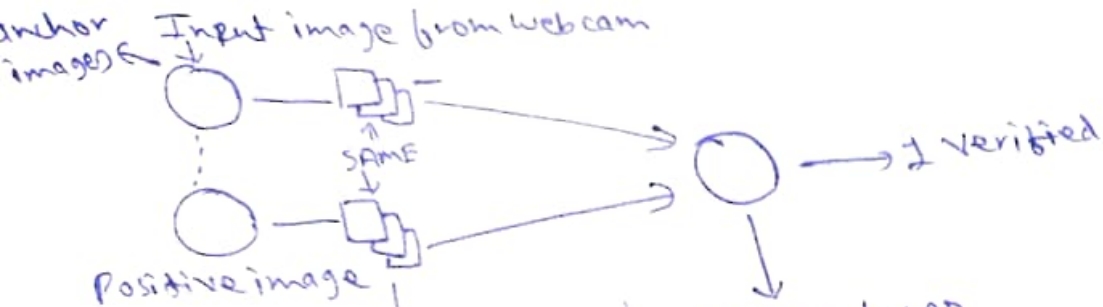
Positive example:-

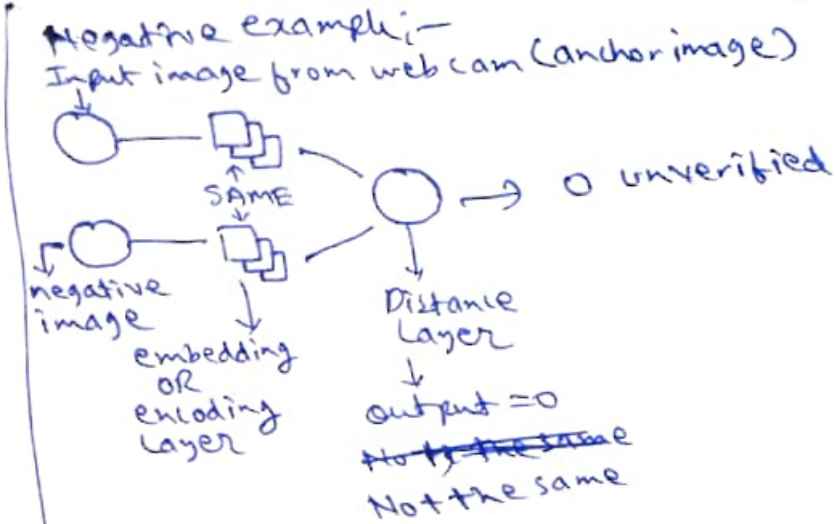


Embedding or encoding model
↓
We are gonna pass these Input & positive images into the encoding model It is gonna convert our anchor representation images coming from the web cam to our data representation, and then when we actually go and build our model what we are actually gonna be doing is trying to determine the difference between our anchor and our positive Input Images convert → Encoding

Here we are trying to see how similar our anchor images & our positive images are. and If they are very similar then what our model is going to do is it is gonna output a 1 → i.e we are verified. hence if the output is 1 then we can say that the person in the +ve image is the same person in the anchor image.

encoding model is going to learn how to best represent the input images to be able to ensure that we actually go to do a similarity analysis in our Distance layer that we all actually accurately classifying them either +ve or -ve

Negative example:-



SIDE NOTE: Now The cool thing about using "Siamese network architecture" in CNN is that if you want to go ahead and implement this on other people then you definately could. all you need to do is pass through a different positive image and pass through the same anchor image or pass through a different anchor image and it will be able to verify against whole range of people