Deep Learning & Biomedical imaging

November 2025

1 What is deep learning?

Deep learning is a kind of machine learning, which is a kind of artificial learning (AI). Deep learning is a way for computers to learn from examples and experiences by mimicking the structure of the human brain, allowing them to recognize patterns and make decisions on their own. In detail, deep learning always has a neural network, which is the core part of the simulation of the human brain. It has a couple layers of interconnected nodes, namely input layer, hidden layer and output layer. Input layer receives the input data, hidden layers do the calculation and guessing, and then output layer produces or visualizes the output.

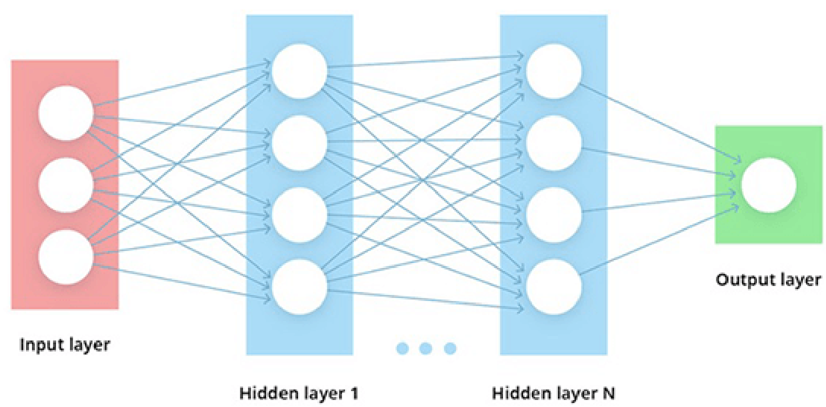
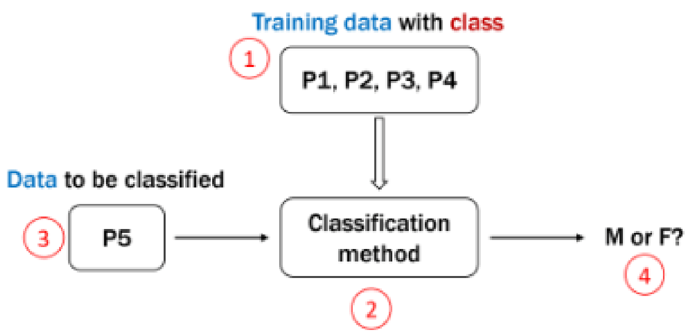


fig.1 general structure deep learning

For example, human can recognize a cat by pointing out that it has fur and tail or it has "meow" sound. But for deep learning model, we need to feed them with lots of photos, some with cats and some without. The input layer will break them into pixels and deliver to hidden layer. The hidden layer will do calculation and generalize similarities between images with cats. The output layer will output which images are cats while which are not. This is the training

process. When we input a new unknown image. The hidden layer belike "Ahh, it has the similarities of those images with cats", and the output layer will return "It is a cat." We check if it is really a cat and find the accuracy. This is the testing process.
In our lecture, one of the usage of deep learning in biomedical field is shown–disease screening

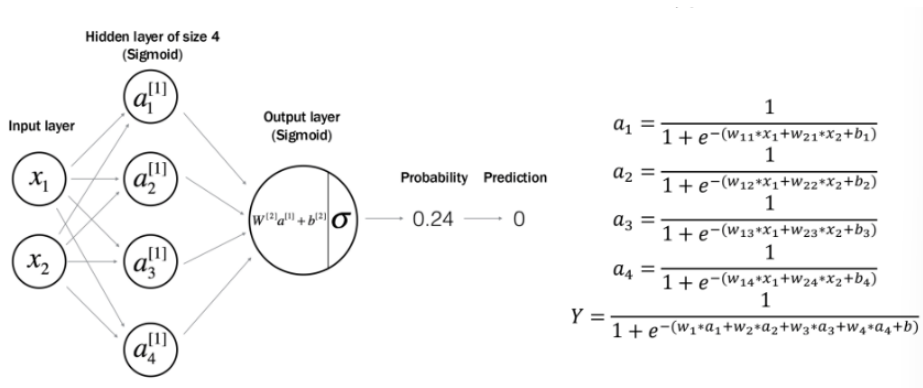
2 how does deep learning help in disease screening



Here is the simplified version of using AI to do classification. This is a supervised learning.

Person	Height	Weight	Gender
P1	0.625	0.875	M
P2	0	0	F
P3	0.25	0.375	M
P4	1	1	M
P5	0.4583	0.6667	??

Take the above tableas example, P1-P4 are labelled, which means that they are the answer of training. The model can understand "Oh this combination of height and weight leads to gender M while that leads to gender F". But then what method (model) we should use to do the prediction of P5? Of course we can use logistic regression which we learnt from previous lesson to do the prediction, but there will be a problem. The relationship between height and weight (input) may not be linear. Here we imply deep neural networks. Let say we have a hidden layer of size 4. We can have structure and calculation like this:

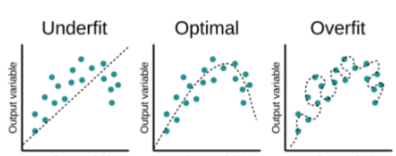


Where x_1 and x_2 are the input (height and weight), a_1, a_2, a_3 and a_4 are the weights, Y is the final output with highest probability. In this case, number of parameter from input layer to hidden layer is equals to $2 * 4 + 4 = 12$.
The number of parameters between 2 layers can be generalized as:

$$no.of parameters = previouslayer * currentlayer + bias$$

3 Potential problem of fully connected network

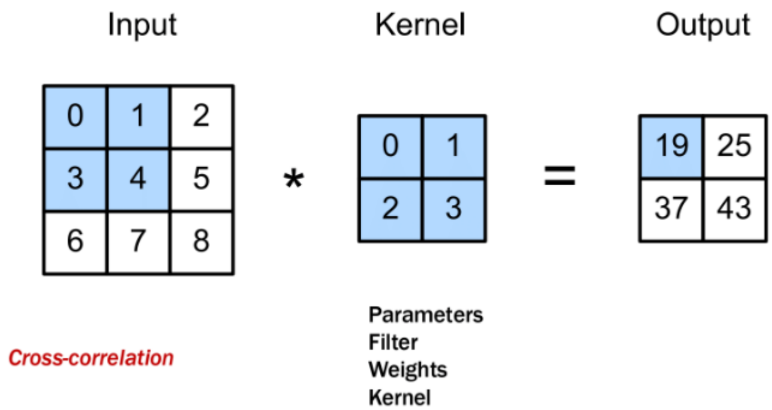
- layers are fully connected so we can do complicated calculation and get any function (relation)
- Suppose we have an image $(256 * 256 * 3)$ and three layers for binary classification. The internal layer has 1000 nodes.
- We can have 196,610,001 parameters, which is super complex.
- Complex model may leads to overfitting.



- Complex model occupy many storage.
- Complex model trains for a long, long time.

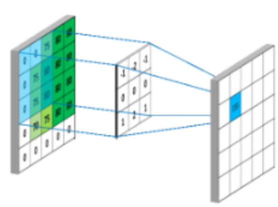
Finding a suitable size of model is very important in practical cases.

4 Convolution operation



4.1 How to do convolution?

Share parameters. Instead of each position in the input having its own unique weights, the same filter is reused across the entire input. One small set of weights (the kernel) is shared across all spatial positions.



- Alleviate the overfitting issue by reducing parameters.
- Features can be recognized regardless its position in input because the same filter detects the same pattern anywhere.
- Each filter only looks at a samll local region of the input at a time. This increase computational efficiency by processing small patches instead of entire image.

4.2 Padding

Padding is adding extra pixels around the border of the input image before applying convolution. It acts like a photo frame. Adding corresponding number of row or column to the image. It preserves the input size. There are 3 types of padding, valid padding (padding = 0), same padding (padding = 1) and full padding (padding ¿=2).

4.3 Stride

Stride determines how many pixels the filter moves each time it slides across the input. It controls output size reduction.