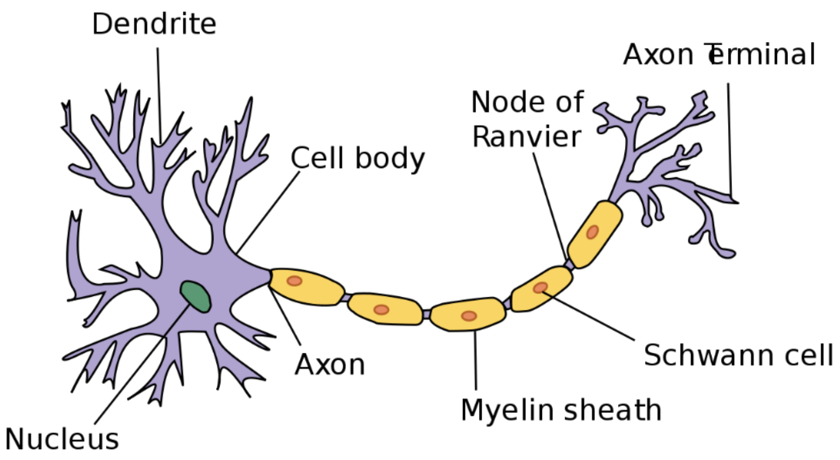
Convolutional Neural Networks

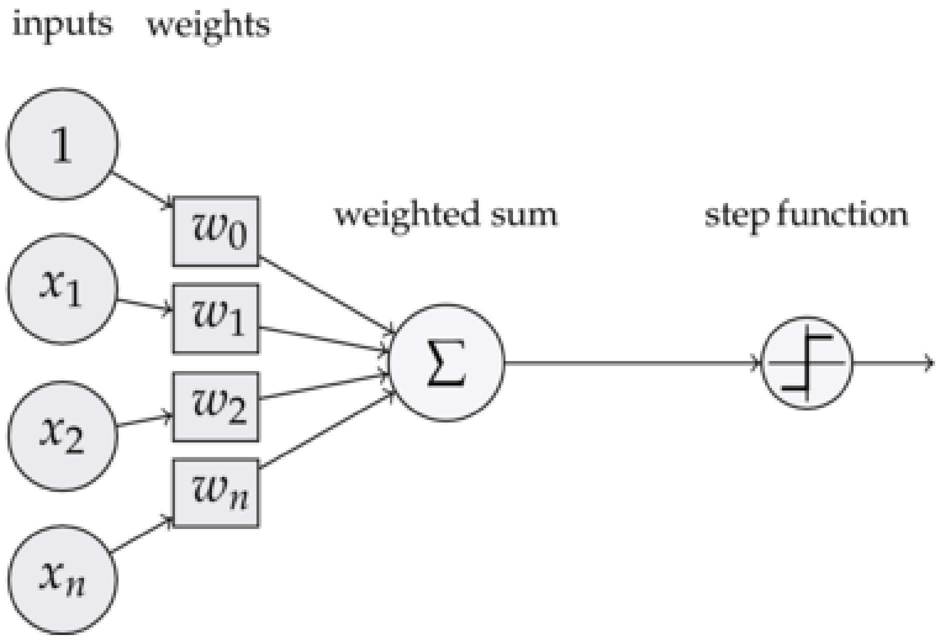
How to build a machine that recognizes objects?



Convolutional Neural Networks

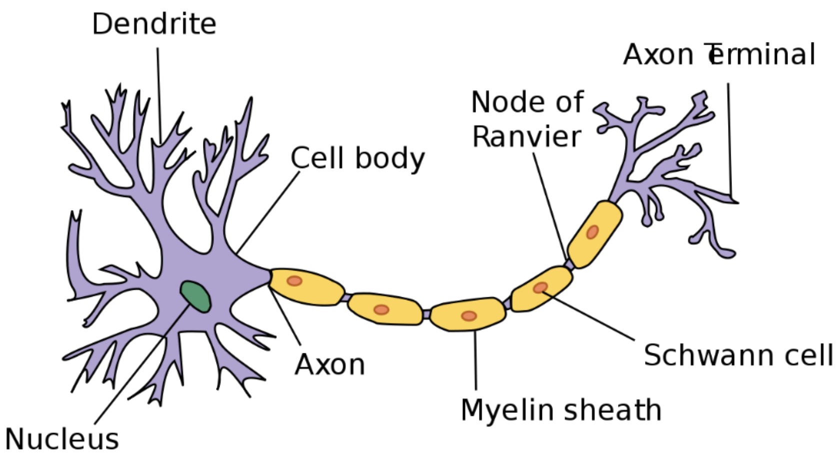
We model the human brain - neurons

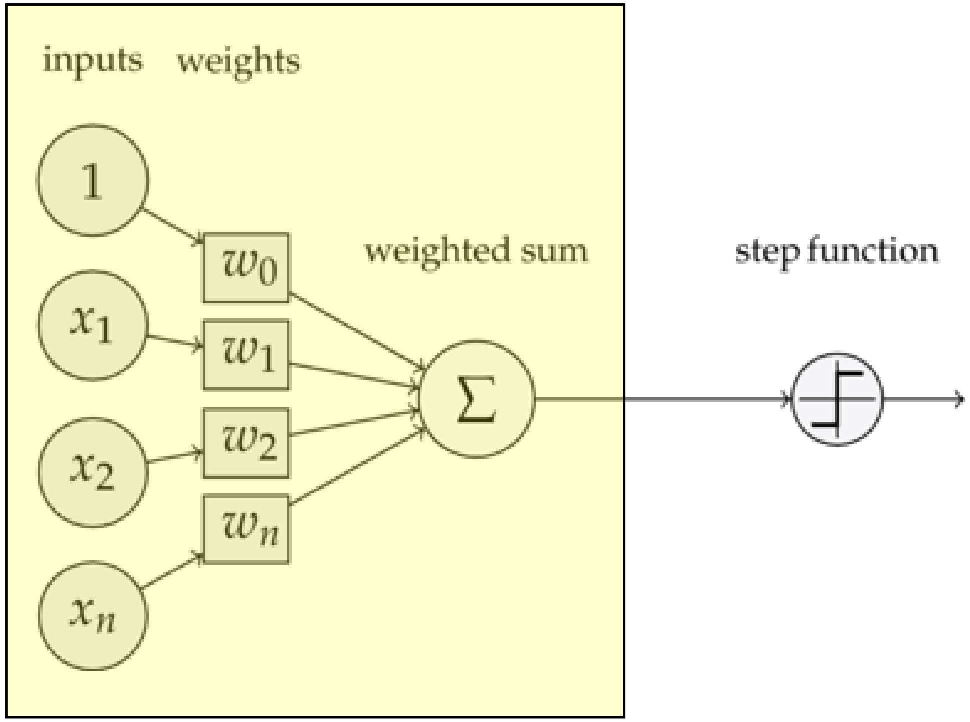




Convolutional Neural Networks

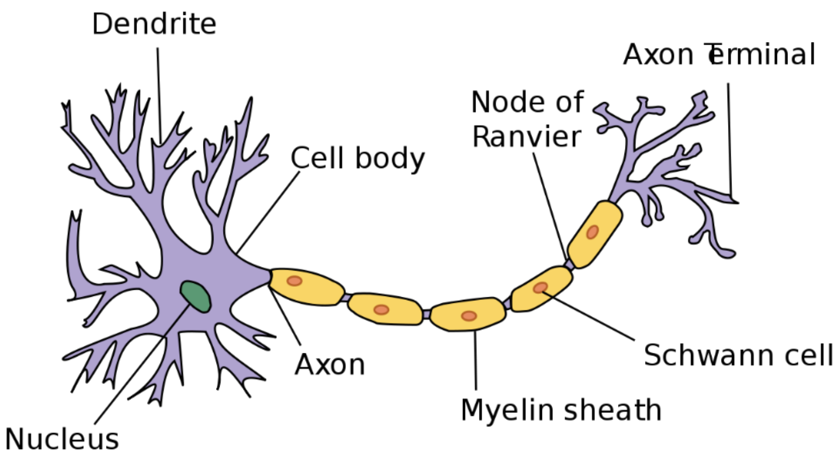
We model the human brain - neurons

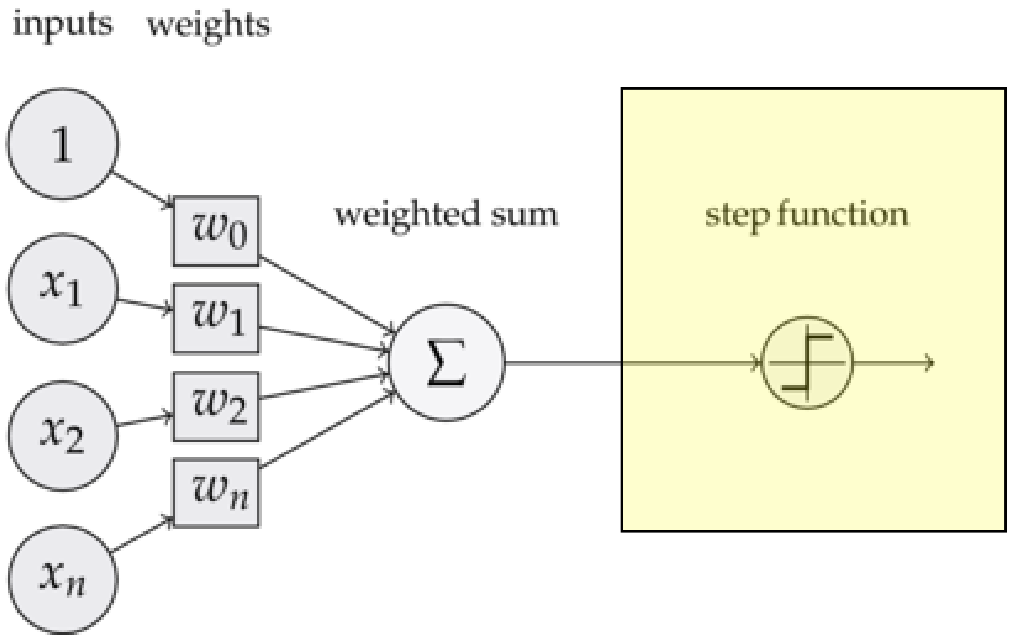




Convolutional Neural Networks

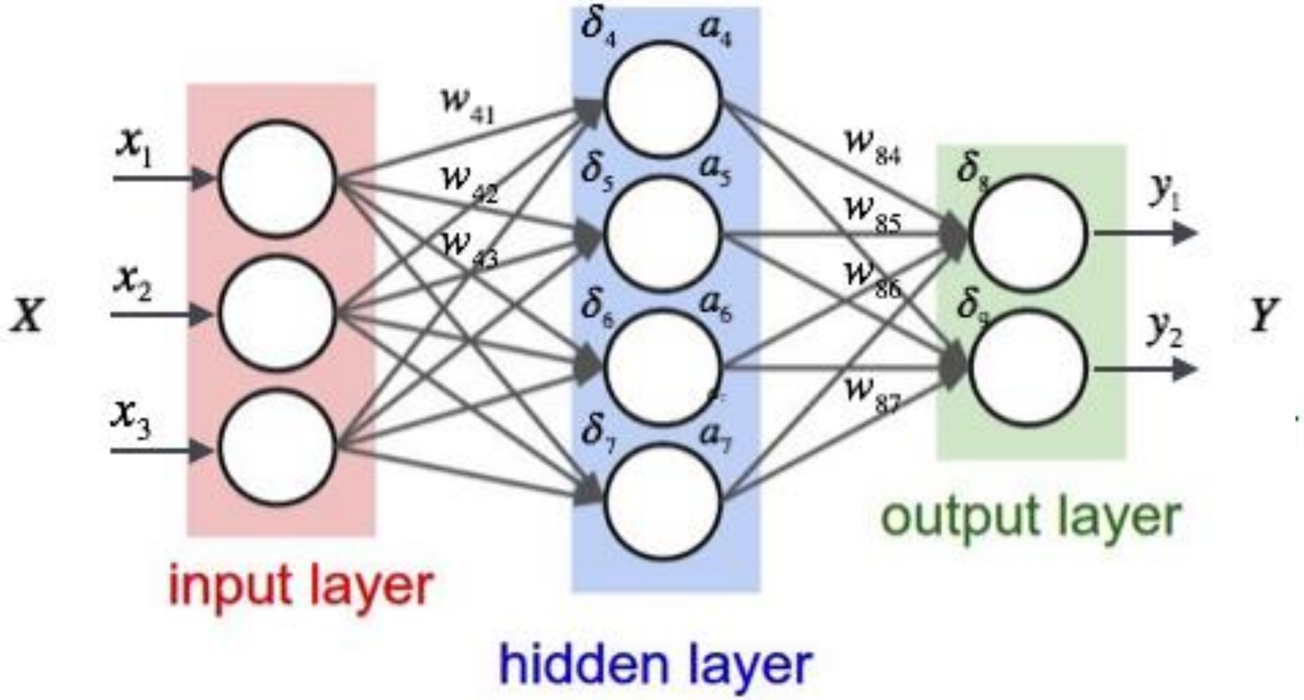
We model the human brain - neurons





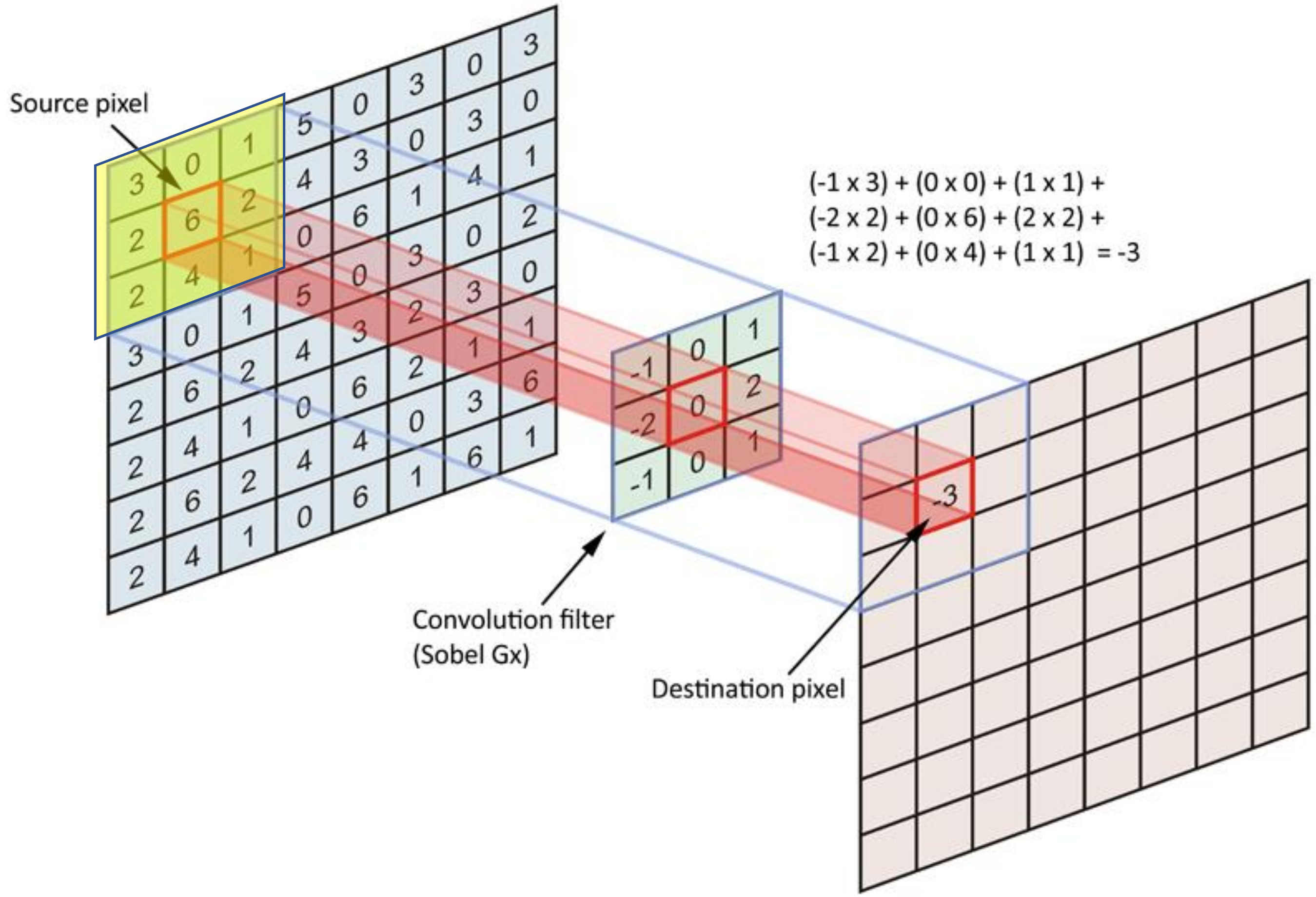
Convolutional Neural Networks

A neural network



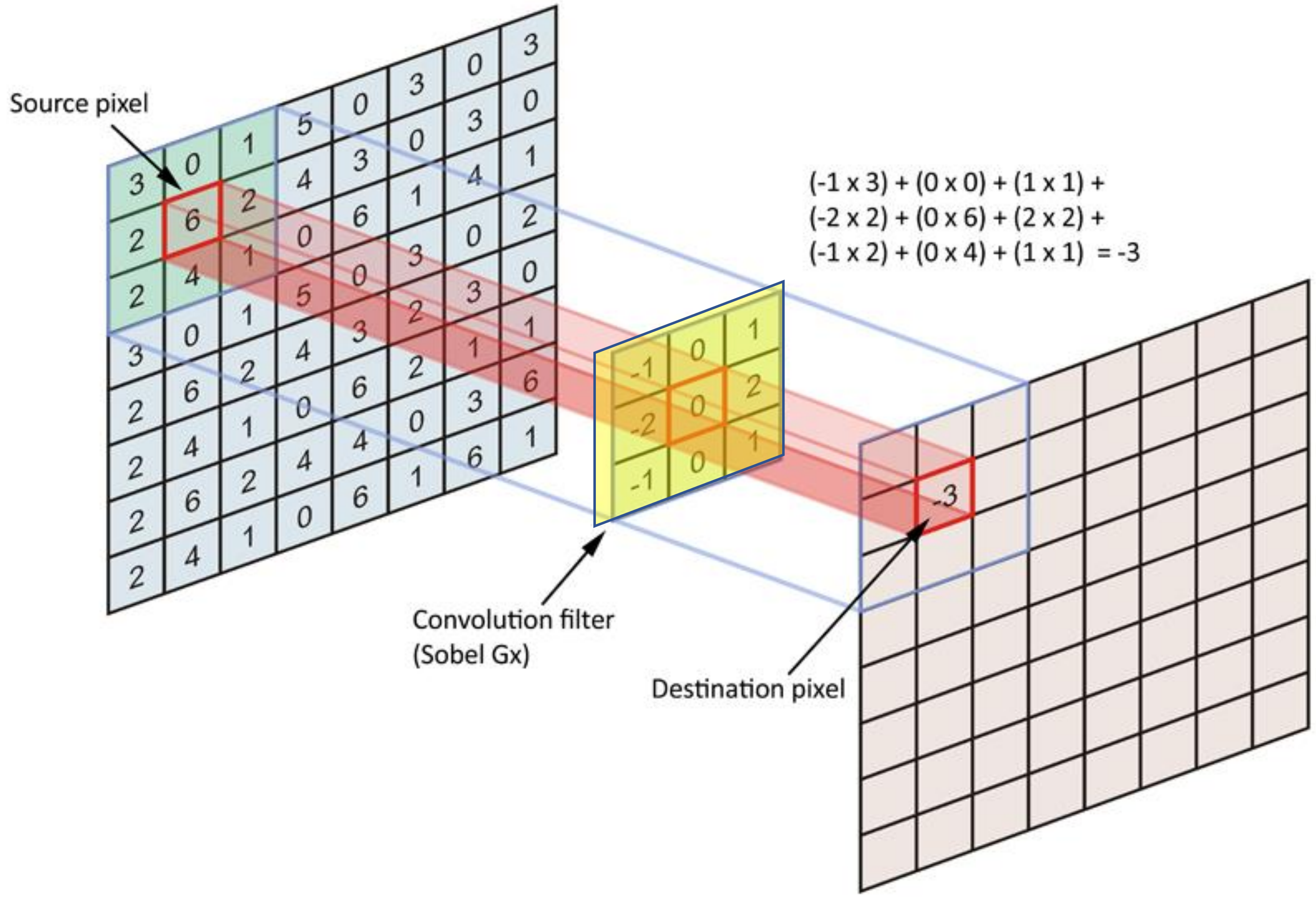
Convolutional Neural Networks

Approximating human vision - Convolution



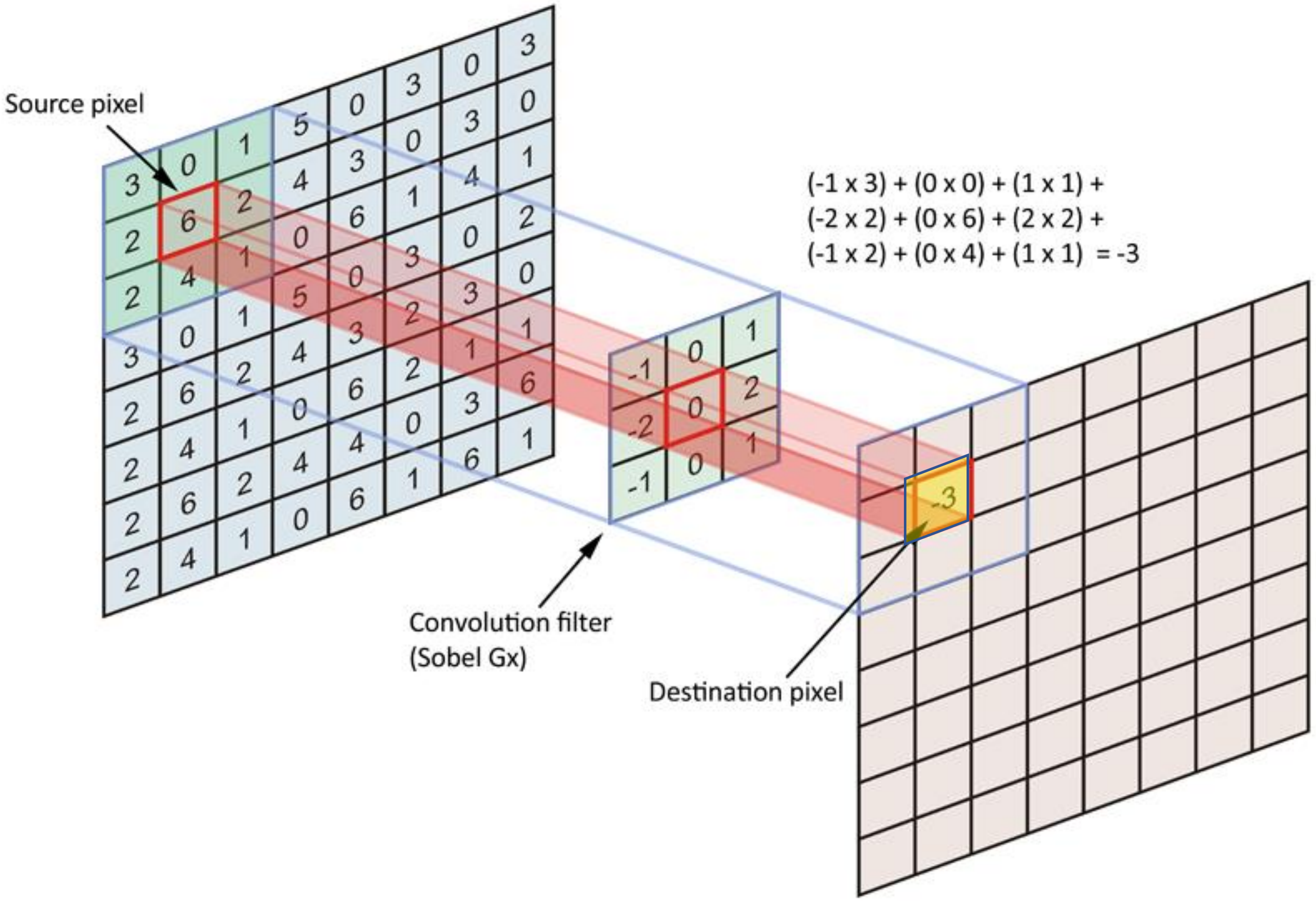
Convolutional Neural Networks

Approximating human vision - Convolution



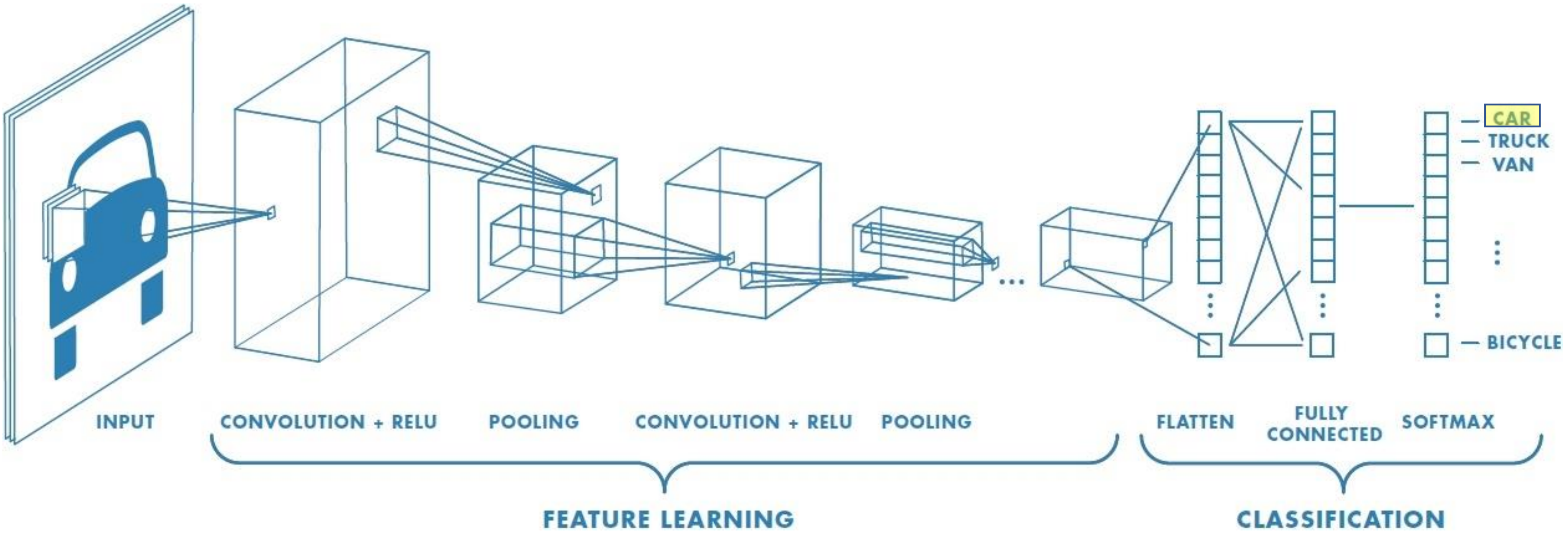
Convolutional Neural Networks

Approximating human vision - Convolution



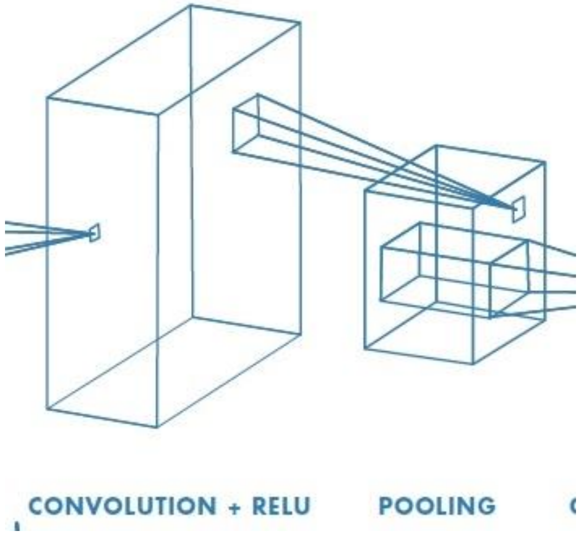
Convolutional Neural Networks

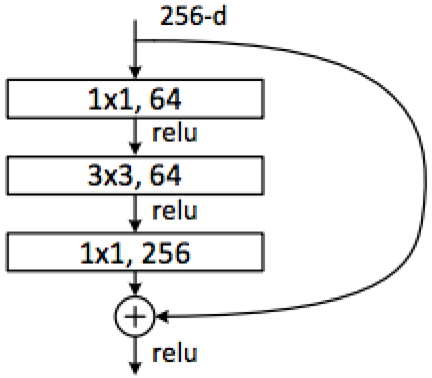
Scale up the model



Outline

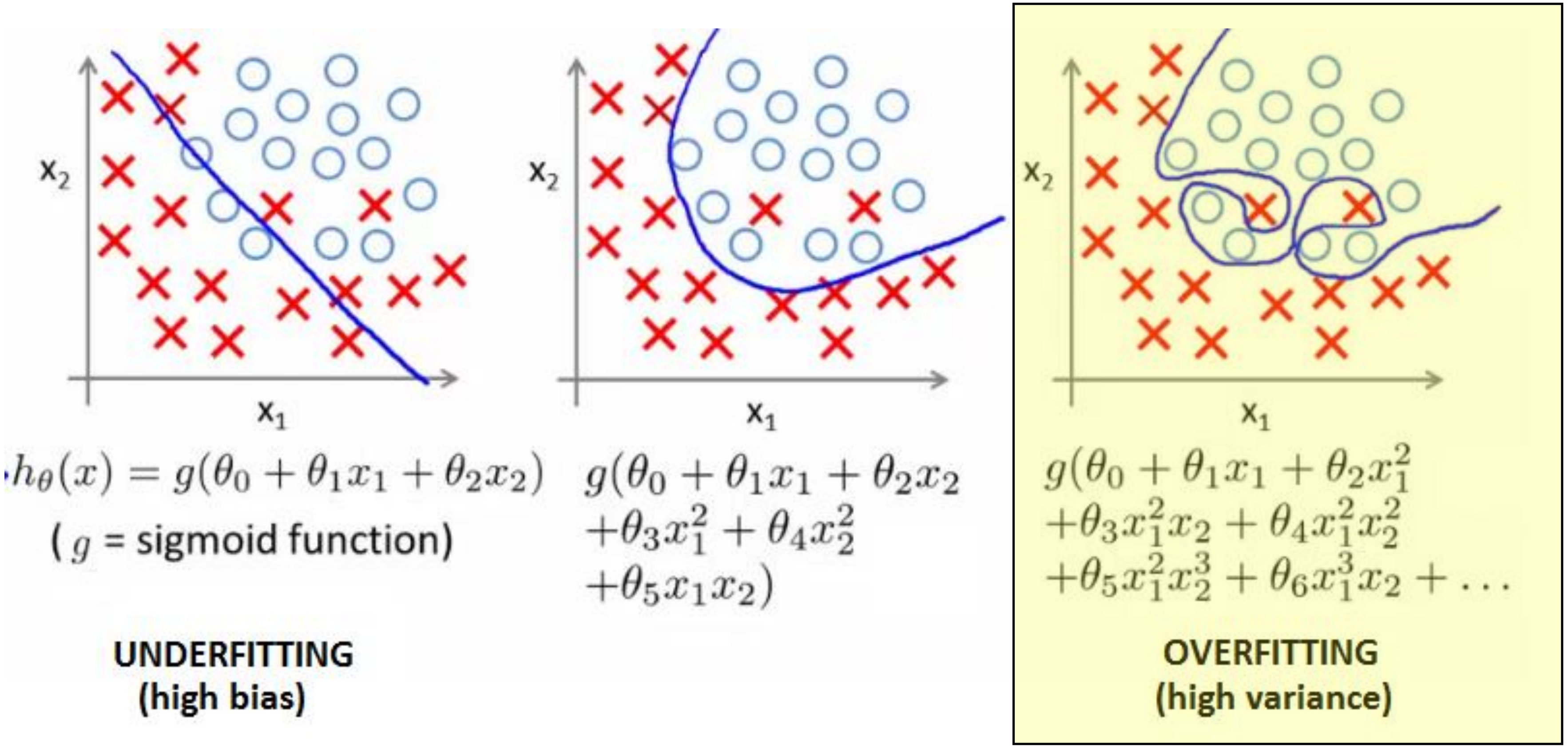
- Digital Pathology
- Convolutional neural networks (CNN)
- **From CNN to Residual Network**
- First prototype
- Second prototype
- Future direction





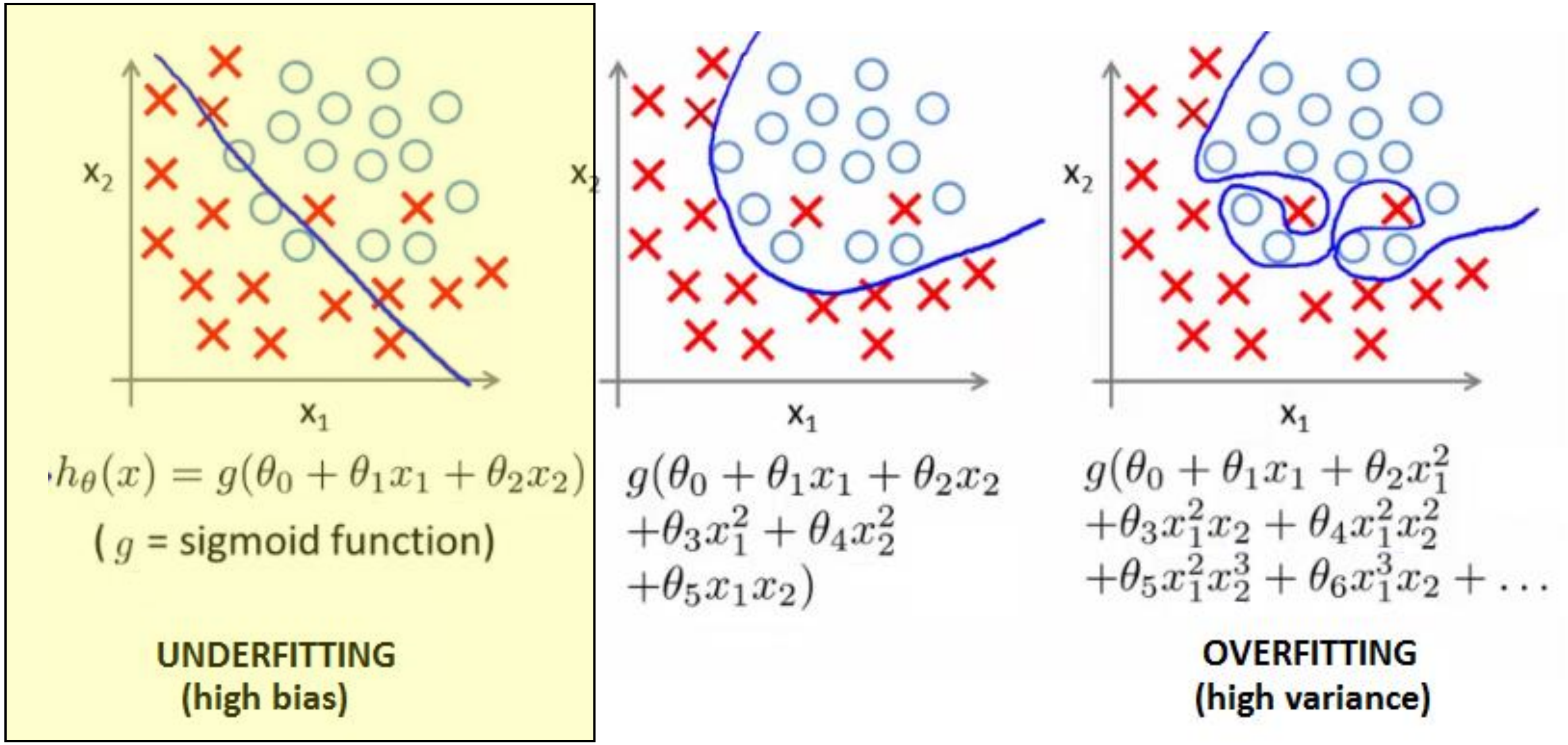
Overfitting and Underfitting

Difficult to find the right depth



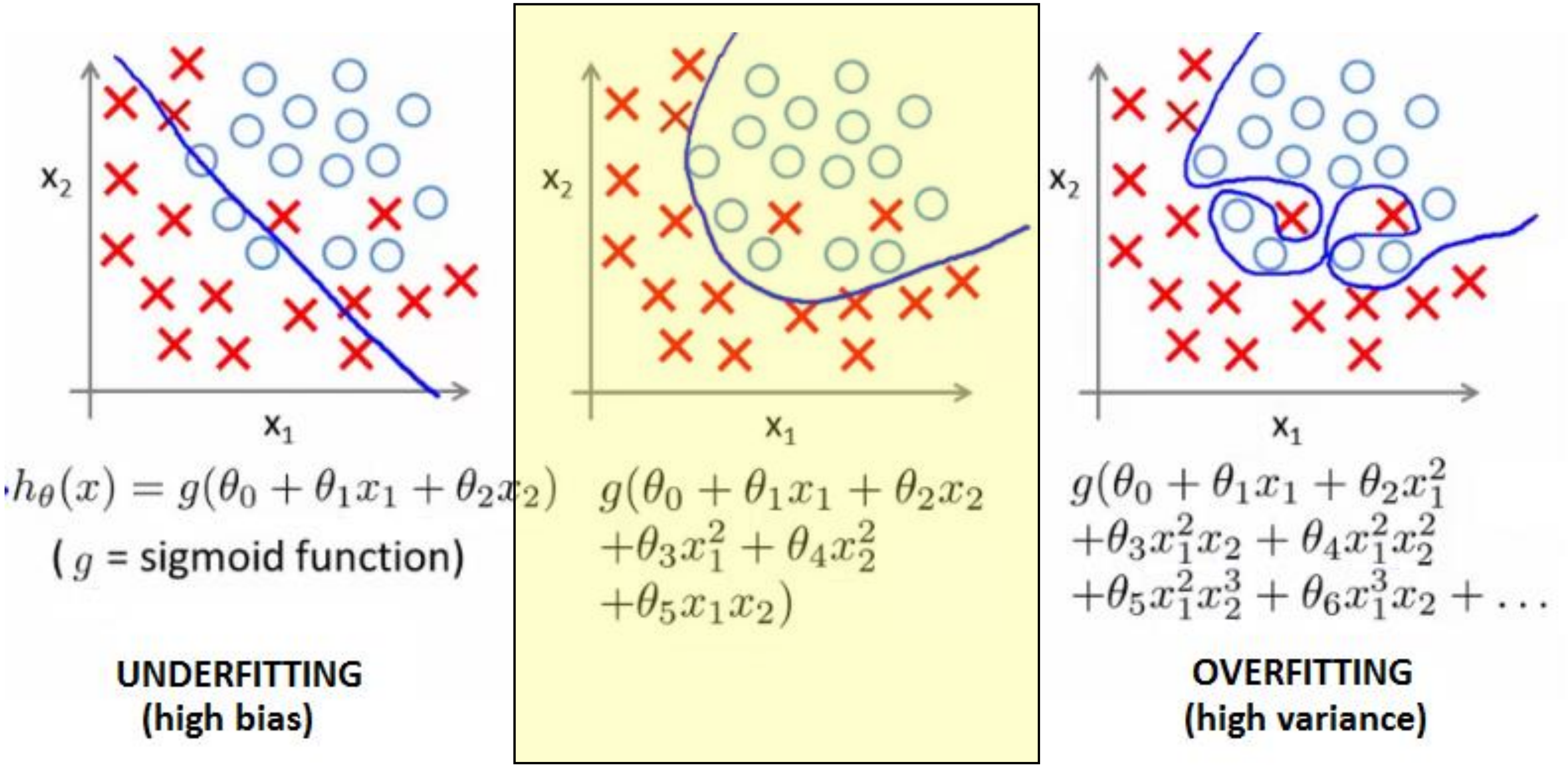
Overfitting and Underfitting

Difficult to find the right depth



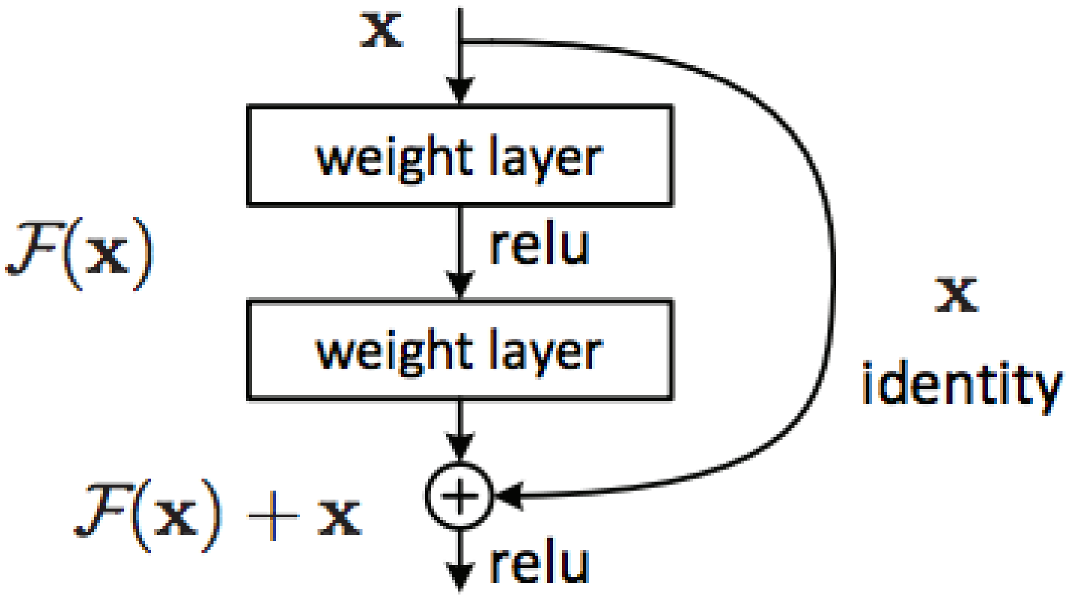
Overfitting and Underfitting

Difficult to find the right depth



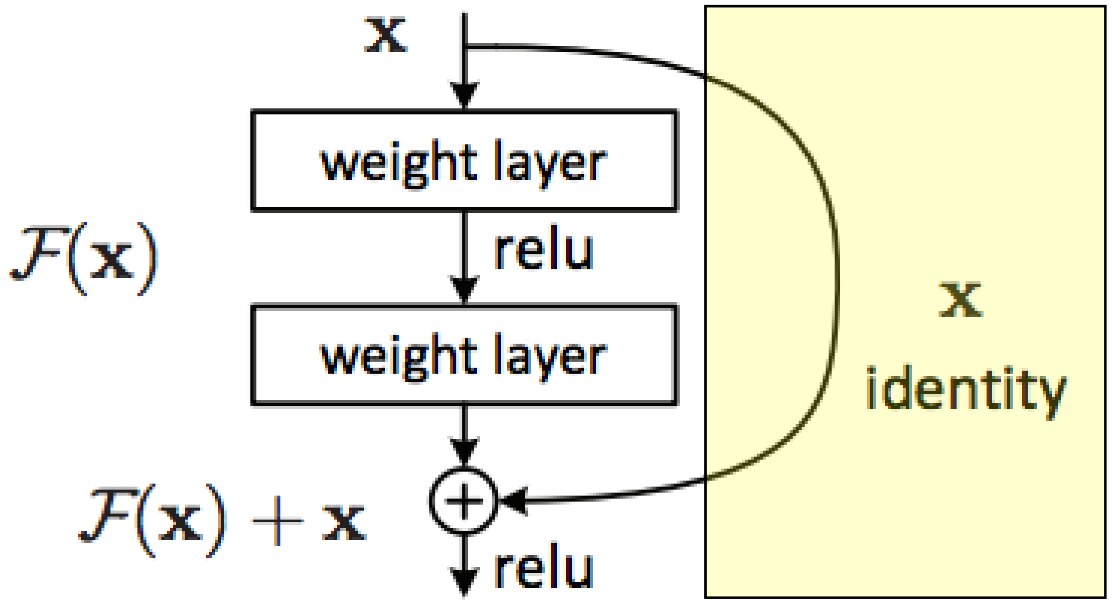
Skip Connection

Adaptive Depth



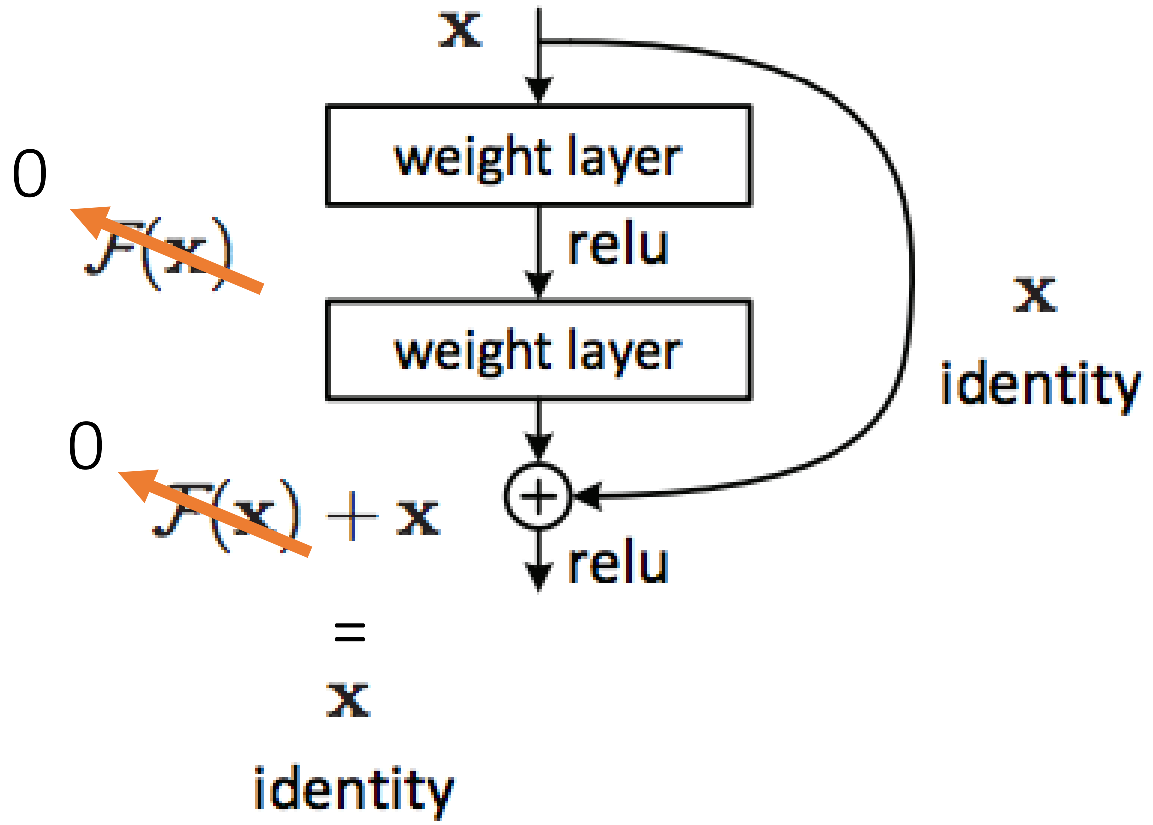
Skip Connection

Adaptive Depth

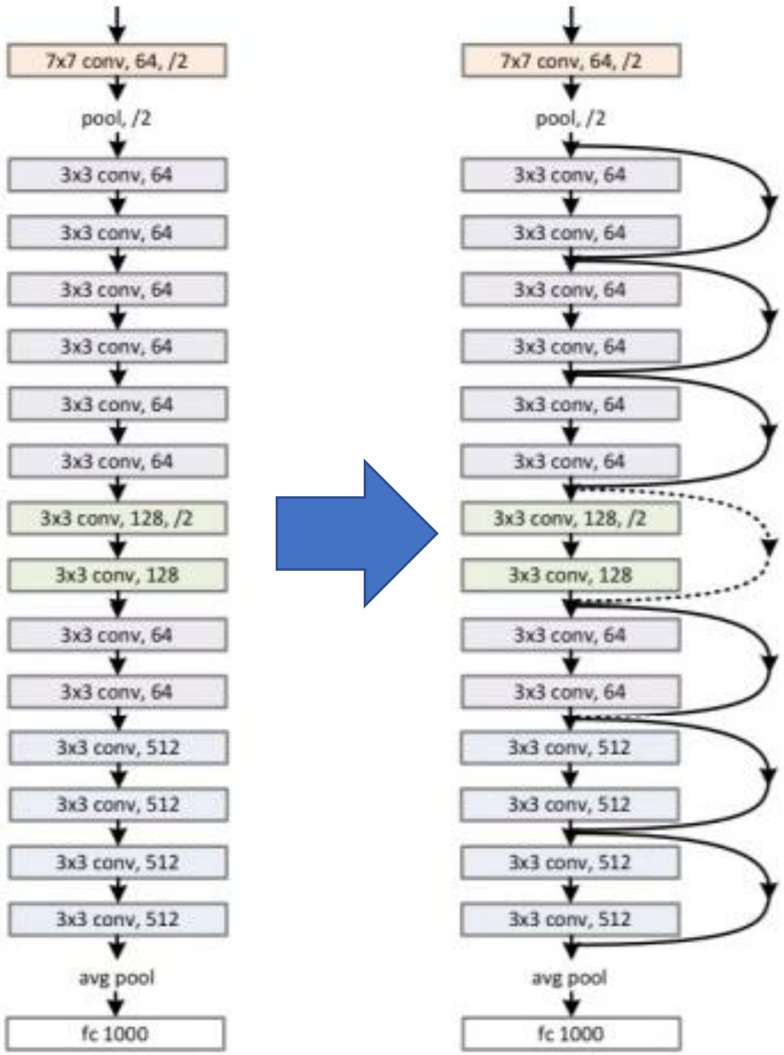


Skip Connection

Adaptive Depth

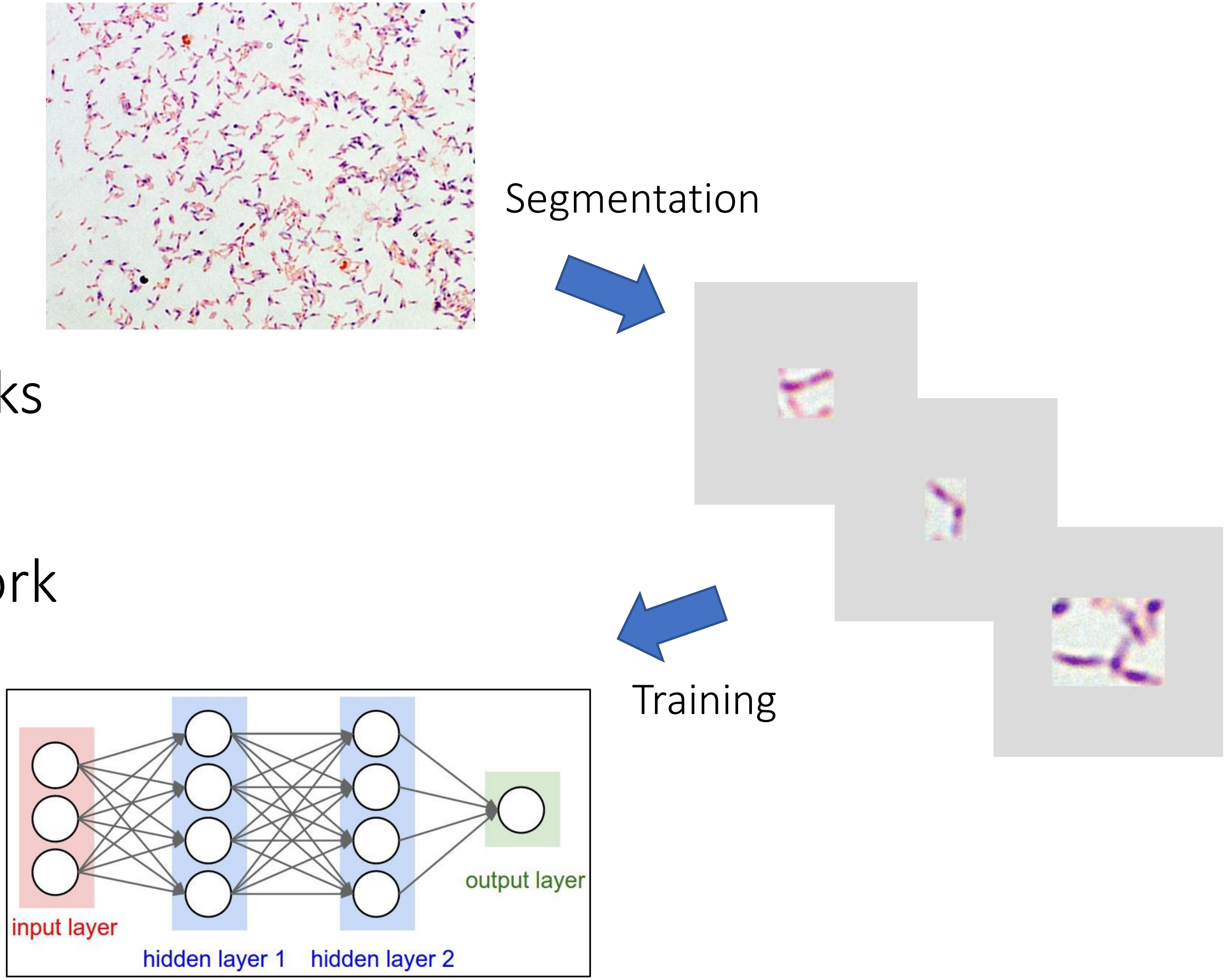


CNN to Residual Network



Outline

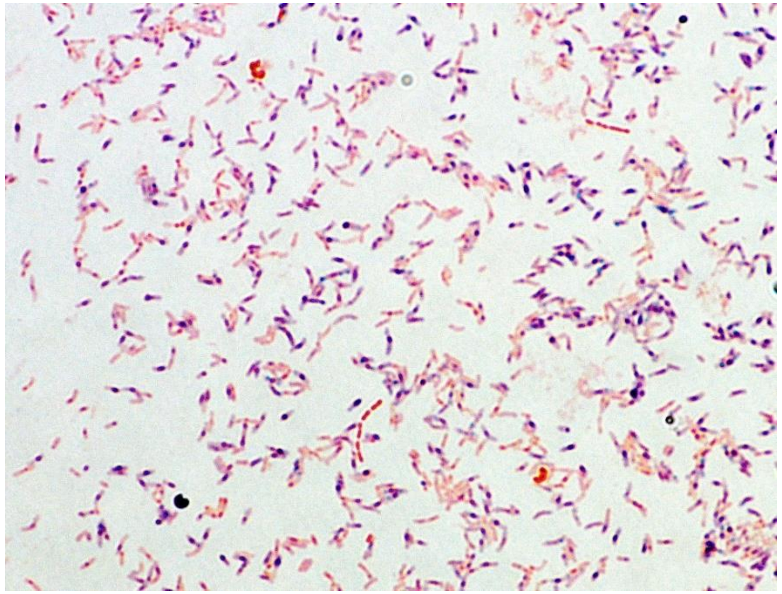
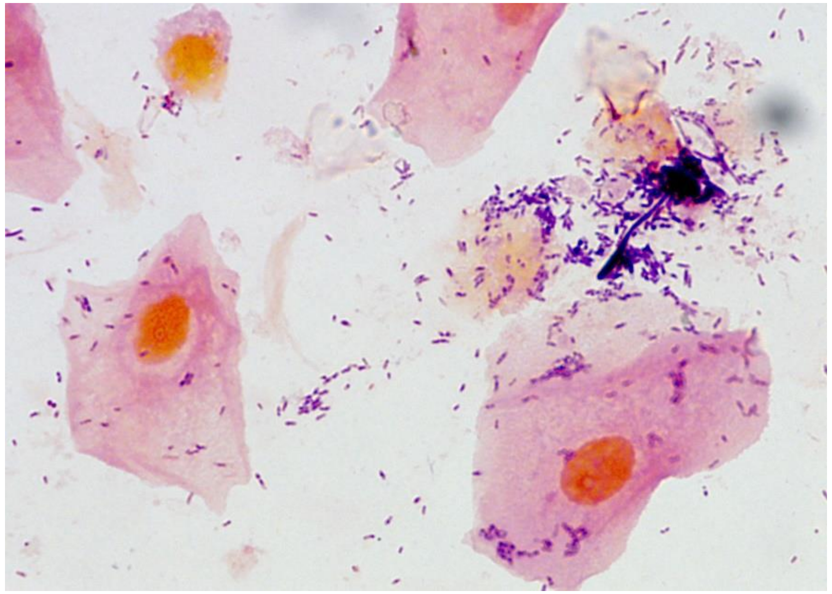
- Digital Pathology
- Convolutional neural networks (CNN)
- From CNN to Residual Network
- **First prototype**
- Second prototype
- Future direction



Machine Learning Model

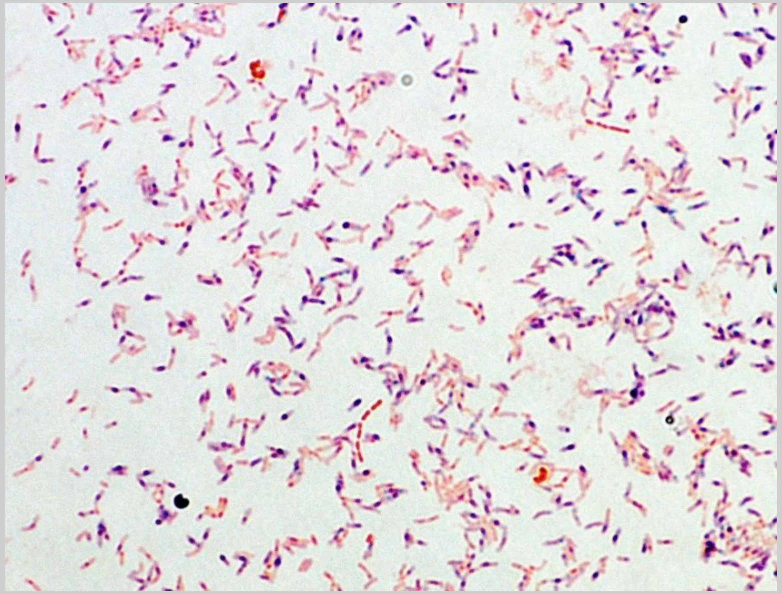
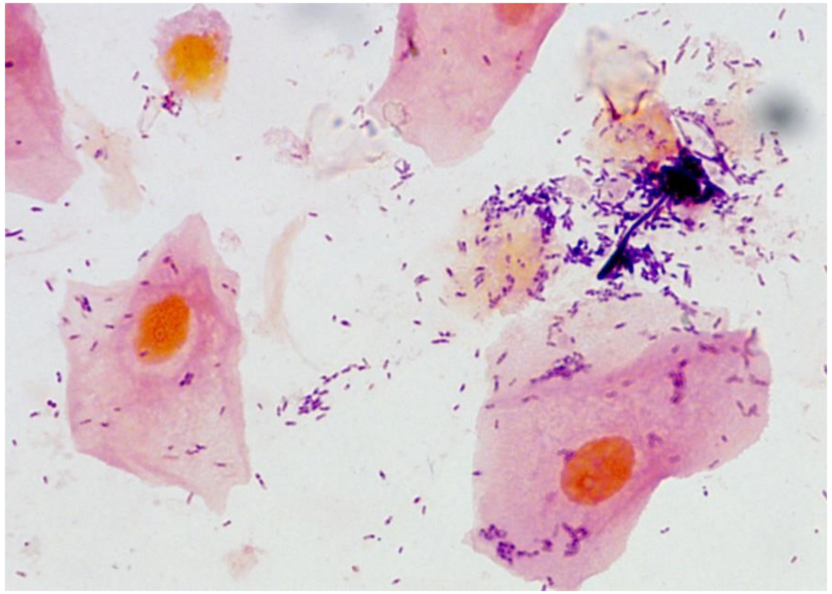
The First Prototype - Data Pipeline

Step 1: Separation of training and testing Data

Training	Testing
 Smear Images of bacterial colonies	 Smear Images from patients

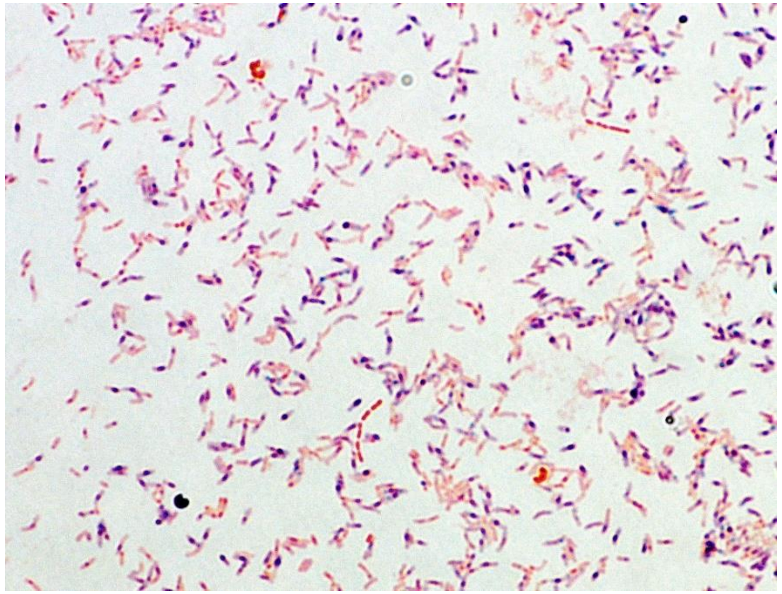
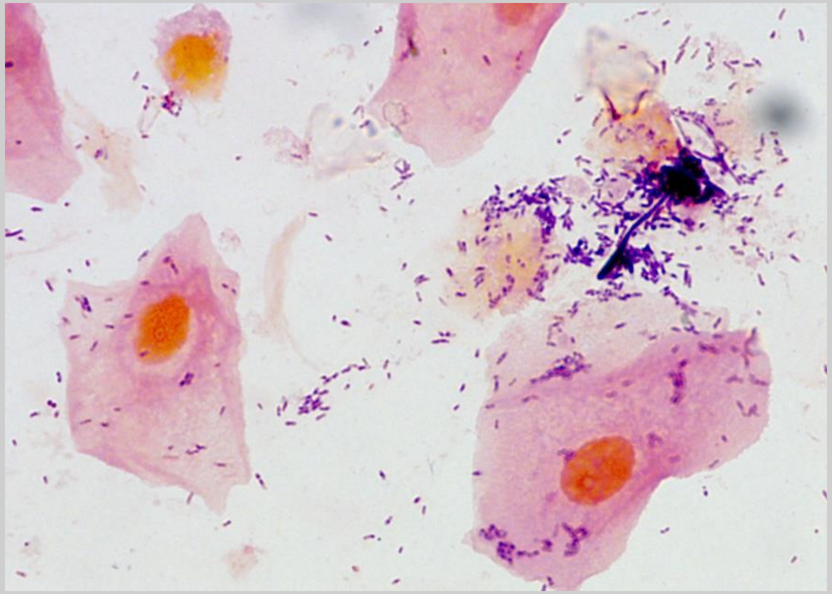
The First Prototype - Data Pipeline

Step 1: Separation of training and testing Data

Training	Testing
	
Smear Images of bacterial colonies	Smear Images from patients

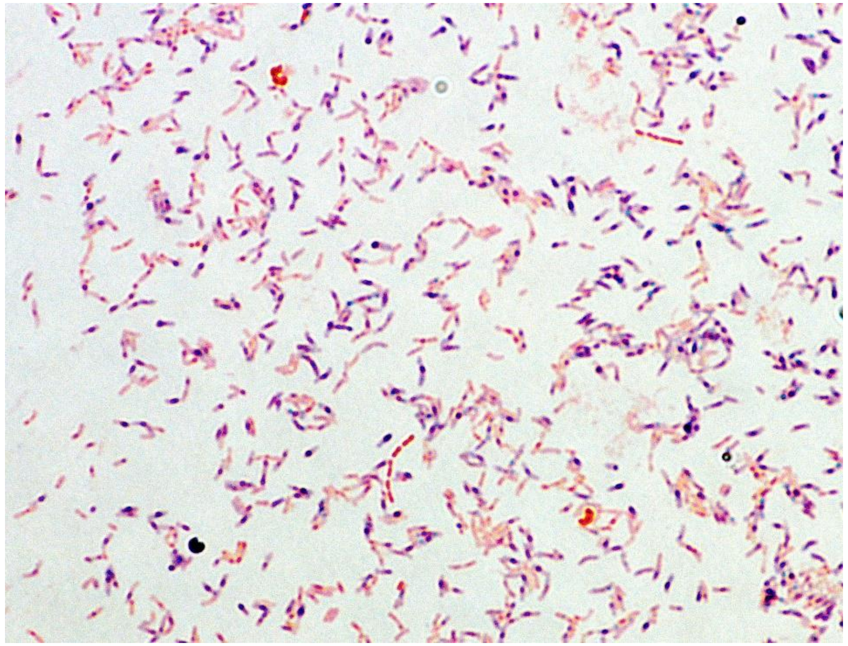
The First Prototype - Data Pipeline

Step 1: Separation of training and testing Data

Training	Testing
 Smear Images of bacterial colonies	 Smear Images from patients

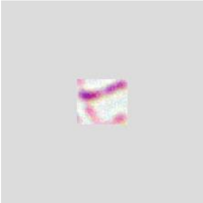
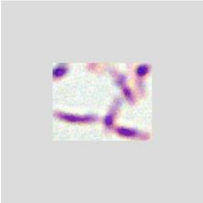
The First Prototype - Data Pipeline

Step 2: Segmentation of images

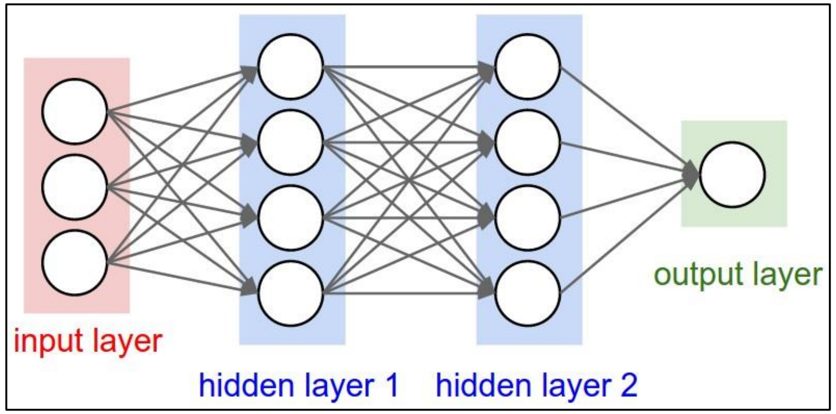

Gardnerella Vaginosis colony

The First Prototype - Data Pipeline

Step 3: Training using the images

Input	Output
	Lactobacillus
	Gardnerella
	Gardnerella

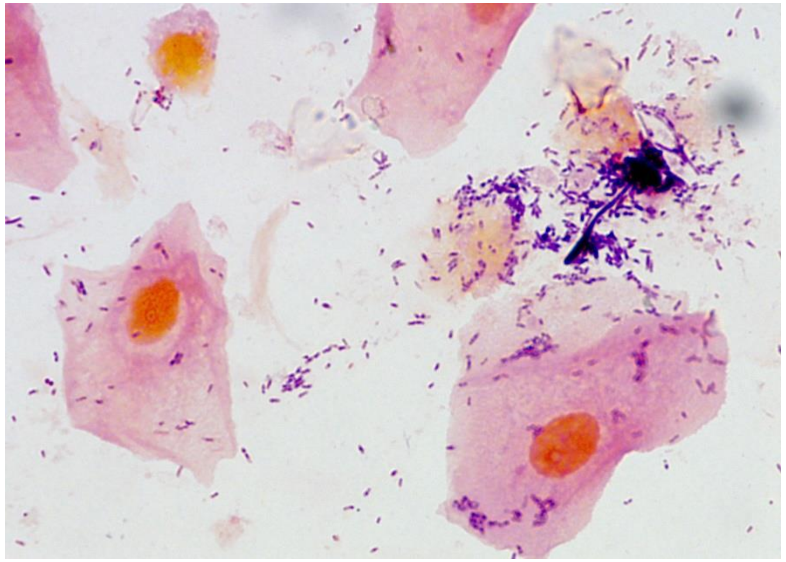
Training



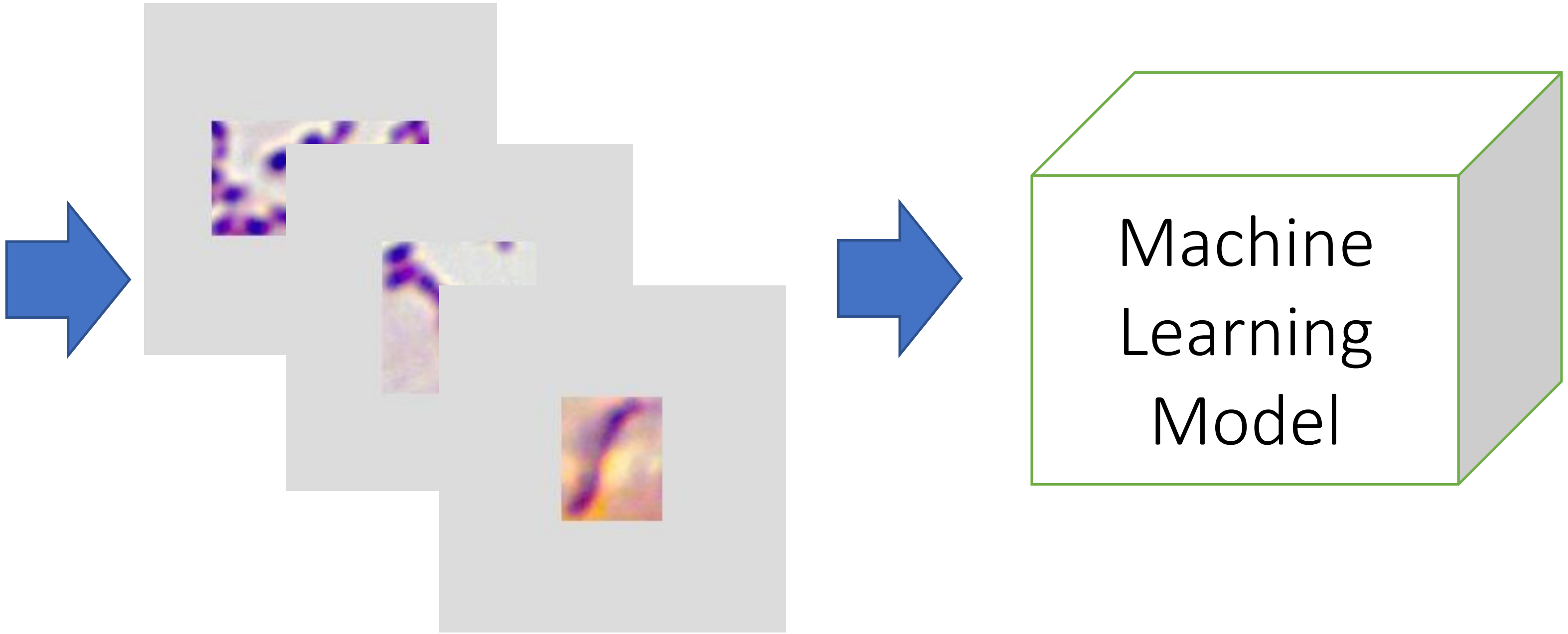
Machine Learning Model

The First Prototype - Data Pipeline

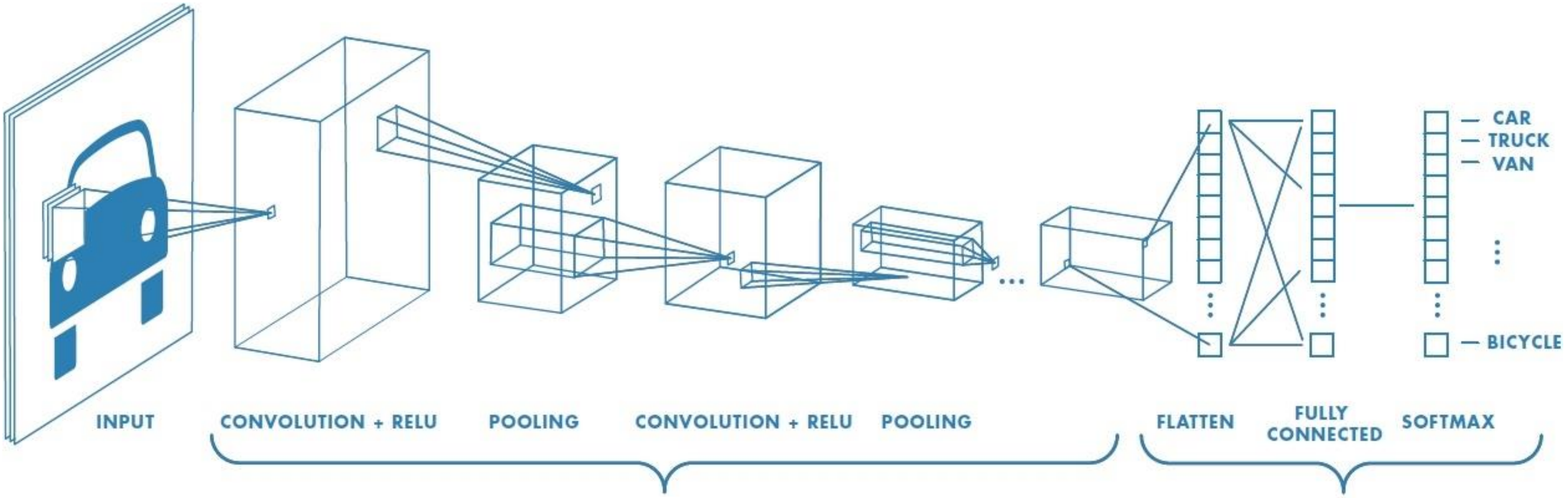
Step 4: Testing using the patients images



Patient Smear Image



The First Prototype – Model Architecture



7 Convolutional Layers

3 Fully Connected Layers

The First Prototype - Performance

Hold-out validation of 30% of unseen training data:

98.7% accuracy

		Prediction by the model			
		Lactobacilli	Gardnerella	Curved rods	Other
Actual type of bacteria	Lactobacilli	**499**	6	1	2
	Gardnerella	0	**559**	3	4
	Curved rods	0	2	**589**	0
	Other	2	5	0	**63**

The First Prototype - Performance

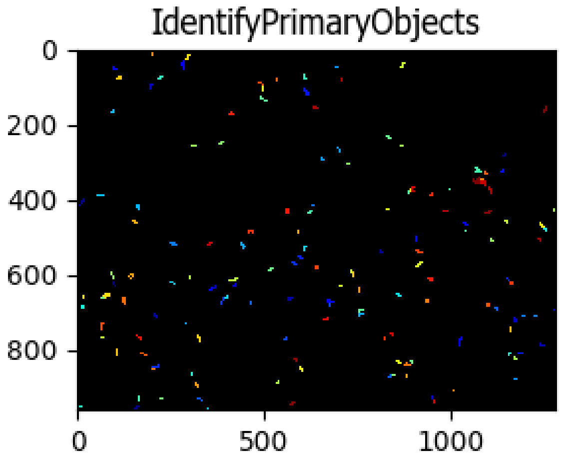
Testing on 31 patient images:

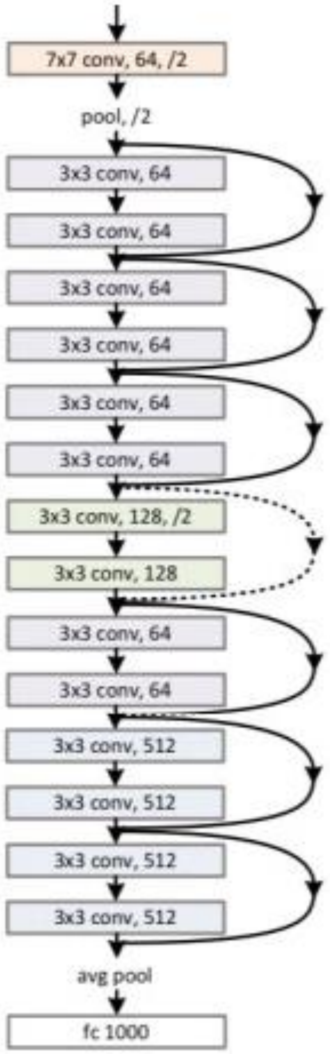
45.1% accuracy

		Prediction by the model		
		Normal	Intermediate	BV Infection
Actual degree of infection	Normal	**8**	0	0
	Intermediate	5	**3**	0
	BV Infection	6	6	**3**

Outline

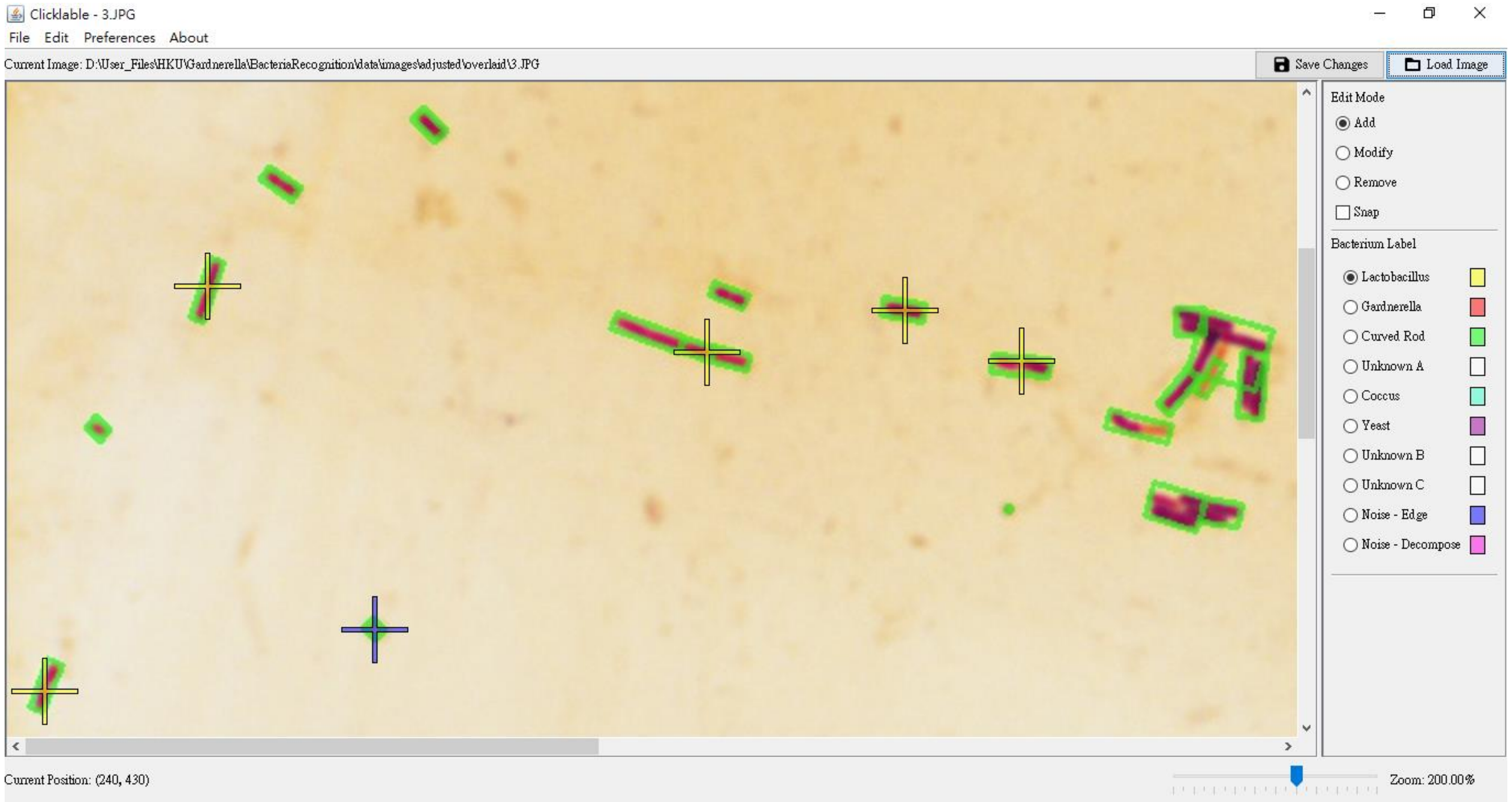
- Digital Pathology
- Convolutional neural networks (CNN)
- From CNN to Residual Network
- First prototype
- **Second prototype**
- Future direction





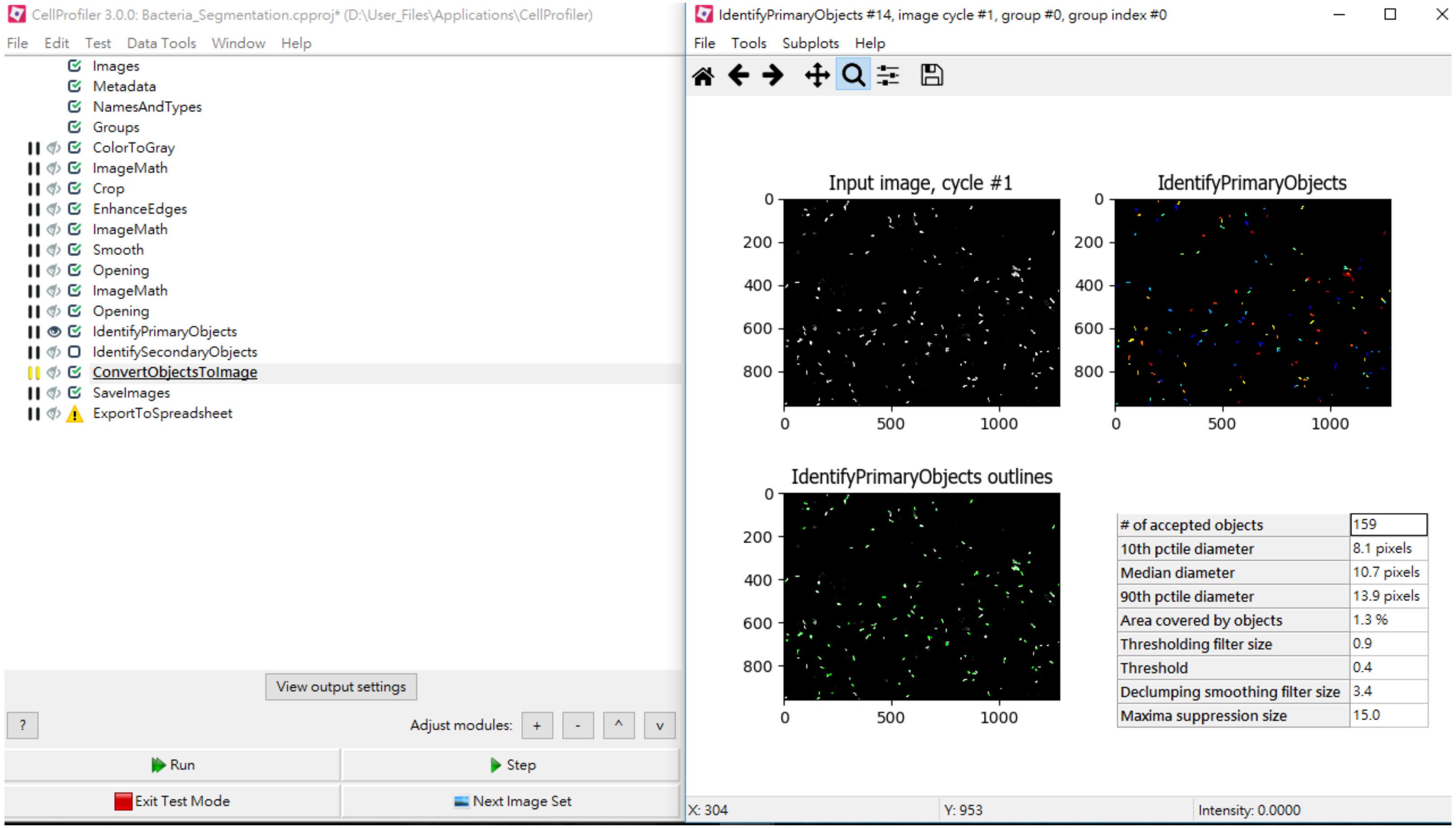
Second Prototype – Data Pipeline

Step 1: Collection of Data by Customized Tool "Clicklable"



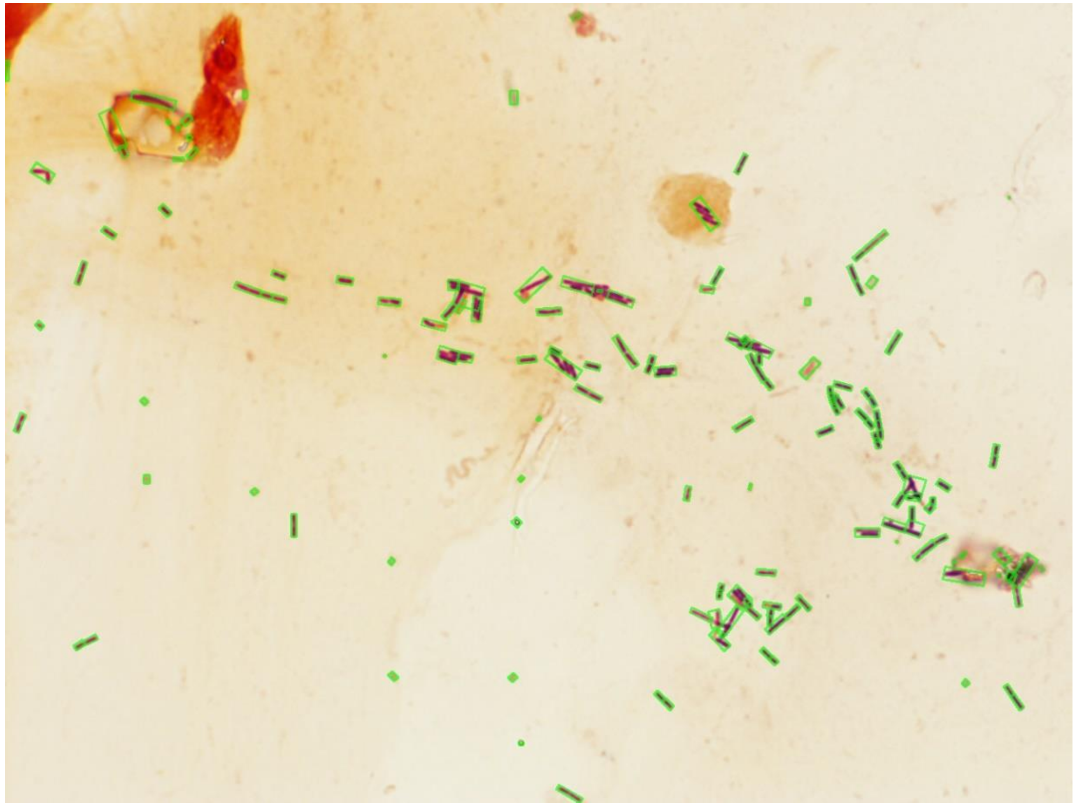
Second Prototype – Data Pipeline

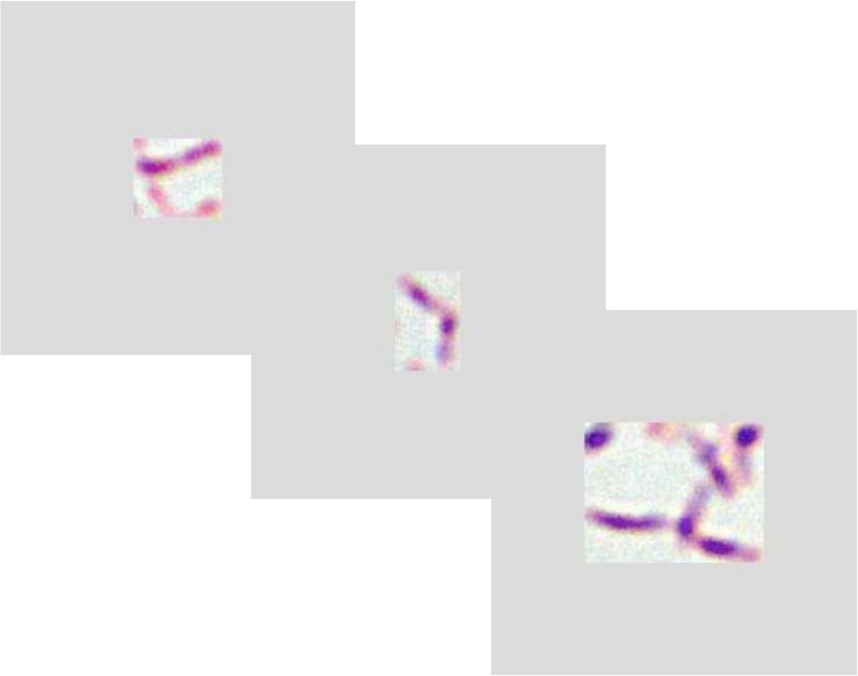
Step 2: Regions of Interest (ROIs) Identification by Image Processing



Second Prototype – Data Pipeline

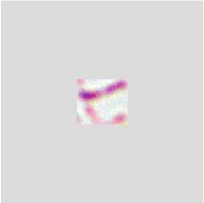
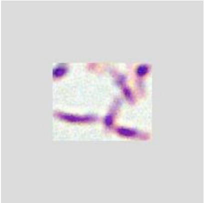
Step 3: Image Segmentation

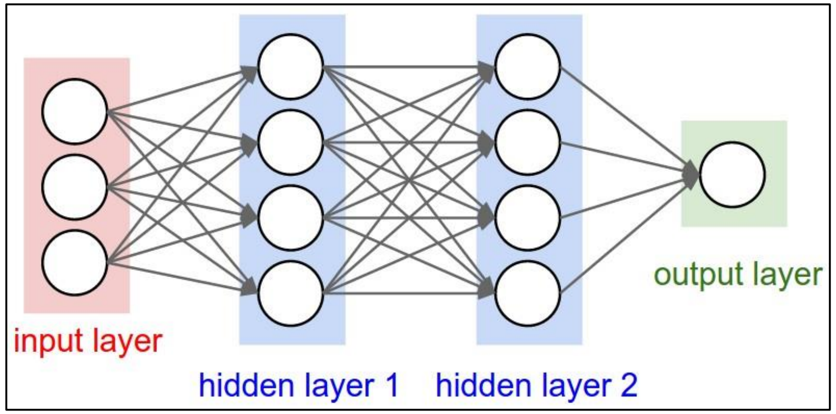




Second Prototype – Data Pipeline

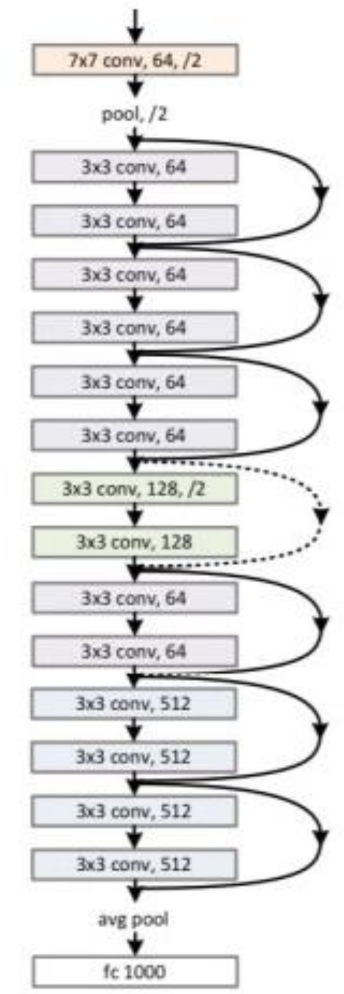
Step 4: Training and Validating using the images

Input	Output
	Lactobacillus
	Gardnerella
	Gardnerella



Machine Learning Model

The Second Prototype – Model Architecture



1 Standard Convolutional Layer

7 Residual Blocks

1 Fully Connected Layer

The Second Prototype – Performance (Provisionary)

Hold-out validation of 10% of unseen training data:

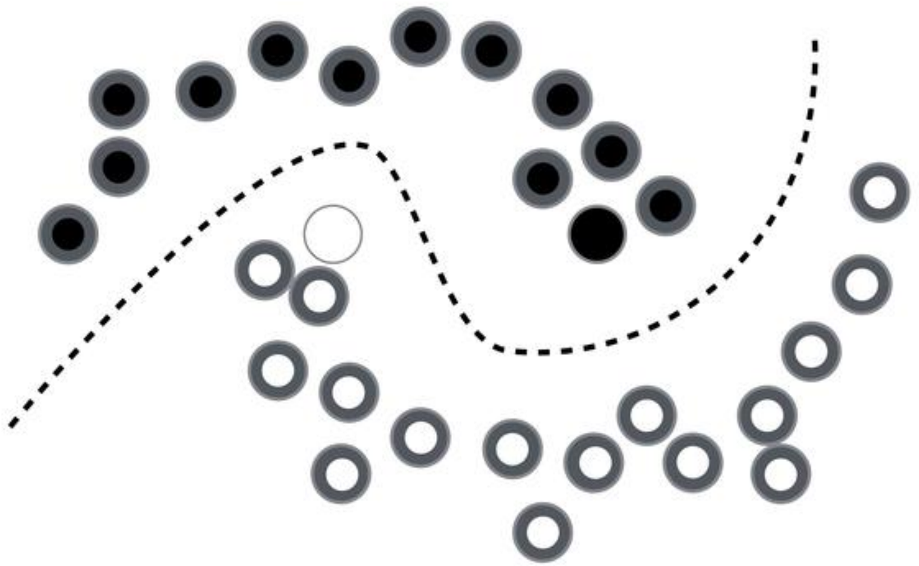
81.4% accuracy

Testing on unseen image data (not used in tuning):

72.9% accuracy

Outline

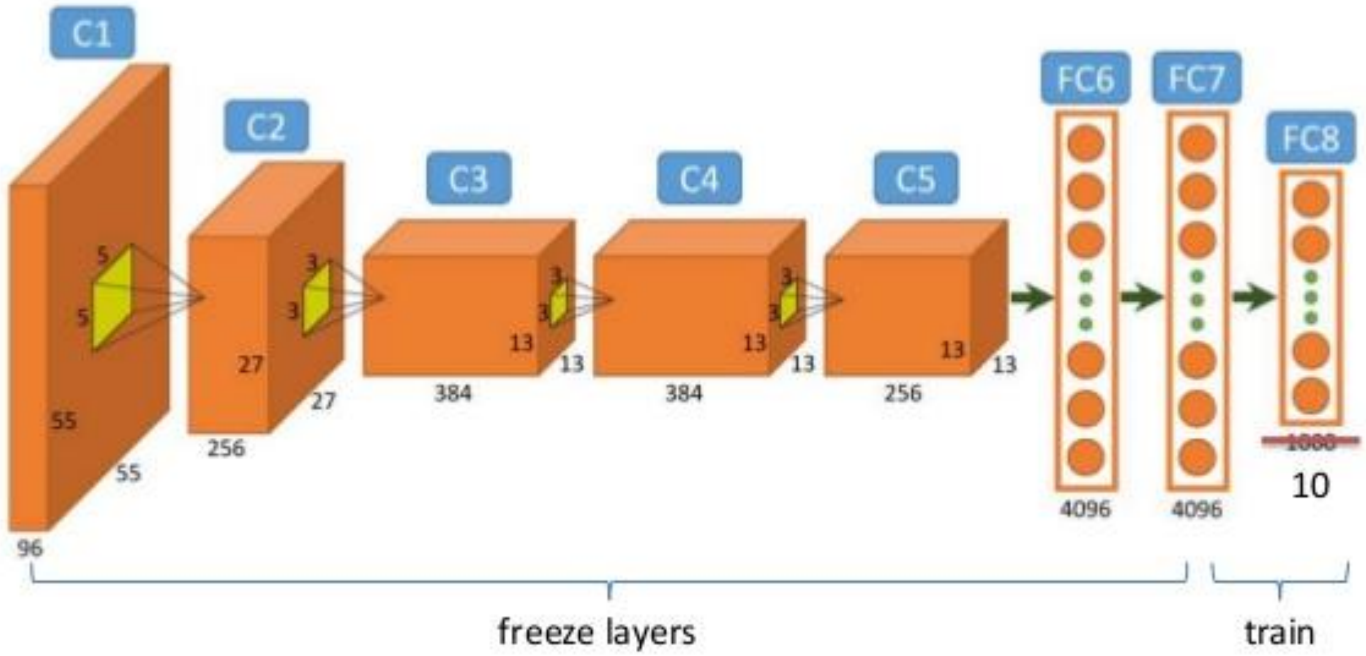
- Digital Pathology
- Convolutional neural networks (CNN)
- From CNN to Residual Network
- First prototype
- Second prototype
- **Future direction**



Future Direction

Validation accuracy < 90% →

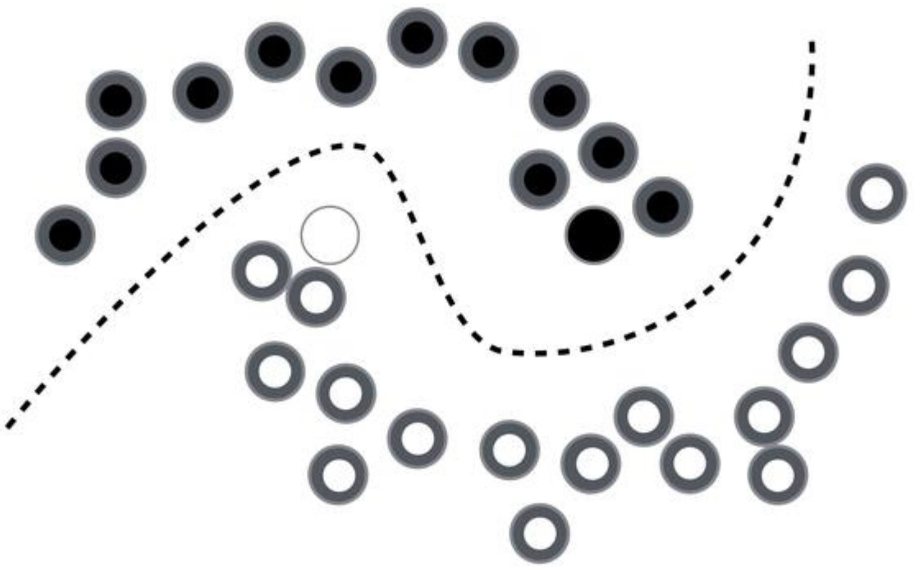
Increase complexity, fine-tuning and further training



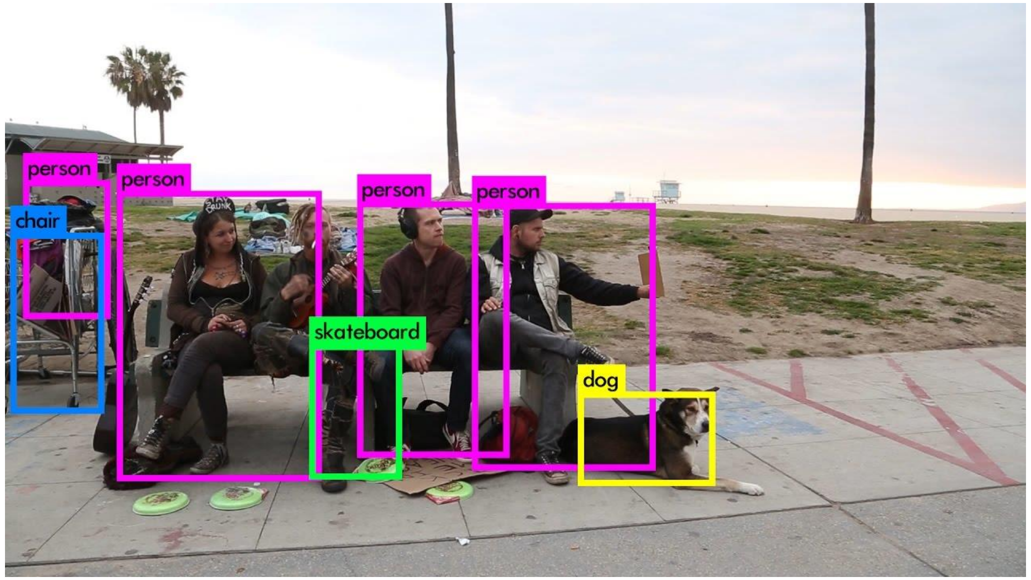
Future Direction

Large number of unlabelled data →

Exploration of other learning algorithms / architecture



Semi-supervised learning



YOLO algorithm for object recognition

Q & A