Self-supervised Visualisation of Microscopy Datasets

Ifeoma Veronica Nwabufo, Jan Niklas Böhm, Philipp Berens, Dmitry Kobak

Hertie Institute for AI in Brain Health, Tübingen AI Center

ifeoma-veronica.nwabufo@uni-tuebingen.de

Introduction

Self-supervised learning methods such as SimCLR and BYOL are used to obtain meaningful representations of image datasets and are widely used for pre-training.

A recent method, t-SimCNE[1], uses contrastive learning to train 2D representations for visualisation. When applied to medical image datasets, t-SimCNE produced 2D visualisations that revealed semantically meaningful clusters.

We investigate the effect of augmentations and show that adding arbitrary rotations to data augmentations improved class separability in medical images, highlighting medically relevant structures that aid in data exploration and annotations.

Methods

Architecture

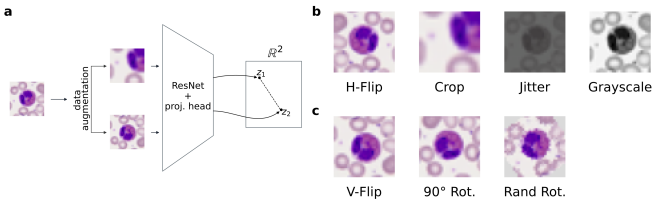

Figure 1: Study Overview

An image is augmented twice and passed through a ResNet backbone and an MLP projector to obtain representations z_1 and z_2. The network is trained the minimize the InfoNCE los via a Euclidean distance and a Cauchy kernel. (b) Augmentations for natural images. (c) Additional augmentations for medical images.

Loss Functions

$$\ell_{\mathbf{SimCLR}} = -\log \frac{\exp\left(\mathrm{sim}(\mathbf{z_i},\mathbf{z_j})/\tau\right)}{\sum_{\mathbf{k}\neq\mathbf{i}}^{\mathbf{2b}} \exp\left(\mathrm{sim}(\mathbf{z_i},\mathbf{z_k})/\tau\right)} \quad (1)$$

$$\ell_{\mathbf{t-SimCNE}} = -\log\frac{\mathbf{1}}{\mathbf{1}+\|\mathbf{z_i}-\mathbf{z_j}\|^2} + \log\sum_{\mathbf{k}\neq\mathbf{i}}^{\mathbf{2b}} \frac{\mathbf{1}}{\mathbf{1}+\|\mathbf{z_i}-\mathbf{z_k}\|^2} \quad (2)$$

The SimCLR[2] loss function (1) uses the cosine similarity to calculate how points are similar. The t-SimCNE loss function (2) uses the Euclidean distance and the Cauchy kernel to measure similarity in 2D.

Training

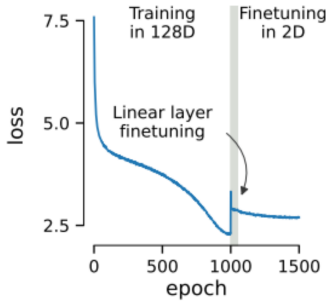

Figure 2: 3 step training process

Pre-training in high dimension to get a 128D output, Randomly initializing a 2D output layer, freezing the rest of the network and finetuning for 50 epochs. Unfreezing the entire network and finetuning for 450 epochs.

Results

Table 1: Silhouette scores of 2D embeddings

Method		Dataset				
		Leukemia	BloodMNIST	DermaMNIST	PathMNIST	PCam16
t-SimCNE	def. augm.	0.13 ± 0.00	0.40 ± 0.00	0.13 ± 0.01	0.45 ± 0.02	0.04
	+ 90° rot.	0.33 ± 0.01	0.44 ± 0.03	0.11 ± 0.00	0.48 ± 0.06	0.05
	+ rand. rot.	0.52 ± 0.02	0.50 ± 0.01	0.13 ± 0.06	0.41 ± 0.03	0.05
t-SNE of SimCLR	def. augm.	0.21 ± 0.01	0.37 ± 0.00	0.14 ± 0.00	0.23 ± 0.01	0.16
	+ 90° rot.	0.23 ± 0.01	0.35 ± 0.02	0.14 ± 0.01	0.25 ± 0.01	0.13
	+ rand. rot.	0.21 ± 0.00	0.37 ± 0.02	0.16 ± 0.00	0.26 ± 0.00	0.06
t-SNE	pixel space	-0.09	0.07	0.08	-0.05	0.02
	ResNet18	-0.11	0.13	0.14	0.17	0.04
	ResNet152	-0.15	0.03	0.14	0.19	0.05

Table 2: Effects of augmentations for t-SimCNE

Augmentations	Leukemia		BloodMNIST		PathMNIST	
	*k*NN acc.	Silhouette	*k*NN acc.	Silhouette	*k*NN acc.	Silhouette
All	95.1 ± 0.2	0.52 ± 0.02	92.9 ± 0.1	0.50 ± 0.01	97.3 ± 0.0	0.41 ± 0.03
No crops	79.7 ± 0.6	0.14 ± 0.00	76.0 ± 1.1	0.20 ± 0.01	59.8 ± 1.1	-0.02 ± 0.03
No color jitter	82.0 ± 0.2	-0.01 ± 0.01	90.0 ± 0.1	0.45 ± 0.02	94.3 ± 0.3	0.24 ± 0.02
No grayscaling	95.6 ± 0.4	0.52 ± 0.02	92.1 ± 0.3	0.44 ± 0.01	98.5 ± 0.0	0.39 ± 0.05

Table 3: Linear evaluation on SimCLR.

Augmentations	Dataset		
	BloodMNIST	DermaMNIST	PathMNIST
def. augm.	$94.5 \pm 0.0\%$	$82.3 \pm 0.3\%$	$91.3 \pm 0.3\%$
+ 90° rot.	$\mathbf{96.5 \pm 0.2\%}$	$83.6 \pm 0.5\%$	$91.9 \pm 0.0\%$
+ rand. rot.	$\mathbf{96.5 \pm 0.1\%}$	$\mathbf{84.5 \pm 0.4\%}$	$\mathbf{92.7 \pm 0.2\%}$

Datasets and Visualisations

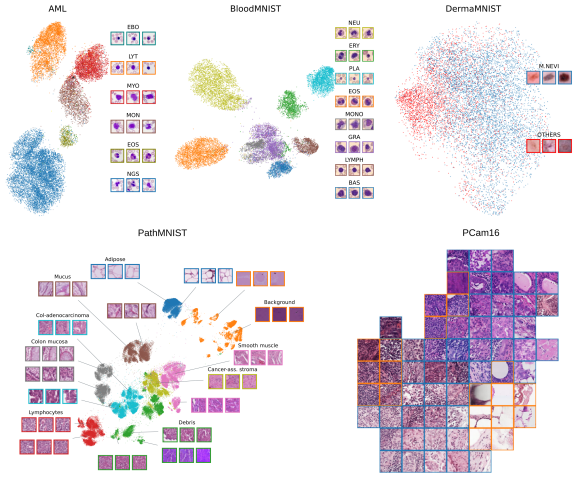

Figure 3: t-SimCNE visualisation of five publicly available medical image datasets with sample sizes ranging from 10k to over 300k.

Additional Experiments

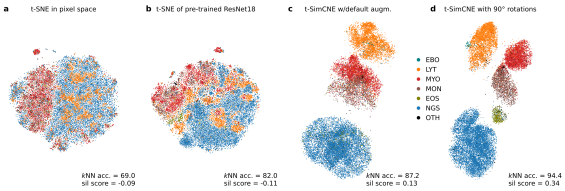

Figure 4: Visualisations of the Leukemia dataset.

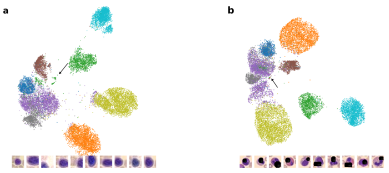

Figure 5: t-SimCNE visualisations of the modified BloodMNIST dataset. (a) With added 100 duplicates of a single image, with random perturbations. (b) With artefacts added to 50 images.

Conclusion

1. t-SimCNE is useful for visualisation and can be used for clustering, highlighting artefacts in data and summarizing dataset.
2. Parametric nature of t-SimCNE allows to embed new (out-of-sample) images into an existing embedding.
3. Both SimCLR and t-SimCNE benefit from rotational data augmentations, leveraging rotational invariance of microscopy images.

References

[1] Böhm et al. Unsupervised visualizations of image datasets using contrastive learning. In ICLR, 2023.

[2] Chen et al. A simple framework for contrastive learning of visual representations. In ICLR, 2020.

HERTIE INSTITUTE FOR AI IN BRAIN HEALTH

Funding

This work has been funded by the DFG.

Contact

philipp.berens@uni-tuebingen.de

www.hertie.ai/data-science

GEFÖRDERT DURCH DIE GEMEINNÜTZIGE Hertie Stiftung