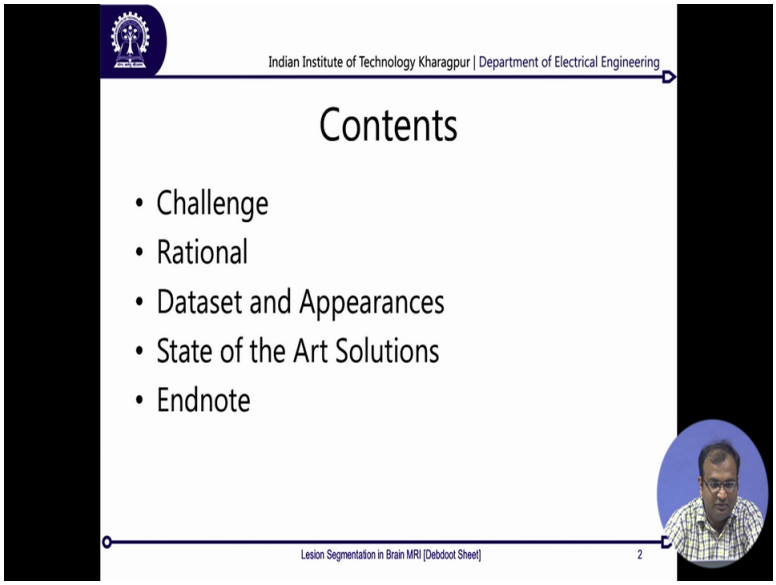
Introduction to Medical Imaging and Analysis Softwares
Professor Debdoot Sheet
Department of Electrical Engineering
Indian Institute of Technology Kharagpur
Module 4
Lecture No 18
Lesion Segmentation in Brain MRI

Welcome to today's lecture which is on another application area and we are going to discuss about Brain MR imaging and its application for medical image analysis. So I am going to speak specifically on Lesion Segmentation within Brain MRI and the kind of Lesion which we are looking down is from a particular disease, which is called as Multiple Sclerosis. We will come down to what is the disease pathology just a basic introduction onto the pathology and then due to that pathology what can be the problem caused by that particular disease and then eventually I would be discussing about on images on MR images how it is visualized and what is the visual appearance model and not just only in one of them but we will be taking down 4 different kind of structural MR images.

So like you have already leant in your MRI physics lectures about T1 and T2 weighted images being 2 different kinds of structural weighted images. We also have other kind of structural weighted images like proton density imaging or flair. So I will be coming down to those particular imaging modalities and just a brief introduction and what these Multiple Sclerosis locations on these Lesions actually look like on multi model images together.

So and there would be one fun aspect about this particular problem, because here we have a Lesion which is again non stationary in nature. So it would mean that there would be a Lesion if you the day of probing, and wherever you see the Lesion if after a week also you are again probing you would be seeing a Lesion in a different spot. The only thing which will remain constant is if there is a Lesion then it should be visible somewhere, else but then location does not remain constant.

(Refer Slide Time: 2:08)

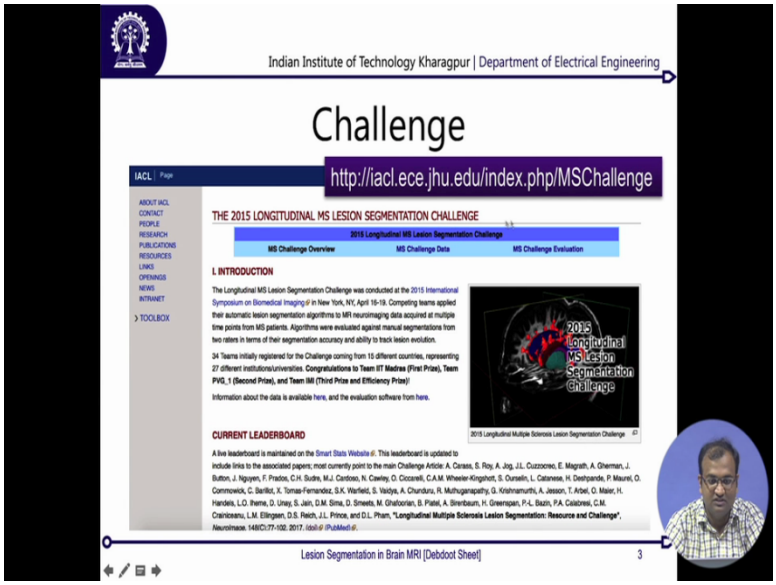


So it is sort of a moving around Lesion over there and then that has a quite a big clinical aspect implication as well. So the way in this lecture is organized is, I will be initially introducing you to the challenge, so although I have told briefly about what they solve problem is around over there. But then how this was crafted as a challenge for medical image analysis in one of the conferences and it is part of the grand challenge is still going on so I will be speaking about that.

Then enter into rational, as to what is the actual pathology as to how it behaves, and then what is the diagnostic significance associated with being able to do medical image analysis is what I would discuss in the rational. Followed by that, I will be making you aware about the dataset and the sort of variabilities which exist in that dataset including variations in the appearance models of these Lesions and the dataset over there.

So this would be variations across modalities, it would be variations on the same person, but so it will be variation of the same person across different modalities as well as the other side of variation is when you have the same modality but different subjects being imaged over there and all of them with a different grade of Multiple Sclerosis. So following that I will be discussing about few of the state of art solutions, which are proposed in the particular challenge paper which was released and then on endnote I will be ending it.

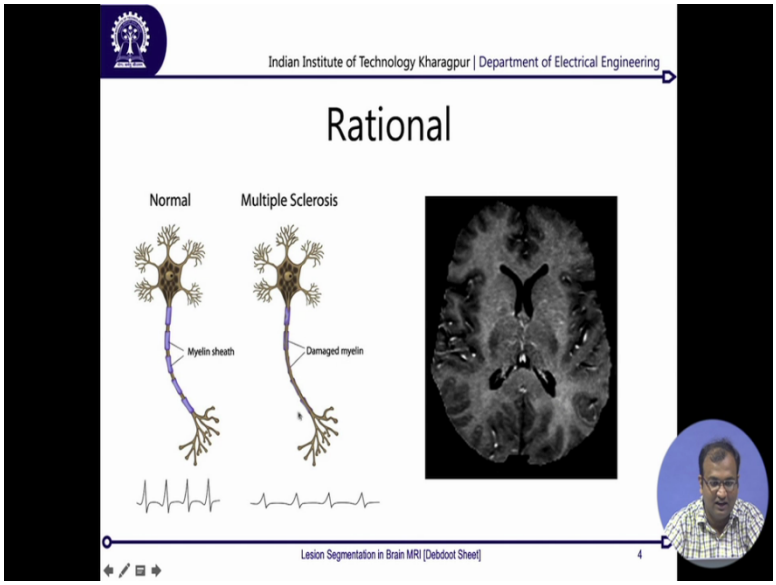
(Refer Slide Time: 3:31)



So this challenge is what is called as the 2015 Longitudinal Multiple Sclerosis Lesion Segmentation Challenge and this was held at the International Symposium on Biomedical Imaging in New York in April of 2015. And if you look over here on this one, the winning team was actually from IIT Madras and one of those few exceptional guys from India, who make up to the grand challenges podium on these kind of contest, which is predominantly dominated by all of other developed universities and we see quite a less participation as of coming from India.

And the hope is that a lot of people who are taking these courses based out of India you would also be getting more enthusiastically interested in going up and participating in these kind of challenges which is going to boost up the whole field as such domestically as well as on the international sphere together. So with that so that is a just a bit of motivational ones and then let us enter into what this whole thing is.

(Refer Slide Time: 4:33)



So on the rational side of it generally what this comes down as a challenge is on the pathology, this kind of a Multiple Sclerosis as a disease is actually associated with damaging of or some sort of a damage caused down to the Myelin sheet. Now, if you look at a neuron of your body so your brains and then your main communication path within your spinal cord, then your sensory conducting pathways from the skin to the brain and then form your eyes to the brain and then back from the brain onto your vocal cords which are going to make you a vocally excited to speak or say your eyelids blinking.

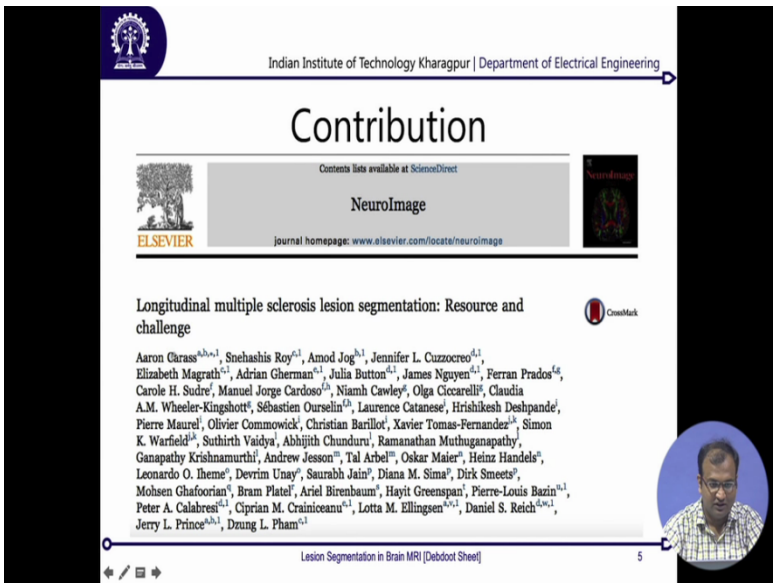
So all of them are connected through some sort of communication channel within our body and this communication channel within our body is what is made out of this fundamental units called as neurons. Now these neurons have some Axons Dendron and in between over here beside one another. Now this Multiple Sclerosis is basically associated with a damage being caused down to these ones and this is so the exact reason is still varied over there, but majorly it is a genetic disorder and due to one of the genetic problems, which gets carried down from generations to generation it keeps just happens out over there.

Now on the image side, what happens is so when somebody has a Multiple Sclerosis, generally they start exhibiting a lot of fatigue and uncontinuous fatigue then there would be a breathlessness when you are trying to do some work and then you will have consistent loss of appetite and this kind of problems, so but do not just get worried about whatever I am telling, a lot of us exhibit experience that but that does not really mean that we have Multiple Sclerosis, so there has to be conclusive diagnosis for that.

And the best way of diagnosing is we actually get down brain MR scans done down and so this one is what you see over here is a T1 weighted MR because your ventricles are appearing black, so it is obviously going to be a T1 weighted MR and on this particular T1 weighted MR you would be seeing that at some spots, so somewhere over here wherever this arrow keeps on coming down so you would be seeing down additional bright spots coming down and this one is basically of the same person who had some pre symptoms of he being affected by a Multiple Sclerosis and there were longitudinal scans taken down, which means that the person is coming for the first time you take one scan, then the person comes after six months you take another scan, then again after six months you take another scan.

So every six months you keep on repeating and you are looking at this whole study across time, which is what is also called as a longitudinal study. Now for this challenge they had actually released out longitudinal datasets for all subjects and that made it much more fun because now you have a 3D data of the same persons brain who is affected over different periods of time and we will be looking into how this was distributed, whether it was cross six months, or a year, or a year and a half, so all of these things are there.

(Refer Slide Time: 7:58)

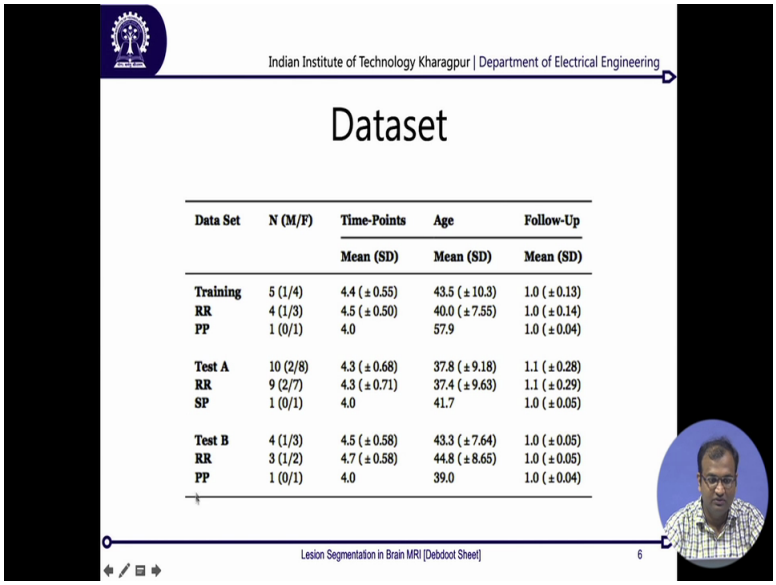


And then interestingly, you can actually try to localize how this Multiple Sclerosis Lesion has being moving around in space and that makes it a quite challenging and interesting problem. So if you look into this paper which was published which is recently published in this 2017 itself. Now on this one you would be seeing a lot names, now there were actually a lot of groups of people who were contributing to this one, multiple institutions who were doing and

together from all of them is who contributed to writing out this one single paper which got published out in neuroscience.

So it is about 23 or 24 pages roughly in length and is quite detail, starting from pathology to systemic data collection on a perspective scale and then trying to design evaluation methodologies and then how do you create a whole challenge over there and then integrate multiple solutions from multiple parties who are contesting on the challenge in order to come down with a much consistent solution, then would be done by human observers.

(Refer Slide Time: 8:50)



So they also provide whole summary of the dataset, now if you look over here they have the dataset which is divided into training and testing and they have testing datasets from two different centers and that is why you have a Test A set and a Test B set. So on one thing this would also be making it interesting because you can now look into domain adaptation problems over here, if I am training on one domain on one centers data can I also try to look into other centers data with the same equivocal performance coming down.

Now, on that they have two different kind of distributions on the dataset, one of them is called as RR and this is basically a collection of patients who already have a recurring history of Multiple Sclerosis, which means that when the first data was taken it was not actually the first time the person was coming down to a clinic. So the person might have come down to a clinic earlier as well and these persons, but the earlier data is not available for the challenge purposes.

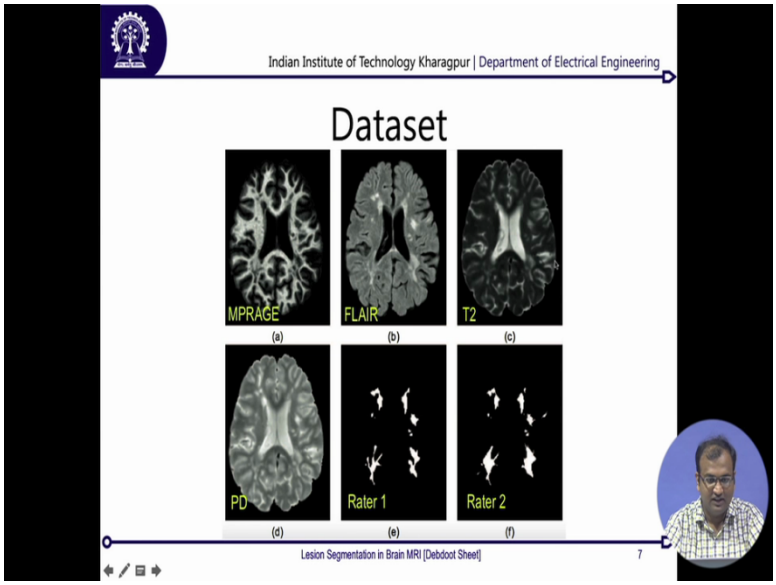
So it is a ongoing part and you are just taking a small snapshot over there and PP is basically all the patients who had come down over here for the with the first time of an appearance of Multiple Sclerosis. Now so the training has 5 datasets of which 4 of them are RR and one of them is PP and they also have a distribution of male and female perfectly balanced out over there, not exactly perfectly balanced but you have both constituents of male and female.

So what they have is basically in RR there is 1 male and 3 female and in PP there is there are no males and there is 1 female who is over there. Now this other things over here what they show is Time-Points, which is basically and then you have the age so this Time-Points over here are basically in months. They try to specify as to after what interval where image is taken.

So the average over here is 4.4 months with the standard deviation of 0.55, this is because all subjects were not imaged on the longitudinal scale with a same kind of an interval. So everybody had a different interval across which they were imaged over there. So the average duration comes down to 4.4 months and the mean age and standard deviation of the age of all the subjects who were imaged is also provided. Along with that we also have a duration to the follow up and this follow up is basically like after how many so one you had the first appearance over there for the patient, you take after a few months a group of snapshots and then you see that this kind of a Lesion keeps on rotating so that is a significant proof that there is Multiple Sclerosis, then you wait for a significant period of time may be another four months, six months or something.

And then you take another so that is what comes out as a follow up over there and then you repeat this kind of a longitudinal scan again. So this follow up over there typically is about a year which is over there for all of this. So you have the same thing for your testing cases as well given and this gives you a very clear idea as to what was the way in which your time stamp data was distributed longitudinally. Now given that we do understand about the importance on the datasets which we had studied in the earlier weeks in the first week about systematic data collection for evaluation. I would be showing you about the visual appearance and their manifestations over there.

(Refer Slide Time: 12:11)



Now what they have is they basically have four different sequences of structural data collected one of them is MPRAGE, and then you have FLAIR, T2 and a PROTON DENSITY image over there. Now, the point was why these were selected is basically based on clinical acumen and practicing radiologist who were specialist on MR and specifically on neuro-imaging based on their experiences of trying to locate down Multiple Sclerosis, they had suggested that these 4 modalities is what would be the best modalities to show down Multiple Sclerosis occurrences.

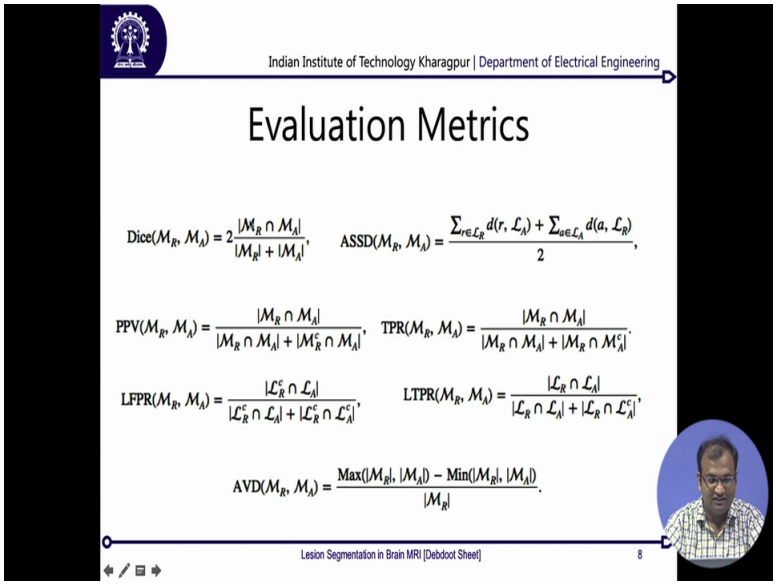
Now on top of that what they do is, since all of these four modalities are registered across each other, you have the same machine on which you are going to acquire all of them and more likely most likely it is basically one single acquisition of a raw case space MR data, which is eventually processed down to create this four different modalities of imaging data.

Now, since all of the slices and the whole volume is registered one across the other, so what you would typically have is that on all of them the same kind of a Lesion would be visible at that the same locations, but their visual appearances would be different. So somewhere it might be bright, somewhere it might be dark, and somewhere the textures would also be different. So based on that they have basically two different radiologists they are called as Raters, so we are the first Rater and the second Rater and both of them were asked to manually delineate and mark where this Lesion is present on a particular frame.

Now since all the frames are registered across each other, so given that you have this binary over there you can take this one and extend across all the different modalities and since you

have the binary marking available on volume space on the 3D data, you can take it consolidated together and then propagate it over the whole volume as well.

(Refer Slide Time: 14:18)



On top of that, since it is a same person of whom you have longitudinal data and on the longitudinal scale, which means that for every different time stamp where the data was acquired, you also have the markings or the ground truths available. And this creates a huge treasure troop for what you can work on. Now from there we enter into evaluation matrix so what they use is some of the standard matrix which we had already done but except for one major thing that this is extended onto the 3D space.

So earlier what we had done on the matrix they were all on 2D space, where I was discussing it in systematic evaluation and validation. So we did learn about what are positives, what are negatives, what is a dice coefficient, what is a house of distance. So we have extensions to them on the 3D space, because when you have say a 2D space, you would be getting down a contour as a boundary of an object. But when you are on a 3D space you would be getting a surface as a boundary of an object in a 3D space.

So you would have to extend out and create down new kind of measures in order to find out how good is the boundary delineation for segmentation. So over here we see that one of the measures which they have used is called as dice, so it is a standard same dice coefficient over there, which is giving you just a cardinality or the amount of overlap divided by the total amount the sum of total volume over there but it is no more area now you are going to count down voxels exactly on volume space and then this mod of MR intersection MA is the rater R

is a rater and A is one of these algorithms which has given out the results. So you take a intersection of wherever both in both of them the same voxel is marked over there divided by the count of both the voxels and there is a multiplier factor to which is the standard for dice in order to make it balanced in a 0 to 1 range scale.

So from there you have ASSD which is absolute sum of squared differences, so what this does is that you try to find out from each voxel to the closest voxel over there on the ground truth and then find out what is the squared of the difference coming down over there. So next we go down into positive predictive value and true positive grades and all of them will now be on the 3D case, so instead of pixels you will now be looking at voxels and comparing it with ground truth.

On top of that there are two interesting measures which are introduced over there and one of them is called as Lesion False Positive Rate LFPR, ok and the Lesion True Positive Rate, what means over here is typically you would see that the Lesion occupies a much smaller area as compared to the whole volume over there. Now if I am trying to look at accuracies all of them, then I have a classing balance problem, so majority of my classes is just background tissue and the minority is just my Lesion over there.

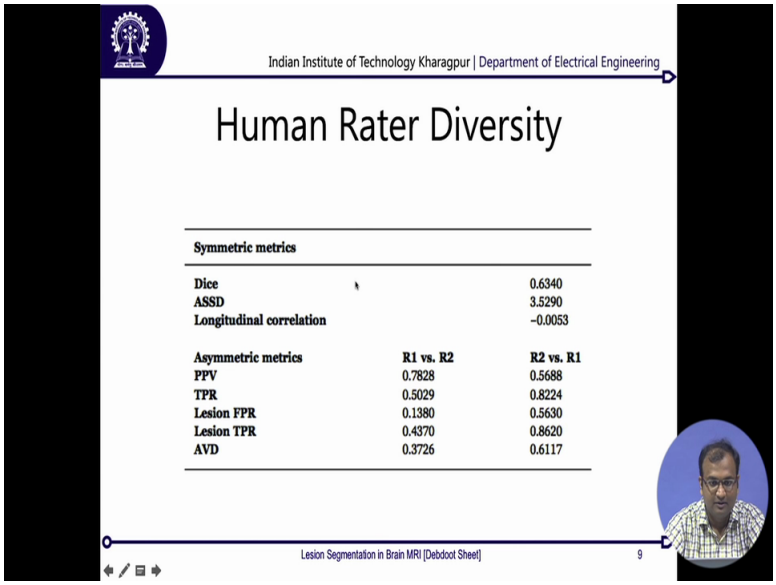
Now, I want to be very accurate about how good I am segmenting the Lesion, if say I am not segmenting even any Lesion, till my accuracies would not be figuring out the major difference because majority of my background is something which is devoid of my Lesion. And for that reason what they decided to do is, they decided some figures which are very specific to Lesions itself.

So they are not concerned about what happens to the background, but are very much concerned about whether the Lesion got properly segmented or it was under segmented over segmented, so these are Lesion Specific False Positive Rates and the Lesion Specific True Positive Rate. On top of that there is another one which is called as the Absolute Volume Difference and this is for the first time when they make use of volumetric concept over here.

So earlier, when you had looked into concepts of area, then you had absolute sum of area differences coming down and over here now it will be counting down voxels, so they end up having this as a Absolute Volume Difference between the predictions. Now, based on all of this the first thing which come down to our mind as to what will be the bench marks if we are trying to do a computer assisted diagnosis over there because the main goal of computer

assisted diagnosis is that it has to be much more consistent than human observers are or human raters are when trying to annotate and detect out the same Lesions.

(Refer Slide Time: 18:16)



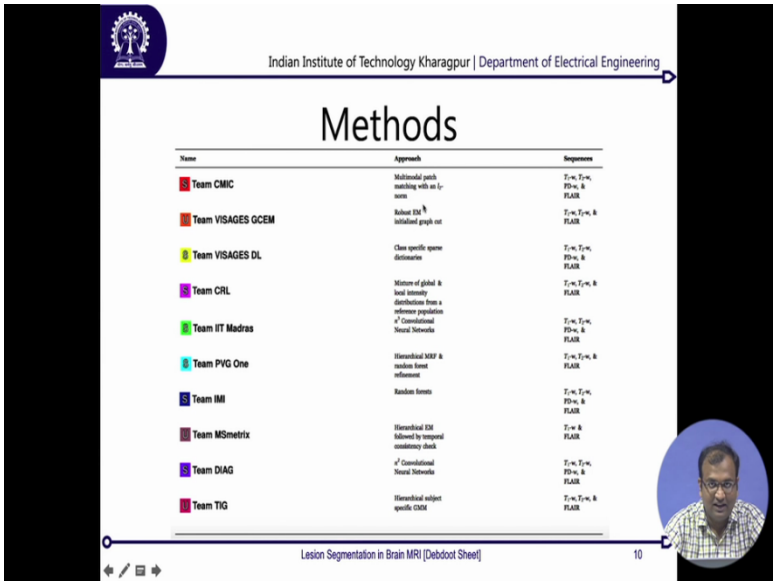
So for that they had run a basic evaluation study which was called as the Human Rated Diversity Study. So in this Human Rated Diversity Study they have taken down the same kind of matrix over there as we had done in the earlier case and what they do is they do a repetitive test for all of these PPV and TPR where you need a reference and the other one you are comparing.

So once they take R1 as the reference, the first Rater as the reference saying that Rater 1 gives the ground truth and Rater 2 is who is being tested against Rater 1. In the second case what they do is the Rater 2 is who gives the ground truth and Rater 1 is being tested against them. Now since each of these measures they are not symmetric matrix over there, so that would mean that if I am changing R1 with R2 then the value changes and that is what you observes typically over here.

Now you would see that when R1 is the reference and R2 is giving out some results, then you would see that the scores are much higher as compared to others for PPV whereas for TPR you would see the inverse trend over there. Now from this one it does come down that R1 and R2 are biased towards to different conditions, they are not biased towards the same kind of a condition either they are over emphasizing with respect to the other or under emphasizing with respect to the other and that is the major problem, which will come down when always trying to do it with human raters. Now in order to get rid of that is when

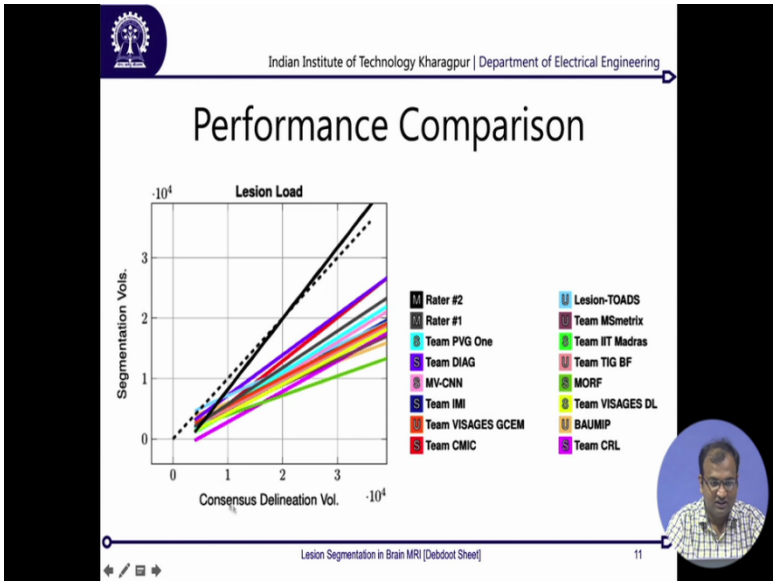
machines are brought into play. So we have a comparison of all the methods which were over here.

(Refer Slide Time: 19:36)



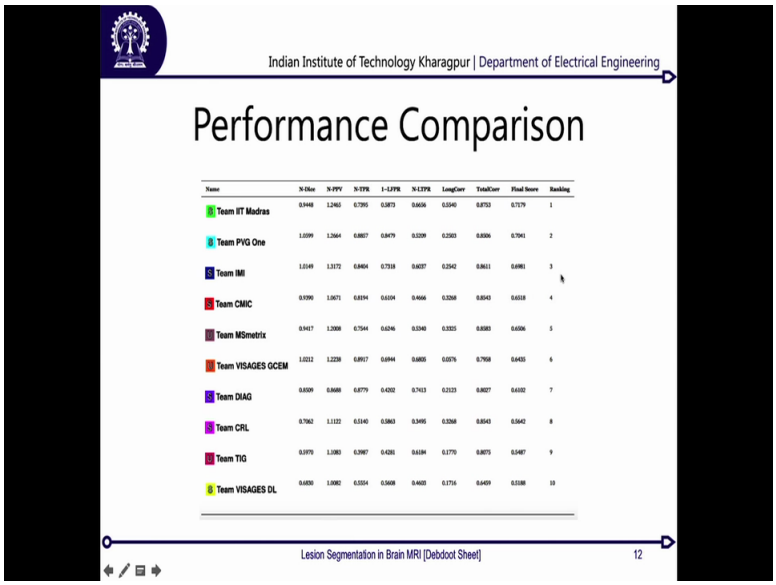
So now the slides it will be really hard to look over here but the details are there on the paper and on the slides which will be presented you can just have a look through them. Now there were multiple teams over there and each teams was given down a name and specifically I would draw down that all of these teams had different kind of contributions, so that included texture analysis to voxel morphometry, from there going down to use of commercial softwares like free surfer, then convolutional neural networks on 2D on two and a half dimension to fully 3D convolutional neural networks and then all of them were being used in order to segment out these Lesions in the best possible way over there.

(Refer Slide Time: 20:24)



And together what they come down is a quite a good performance, so what we have over here is a study about the Lesion load and this is about Consensus on the Delineation Volume versus the Segmentation Volume. So what this curve basically lies down over here is that if the Delineation Volume and the Segmentation Volume is equivocally given down over there, then you would be getting down an isotropic line over there at 45 degrees, whereas for others based on taking a ground truth and the other one being compared as to with respect to that how much do I have a difference coming down.

(Refer Slide Time: 21:08)



So this was just a fanciful illustration of that one given that you have just soft probabilities and then you make hard thresholds and then you are going to find out over there. Now, from

there when we enter into some actual numbers because this was just a fanciful way of showing whether everybody has a linear nature of performance or there are lot of nonlinearities. But when you come down to performance the figures actually have quite a conflicting results, because some of them might have a very good positive predictive rate, some of them might have a very good negative predictive rate, but then how do you combine all of them. Now the catch over here is they had designed a weighted scheme for ranking and the ranking weighted scheme was that methods will have to be consistently performing, the method which has the best consistence performer that is the one which wins over there.

So by that what they had done is, they had taken down all the team performances over there and then based on a particular score they had ranked out all the teams, ok. So each team now has a rank and then you can take a summation of all the ranks. Now, a consistent performer so may be a consistent performer is always getting a consistent rank of 3 or 4, but his but that method never gets a consistent rank of 1.

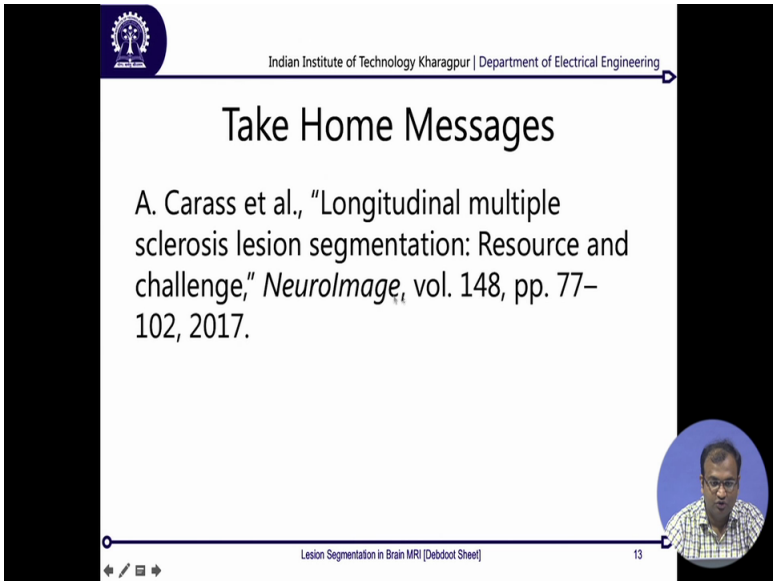
Other methods may sometimes get a rank of 1, but most of the cases it might get a higher rank. Now if you take down a total summation or say a multiplication of all of these ranks over there, than the one which is consistently performing on the higher side is the one which is going to get down the lowest sum or the least multiplication product over there and this is the queue which they use in order to rank it out total and what was found us that the contribution by this Team at IIT Madras, who were using actually a convolutional neural network in order to do Multiple Sclerosis Segmentation on volume, so it was a 3 dimensional CNN which was being used.

So all of your kernels in your CNN which you have implemented till now which were actually 2D kernels, now over here they become 3D kernels and today is current state of the art tools, say we had done it with torch. So you can actually implement these kinds of 3D CNNs over there, so the command over there is actually volumetric special convolution and you can easily implement a volumetric convolutional neural network on the 3 dimensional space over there.

So I would definitely encourage you to go through more details about on this paper for all the other methods and do not just restrict yourself to think believing that only CNNs can work over there because there are performers since by Random Forest based approaches and by Classical approaches, which also make use of textures and follow down with support vector

machines and they are quite well along in line and sometimes even out perform for certain kind of Lesions these CNN based methods.

(Refer Slide Time: 23:49)



So with that while we come to an end I would just point you down to this particular paper which is now recently published out in NeuroImage as well, so do make a note of this particular paper which has much more details and about all the methods and I belief you would be really interested to try them out as well. And do keep an eye on the upcoming challenges this year and although Multiple Sclerosis is not there, but there are many more interesting ones which are present and still ongoing in the field of Brain MR as well, so with that thanks.