# Geometry of Disentangled Represenatations of Deep Generative Models

**Ankita Shukla[1], Shagun Uppal[1], Sarthak Bhagat[1], Saket Anand[1], Pavan Turaga[2]**

[1] IIIT- Delhi, [2]Arizona State University

*ICVGIP 2018, 18-22 December, Hyderabad, India*

## Need for geometry awareness in latent space models

➢ Latent space models learn a mapping from low dimensional latent space to high dimensional data space

➢ Generator mapping approximates the data manifold reasonably well, allowing to generate novel data samples

➢ Metric in the latent space deviates from Euclidean distance

➢ Generator mapping function $g : Z \rightarrow X$

$Z \subseteq \mathbb{R}^d$

$X \subseteq \mathbb{R}^D$

$d \ll D$

➢ *'g' is a composite function formed by*

➢ *affine maps and activations*

- ***Activation functions are smooth***
- ***Weight matrices have maximal rank***

## What we do?

- ***Study and compare the latent space geometry of deep generative models that learn disentangled representations***
- ***Establish the presence of curvature in latent space using several metrics***
- ***Establish the relevance of geometry aware metric over Euclidean metric***

## Metrics

➢ **Residual**

➢ **Cross Correlation**

$$c_k = 1 - \frac{(r_M(k) - \mu_{r_M})(r_E(k) - \mu_{r_E})}{\sigma_{r_M}\sigma_{e_E}}$$

$$\hat{c} = \frac{2}{N(N-1)}\Sigma_k c_k$$

Highlights the difference between Euclidean and Riemannian distances

➢ **Normalized Margin**

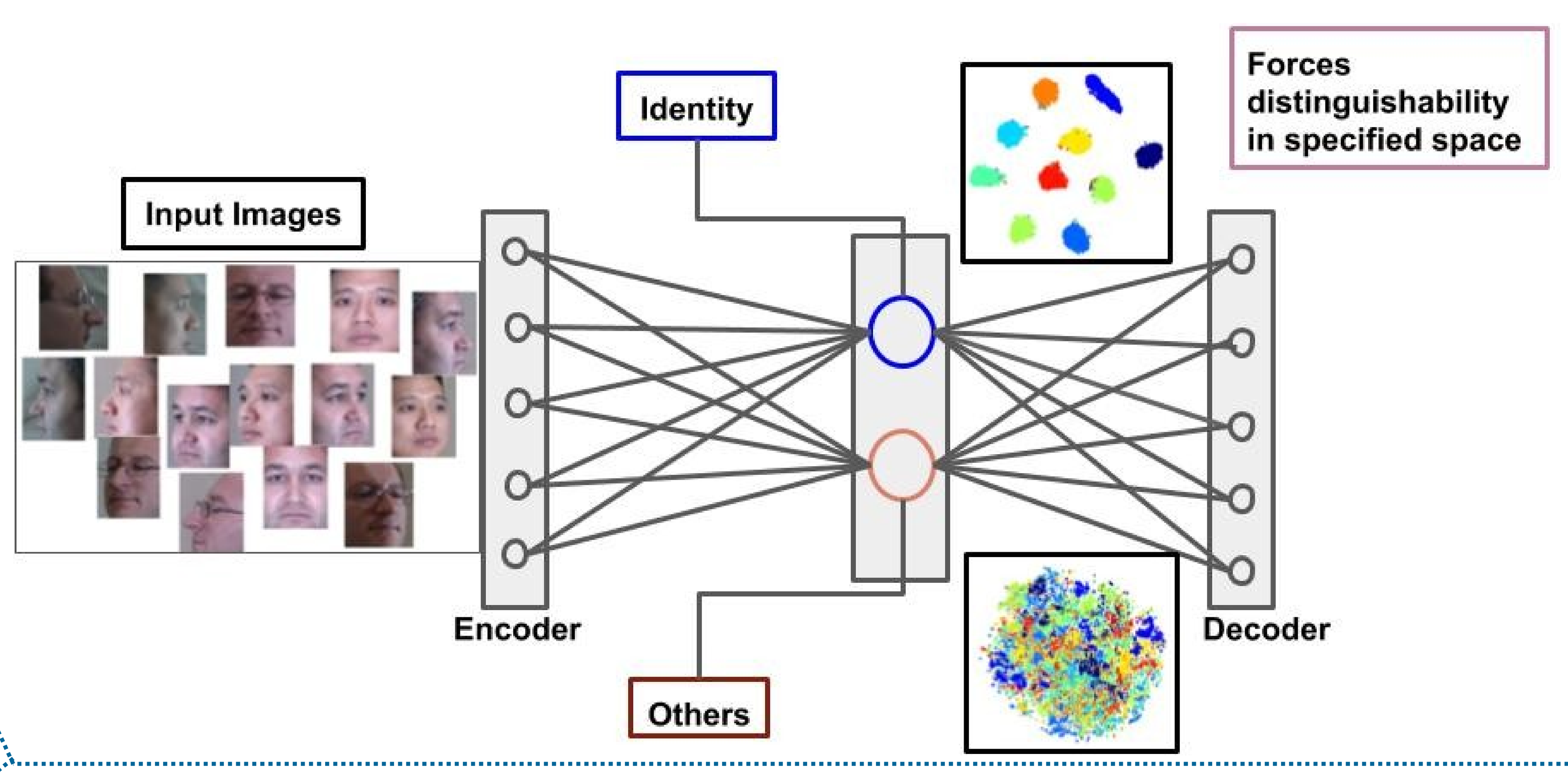
$$m_n = \frac{||x_n - \mathcal{M}(x_n)|| - ||x_n - \mathcal{H}(x_n)||}{||x_n - \mathcal{M}(x_n)||}$$

Measures class seperability

➢ **Tangent Space Alignment**

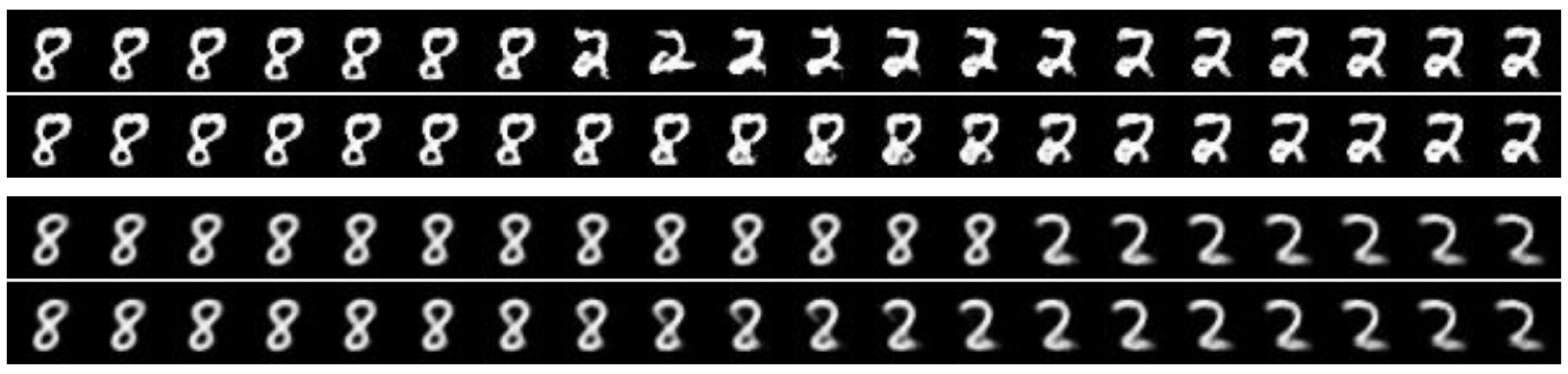
- Principle angle between subspaces defined by the tangent spaces
- Large angle denotes higher curvature

## Disentangled Representation Learning




VAE

Dis. Model



## Experimental Results

**Residual Cross Correlation**

| Dataset | VAE | Szabo *et al.* | Mathieu *et al.* | Jha *et al.* |
|---|---|---|---|---|
| MNIST | 0.071 | 0.142 | 0.178 | 0.167 |
| MultiPie | 0.65 | - | 0.72 | 0.71 |
| 3D chairs | 0.162 | 0.262 | 0.311 | 0.315 |

**Comparison of $\hat{c}$ values for different disentangling models with VAE for MNIST digits, MultiPIE and 3D chairs dataset.**

**Tangent Space Alignment**

| Datasets | VAE | Szabo *et al* | Mathieu *et al.* | Jha *et al.* |
|---|---|---|---|---|
| MNIST | 21.45 | 34.12 | 32.25 | 32.76 |
| MultiPIE | 27.95 | 37.42 | 36.88 | 36.96 |
| 3D chairs | 23.45 | 36.77 | 35.86 | 36.50 |

**Approximate Curvature estimated with principal angles between Tangent Spaces.**

| | Models | VAE | Mathieu *et al.* | Jha *et al.* | Szabo *et al* | Mathieu *et al.* | Jha *et al.* | Szabo *et al* | Mathieu *et al.* | Jha *et al.* |
|---|---|---|---|---|---|---|---|---|---|---|
| Distances | Euclidean | 0.312 | 0.346 | 0.332 | 0.158 | 0.160 | 0.156 | 0.114 | 0.112 | 0.116 |
| | Riemannian | 1.142 | 1.784 | 1.602 | 0.344 | 0.376 | 0.365 | 0.297 | 0.355 | 0.336 |
| Clustering | Euclidean | 82.98 | 89.37 | 90.06 | 91.16 | 94.33 | 94.24 | 91.12 | 94.32 | 92.22 |
| F score | Riemannian | 89.04 | 94.45 | 95.60 | 95.22 | 96.34 | 96.44 | 94.56 | 98.00 | 96.60 |

Effect of curvature on distances and clustering performance

MultiPIE

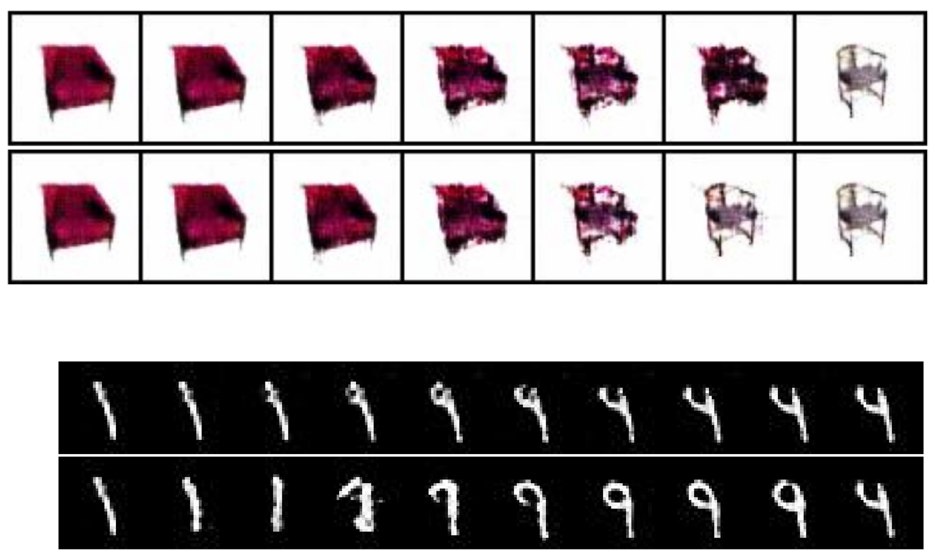
3d Chairs

MNIST

Interpolations

**Interpolation between two samples from different classes in the latent spaces of VAE and specified space of Mathieu et al.[15] with fixed unspecified using Euclidean (Left) and Riemannian Metric (Right).**

Randomly sampled unspecified component

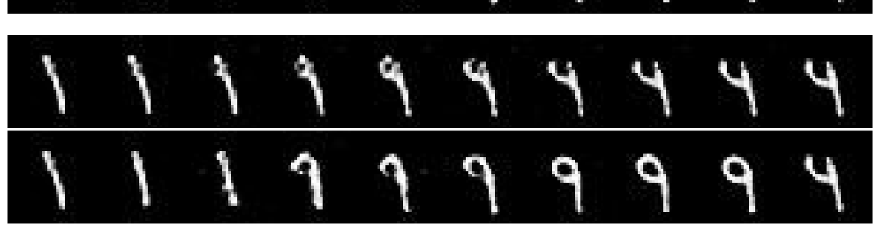


| Dimension | 16 | 64 | 80 |
|---|---|---|---|
| $\hat{c}$ | 0.065 | 0.069 | 0.071 |
| F score (Euclidean) | 83.32 | 85.22 | 87.38 |
| F score (Riemannian) | 91.74 | 92.33 | 96.23 |

Dimensionality vs Non-Linearity

| Activation function | VAE | Mathieu *et al.* |
|---|---|---|
| ReLU | 8 | 15 |
| ELU | 16 | 16 |

Rank of Jacobian



ReLU (top )vs ELU (bottom)

## Conclusion

➢ Riemannian metric should be used as opposed to Euclidean distance

➢ Better image synthesis, interpolations and clustering is achieved with Riemannian metric

➢ Disentangling models impose higher curvature as opposed to VAEs

## References

[1] Attila Szabó, Qiyang Hu, Tiziano Portenier, Matthias Zwicker, and Paolo Favaro. Challenges in disentangling independent factors of variation. CoRR, abs/1711.02245, 2017.

[2] Michaël Mathieu, Junbo Jake Zhao, Pablo Sprechmann, Aditya Ramesh, and Yann LeCun. Disentangling factors of variation in deep representations using adversarial training. In NIPS, 2016.

[3] Ananya Harsh Jha, Saket Anand, Maneesh Kumar Singh, and V. S. R. Veeravasarapu. Disentangling factors of variation with cycle-consistent variational auto-encoders. In ECCV 2018

[4] Brahma, P.P., Wu, D. and She, Y., 2016. Why Deep Learning Works: A Manifold Disentanglement Perspective. IEEE Trans. Neural Netw. Learning Syst., 27(10), pp.1997-2008.

## Acknowledgements

This work is supported in part by Infosys Center for Aritifical Intellegence at IIIT Delhi, India

## Contacts

Ankita Shukla (ankitas@iiitd.ac.in), Shagun Uppal ( shagun16088@iiitd.ac.in), Sarthak Bhagat ( sarthak16189@iiitd.ac.in), Saket Anand (anands@iiitd.ac.in), Pavan Turaga (pturaga@asu.edu)