universität freiburg

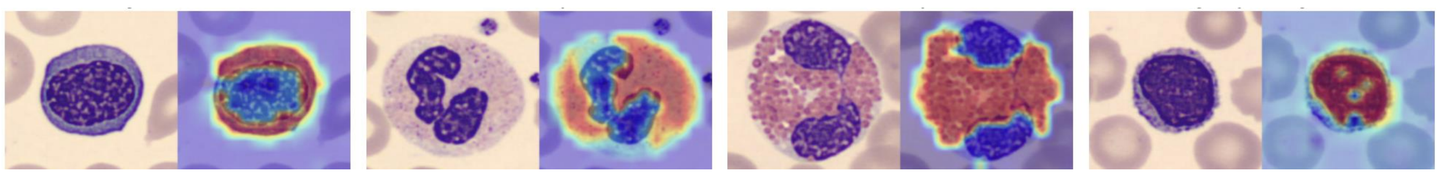
NR#4: Explaining blood cells attributes in style

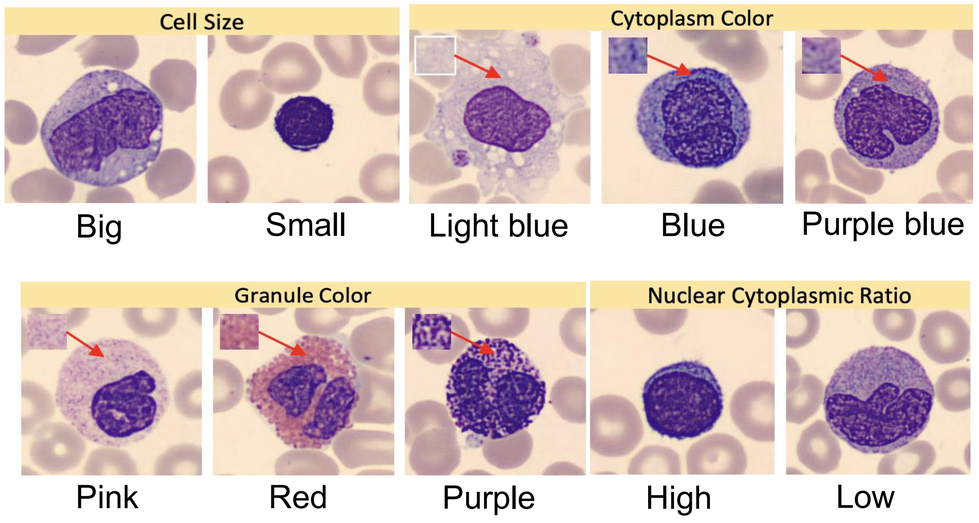
Johannes Effenberger, Huy Bao Tran, Maren Wolffram

Mehdi Naouar

Motivation

- SOTA CNNs classifiers and previous explanation approaches (e.g. Grad-CAM) only paint where the network looks, not what attribute (color, shape, etc) drives the decision. **The causal visual attributes are hidden**
- We pair StyleGAN models that **generate disentangled representations** of images with a Classifier-Encoder framework **StylEx** to explain **one attribute at a time** to enable richer **XAI** use cases in clinical setting
- Most of StyleGAN usecases have been to explain human faces, we aim to generate and test its capability **on complex medical images**.





Satoshi Tsutsui and Winnie Pang and Bihan Wen, WBCAtt: A White Blood Cell Dataset Annotated with Detailed Morphological Attributes

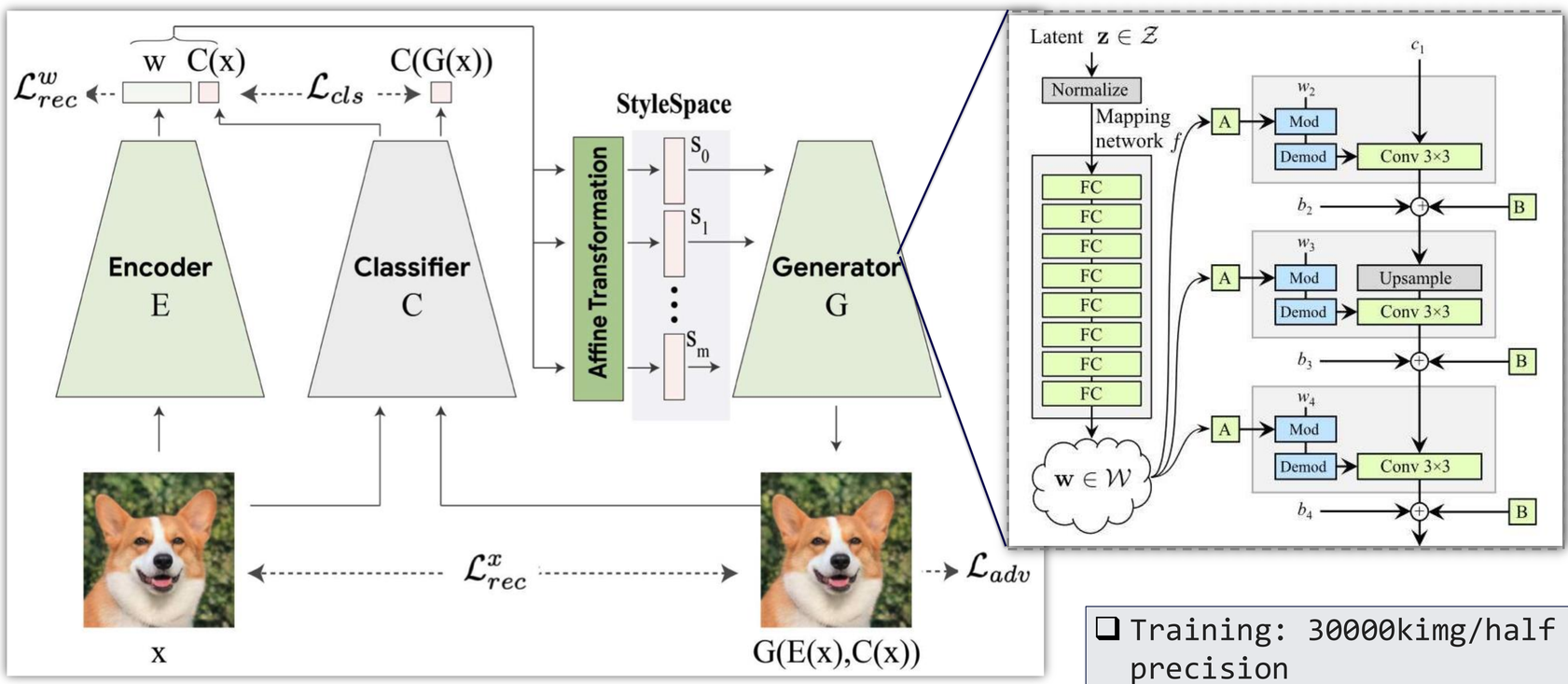
Method

StyleGANs on WBCAtt images → StyleGAN-ADA

Classifier to recognize 5 cell types → ResNet 18

Reimplement and trained with StylEx → FC Map. Net

Extract and visualize disentangled coordinates



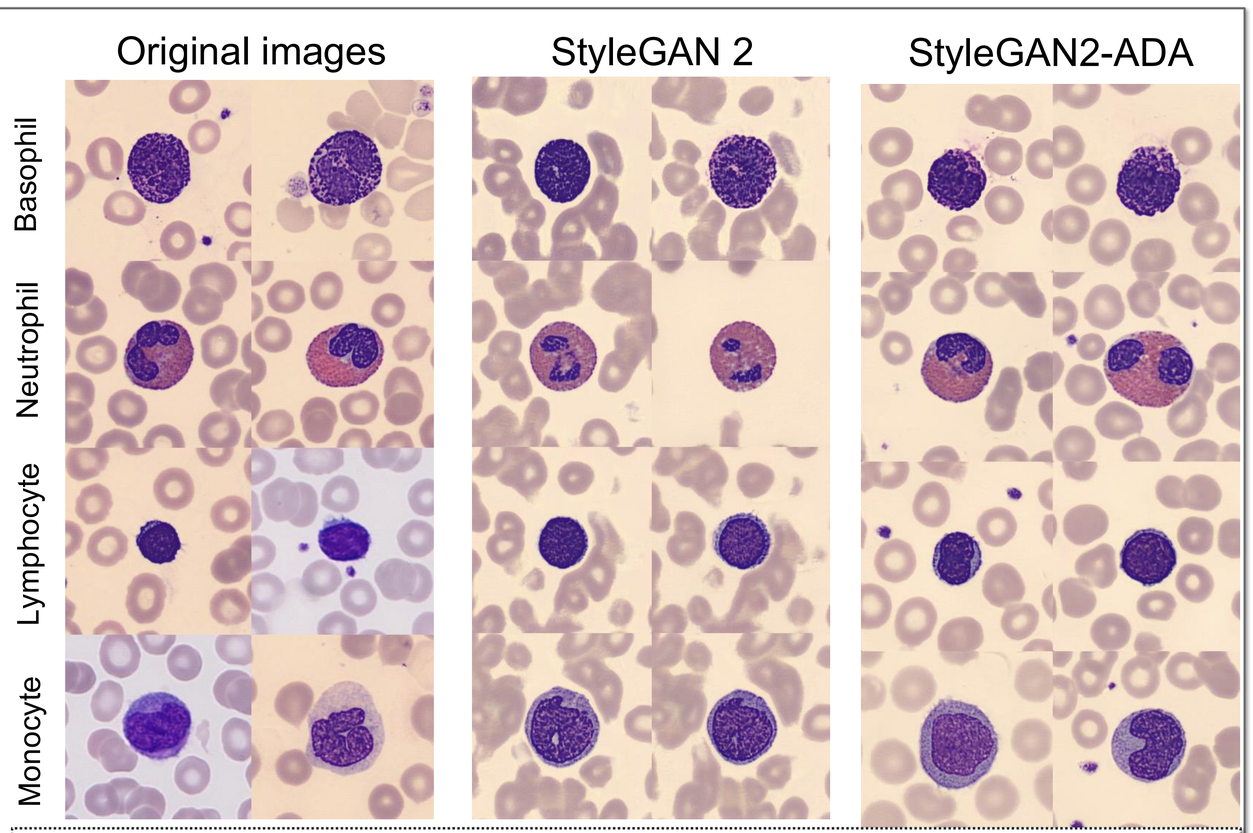
Oran Lang et. al, Explaining in Style: Training a GAN to explain a classifier in StyleSpace,

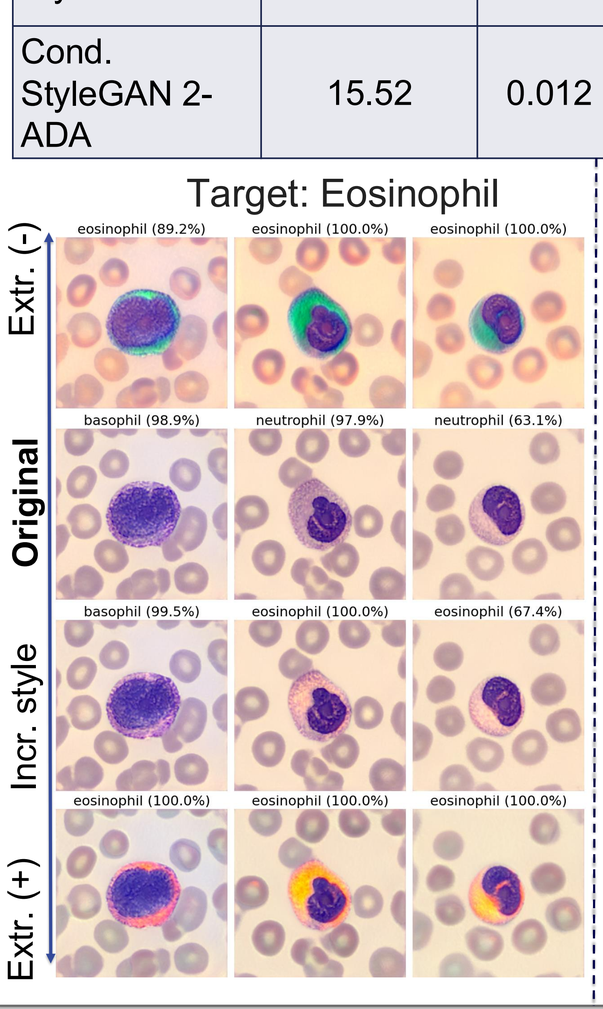
- Training: 30000kimg/half precision
- 55 hours on 4xA100 GPU

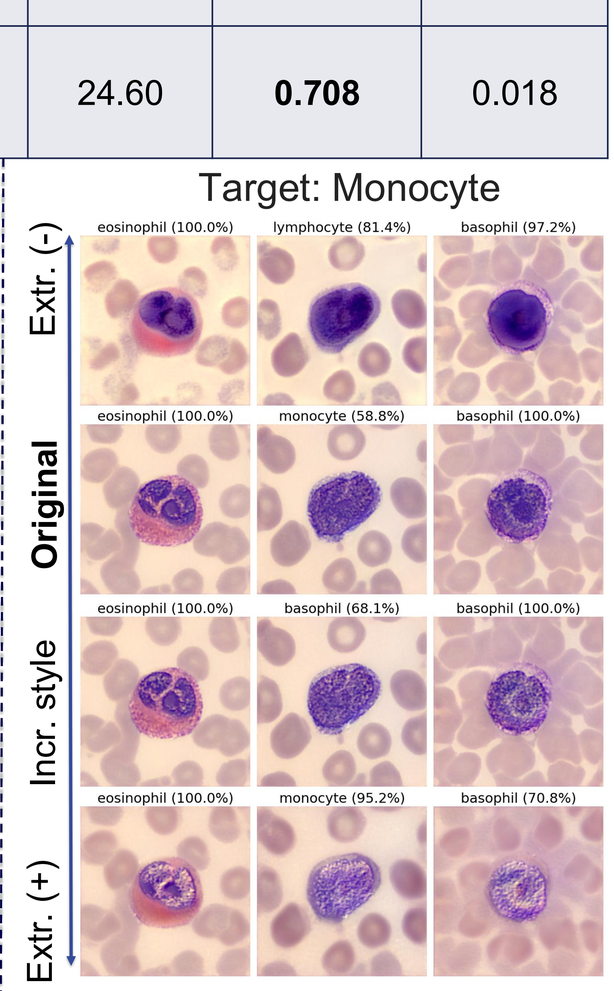
Empirical results

Table 1. Main results on 10,298 white blood cells image - **256x256**

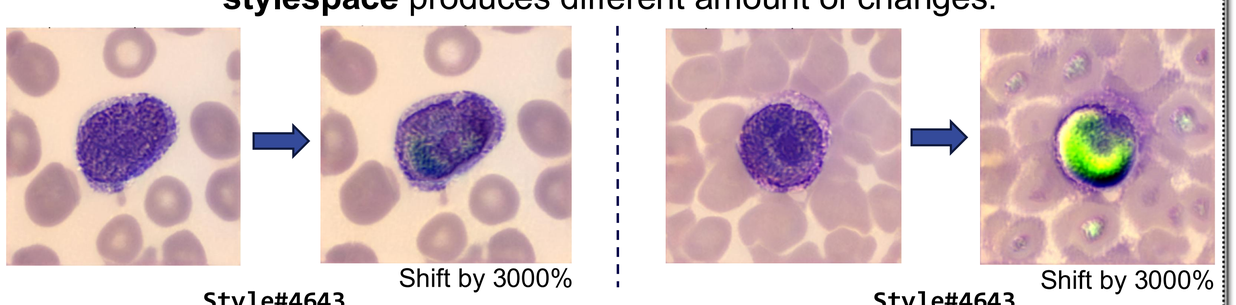
Configuration	FID ↓	KID ↓	PPL ↓	Precision ↑	Recall ↑
StyleGAN 2	11.81	0.012	146.47	0.646	0.491
StyleGAN 2-ADA	**7.86**	**0.008**	30.75	0.689	**0.573**
StyleGAN 3	29.8	0.036	139.30	0.371	0.024
Cond. StyleGAN 2	77.19	0.082	**14.79**	0.312	0.0
Cond. StyleGAN 2-ADA	15.52	0.012	24.60	**0.708**	0.018







Style visual effects depend on the cell's original state. A fixed shift in **stylespace** produces different amount of changes:



- ***StyleGAN2-ADA*** *works best because of* ***limited image samples***
- ***Caveat***: top-ranked attributes aren't always the most human-perceptible. Some style coordinates in multi-class setting blend along a continuum instead of aligning at the two ends of the spectrum.