# Image recognition (For characters and Image)

Below demonstrate how to use AI for character recognition (This also applies to Image recognition). First, we need to train each character from A to Z. We will use the letter “a” for an example.

First thing we need to know that letter “a” is a combination of different edges **Figure 1**, shape a, b and c correspond to letter “a”. We need to train the letter “a” in correspond to the edges, that is **Figure 2**. When we train the neural network for the letter “a”, all the **weights** and **bias** will be set for prediction. For prediction **Figure 3,** we use the trained weights and bias to predict the image.

Why do we use ReLu for the activation function? It is because there will be no negative values only zeros and maximum positive values, if the maximum value is large, this means it is likely to be the shape of the edge and when passed to the sigmoid activation function the output will be close to 1 which is the more likely to be the letter “a” as shown in **Figure 3**.

This type of approach also can be applied to images, all you need is to train the different types of edges together with the image given for example an image of a “**Lion**”. All you need is to convert the “Lion” image to grey scale and train it with different edges, the more the edges the more the accurate. But the problem with this approach is that the “Lion” must be a self-portrait. If there is dog together with a Lion, this approach fails to predict. To handle this problem, we can use a sliding window approach that is the image “Lion with the dog” size is cut to different grid sizes, for example 5 x 5 grid, 9 x 9 grid, 12 x 12 grid…etc. The more the choice of grid sizes, the more the chance to capture the “Lion” in the image. You just need to move the grid from left to right and top to bottom. Then use the different “Lion with the dog” grid to train with the edges. For sure, you with find the “Lion” next to the dog.

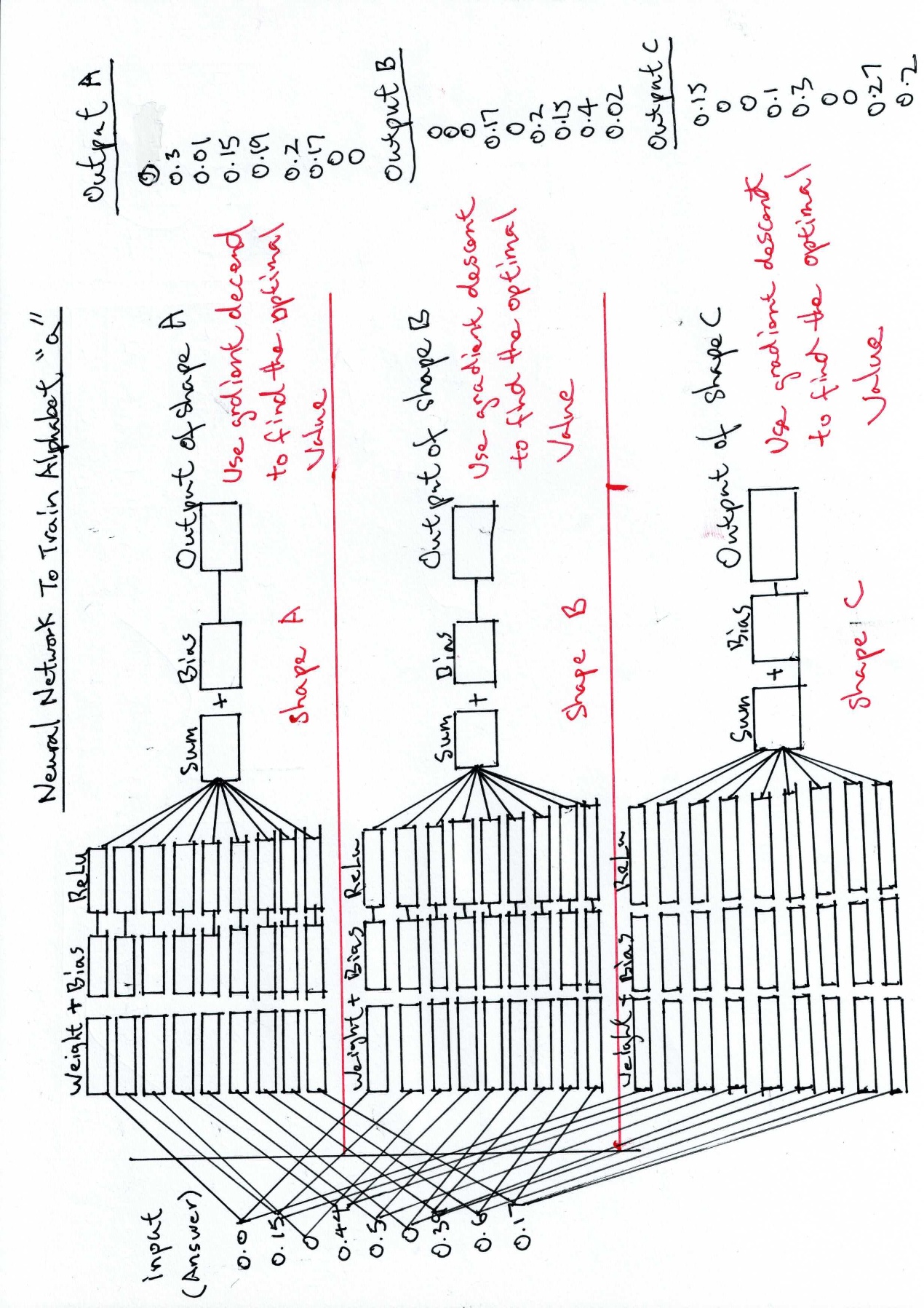
Not only images even video content searching can use this approach, but you need to covert the video to frames of images. For example, you want to search a particular “Flask with flower” that appeared in the video. You just need to convert the video to frames of images and then search the “Flask with flower” just like you search the “Lion with the dog”. Even police want to search a suspect from a surveillance video that wear a hat or eyeglasses. You just need to train the “hat” or “eyeglasses” with the edges.

Some interesting facts, if the sliding window cut half of the face of an “Lion”, it will be possible to search “Lion” that the face is turned to left or right. Check it out from **Microsoft Bing** image search.

A paper with writing on it

AI-generated content may be incorrect.

***Figure 1.***



***Figure 2***

***A diagram of a building

AI-generated content may be incorrect.***

***Figure 3.***

As said, we can also apply for video content searching. However, one might say a one second film consists of 70 separate image frames, if a 10 second film, that will be 700 images frames. Not talking about a one-and-a-half-hour movie, a 10-minute film will be a lot of images. That is right, however, we are talking about different images frames, a one second video if not much motion, perhaps not 70 images but may be 10, because most of the image frames are duplicated. For video content searching, we need to remove duplicates before apply searching. Therefore, a one-and-a-half-hour movie will be possible for video content searching.