Experiments

The GAN is not detailed in the paper, we know that dropout is used and that the non-linearities in G are a mix of sigmoid and relu while non linearities in D are maxout units.
This architecture was run on 3 of the common public datasets:
1. **MNIST** : a popular database of 70000 handwritten digits.
2. **Toronto Face Database** : a large aggregate face database , cropped and centered.
3. **CIFAR**: 32×32 natural image dataset with 10/100 categories

It's performance was **compared to**:
1. A **Deep Belief network**, a hybrid hierarchical markov chain estimated representation.
2. **Stacked CAE** convolutional autoencoders a spatially aware autoencoder with allowed for greedy pre-training.
3. **Deep GSN** , backpropagation trainable Markov transition estimation
For data distributions.

The **evaluation** of the model was estimated using a **gaussian parsen window** (explained on next slide..)

Parsen window evaluation

- The method proceeds by computing the log-likelihood is given to the test set by the nonparametric density estimator trained on the generated samples (parsen gaussian estimator)
- The parsen gaussian method estimates a density by equal weighting of a set of gaussians centered on each output datapoint of G, in principle it is a generalisation to estimating a pdf by computing for each point at a small region the proportion of points that fall in that space. $p(x) = (k / n) / V$.
- This estimated density function is then assessed against the test-set through a log-likelihood: $\int p(y) \log q(y) \, dy$
- The variance of the gaussians are learned through cross-validation.

Results

Model	MNIST	Torronto Face database
DBN	138 +- 2	1909 +- 66
Stacked CAE	121+-1.6	**2110 +- 50**
Deep GSN	214 +- 1.1	1890 +- 29
Adverserial nets	**255 +- 2**	**2057 +- 26**

As you can see in Table 1 in paper the adverserial net has the best log-likelihood score on the MNIST set and the second best on the face database, variation are the standard error of the mean.
The results are on par with most GAN but no claim is made of superiority in output, below you can see the digits and faces generated, the two strips at the bottom shows interpollated values in the representation space.