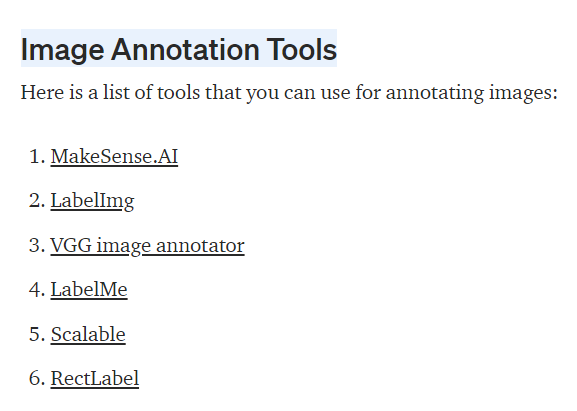
**Problem 1 (40 points) Suppose you are implementing a deep learning system for image classification task. Please describe how to build a deep learning system step-by-step. (a) (i) Describe how to annotate your dataset for image classification. What kind of tool will you use? (ii) How do you save each label as files? (iii) How would you split the dataset as training, validation, and test sets? Why? (15 points)**



Bounding boxes: Bounding boxes are the most commonly used type of annotation in computer vision. Bounding boxes are rectangular boxes used to define the location of the target object. They can be determined by the 𝑥 and 𝑦 axis coordinates in the upper-left corner and the 𝑥 and 𝑦 axis coordinates in the lower-right corner of the rectangle. Bounding boxes are generally used in object detection and localisation tasks.

1. Training sets are used to adjust the weights of the neural network during training;

2. Validation sets are not used to adjust weights, but to prevent overfitting.

3. The test sets are only used to test the model. To see how good the model is, it is to evaluate the generalization ability of the model.

**(b) (i) Please choose a convolutional neural network (CNN) architecture (e.g., AlexNet, GoogLeNet, VGGNet, ResNet ....). Justify your answer why you chose the architecture. (ii) How can you be sure that your model is implemented correctly? Explain that in terms of loss. (10 points)**

The advantages of AlexNet are:-Using the non-linear activation function ReLU

GoogleNet uses auxiliary classifiers (auxiliary classifiers), the output of a certain middle layer is used as a classification, and a small weight is added to the final classification result. This is equivalent to model fusion, and at the same time adds a back-propagation gradient signal to the network, and also provides additional regularization, which is of great benefit to the training of the entire network.

ResNet, deep residual network, prevents over-fitting