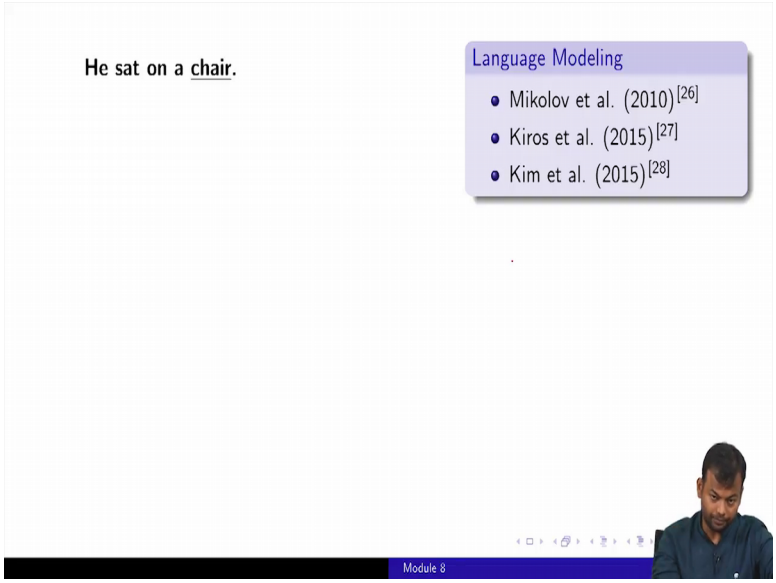
Deep Learning
Prof. Mitesh M. Khapra
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture – 01
Chapter 8: The Madness (2013-)

So, this was all happening where deep learning now started showing a lot of promise in a lot of fields N L P, vision speech and again this deep reinforcement learning and so on, which led to this complete madness starting from 2013.

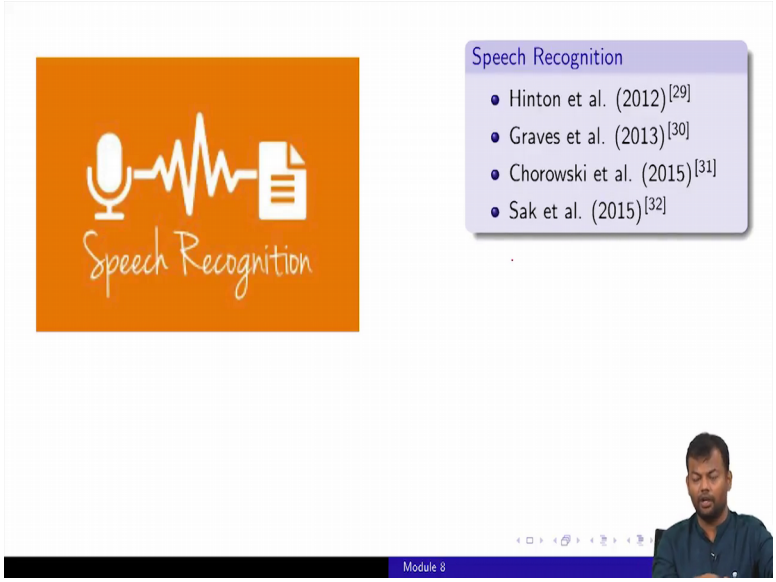
(Refer Slide Time: 00:28)



Well almost for every application the traditional methods were then overwritten or kind of beaten by deep neural network based system. So, something like language modelling, which has been around since probably 1950s or so.

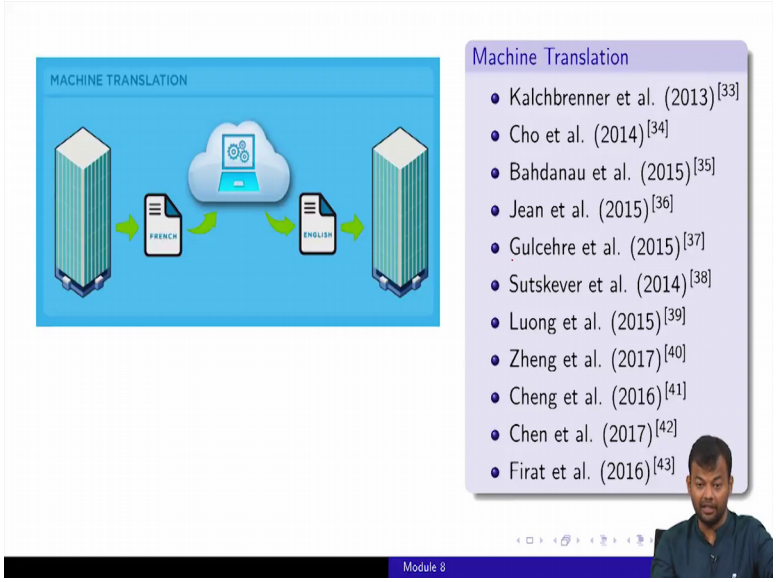
Now the reining algorithm or the better algorithm for language modelling is now something which is based on deep neural networks.

(Refer Slide Time: 00:50)



Then similarly for speech recognition, lot of work, a lot of probabilistic, lot of work based on probabilistic models was done in this or in the speech area or the speech literature for the past 30 40 years, and now all of that has been overcome by deep neural network based solutions.

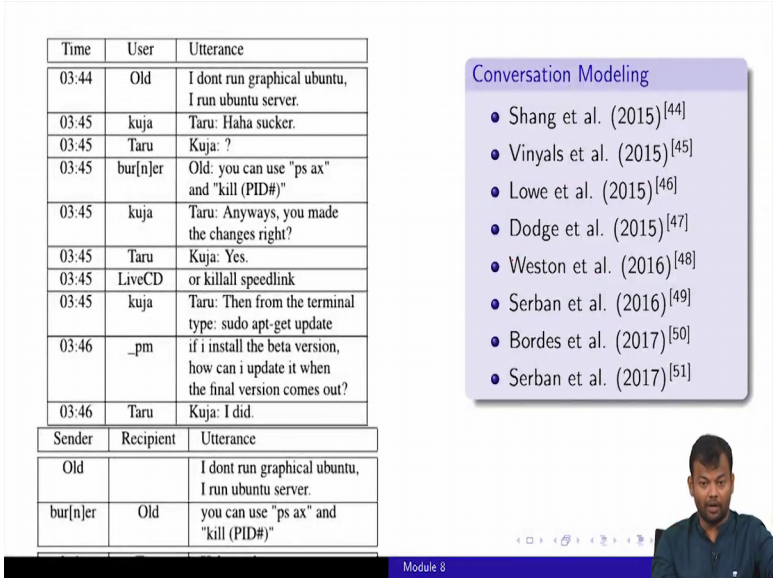
(Refer Slide Time: 01:08)



Same for machine translation, a lot of interest in this field, a lot of companies now have their machine translation systems based on deep neural networks as opposed to the

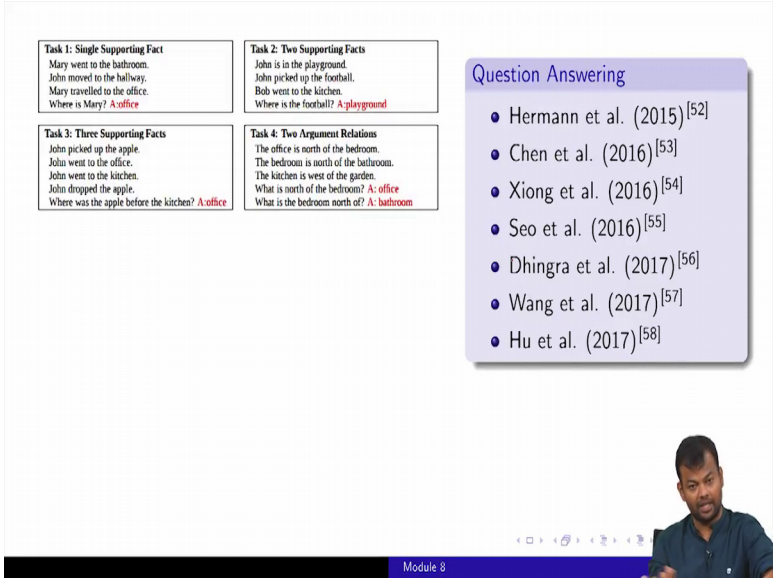
earlier phrase based statistical machine translations or the probabilistic models, which were used earlier.

(Refer Slide Time: 01:23)



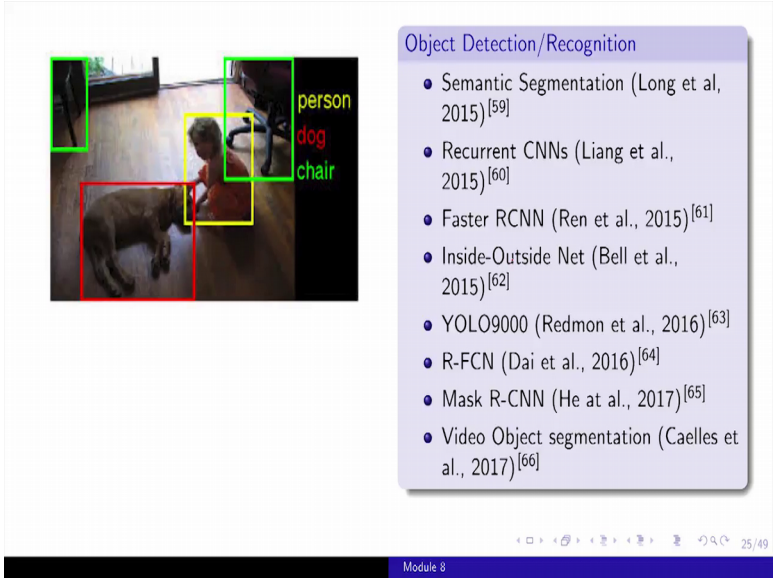
Similarly, for conversation modelling dialogue, a lot of new work started in dialogue post a deep learning era, where people now realize that if you have a lot of sequences of conversations, you could actually try to train a deep neural network to learn from this sequence and have conversations with humans. Of course, you are nowhere close to human level conversations, we are very very far off from them, but in limited domains these bots are showing some success now.

(Refer Slide Time: 01:50)



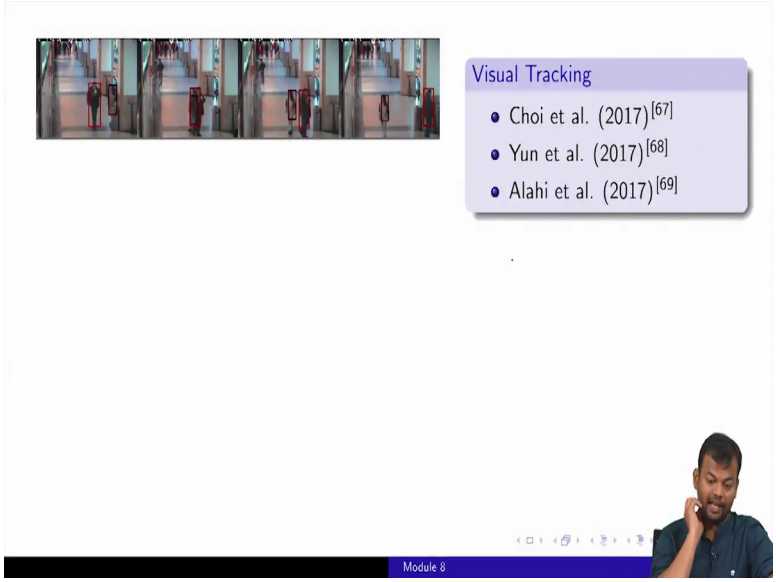
Same for question answering where you are given a question and you want to answer it, either from a knowledge graph or from a document or from a image and so on.

(Refer Slide Time: 01:57)



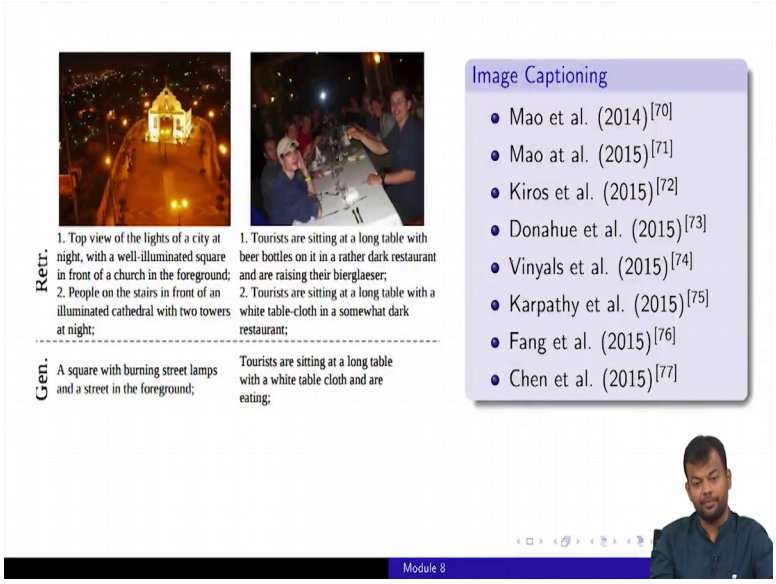
And in the field of computer vision things like object detection, most of the raining systems or the best performing systems, nowadays are deep neural network based systems, a lot of advances are being made on these systems over in the last few years.

(Refer Slide Time: 02:17)



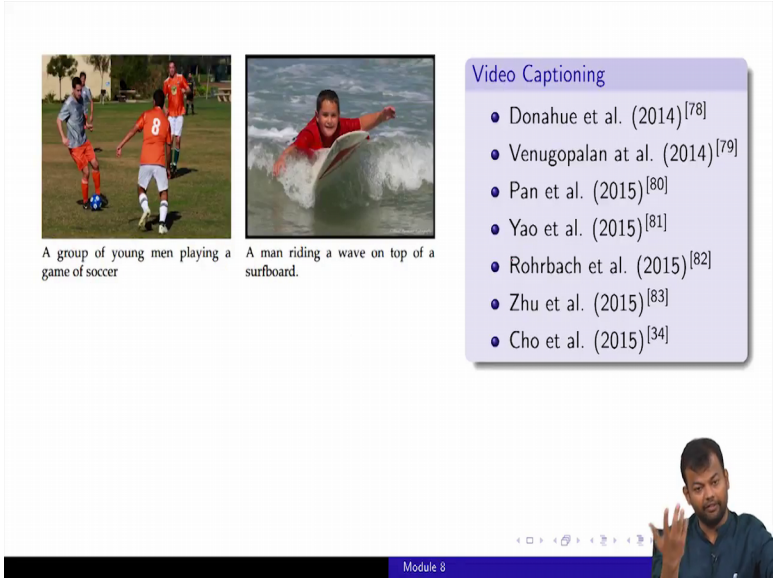
Same for visual tracking where you want to track the same person in a video or image captioning, where you want to generate captions for images. For example, people upload a lot of images on Facebook.

(Refer Slide Time: 02:21)



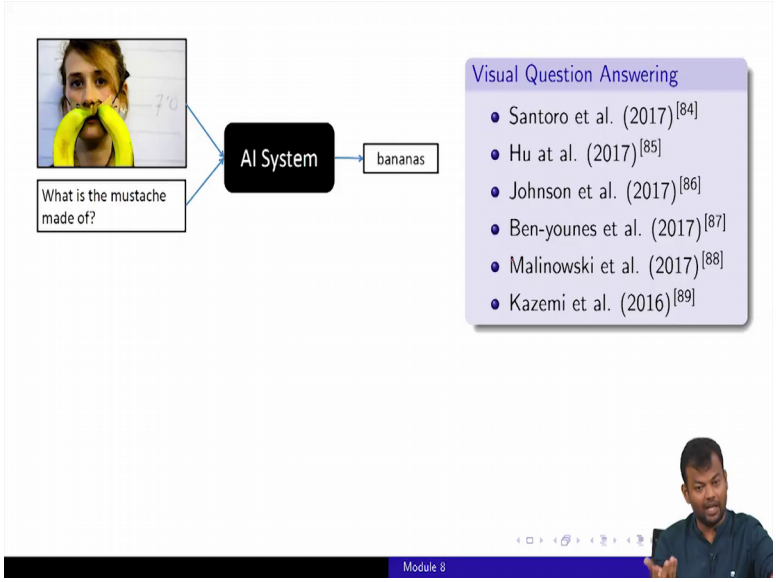
And if you want to automatically caption them or imagine you are on a reselling site right, something like O L X where you upload your furniture, and you do not provide a description from that, but can the machine already automatically generate a description for it. So, it is easier for the human to read what that product is and so on right.

(Refer Slide Time: 02:45)



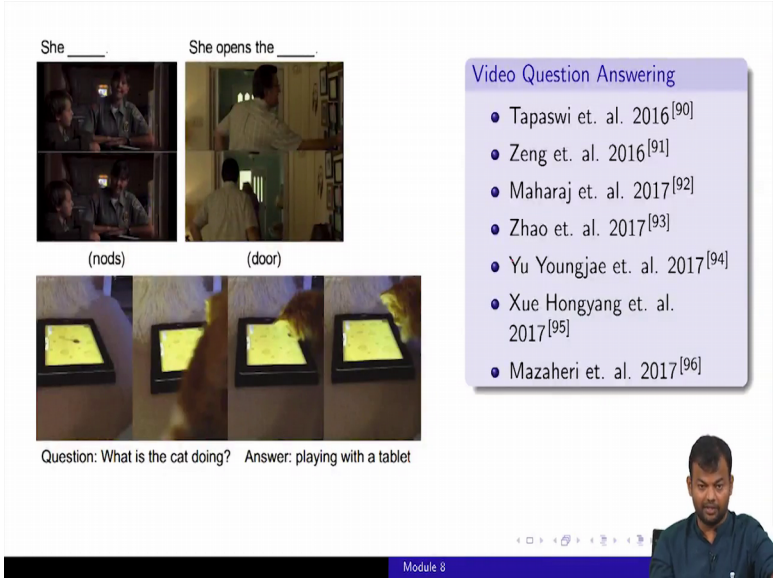
So, similarly video captioning, I given a video anyone to caption the main activity which is happening in that video; all of these problems are being solved using deep learning based solutions, using a combination of something known as feed forward neural networks or convolutional neural networks or recurrent neural networks and so on.

(Refer Slide Time: 03:03)



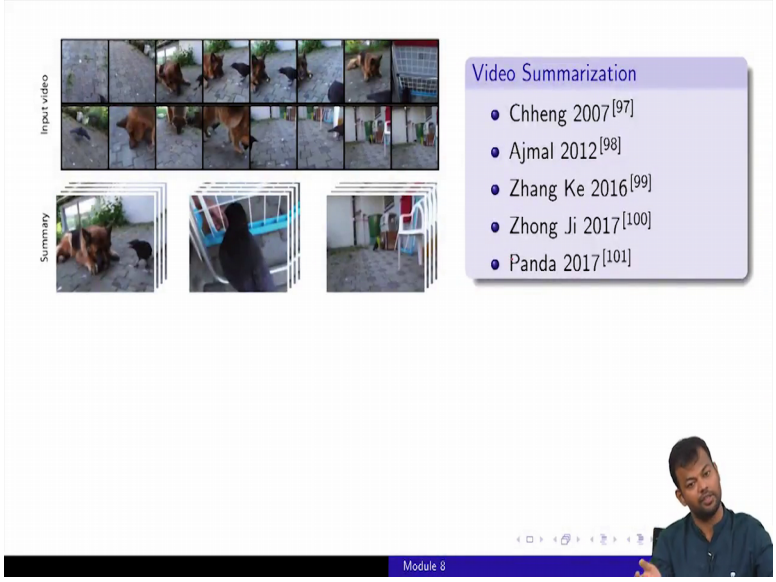
Visual question answering, you are given an image and a question and you want to answer that question.

(Refer Slide Time: 03:08)



Video question answering; answering questions from videos.

(Refer Slide Time: 03:11)



Video summarizations; if you are given a large video and you want to generate a trailer, a sort of a trailer for that video contains, which kind is the most important frame for that video. Even these systems are based on deep learning.

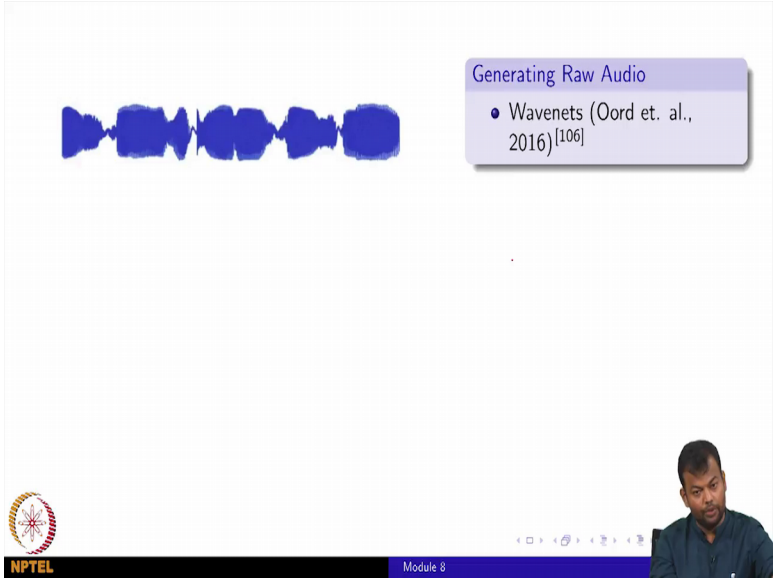
(Refer Slide Time: 03:22)

Then this was all about classification recognition and so on, but now people started getting more ambitious that can, we humans are very good at creativity. So, can we use machines to be creative right to generate images? So, now, if I have seen a lot of celebrity faces, can I generate new celebrity faces or if I have seen a lot of bedroom images.

And I am if a fireman architect. Now can I generate new bedroom images can i, can we train a machine to generate new bed bedroom images. So, a lot of phenomenal progress or work has happened in this field in the last 4 5 years, starting with things like generative adversarial networks, we reached an auto encoders and so on.

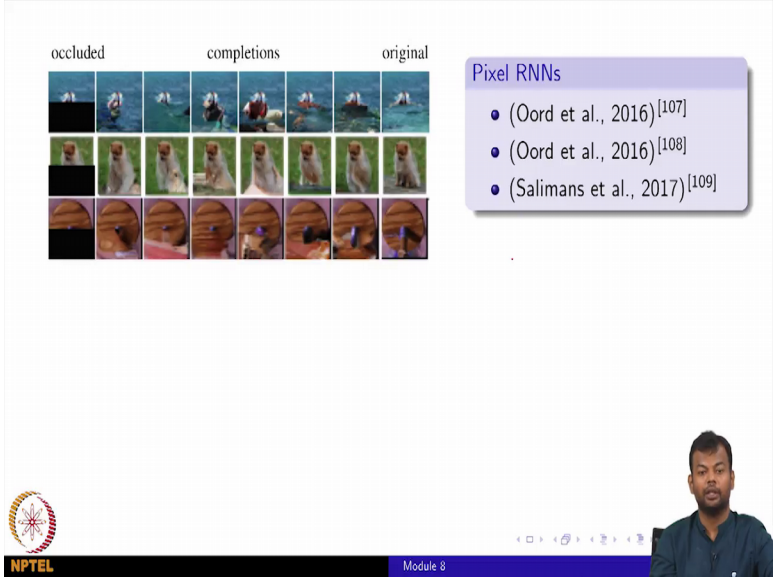
And people are now starting to seriously invest into creativity that how to make machines creative, again we are far off from where the desired output, but there is still significant progress happening in this field generating audio.

(Refer Slide Time: 04:15)



So, that was about generating images, you can generate music also.

(Refer Slide Time: 04:21)



And this is again about generating images and so on.