Quantify Uncertainty in Deep Neural Networks

Jose Jaita Aguilar
Advisor: Rensso Mora Colque

San Pablo Catholic University

28th January 2019

What is Uncertainty?

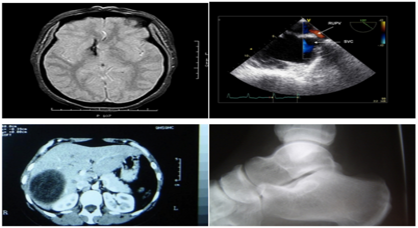
- Uncertainty is a situation which involves imperfect or unknown information
- Something that is uncertain or that causes one to feel uncertain

(a)

(b)

(c)

Motivation: Why is uncertainty important?



(a) Medical diagnostics

(b) Financial prediction

(c) Self-driving car

Problems

- Standard deep learning tools are poor at quantify uncertainty (deterministic approach)
- Bayesian probability theory offers us mathematically grounded tools to study uncertainty
- Bayesian inference comes with a prohibitive computational cost
- Stein Variational Gradient Descent (SVGD)[1] is a new variational inference algorithm
- SVGD in high dimensions, the convergence is slow

Goals

- Accelerated-SVGD algorithm to accelerate the convergence
- A method based in VAE for task out-of-distribution detection

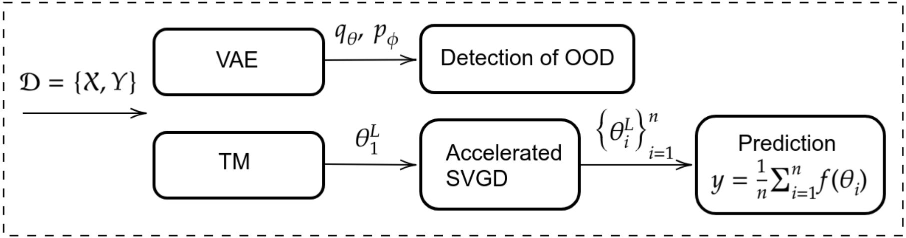


Figure 4: Pileline

Out-of-Distribution (OOD)

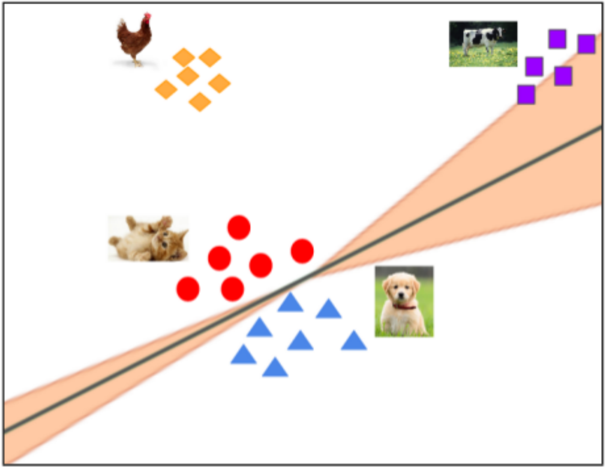


Figure 5: OOD

OOD Detection

Goal:

- Detect out-of-distribution images

Start:

- We have a trained network (DenseNet, ResNet)
- The network only view train data (cifar10 50000)

Test:

- ID (images that belong to the classes of training data)(cifar10 10000)
- OOD (images that don't belong) (LSUN 10000)

OOD Detection (at train time)

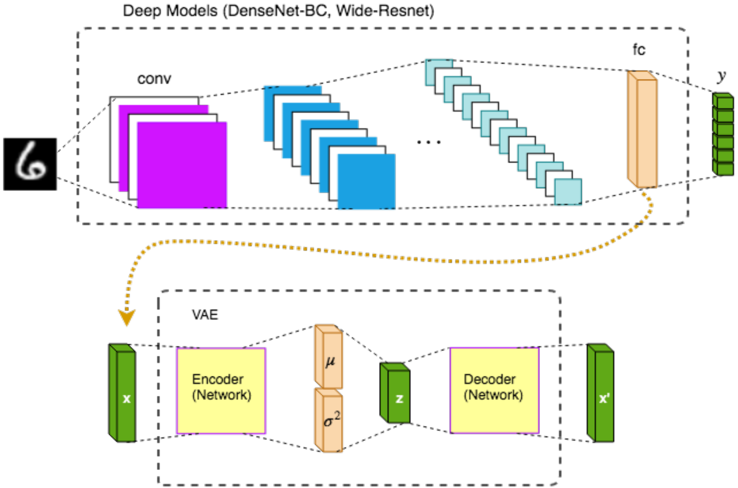
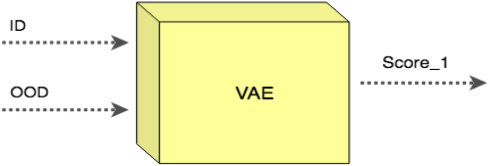
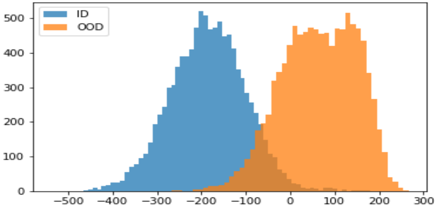


Figure 6: Feature extraction

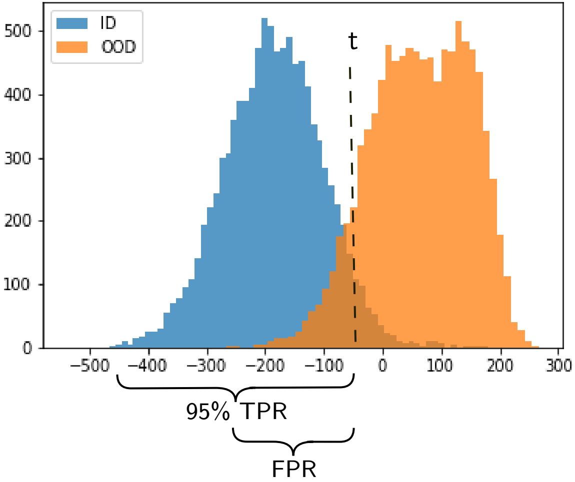
OOD Detection (at test time)



$$Score\_1 = \mathcal{L}_i(\theta, \phi) = -\mathbb{E}_{z \sim q_\theta(z|x_i)}[log p_\phi(x_i|z)] + KL(q_\theta(z|x_i)||p(z)) \quad (1)$$

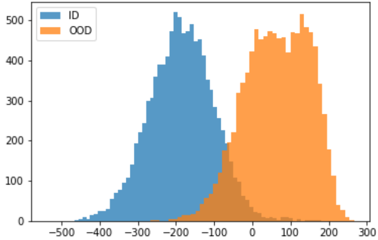


OOD Detection

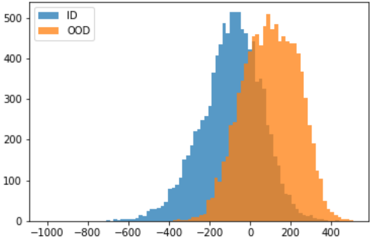


OOD Detection

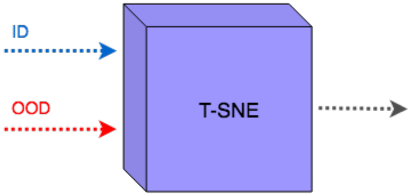
ID: CIFAR 10
OOD: TinyImagenet

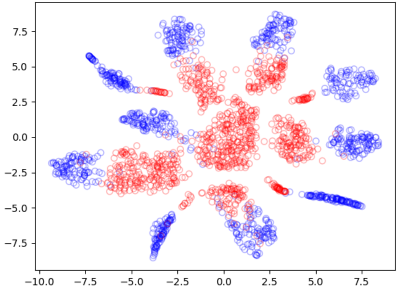


ID: CIFAR 100
OOD: TinyImagenet



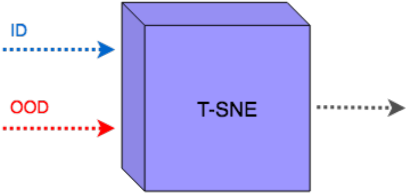
OOD Detection

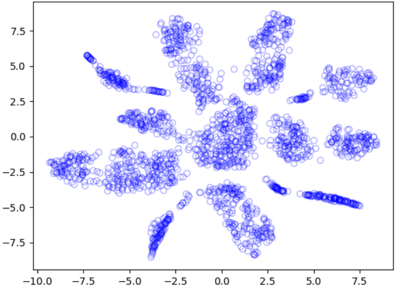




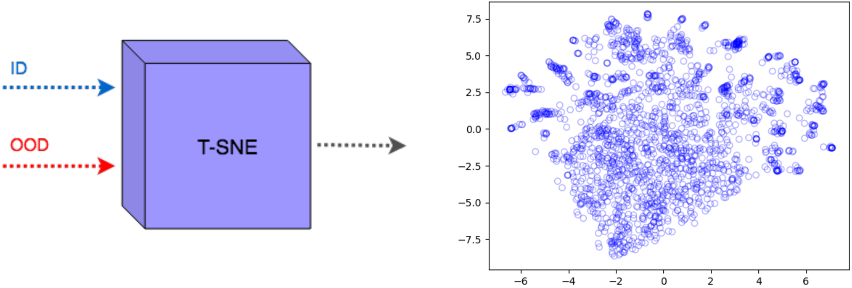
Clustering!

OOD Detection

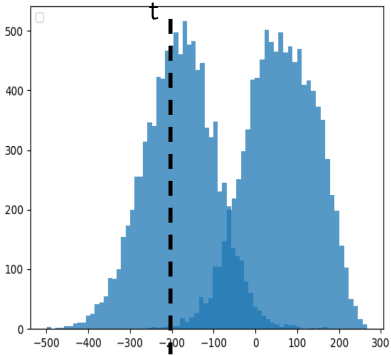




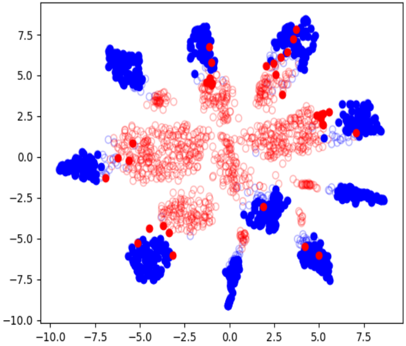
OOD Detection



OOD Detection - VAE+T-sne

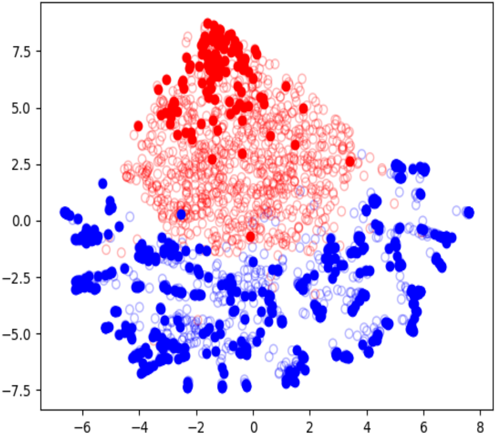


ID OOD



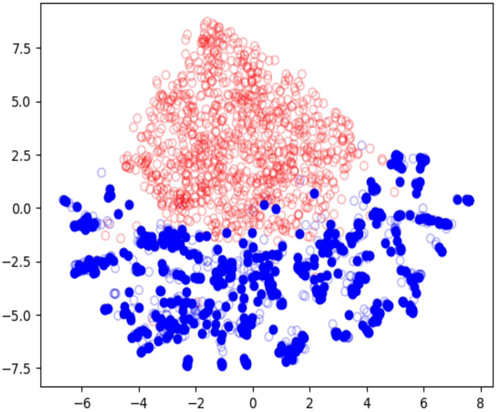
Clustering! GMM, KDE

OOD Detection - VAE+T-sne



OOD Detection - T-sne

- For train, 1000 samples from ID dataset are used



OOD Detection

	Out-of-distribution dataset	FPR (95%TPR)	Detection error	AUROC	AUPR In	AUPR Out
			ODIN[2] / Vyas[3] / VAE+T-sne			
Dense-BC CIFAR-10	TinyImagenet(c)	4.3/1.23/**0.54**	4.7/2.63/**2.05**	99.1/**99.65**/99.53	99.1/**99.68**/99.61	99.1/**99.64**/99.43
	TinyImagenet(r)	7.5/2.93/**1.19**	6.1/3.84/**3.40**	98.5/**99.34**/98.88	98.6/**99.37**/99.22	98.5/**99.32**/98.75
	LSUN(c)	8.7/3.42/**1.18**	6.0/4.12/**3.90**	98.2/**99.25**/99.00	98.5/**99.29**/99.10	97.8/**99.24**/99.00
	LSUN(r)	3.8/0.77/**0.28**	4.4/2.10/**1.96**	99.2/**99.75**/99.52	99.3/**99.77**/99.60	99.2/**99.73**/99.41
	iSUN	6.3/ - /**0.71**	5.5/ - /**2.66**	98.8/ - /**99.35**	98.9/ - /**99.36**	98.8/ - /**99.31**
Dense-BC CIFAR-100	TinyImagenet(c)	17.3/8.29 /**3.23**	11.2/6.27/**5.67**	97.1/**98.43**/98.18	97.4/**98.58**/98.10	96.8/**98.3** /98.18
	TinyImagenet(r)	44.3/20.52/**12.26**	24.6/9.98/**7.64**	90.7/96.27/**97.53**	91.4/96.66/**96.86**	90.1/95.82/**97.93**
	LSUN(c)	17.6/14.69/**11.96**	11.3/**8.46**/10.37	96.8/**97.37**/97.02	97.1/**97.62**/97.54	96.5/**97.18**/97.11
	LSUN(r)	44.0/16.23/**2.07**	24.5/8.77/**5.07**	91.5/97.03/**97.80**	92.4/97.37/**98.37**	90.6/**96.6** /96.35
	iSUN	49.5/ - /**16.1**	27.2/ - /**10.8**	90.1/ - /**96.34**	91.1/ - /**96.26**	88.9/ - /**96.76**

Table 1: Comparation between VAE+T-sne and others methods

OOD Detector

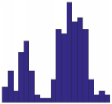
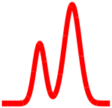
	Out-of-distribution dataset	FPR (95%TPR)	Detection error	AUROC	AUPR In	AUPR Out
				ODIN[2] / Vyas[3] / T-sne		
Dense-BC CIFAR-10	TinyImagenet(c)	4.3/1.23/**0.55**	4.7/2.63/**2.60**	99.1/**99.65**/99.44	99.1/**99.68**/99.45	99.1/**99.64**/99.37
	TinyImagenet(r)	7.5/2.93/**1.60**	6.1/3.84/**3.40**	98.5/**99.34**/98.77	98.6/**99.37**/98.75	98.5/**99.32**/98.85
	LSUN(c)	8.7/3.42/**0.78**	6.0/4.12/**3.90**	98.2/99.25/**99.45**	98.5/99.29/**99.46**	97.8/99.24/**99.42**
	LSUN(r)	3.8/0.77/**0.39**	4.4/2.10/**2.08**	99.2/**99.75**/99.53	99.3/**99.77**/99.61	99.2/**99.73**/99.43
	iSUN	6.3/ - /**0.76**	5.5/ - /**2.43**	98.8/ - /**99.36**	98.9/ - /**99.34**	98.8/ - /**99.34**
Dense-BC CIFAR-100	TinyImagenet(c)	17.3/8.29 /**5.40**	11.2/6.27/**6.90**	97.1/**98.43**/97.80	97.4/**98.58**/97.50	96.8/**98.3** /98.10
	TinyImagenet(r)	44.3/20.52/**2.19**	24.6/9.98/**4.06**	90.7/96.27/**98.96**	91.4/96.66/**98.74**	90.1/95.82/**99.11**
	LSUN(c)	17.6/**14.69**/17.69	11.3/**8.46**/11.64	96.8/**97.37**/93.64	97.1/**97.62**/92.90	96.5/97.18/**93.63**
	LSUN(r)	44.0/16.23/**0.45**	24.5/8.77/**1.25**	91.5/97.03/**99.75**	92.4/97.37/**99.66**	90.6/96.60/**99.79**
	iSUN	49.5/ - /**1.68**	27.2/ - /**2.69**	90.1/ - /**99.38**	91.1/ - /**99.33**	88.9/ - /**99.44**

Table 2: Comparation between T-sne and others methods

In	Out-of-distribution dataset Out	FPR (95%TPR)	Detection error	AUROC
			Mahalanobis[4] / VAE+T-sne / T-sne	
Dense-BC CIFAR-10	TinyImagenet(r)	5.0/1.60/**1.19**	5.0/4.17/**3.40**	98.8/98.77/**98.88**
	LSUN(r)	2.8/0.39/**0.28**	3.7/2.08/**1.96**	99.3/**99.53**/99.52
Dense-BC CIFAR-100	TinyImagenet(r)	13.4/**2.19**/3.38	7.8/**4.06**/7.64	97.4/96.27/**97.53**
	LSUN(r)	8.6/**0.45**/1.41	6.1/**1.25**/5.07	98.0/**99.75**/97.80

Table 3: Comparation between T-sne , VAE and Mahalanobis

Stein Variational Gradient Descent (SVGD)

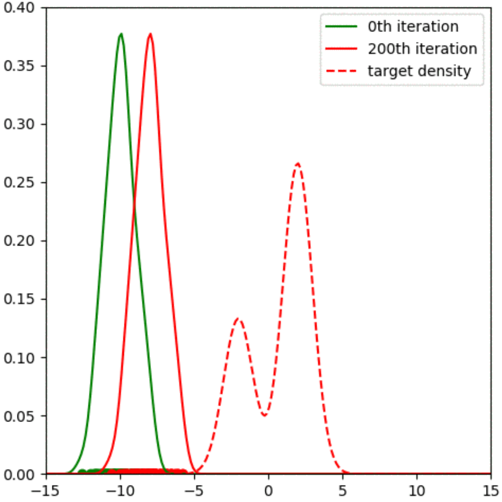
$\mathbb{D}($  ,  $)$

$\{\theta_0\}_{i=1}^{n}$ $\quad$ *goal* : $p(x)$

Inference (Sampling): Given p, find optimal $\{\theta\}_{i=1}^{n}$

$$argmin\ \mathbb{D}(\{\theta\} \parallel p)$$

SVGD


0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−15
−10
−5
0
5
10
15
0th iteration
200th iteration
target density

Accelerated-SVGD

Normal

for iters :
    for 1 :
        $\theta \leftarrow \theta + \nabla \text{loss}(\theta)$

SVGD

for iters :
    for i = 1,...,n :
        $\theta_i \leftarrow \theta_i + \phi(\theta_i)$

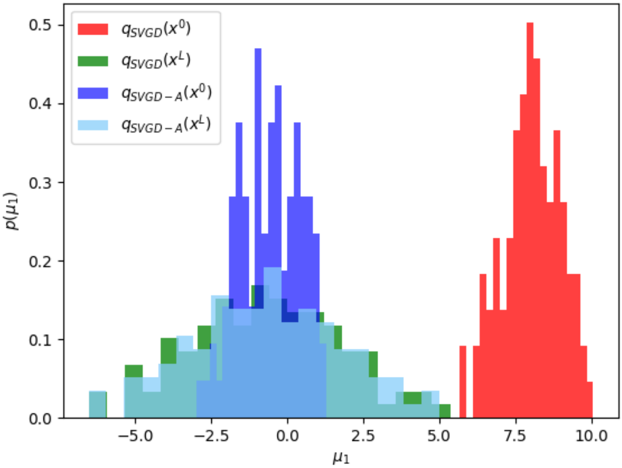
Acelerated-SVGD

for iters :
    for 1 :
        $\theta_1 \leftarrow \theta_1 + \nabla \text{loss}(\theta_1)$
for i = 2,...,n :
    $\theta_i \leftarrow \theta_1 + \eta_i$
for iters :
    for i = 1,...,n :
        $\theta_i \leftarrow \theta_i + \phi(\theta_i)$

Figure 11: Comparation

Experiments: SVGD vs Acelerated-SVGD

Problem: Multivariate Normal, $Rate = 2000ite/400ite = \mathbf{5x}$



Experiments: Accelerated-SVGD in Neural Network

Problem: Classification in toy-example

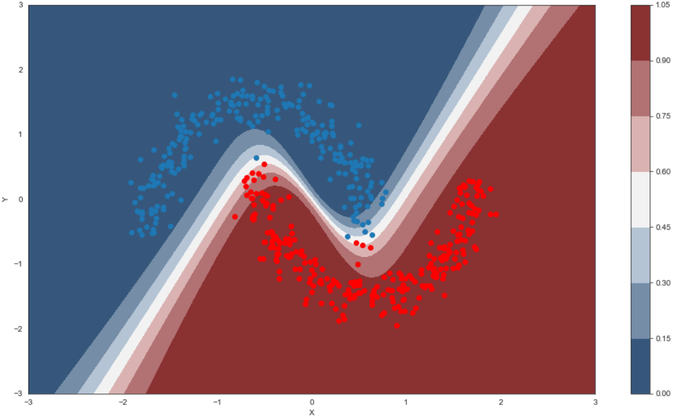


Figure 13: Bayesian Neural Network

Conclusions

- We observe significant computational gains of Accelerated-SVGD over the original SVGD algorithm
- It is difficult to compute the true distribution for deep models
- We show interesting results for ood detection
- Study of other clustering methods, we use gmm and kde
- The VAE can be replaced for GANS

References

- [1] Liu, Q. y Wang, D. (2016). Stein variational gradient descent: A general purpose bayesian inference algorithm. In *Advances In Neural Information Processing Systems*, pages 2378–2386
- [2] Liang, S., Li, Y., Srikant, R. (2017). Enhancing the reliability of out-of-distribution image detection in neural networks. *arXiv preprint arXiv:1706.02690*
- [3] Vyas, A., Jammalamadaka, N., Zhu, X., Das, D., Kaul, B., Willke, T. L. (2018). Out-of-distribution detection using an ensemble of self supervised leave-out classifiers. In *Proceedings of the European Conference on Computer Vision* (ECCV), pages 550-564
- [4] Lee, K., Lee, K., Lee, H., Shin, J. (2018). A simple unified framework for detecting out-of-distribution samples and adversarial attacks. In *Advances in Neural Information Processing Systems*, pages 7167-7177

- Thank You.