SLIDE 2

First of all, I would like to introduce the topic. Since I imagine everyone of you might know about data augmentation I will not take too much time on it. As it has been proven, DNN generalize poorly when they are trained with low amounts of data, what leads to overfitting.

The standard solution to reduce overfitting is data augmentation that artificially enlarges the dataset by means of transformations on it that preserve the class labels.

However, it has its limitations firstly because not every technique can be applied to every dataset and because the increment in generalization is limited as it is proven in this paper.

SLIDE 3

Of course the ideal would be to have more data independent from the current examples so here is where GANs make act of presence. They consist on two models:

The first one is the discriminative model which simply tries to guess if an input sample is real or fake.

Ans the generative model which tries to fool the discriminative model by creating examples that look like real ones.

SLIDE 4

Entering to the papers, I have analyzed 2 in order to contrast them and I have seen that there are 2 approaches of the GAN. The first one is the simple approach, from a z noise distribution both G and D are trained and in Z appears a latent space.

The second approach allows external information in order to improve the quality of the generated samples. For example, generating labeled samples. The idea is to add an auxiliary network in the decoder to guess the label.

SLIDE 5

In the results section i Will focus more in one paper since there both GAN methods are compared. The paper consists on a medical application for classifying tumors. First of all, the approach presented in the paper was really interesting, they trained and evaluated the model with different amounts of data augmentation and shown that increasing the amount of data increased the performance up until a point, when they started the synthetic data augmentation with the GANs.

SLIDE 6

These are the results achieved with both, the DCGAN and the ACGAN. They are a bit surprising since in all the other references that used the ACGAN they got better results with it. Here there is one point that might be taken into consideration with the other paper and it is that they only used the ACGAN without taking into account that the DCGAN could achieve better results.

SLIDE 7

Finally, the results of the DCGAN where also compared to the performance of a variation of the Bag of Wards algorithm, which is the state of the art in medical TAC analysis, and it is shown that the GAN outperformed it. So that makes me think if the new possibilities of augmenting the amount of data that GAN will bring, will make that DL algorithms outperform the others in every field.