EECS 6322 Final Project

By Ali Maleky, Shayan Kousha, and Vincent Sham

SurVAE Flows: Surjections to Bridge the Gap between VAEs and Flows

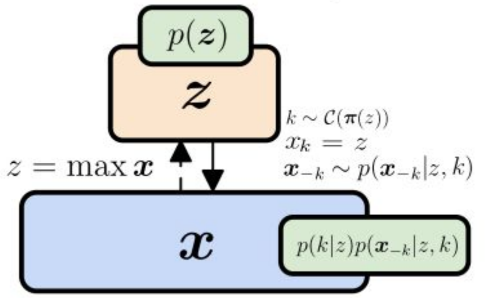
- Normalizing flows lack of ability to alter dimensionality, model discrete data and distributions with discrete structure or disconnected components
- VAEs lack of ability to do exact likelihood evaluation and Only providing lower bound estimates of the marginal density

SurVAE Flows: Surjections to Bridge the Gap between VAEs and Flows

- Unifying Framework:

$$\log p(\boldsymbol{x}) \simeq \log p(\boldsymbol{z}) + \mathcal{V}(\boldsymbol{x}, \boldsymbol{z}) + \mathcal{E}(\boldsymbol{x}, \boldsymbol{z})$$

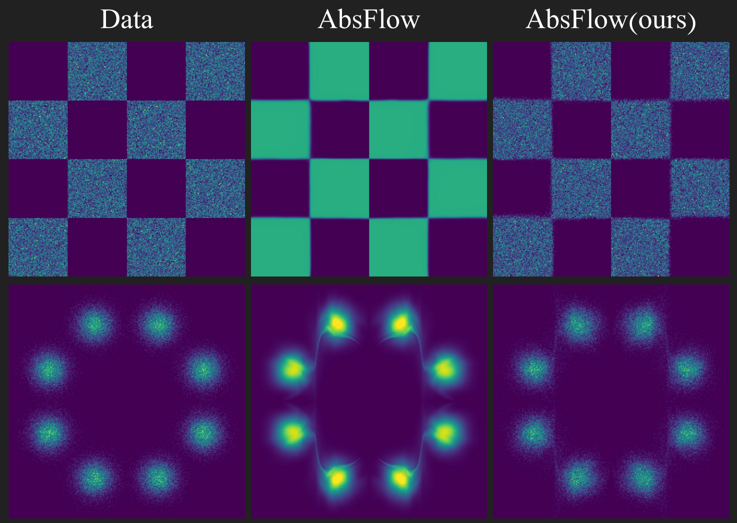
- Surjections:
  - Max
  - Slice



(c) Inf. max

AbsFlow experiments

Abs Experiments claim that certain SurVAE structures are able to help normalizing flows capture structures that are hard for normalizing flows to model; such as symmetric datasets.



Dataset	AbsFlow	AbsFlow(Ours)
Checkerboard	3.49	3.49
Corners	3.03	3.03
Gaussians	2.86	2.86
Circles	2.99	3.03

Max Pooling Experiment

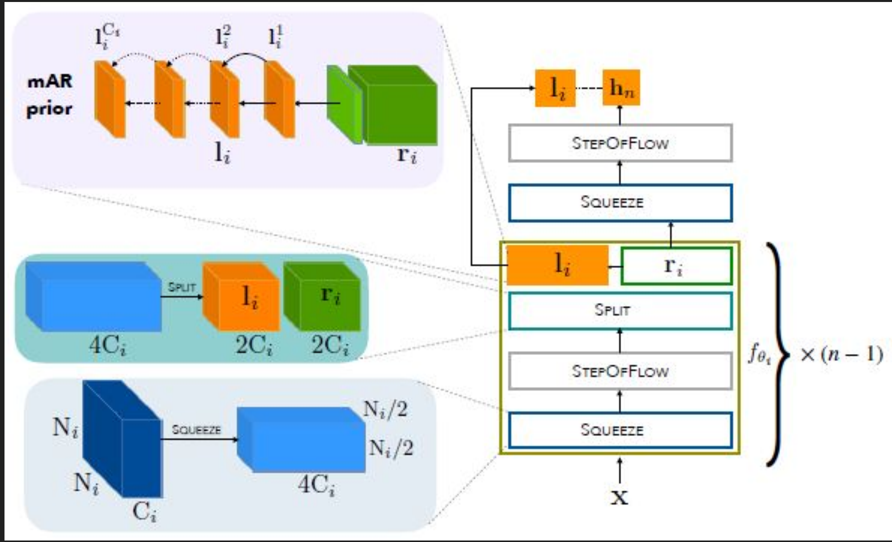
Claim: The performance of the max pooling surjection is similar to the slice layer introduced by (Dinh et al., 2017) when used to create a multi-scale architecture.

Models	From SurVAE Paper	Ours
RealNVP (Dinh et al., 2017)	3.49	-
Glow (Kingma & Dhariwal, 2018)	3.35	-
Flow++ (Ho et al., 2019)	3.08	-
Basline	3.08	3.16
MaxPoolFlow	3.09	3.19

* Dinh, L., Sohl-Dickstein, J., and Bengio, S. Density estimation using real nvp, 2017.

Normalizing Flows with Multi-Scale Autoregressive Priors - summary

- Multi-scale structure of the latent variable to reduce computational cost
- A combination of `Squeeze`, `StepOfFlow`, and `Split` operations.
- Channel-wise autoregressive dependencies in the latent space.
- Enables Conv-LSTMs for faster inference and sampling.



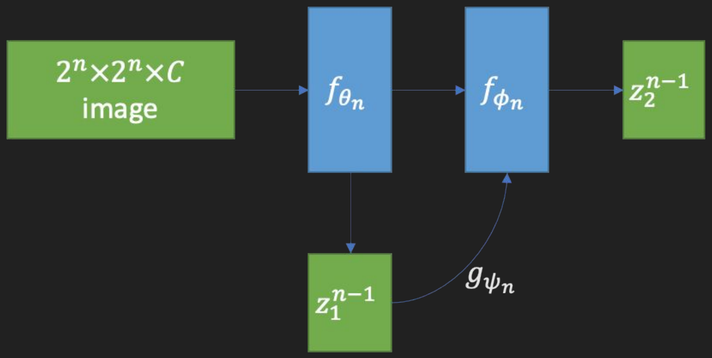
mAR-SCF Experiment

- The authors claimed that the channel-wise AR prior can help improve representational power.
- We couldn’t fully replicate the result, possible explanation: This might be caused by the built-in Conv-LSTM layer in Flax, which lacks appropriate documentation and we suspect we might have been using it the wrong way.

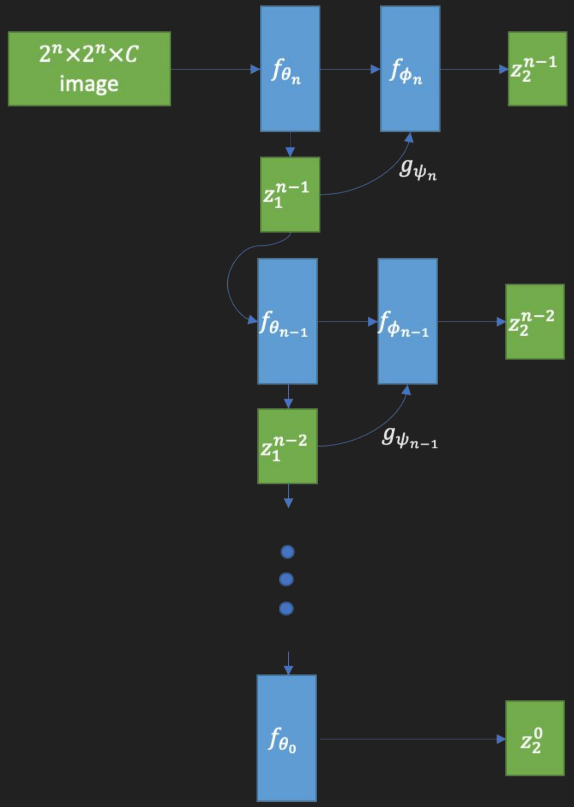
Models	From Paper	Ours
mAR-SCF (Affine)	3.33	3.39

CIFAR10 bits/dim

Stretch goal - Progressive Normalizing Flow



ProNF with an image scale of $2^n \times 2^n \times C$



Chain-up all ProNFs together

ProNF - Objective

$$\log p_{\mathbf{I}_n}(I_n|I_{n-1}) = \log p_{\mathbf{Z}_1}(f_\theta(I_n)|I_{n-1})|\det J_\theta| + \log p_{\mathbf{Z}_2}(f_\phi(I_n)|g_\psi(f_\theta(I_n)))|\det J_\phi| \tag{8}$$

ProNF - super-resolution

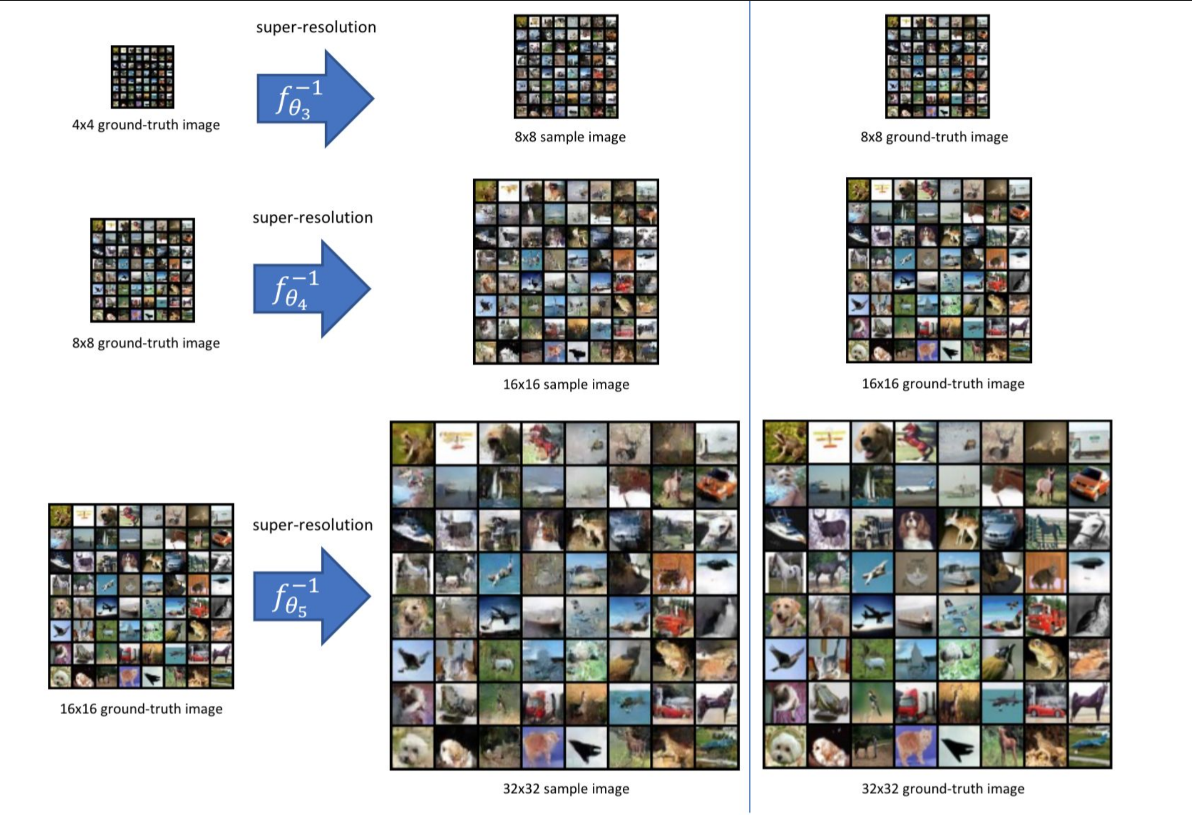


Figure 3: ProNF super-resolution samples

ProNF - sampling

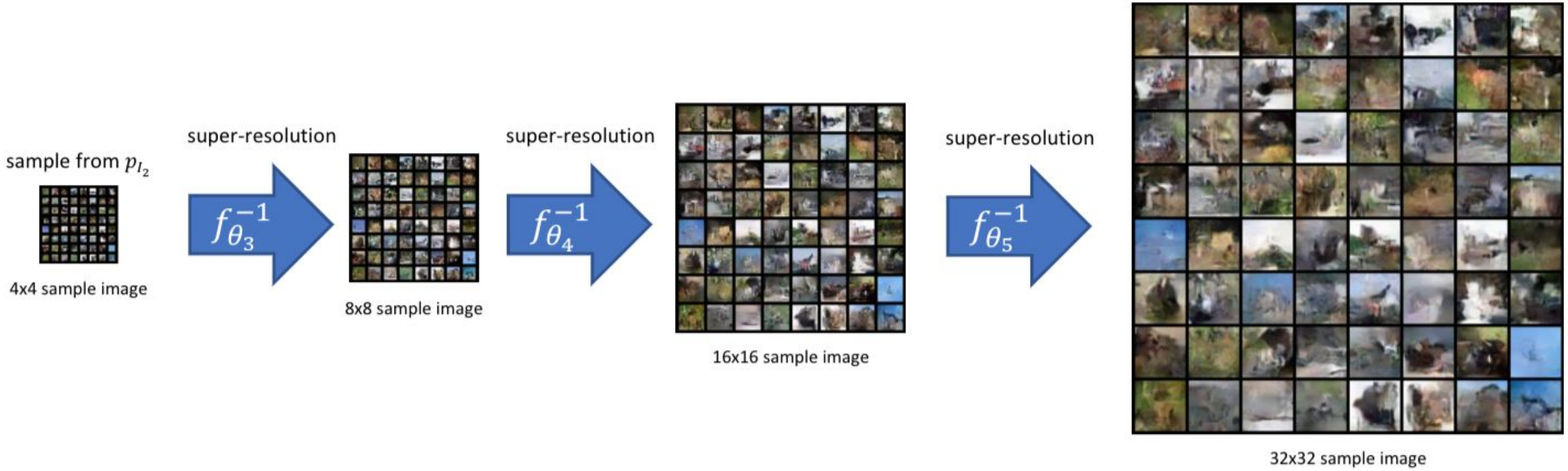


Figure 4: ProNF samples

ProNF - quantitative result

Models	Ours
ProNF (16×16 **to** 32×32)	3.29*
ProNF (8×8 **to** 16×16)	4.55*
ProNF (4×4 **to** 8×8)	5.30*
ProNF (4×4 **unconditional**)	6.57
ProNF (chain-up unconditional)	259.42

Table 4: CIFAR-10 image modeling results in bits/dim. (*) is the bpd of conditional negative log-likelihood (8).

Thank you