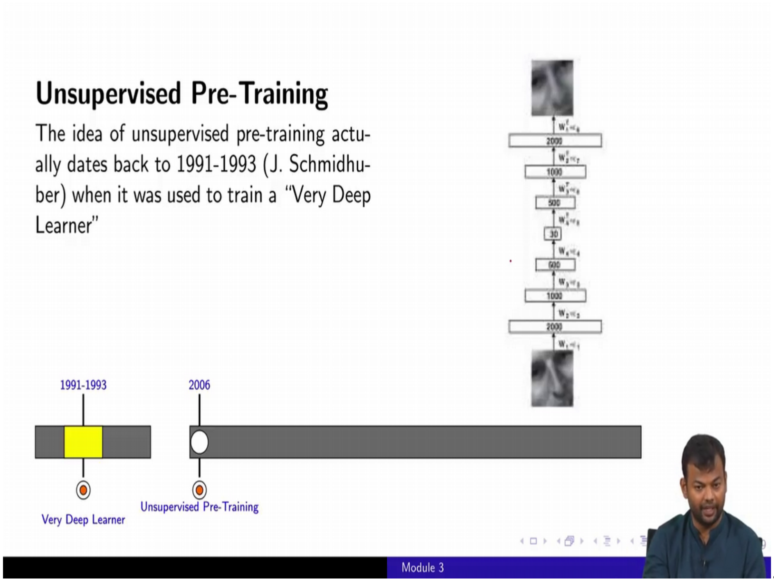
Deep Learning

Prof. Mitesh M. Khapra
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture - 01
Chapter 3: The Deep Revival

When this deep revival happened, right so in 2006 a very important study was or a very important contribution was made by Hinton and Salakhut Dinov.

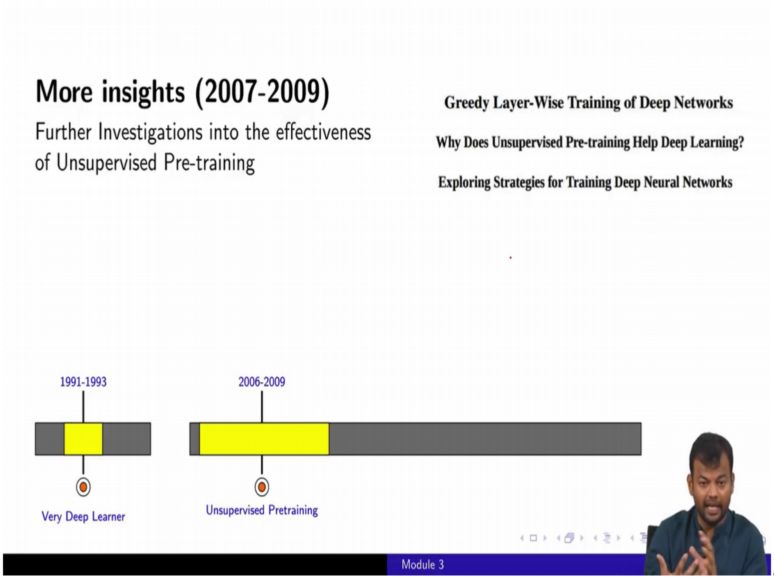
(Refer Slide Time: 00:15)



Sorry if I have not pronounced it properly and they found that a method for training very deep neural network effectively. Now, again the details of these are not important. We will be doing that in the course at some point, but what is important take away here is that while from 1989 to 2006. We knew that there is an algorithm for training deep neural networks and they can potentially be used for solving a wide range of problems because that is what the universal approximation theorem said, but the problem was that in practice we were not being able to use it for much, right.

It was not easy to train these networks, but now with this technique there was revived interest and hope that now actually can train very deep neural networks for lot of practical problems, this part of the interest again. And then, people started looking at all such of thing, right that even this particular study which was done in 2006 will actually be very simple to something done way back in 91-93 which again showed that you can train a very deep neural network. But again due to several factors may be at that time due to the computational requirements or the data requirements or whatever I am not too sure about that, he did not become so popular then, but by 2006 probably the stage was much better for these kind of networks or techniques to succeed. So, then it became popular in 2006.

(Refer Slide Time: 01:41)

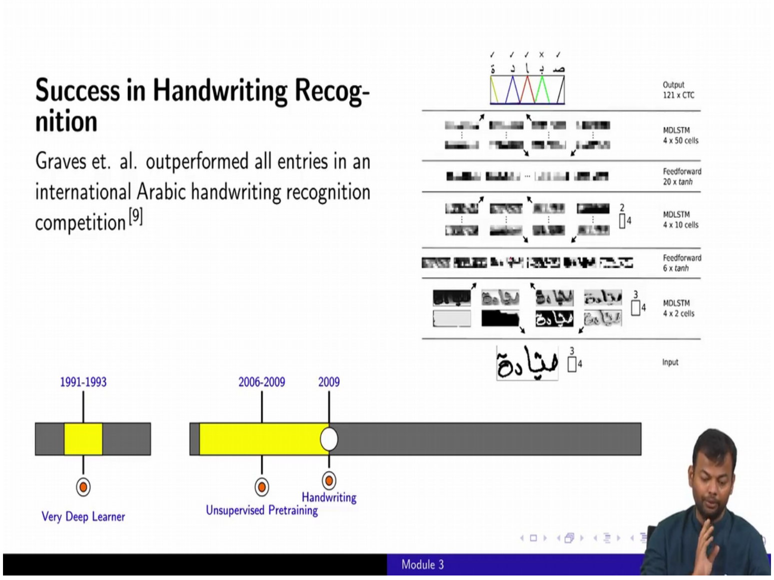


Then, this 2006 to 2009 people started gaining more and more insights into the effectiveness of this discovery made by Hinton and others which is unsupervised pre-training, right. That is what I spoke about on the previous slide unsupervised pre-training.

They started getting more and more insights into how you can make deep neural networks really work, right. So, they came up with various techniques, some of which we are going to study in this course. So, this was about how do you initialise the network better, what is the better optimization algorithm to use, what is the better regularization algorithm to use and so on. So, there were many things which were started coming out at this period 2006 to 2009 and by 2009, everyone started taking note of this and again deep neural networks of artificial neural networks started becoming popular.

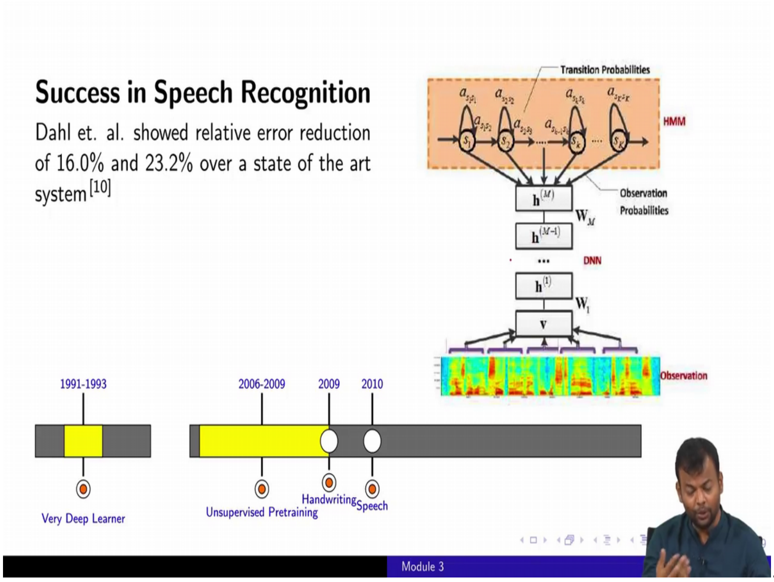
That is when people realised that all this, all the negative things that were tied to it that you are not able to train it well and so on helps slowly. People have started finding solutions to get by those and maybe we should start again focusing on the potential of deep neural networks and see if they can be used for large scale practical application, right. So, this 2006 to 2009 was again a slow boom period were people were again trying to do a lot of work to popularize deep neural networks and get rid of some of the problems which existed in training them.

(Refer Slide Time: 02:58)



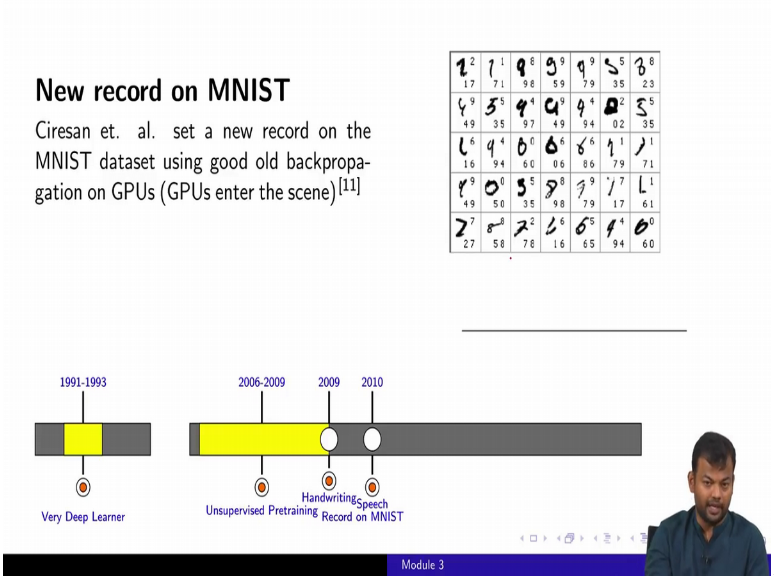
Now, from 2009 onwards there was this series of success is which kind of caught everyone which made everyone to stand up and take notice, right that this is really working for a lot of practical applications starting with handwriting recognition. So, around 2009, these guys won handwriting recognition competition in Arabic and they did way better than the computer systems using a deep neural network and then, this was a success.

(Refer Slide Time: 03:23)



So, this was an handwriting recognition and then, there was speech. So, this showed that various existing systems, the error rate of these system could be seriously be significantly reduced by using deep neural networks or plugging in a deep neural network component to existing systems, right. So, this was handwriting and then, speech.

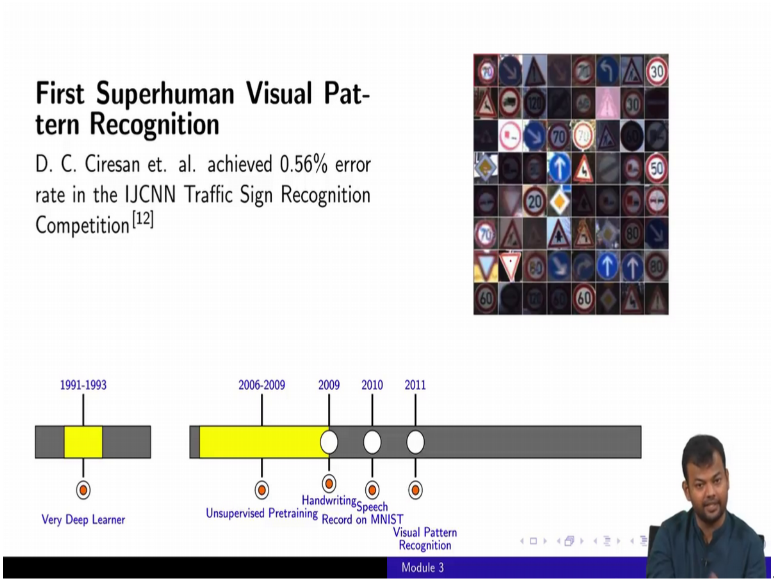
(Refer Slide Time: 03:45)



Then again some kind of speech recognition which was on android handwritten digit recognition for MNIST; this is a very popular data set which had been around since 98 and a new record was set on this data. So, this is the highest accuracy that was achieved on this data set around that time in 2010, sorry and this is also the time when GPUs entered the same, right. So, before that all of the stuff was being done on CPUs, but then people realised that very deep neural networks require a lot of computation and lot of this computation can happen very quickly on GPUs as supposed to CPUs.

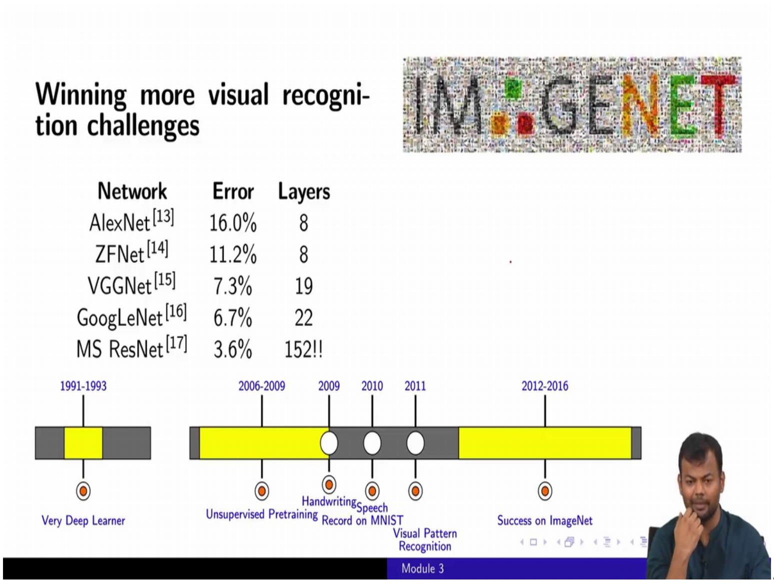
So, people started using GPUs for training and that drastically reduced the training and inference time. So, that was again something which sparked a lot of interest, right because even though these were successful, they were taking a lot of time to train, but now the GPUs could even take care of that and this success continued.

(Refer Slide Time: 04:38)



So, people started gaining or getting success in other fields like visual pattern recognition. So, this was a competition on recognising traffic sign boards and here again a deep neural network did way better than its other competitors.

(Refer Slide Time: 04:53)



Then, the most popular or one thing which made neural networks really popular was this Image Net Challenge which was around since 2008 or 2009 and before 2012 when this AlexNet was one of the participating systems in this competition, most of the systems were non neural network based systems and this competition was basically about classifying a given image into one of thousand classes, right.

So, this could be an image of a bird or a dog or a human or car, truck and so on say you have to identify the right class of the main object in the image, right. So, in 2012 this AlexNet which was a deep neural network or a convolutional neural network based system was able to actually outperform all the other systems by a margin of 67 percent, right. So, the error for this system was 16 percent and this is a deep neural network because it had 8 layers.

The next here this was improved further and something known as ZF network propose which was again 8 layers, but it did better than AlexNet. The next here even a deeper network with 19 layers was proposed which did significantly better than AlexNet. Then, Google entered the scene and they proposed something which is 22 layers and again reduced the error, then Microsoft joined in and they proposed something which had 152 layers and the error that you see here is actually better than what humans do, right.

So, even if a human was asked to label this image because of certain law, certain noise in the image and so on, even a human is bound to make more errors than 3.6 percent, right.

That means, even if you show hundred images to humans, he or she is bound to make go wrong or more than three or four of these images right whether there is this system was able to get an error of 3.6 percent over the large.

So, this 2012 to 2016 period were there was this continuous success on the Image Net Challenge as well as successes in other fields like Natural Language Processing, Handwriting, Recognition Speech and so on. So, this is the period where now everyone started talking about deep learning and lot of company started investing in it. A lot of traditional systems which were not deep neural network based was now started, people started converting them to deep neural network based system, right.

So, translation system, speed systems, image classification object detection and so on, there were lot of success in all these fields using deep neural networks, right. And this particular thing that we are talking about which is image net and the success in this was driven by something known as Convolutional Neural Networks.