Affective Computing
Dr. Abhinav Dhall
Department of Computer Science and Engineering
Indraprastha Institute of Information Technology, Delhi

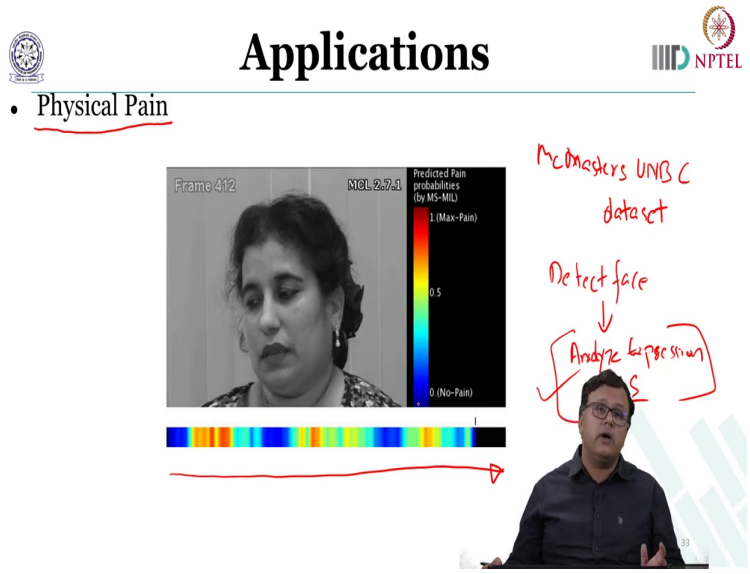
Week - 04
Lecture - 04
Automatic Facial Expression Recognition Applications

Welcome back, I am Abhinav Dhall and we are going to discuss about the Applications of Automatic Facial Expression Recognition as part of the Affective Computing lecture. So, friends we have been talking about different aspects of automatic facial expression recognition, how we create the different blocks, the aspects of the data and the recent things such as group level emotion recognition.

So, now I will be discussing with you some of the applications. So, it is like asking the question that you create a system and that system is able to identify the expression of a user. Now, once you know the expression, what do you do with it? What type of meta information can be extracted so, that we can solve a real-world problem and also how to use that information in driving a user interface? student:

o, we will start with different areas and I would like to draw your attention to the problem of digital health. Since, when we are talking about facial expressions, we are looking at the different movements in the face. So, we can use that information for different indicators about the health and wellness of a person.

(Refer Slide Time: 01:57)



So, here is the first example. How can facial expressions be used to predict if a person is feeling pain or not? Now, the assumption is that the understanding of a painful event is going to be on the basis of the expressions of the person. Of course, we know that there is a lot of variability. Some people express pain high; some would express similar event as low and this is about the self-pain labelling.

If there is a third person who is observing the patient, he or she will have their own different observation because there is a bias. So, there is a subjective bias which comes in when pain is described. Therefore, from my analytics perspective, understanding the pain objectively is useful for better care of a patient.

Now, imagine a scenario, there is an intensive care unit, there is a patient who is unable to communicate well and we want to understand that if the patient is feeling pain, how many

times have they felt pain? What were the intensity of those painful events? And, we are going to use the facial expression as the cues and then we can have a summary across time and that will give very rich information to the physician.

Now, what we have here on the slide is video from the McMasters, UNBC, University of North British Columbia data set. Now, this subject has gone through a shoulder reconstruction surgery ok. And, what they have is there is a camera in front of the person and there is a research assistant who is asking this person to move the shoulder, the place where they had the reconstruction done a few weeks ago.

Now, what we want to do is we want to detect the face, analyse the expression and this analysis could be your universal or could let us say be your facial actions, the facial action coding system. And based on this analysis, we want to plot let us say graph like this where on the timeline here, we have overlaid the score which we are getting from our classifier based on the expression analysis, where the red score means the classifier said there is a high probability that this person is feeling pain.

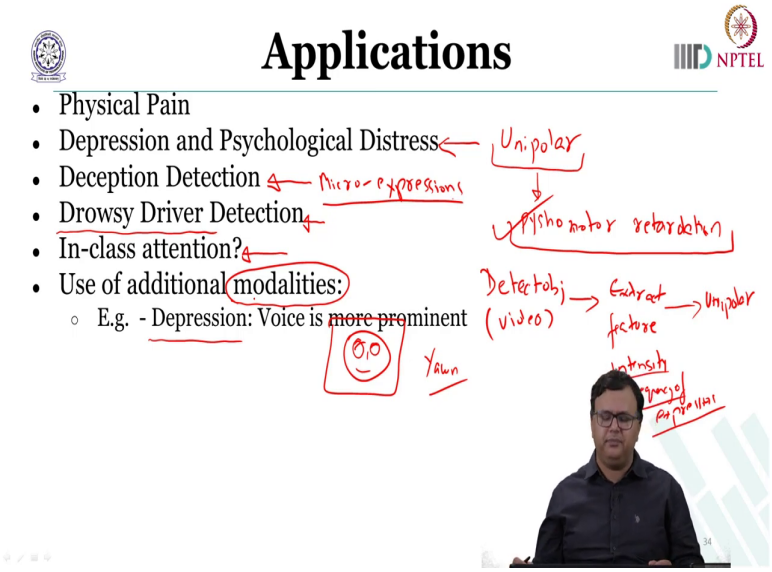
And, blue means there is a very low probability that the person is feeling pain so, this is like no pain event ok. Now, let us play the video. So, the person is moving their arm and you can clearly tell from the facial expression that they are feeling pain right. Again, they are moving there is a painful event. They will now rest down a bit so, expression is relaxed.

Now, again a painful event and you can tell that the person is feeling pain right. So, we can now have these information on a longer time duration and this can be very vital for the treatment. From a machine learning perspective friends, you could also ask this question that well when I am going to train such systems from where will I get such rich label data that at every time stamp, what was the painful event and what was the intensity of the painful event?

So, in that case what we typically do is we model these kind of digital health based problems as weakly supervised learning problems ok. Now, these problems would use these rich

information from the features and then we will use this classifiers to tell if the person is feeling pain or not.

(Refer Slide Time: 06:27)



The other one from the digital health one which I would like to discuss with you is depression and psychological distress. Now, depression is one of the biggest disablers worldwide and this problem is very prevalent, but there is a lot of social stigma around this problem. Typically, when you go to psychologist, they would look at certain cues and they will get some forms filled in by the patient in order to understand if the person is feeling depressed, if the person is going through a depressive episode.

Now, from the perspective of automatic facial expression recognition, what is being done is we can try to map the facial expressions of a patient onto a model and then tell if there is a

probability of this person being depressed. And, also in the same exercise predict the intensity of depression.

So, an example of that is the work which is being done at the University of Canberra in Australia. So, there is Professor Roland Goecke's lab. What they are doing is they are saying well if we are considering unipolar depression, we can make an assumption here. The assumption is that it is commonly observed that in the case of unipolar depression, a person may go through a psychomotor retardation.

Now, in the context of facial expression friends, what that will mean is that the frequency of occurrence of facial expressions, the intensity of facial expressions that will be mellowed down, that will be slow and less intense in patient suffering from unipolar depression. Now, what that means, is if we can get clinically validated data then we can try to find this observation by analysing the facial expressions.

So, the pipeline of course, here is also similar. You are going to detect the object which is the face in your video. So, you have now series of faces, you are going to then extract features which will be taking care of let us say the intensity and the frequency of expressions.

So, yes you have clearly guessed it right. You can actually now use facial action coding systems here right. You can say well the frequency of occurrence of action units and the intensity of action units; if I can have a feature representation perhaps, I can train a system which can predict unipolar depression.

Now, similar to this depression, you can use facial expression for understanding states such as confusion. So, there are lot of neurological diseases where confusion is one of the attributes. So, how can we assess if a person is confused? Now, within a conversation when someone is confused, you would be looking at their facial expressions and speech cues right. So, a multimodal system can be created in this.

Now, moving from digital health examples, let us look at some forensics right. So, there is deception detection, you would have also heard of lie detection right. So, there is let us say an

interrogator who is asking questions to a person and what we have is along with that interrogating asking question, there is a camera which is facing the person who is being interrogated.

So, one can look for micro expressions in that case and these micro expressions could indicate if a person is trying to deceive right. Recall, micro expressions are your involuntary movements, the short duration movements. So, since they are involuntary, only extremely well-trained actors could be able to control part of it. So, one could analyse micro expressions, train a system for deception detection.

Then from smart cars perspective, we have the drowsy driver detection right. Typically, people driving for longer durations would be feeling drowsy, tired. So, how can a camera-based system tell that the person is drowsy and of course, then issue an alarm. So, that the driver either takes a rest or you know gets more careful.

So, in this case what you are typically doing is you are saying well, now here is the frame and I have a face here, this is let us say the face of your driver. Along, with the facial expression you would also be looking at the eye region as well. So, you can use a deep learning based method or a traditional handcrafted based feature based method and that could look for sign such as yawning right.

Again, when you yawn you have different facial movement. We can try to detect yawning and that can raise an alarm right. So, that is a safety feature. Now, the next to this is at understanding the engagement and attention of a person ok. An example of that is let us say we are doing a training exercise, there is a group of people and we want to understand if the people, if the attendees they are engaged or not right.

So, in this case we would be looking at cues again such as let us say yawning and a bit of the pose and the expression as well right and this is far varied applications. For example, if you are talking of a human robo interaction, where a robo is trying to interact with a human and it needs to understand subtle cues. An example of that is when does the conversation end?

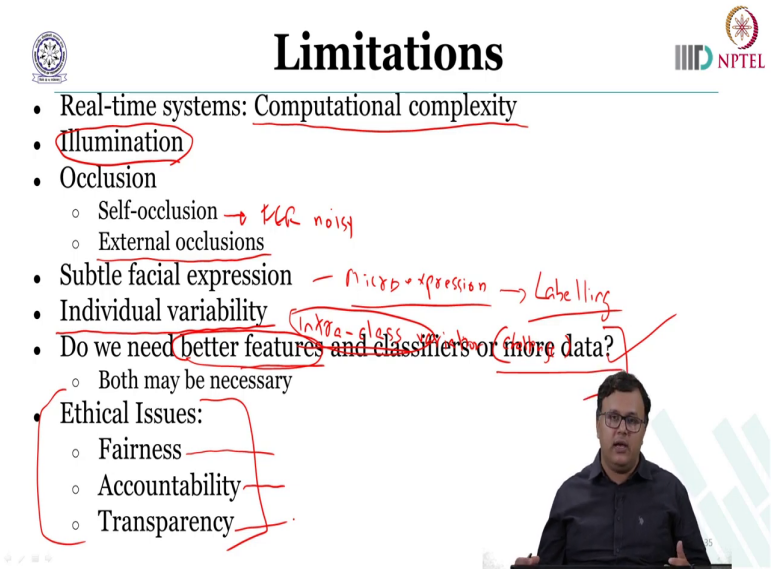
Is the user paying attention to the robo or not and if the user is busy in a task should a robo make some noise, some gesture and try to start the interaction? So, now based on the facial expression and other body cues of the user, one can try to understand the attention. And, that can be used for things such as human robo interaction or for in class attention as well, that you know you are doing a training exercise if people are paying attention or not.

Now, this is not just from the perspective of judging anyone, it is also that sometimes the content needs to be modified right. So, before let us say a content is shared to a wider audience, one can create a facial expression analysis system and do a testing with a smaller number of people. And, when you analyse the time series, you look at the expressions and then you can modify the content before it is released to a larger audience.

Now, one could also do another thing where you can say well you know for these problems such as attention or looking at you know the health and well-being problems, phase in itself might not be enough right. So, for example, in the depression, voice is also a very strong indicator of this psychomotor retardation form observation. So, multi-model systems, you can use modalities to make a more robust system. So, this will give you complementary information.

Now, these are some of the serious application friends. There can be some mobile phone based applications as well. For example, one could use a facial expression analysis technique to analyse the faces in a gallery app of a smartphone and the images which are presented to the user could be for example, sorted by expression right. Just another way of visualizing information.

(Refer Slide Time: 15:17)



Now, from the applications I would like to now take you to the last part which is essentially the limitations and where are the scope for improvement in automatic facial expression recognition. Now, one of the challenge which comes with automatic facial expression recognition systems, typically when you are using low power devices the computational complexity.

Here you are analysing the face. When you are analysing a face in an image that is typically a matrix which is very rich in data. So, the facial expression recognition system needs to be computationally light so, that we can do real-time analysis. The other challenge which comes now for automatic facial expression recognition is illumination.

Now, what that would mean is let us say you were doing a recording outside right and all of a sudden a very sunny day becomes cloudy. Now, what is going to happen? You are going to see the effect of change in illumination on the faces which are being recorded right.

Now, from computer vision and affective computing perspective what that could mean is if the automatic facial expression recognition system is not trained to be agnostic or to be robust to illumination changes, one could see noisy outputs. The third one is guys occlusion.

So, facial expression recognition under occlusion, when let us say there is self-occlusion; self-occlusion could be for example, a very heavy beard ok. So, you are not able to see parts of the face very clearly. So, in that case as well you could see that the facial expression recognition that could be a bit noisy right.

So, that is there is an opportunity here to improve facial expression recognition when there is a heavy beard or there are dark glasses. The other is external occlusion. This clearly, we have seen in the case of group emotions right. You have a group of people; it is possible let us say that the face of one person is partially occluded by the body part of another person right.

In that case, we have incomplete information right. So, if you have incomplete information, that incomplete information could lead to a noisy output right. So, there is an opportunity to do this partial data based facial expression recognition as well. The next one is subtle facial expressions right. So, these subtle facial expressions again are your micro expressions.

Now, in this case there is a considerable line of work where researchers are looking at micro expressions, but there are lot of challenges in micro expressions right. Of course, the first is you need very sophisticated equipment because we are looking at recording let us say 200 frames per second right. So, the camera and the hardware that needs to be quite sophisticated.

The other is the labelling problem right. So, labelling can be a bit noisy in the case of micro expressions. Now, if your labelling is noisy, if you get noisy labels during your training then it is going to affect the overall performance right of the facial expression recognition system.

Then friends, is the individual variability. What that means, is here you have two friends; let us say one is very open, the other is a bit introvert.

It is possible for the same joke; both of them may show different intensities of happiness or smile right. So, the laughter could be very high in one, could be low in other one. Now, from the perspective of training or automatic facial expression recognition system that would mean that there is a lot of intra class variation right. Now, this is again a challenge right. So, you would like to model this intra class variation.

Now, from the data perspective, you would like to have large intra class variation so as to cover as much type of expressions and the variability in expressions. From the machine learning model perspective, you would like to learn a generic representation which is agnostic to the intra class variation right.

So, if you have very high intra class variation in your data, but you do not have a sophisticated facial expression recognition system that will lead to noisy outputs. Now, here is another interesting question right. For a ideal facial expression recognition system which is more important, the higher quality features or from a deep learning perspective a more complex deep network or more data right?

Now, in more data it could mean both labelled or unlabelled data. Well, typically we will need both. So, now as you are increasing the amount of your data that would also mean that if enough care is taken during the recording of the data, you would have these individual variability which would be more intra class variation.

So, we would like to have intra class variation so, that we can capture the different facial expressions and your facial expression detector is not person specific, but it is generic right. So, with more data you would need better features and more complex networks. So, both go in hand in hand. The other is extremely important friends.

So, the ethical issues with automatic facial expression recognition. Recently, it has been observed that facial expression recognition systems in part have been used for things such as

automatic job interviews right. Now, as I have been discussing from the beginning that facial expression is just one indicator of the emotion.

And, it is possible that without context the facial expression which you see for a person does not really tell their real emotional state. Therefore, there are these issues, ethical issues which are you know asking these very serious questions that which are those areas where we should exercise caution when using automatic facial expression recognition.

Then, along with this is fairness. What is fairness? You have a automatic facial expression recognition system which is trained let us say on Caucasian data and Asian data. Now, during the test time if you get data from a subject of let us say African ethnicity, how will the system behave? Will it be fair? Will it do correct recognition?

Because, in the past in some studies it has been found that some of the commercial systems have been biased against women of color right. So, if you are going to use this facial expression recognition system output for serious things in such as health and well-being, then you have to be very careful that if system is fair.

It is not giving biased outputs based on things such as age, gender and color. Then, friends there is accountability. How did the system reach at a particular output? What was the type of data which was used? What were the design elements let us say of the network? So, we need all this information because ultimately facial expressions are used for serious applications.

And, then there is transparency as well. Now, transparency not only is from the perspective of the model or the system, but transparency in terms of let us say how the data has been collected. What were the rules and regulations, was appropriate approval taken from let us say an ethics committee?

Were the participants who were part of the data recording, were they aware that where the data will be used? So, all these are very serious concerns which are now being addressed

which are now being discussed openly, because facial expression recognition that has reached a healthy state.

And, this healthy state is now helping things such as you know in health and well-being and so forth. So, friends with this we reach to the end of our discussion in automatic facial expression recognition.

Thank you.