CS-GY 6643  Assignment 3
Runze Li

# 1. Deep learning

## 1.1 The difference between classical shallow learning and deep learning

Classical learning :
- One (or few) processing layer on input features
- One step decision
- Manually choose the decision function
- Feature engineering to fit with the model

Deep learning :
- Stacking many simple processing layers
- Hierarchically learn to disentangle data such that a simple linear function could separate them
- Large amount of data
- Act like a "human"

## 1.2 Pros and Cons of a CNN over an MLP architecture

MLP :
- Pros
  - Simple architecture : easy to implement and understand
  - Can be used for various types of data and tasks
- Cons
  - MLP doesn't consider the spatial relations and structures within the image data.
- Under translation, MLP is sensitive to image translation.

CNN
- Pros
  - Spatial hierarchy recognition
  - CNN shares weights across spatial locations, reducing the number of parameters.
- Cons
  - CNN requires more computational resources, especially with deeper architectures.
- When translation of the image is applied, CNN can be robust to translation.

# 2. Perspective Projection

2.1  The fundamental problem of 3D reconstruction from a 2D image using the perspective projection equations arises from the loss of depth information when projecting a 3D world onto a 2D plane.

2.2  From the question, we can find that $\frac{f}{x} = \frac{y}{Y}$

where $f = 50mm$, $y = 20mm$, $Y = 3m = 3000mm$
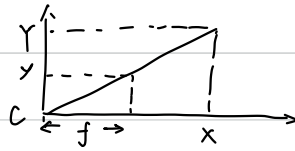
So we can get the distance of the object from the camera: $X = 7.5m$

2.3  For 2.1, we lack sufficient data to determine depth, leading to an indefinite number of possible solutions. But for 2.2, we have object's size, which provides a clear and specific solution.