G1: Autoregressive Models Conclusion

Motivation for autoregressive models

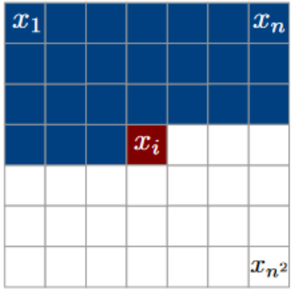
- Goal: estimate underlying distribution of data
- Predict future values in sequential data by using past values (autoregression)

MADE vs PixelCNN: Distribution

- MADE distribution:

$$p(x) = \prod_{d=1}^{D} p(x_d | x_{<d}), \qquad where\ \ x_{<d} = [x_1, \dots, x_{d-1}]^T$$

- PixelCNN distribution:



$$p(x) = \prod_{i=1}^{n^2} p(x_i | x_1, \dots, x_{i-1})$$

$$p(x) = p\left(x_{i,R} \middle| x_{<i}\right) p\left(x_{i,G} \middle| x_{<i}, x_{i,R}\right) p\left(x_{i,B} \middle| x_{<i}, x_{i,R}, x_{i,G}\right)$$

MADE vs PixelCNN: Loss

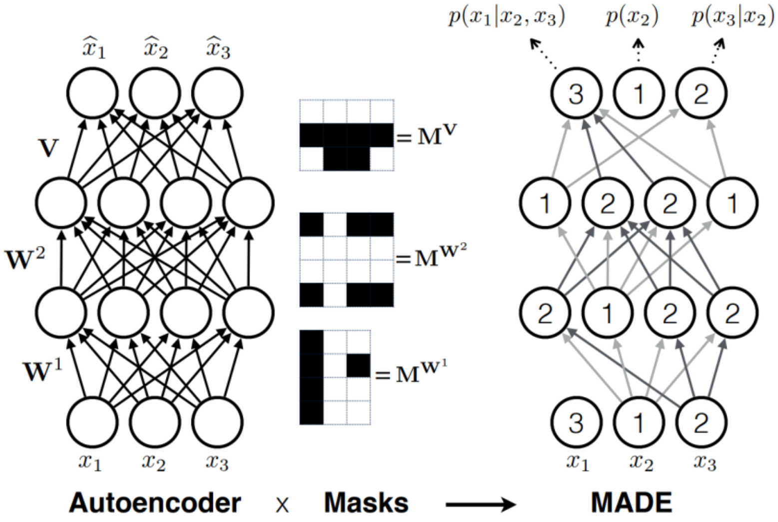
- MADE uses the cross-entropy loss

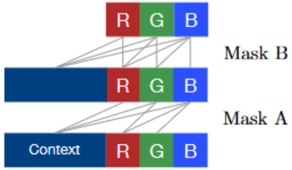
$$\ell(\mathbf{x}) = \sum_{d=1}^{D} -x_d \log \widehat{x}_d - (1-x_d) \log(1-\widehat{x}_d)$$

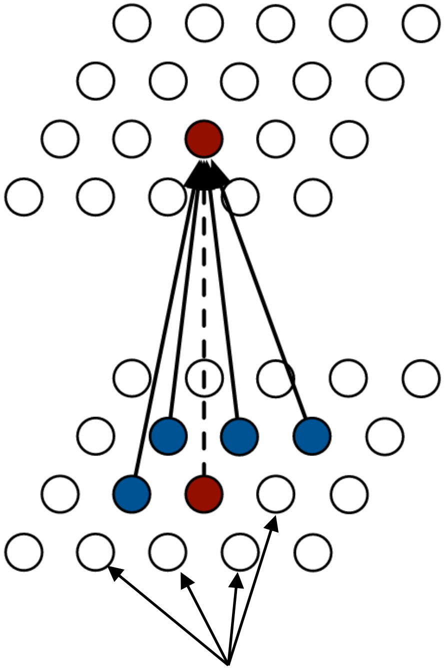
- PixelCNN uses negative log likelihood

$$\mathrm{NLPD} = -\sum_{i=1}^{N} \log p(y_i = t_i | \mathbf{x_i})$$

MADE vs PixelCNN: Masks



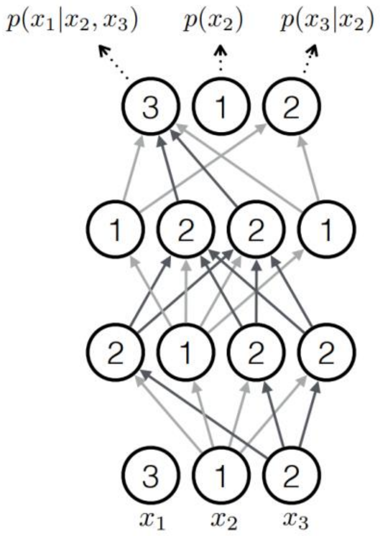




Masked out because they haven‘t been generated yet

Goal: Erase all non-autoregressive connections (logically)

MADE vs PixelCNN: Structure

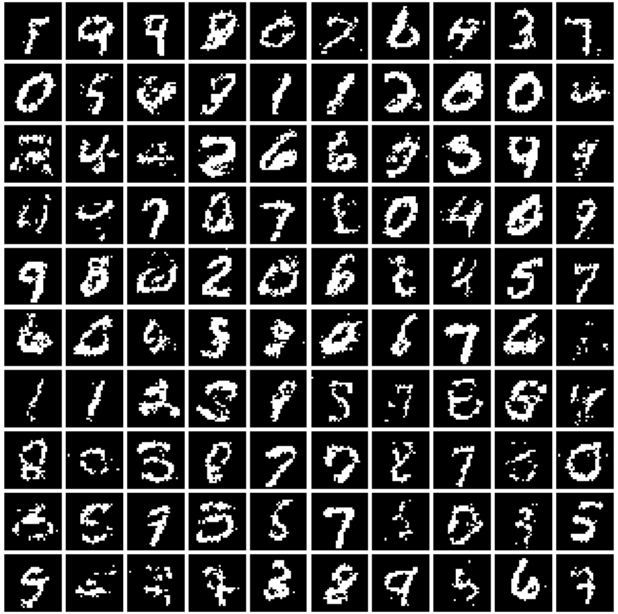


MADE

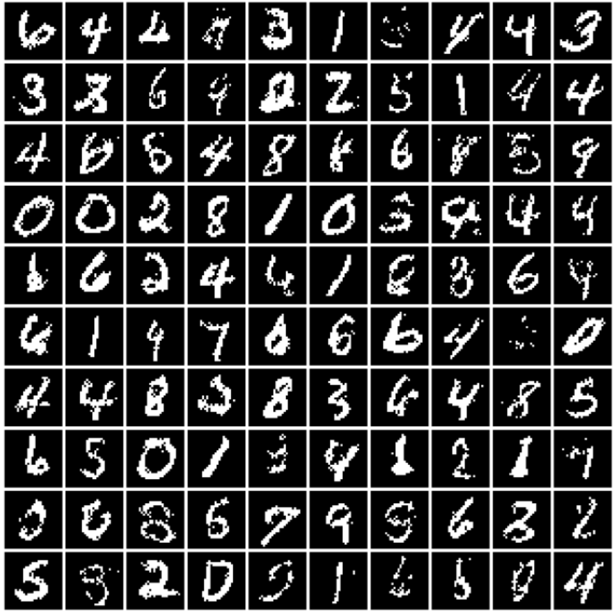
PixelCNN	Row LSTM	Diagonal BiLSTM
7×7 conv mask A		
Multiple residual blocks: (see fig 5)		
Conv 3×3 mask B	Row LSTM i-s: 3×1 mask B s-s: 3×1 no mask	Diagonal BiLSTM i-s: 1×1 mask B s-s: 1×2 no mask
ReLU followed by 1×1 conv, mask B (2 layers)		
256-way Softmax for each RGB color (Natural images) or Sigmoid (MNIST)		

Table 1. Details of the architectures. In the LSTM architectures i-s and s-s stand for input-state and state-state convolutions.

MADE samples

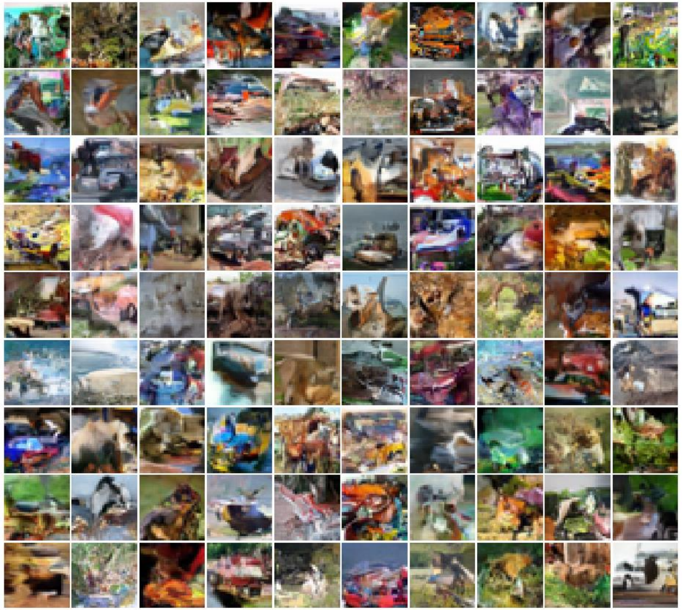


Samples from paper

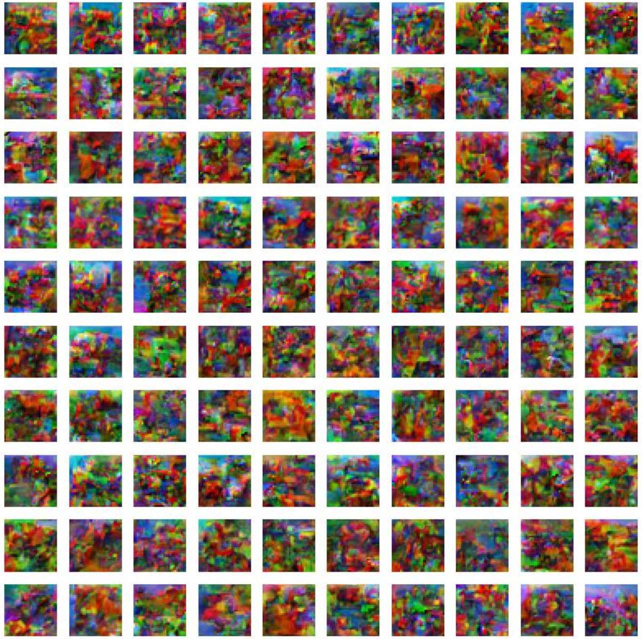


Samples from us

PixelCNN samples

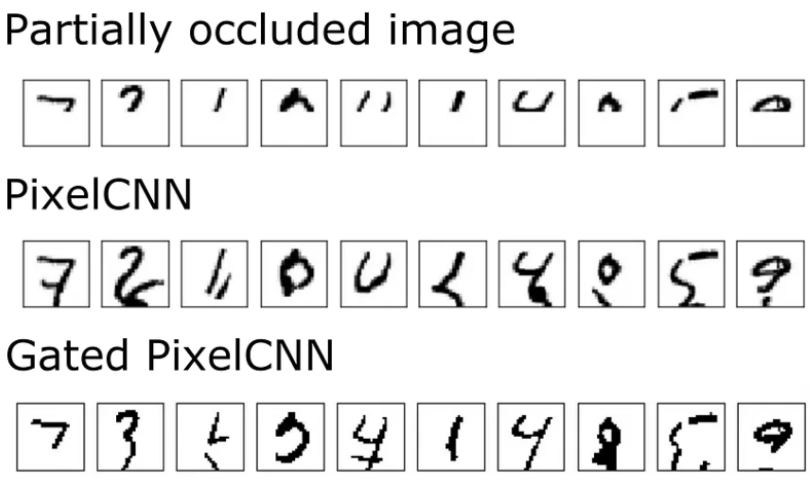


Samples from paper

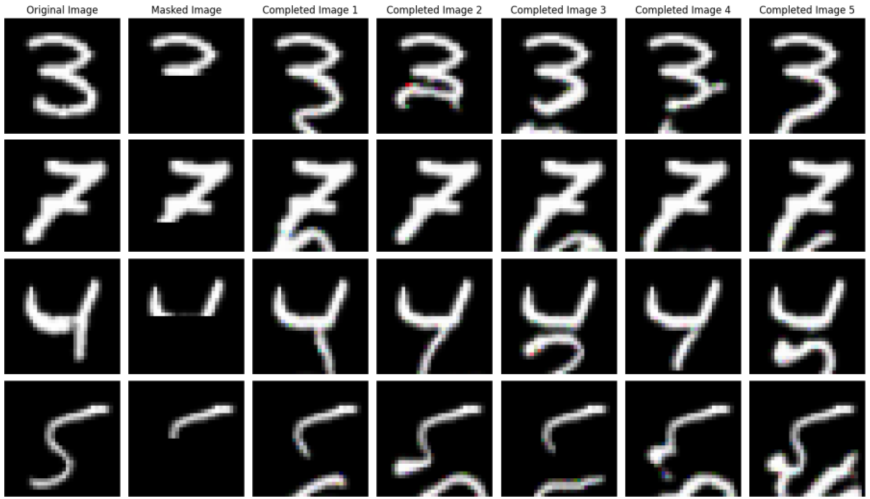


Samples from us

PixelCNN completions



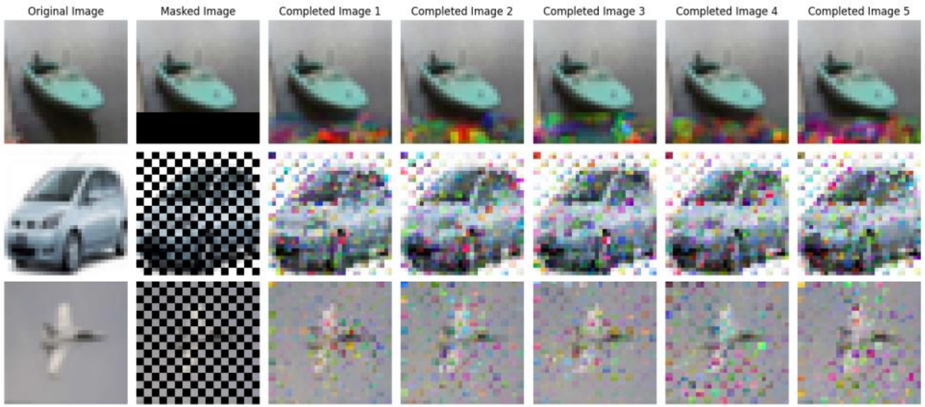
Completions from
https://towardsdatascience.com/pixelcnns-blind-spot-84e19a3797b9



Completions from us

PixelCNN completions

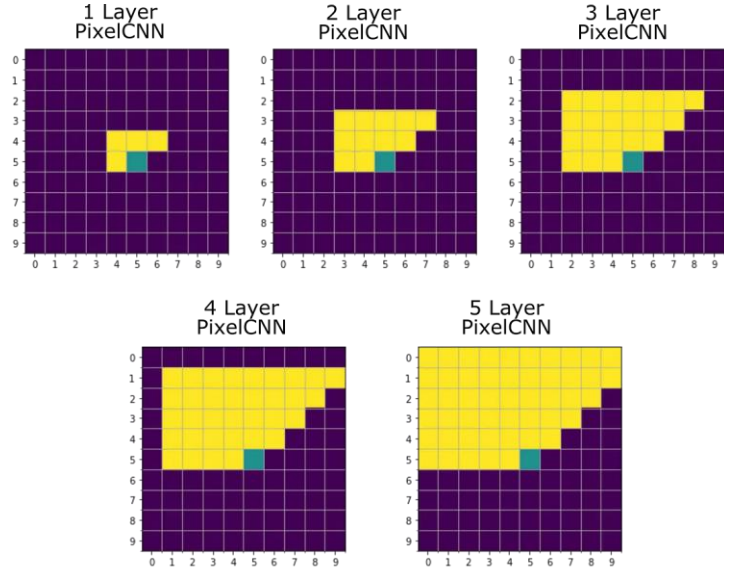
Completions from paper



Completions from us

Problems

- Autoregressive property of pictures is a problem in sampling
- Mistakes in sampling accumulate over time
- Blind spot problem (receptive field)
- Sampling takes long time



https://miro.medium.com/v2/resize:fit:1100/format:webp/1*V0V1blD6mdGkPmYDede3dw.png

Appendix

Sources

- Pixel Recurrent Neural Networks – Google DeepMind
- MADE: Masked Autoencoder for Distribution Estimation - Mathieu Germain
- https://medium.com
- https://miro.medium.com
- https://youtu.be/hfMk-kjRv4c?si=UGt55n13H9mIYcJa - Sebastian Lauge
- https://youtu.be/tIeHLnjs5U8?si=5IEHnfWKft9AT6lP – 3Blue1Brown
- MADE - blog by Kapil Sachdeva
- https://bios691-deep-learning-r.netlify.app/class/04-class
- https://neuroverse0.wordpress.com/2020/08/11/pixelrnn-gated-pixelcnn-and-pixelcnn
- https://images.datacamp.com/image/upload/v1647442110/image2_ysmali.png
- https://towardsdatascience.com/lstm-networks-a-detailed-explanation-8fae6aefc7f9
- https://www.jeremyjordan.me/autoencoders
- https://www.researchgate.net/figure/MLP-Autoencoder-architecture_fig4_373266879