Now, the now the so, if you look at look at the red deep nets right that sort of came out and then the sort of you know timeline right. It is about see 2012 when all this wave started ok. So, we are hardly like you know it is a decade right. So, that is when this AlexNet right when it came out right sort of. So, here right you see this AlexNet. So, you see the jump from 25 to 16 right until then it was like you know 2 percentage, 3 percentage and so on.

And suddenly you see like you know whatever 8 point jump from say 25 to 26 to 16. It is like you know 9 point jump or something then right everybody woke up. And at of course, right this also has to do with the fact that they actually created you know there was an image net you know. So, right this is the main net challenge.

So, of course somebody had to make that kind of a data set which should actually throw up this challenge right which means that you cannot have something simple and be able to solve these kind of problems. So, image net itself was a database with about 1000 classes and lots and lots of images right within it and then it was thrown up as a sort of a tough task to do. And then came this AlexNet which showed you know remarkable accuracy on that. And then right following that VGG came right this will be Oxford group and then a Google net came and then came you know ResNet and all this right. So finally, right you are looking at something 3.

57 this is still old right 2015 right. But this graph just tells you that for example, you know so at this level right 2015 we have already surpassed what humans can do. That means what this means is that if I were to show those images to you already you will end up not doing so well on the average right. So, that sort of gives you a timeline all this happened just too fast right in that sense this time scale is nothing 10 years is nothing right. So, before that if you think about it GABAR filters and all came like 50s and they were still being used and then it people as people are sort of doing things, but then they were still they knew that they are long way to go and then suddenly these things came and then of course the data set part right I will talk about it also ok.

It is not like it is not like right I mean you know everything is all the all the stories being told here on this one slide right. This slide is just to give you an idea that things really started looking up right in that sort of a timeline and then object this one a detection right that also became very very robust before that we did not have this kind of these kind of accuracies over feet then RCNN fast RCNN then YOLO and YOLO itself has several versions now. So, YOLO is like you look you only look once right and things like that some of these right we will talk about ok during the course, but very very briefly ok. This is just to tell you that you know image segmentation which is again at using people used to do using Marko random fields in the you know in the earlier days then came this annotated kind of thing right wherein you know people deep net could actually do a fantastic job and deep lab is one of them right which is which is which is very well known. Then then ok then some of this fun stuff right like style transfer and things like that right.

I mean this you must have seen I am sure you heard about this like for example, right you have you have a picture like this and then and then that you have you have a painting like this by somebody right I mean it is like saying that how would this picture have looked like if that artist were to would actually paint this right you sort of transfer the the style with then keep keep the content to be this you transfer the style of this guy onto this in order to be able to able to red show something like this. But again it this is like fun you know in the sense that there are lots of things which are unexplained here for example, it is not even clear what exactly you mean by style. So, people are just happy with something that just emerged and I think these days I thought I read somewhere this is like 2, 3 years old right that the first image that was created by the deep network was sold for a sold for I do not know it how much for some for a big amount that is like the first picture right generated by this machine. I mean you know some fellow felt felt excited about it and bought it for a whole lot of money I wish it was one of us right after all what does it take to produce some image and then anyway I think that must have been something nice about it. But then the very fact that right so all of this has also as also right you know has also kind of thrown up other kinds of challenges in the sense that right these days you can have images of let us say people right you know who do not even exist right.

So, these generative models which we are not going to cover the generative models and all have become so good that they generate human faces and those faces do not even exist on earth right, but they look very realistic. So, that goes into the forensics part right I mean that is that is even difficult whether an image is real or not right. Here it is about transferring style, but then people can make lives very complicated then artistic applications right. So, ok yeah so translation and then video sort of you know sort of a generation audio generation. So, it is like saying that if I told you the audio right can sort of can sort of create a video corresponding to that audio.

So, it could be like you know like a wave sound of waves and then maybe it is supposed to again it here you know it is like subjective you know you can you can argue you know whether that whether that video produces really good or not because there is no single video that you can associate with it right. If it is give you wave sound of waves right I mean each one of you if I think about if I ask you right what kind of a video would you associate with it each one will associate with a different video right. At the end of the day there is no ground truth and you cannot have a ground truth right because it is like it is like you know one you know which can which can kind of right map to map to several answers. So, so that way right people have gotten away in a sense right showing all this fun stuff people have written papers and all, but I think the real seriousness is coming now in the sense that right what really where really are the weaknesses of deep networks and you know and this kind of domain change and what that entails and how these things right begin to crumble and all that right. For example, right in our own lab I mean we have seen that you know things that you just take off the shelf and if you and you know some other people come let us say you know there is some other you know user that comes to you with a kind of bunch of images and says that see for example we had an example where you know it was actually you know

a defence project.

So, it is like you have a submarine right and it wants to know who is around right on the you know on the surface. So, what it does is it has it has periscope and this just comes out very fast right out of the water and then there is a camera right mounted on and it just spins and spins really fast and the faster you spin the better because you do not want to hang around for too long because the guy up there will know that somebody is there and then he will know that there is probably somewhere in down right. So, you do not want to show yourself off. Idea is that you come out very fast spin go down. The problem is right so the problem is when it spins that fast right if a human looks at that image looks terribly blurred because it is a it is a sweeping rate right and the idea is that is there is there is there some warship or something around you right that is what you want to know.

So, just take a quick sweep and come down and you look at that image looks terribly blurred right and then and then they said ok can you kind of do a de-blurring de-blurring means remove the blur. Now, we thought that oh there are so many de-blurring such a mature area right deep networks and deep networks fantastic scores people have written tons of paper none of them worked it is like that right. So, that is what I am saying right. So, it is not like you know every problem is solved out there the moment it is your turn right you will realize that it can kind of that is when you realize the truth that these all have to be taken with a with a sense of what do you say I mean you know with a sort of a pinch of salt right cannot get swept away.

Alright. So, the data sets right I just wanted to go to it right quickly show you I mean. So, see all of this it is still sort of engineered right and this is that if you did not have data you could not have done anything right none of this could have been could have been done if you did not have data sets right and now only people are talking about self-supervised ways of doing and so on, but really the earlier days were all supervised right annotations annotations and annotations and for that right I thought I thought what I will what I thought right I will do is I am going to share like this kind of document with you right this is a very very nice document I mean which my students actually right you know made I cut off many of the many of the data sets from this which are not probably relevant to you guys ok, but ok. So, if you look at it right. So, so this image net. So, there is Coco there is MNIST, KITI I mean these are all data set that you would have heard about ok for one task or the other something for something for the segmentation something for let us say vehicle surveillance something for object detection something for audio, audio in a AV sort of task right I mean where let us say you have a vision plus audio.

So, lots of things cityscapes, celebrity faces again for face recognition and stuff like that and then low light you know. So, there is something called low I mean and then there is something called I think there is so many of them and I cannot even remember the names of all of them, but there are all these you know LOL and GoPro, GoPro is again something that is very very well known you know people do a lot of deep learning and all using that.

So, all these you see data sets are out there and then one should realize that you know they have also been instrumental right in being able to enable what has happened ok. Without these data sets we could not have we could not have come this far and then there are groups that really work only on this that is also very interesting right you should be should be aware that there are groups that target only making these kind of see data set they never work on the work on work on the deep network per say. The job is to create create read these kind of these kinds of see data sets that that which they then share with the outer world and their citations will go very high because everybody is using that data set right because it is very comprehensively done and that is how they earn their fame.

Not by writing papers that will solve the problem, but by just sharing data sets open in the open right you just throw it open that also goes under goes through a review I mean you cannot just say that take this data set right it also goes through a formal review and then right people know that ok right if this data set was exposed to the public and then it can take the area further up. So right so data sets again right it is another is another big deal ok. So, we should not we should not sort of not not realize that ok. So, all of this I mean if you look at the you know right deep network and all you know ML right machine learning per say all started pretty early ok it is like you know 1940s and so on ok. So, people had started right then and then and then there was a point what is called you know a dark age I will share that slide with you ok.

There is some there is some there is a sort of a timeline right where you know it was considered a dark age because you know when let us say right people were actually excited about about a neuron about about modeling of actually a neuron then there were some people that came along and said that hey look this cannot solve certain kinds of tasks which are still very simple and therefore you know dark age kind of followed right when let us say people sort of felt that hey look all the hype was probably unnecessary. And then again came came a sort of you know you know a golden age when let us say people realized that oh right that problem could still be solved if you did this and then it went on and then this kind of back prop right which is the one that everybody uses now in order to do the optimization right that came and then and then again there was a dark age because the computing resources were not there the the data sets were not there then again the golden age started I do not I think around 2005 or 6 or around that time when data sets emerged right people were you know downloading like people uploading like crazy right all kinds of data visual data multimedia data and then and then came the right GPUs and all and then okay and then of course all of this happened. Okay so this idea of actually neuron right so the earliest was actually what is called this one a perceptron a perceptron okay this is the this was the I mean there are there was also something called you know a mclaughlin pitts model and all I am going to skip all that right because this is going to be a quick review right a perceptron is what I mean let us kind of consider to be a basic neuron right and the kind of form right that it has is the way our kind of neurons work right so what you have is so the way so the way right you would you would kind of read okay let me ignore this so the way right you would kind of think about it is you have a bunch of inputs right which

actually which which sort of enter let us say they are x1 x2 x3 all the way up to some xn okay and then and then and then you do some operations inside here let us say there is some see theta which will just for this kind of put in there and then outcomes outcomes your y okay now this now this y y and then what you would have is you would have you would want to wait the inputs you will have let us say w1 waiting x1 w2 waiting x2 w3 waiting xn wn waiting x wn waiting xn and so on right and the and the and the and the but the basic idea is this right so this y right would actually take a value if your summation wi xi right was actually greater than or or you know equal to theta will be 0 right otherwise okay so so so so this combination right so for example right I mean here you have a non-linearity which is which is which is to say that right so if you could have right I mean you know think about what does what this value of y is right and and if you say that it is along this you have your summation wi xi right then what let us say if 0 is here then what you are saying is if this exceeds a theta value right your y goes to 1 else it is 0 right so that is like non-linear right it is like it is like a step jump okay and then and then prior to that you have this kind of a kind of a right a linear operation going on which is which is a weighted summation of these of the inputs right now this this is this is a perceptron a simple model right which which apparently right mimics mimics what happens but this is still a standalone right the actual thing happens when all these neurons come together interact and so on that is when you have a deep network but at the very very basic level you have what is called really a neuron right and when and right this is where this is where we will start okay you know analyzing what this does and what kind of problems can this solve what kind of what kind of a classification problem can a perceptron just one neuron do and then look at what it cannot do and then see right what else needs to be done in order for you to be able to solve more more complex problems okay I will stop here today.