HoechstGAN

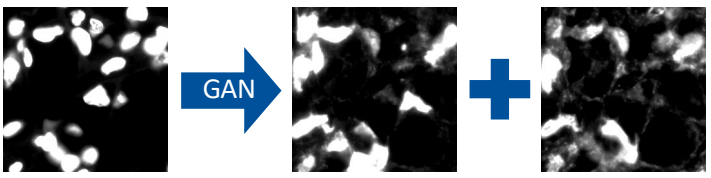
Virtual Lymphocyte Staining Using Generative Adversarial Networks

Georg Wölflein[1] In Hwa Um[1] David J. Harrison[1,2] Ognjen Arandjelović[1]

[1]University of St Andrews [2]Division of Laboratory Medicine, Lothian NHS University Hospitals, Edinburgh

Task

Virtual staining aims to transform a whole slide image from one stain to another.



Hoechst *CD3* *CD8*

Given a kidney cancer tissue slide stained with *Hoechst*, we **synthesise** *CD3* and *CD8* **stains**.

Contributions

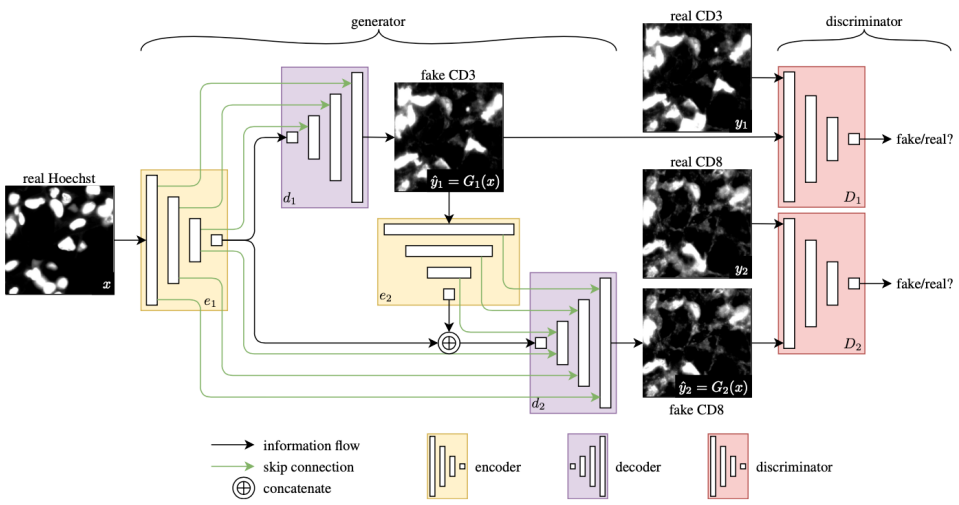
GAN architecture | evaluation metric | dataset

Method

We propose a novel GAN architecture based on **pix2pix** [Isola *et al.*, 2017] that has **two separate output** branches.

The generator's encoder learns **mutually beneficial representations** because it is **shared** across both branches.

Information from the CD3 branch flows back to **assist** CD8 generation as well.



GAN loss: $\mathcal{L}_{\text{cGAN}}(G_1, G_2, D) = \mathbb{E}_x \left[\log\left(1 - D(x, G_1(x), G_2(x))\right)\right] + \mathbb{E}_{x,y_1,y_2}\left[\log D(x, y_1, y_2)\right]$

objective: $G^* = \arg\min_{G_1,G_2} \max_D \mathcal{L}_{\text{cGAN}}(G_1, G_2, D) + \lambda\mathcal{L}_{\text{L1}}(G_1) + \lambda\mathcal{L}_{\text{L1}}(G_2)$

Results

We introduce a novel evaluation metric, the relative **masked intensity ratio** (MIR_{rel}), for assessing virtual staining **quality**.

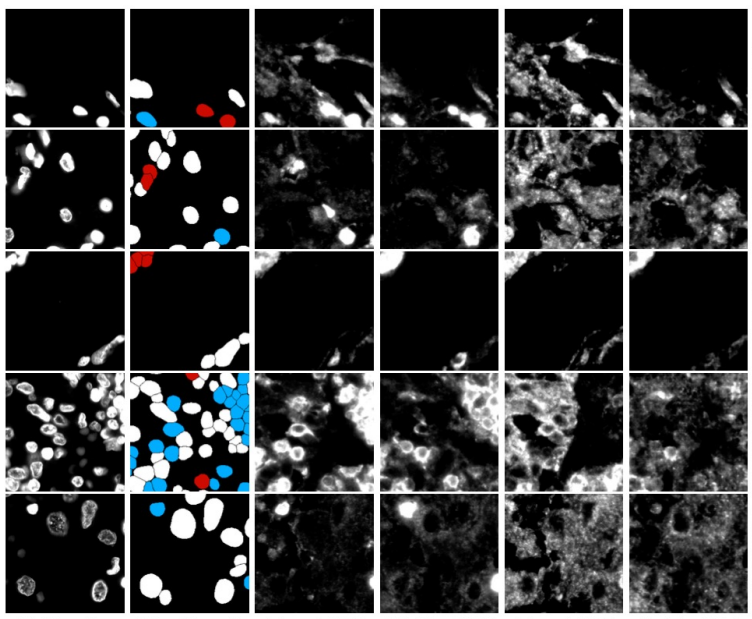
$$MIR = \frac{\textit{mean pixel intensity within mask}}{\textit{mean pixel intensity outwith mask}}$$

$$MIR_{rel} = \frac{MIR_{fake}}{MIR_{real}}$$

Our model **outperforms** U-Net regression and pix2pix baselines.

In an **ablation** study, we find that CD8 stain generation is able to learn from CD3.

We achieve best results with a **joint discriminator**.



(a) Hoechst (b) cell mask (c) real CD3 (d) fake CD3 (e) real CD8 (f) fake CD8

Method	Parameters ↓	Train CD3 MIR_{rel} ↑	Test CD3 MIR_{rel} ↑	Train CD8 MIR_{rel} ↑	Test CD8 MIR_{rel} ↑
HoechstGAN-MCD	2.16×10^8	**1.23±1.21**	**1.48±1.27**	1.07±0.97	1.43±1.04
HoechstGAN-MC	2.19×10^8	0.87±1.07	1.12±1.23	**1.22±0.88**	1.39±1.02
HoechstGAN-MD	2.16×10^8	0.88±1.07	1.15±1.21	1.01±0.93	**1.45±1.03**
HoechstGAN-M	2.19×10^8	0.89±1.26	1.24±1.30	0.84±0.97	1.26±1.06
HoechstGAN-D	$\mathbf{9.20\times10^7}$	0.90±1.01	1.14±1.17	0.87±0.94	1.33±1.05
pix2pix	1.14×10^8	0.89±1.04	1.15±1.23	0.83±0.87	1.16±1.00
Regression-MC	2.13×10^8	0.62±0.74	0.92±0.75	0.57±0.58	1.09±0.73

Discussion

- This paper introduces a **virtual staining model** for ccRCC.
- HoechstGAN architecture is **applicable** to any image-to-images problem.
- We suggest our **evaluation metric** MIR_{rel} as a criterion to benchmark future virtual staining models.
- Future research: investigate **symmetric architectures** where CD3/CD8 information flows both ways.

More information

project page with data, code and more

georg.woelflein.eu/hoechstgan