

Week 1

- General topic field: GANs
- Maybe: generate image by text queries.
- Be familiar with the progress of GANs in these years. Read literature review papers about GANs.
- Install LaTeX and texmaker, play with it.
- Look at the conference, CVPR (annual), ICCV(odd year), ECCV(even year)

Week 2

Proposal

- Problem statement

Medical image datasets are always highly imbalanced due to the rare pathologic cases. We need to generate high quality images of the minority. In this project, we will apply GANs to synthesis images for data augmentation. With the same baseline model, we will compare the improvement of the augmented dataset.

- Dataset

At the beginning, dataset 1 is used to build our model. It is easy to train and tune. Then, if we have time, dataset 2 (over 100GB) would be used. With the leaderboard, it is better to examine the performance of our model.

- 1) Small dataset: Red blood cells (from ML II exams).
- 2) Large dataset: SIIM-ISIC Melanoma Classification. Identifying melanoma in lesion images (from Kaggle ongoing competition).

- Technique and framework

Generative Adversarial Networks (GANs).

- Evaluation and metrics

Evaluation of GAN model:

Inception Score: measure the image quality.

Multi-scale Structural Similarity (MS-SSIM): measure the diversity and avoid mode collapse.

Evaluation of classifiers:

Area under the ROC curve (AUC). In medical diagnosis, we care more about finding out all the positive cases even if some images are misclassified into positive labels. (i.e. sensitivity or recall rate)

- Reference materials

GAN-based Augmentation

Huang, Sheng-Wei, Che-Tsung Lin, Shu-Ping Chen, Yen-Yi Wu, Po-Hao Hsu, and Shang-Hong Lai.

"Auggan: Cross domain adaptation with gan-based data augmentation." In *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 718-731. 2018.

Shin, Hoo-Chang, Neil A. Tenenholtz, Jameson K. Rogers, Christopher G. Schwarz, Matthew L. Senjem, Jeffrey L. Gunter, Katherine P. Andriole, and Mark Michalski. "Medical image synthesis for data augmentation and anonymization using generative adversarial networks." In *International workshop on simulation and synthesis in medical imaging*, pp. 1-11. Springer, Cham, 2018.

Kazeminia, Salome, Christoph Baur, Arjan Kuijper, Bram van Ginneken, Nassir Navab, Shadi Albarqouni, and Anirban Mukhopadhyay. "GANs for medical image analysis." *arXiv preprint arXiv:1809.06222* (2018).

Han, Changhee, Kohei Murao, Tomoyuki Noguchi, Yusuke Kawata, Fumiya Uchiyama, Leonardo Rundo, Hideki Nakayama, and Shin'ichi Satoh. "Learning more with less: Conditional PGGAN-based data augmentation for brain metastases detection using highly-rough annotation on MR images." In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pp. 119-127. 2019.

Other Augmentation (Semi-Supervised Learning)

He, Tong, Zhi Zhang, Hang Zhang, Zhongyue Zhang, Junyuan Xie, and Mu Li. "Bag of tricks for image classification with convolutional neural networks." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 558-567. 2019.

Berthelot, David, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, and Colin A. Raffel. "Mixmatch: A holistic approach to semi-supervised learning." In *Advances in Neural Information Processing Systems*, pp. 5050-5060. 2019.

GANs and Literature Review

Gui, Jie, Zhenan Sun, Yonggang Wen, Dacheng Tao, and Jieping Ye. "A review on generative adversarial networks: Algorithms, theory, and applications." *arXiv preprint arXiv:2001.06937* (2020).

Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." In *Advances in neural information processing systems*, pp. 2672-2680. 2014.

Zhang, Han, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. "Self-attention generative adversarial networks." *arXiv preprint arXiv:1805.08318* (2018).

Gulrajani, Ishaan, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C. Courville. "Improved training of wasserstein gans." In *Advances in neural information processing systems*, pp. 5767-5777. 2017.

Karras, Tero, Timo Aila, Samuli Laine, and Jaakko Lehtinen. "Progressive growing of gans for improved quality, stability, and variation." *arXiv preprint arXiv:1710.10196* (2017).

Shaham, Tamar Rott, Tali Dekel, and Tomer Michaeli. "Singan: Learning a generative model from a single natural image." In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4570-4580. 2019.

- Tentative work
 1. Use GANs to generate a group of images. -> Apply a prior classifier to soft label these images. -> send true data and generated data into the model
 2. Train a GAN with the label-0 data (healthy) -> get the D as a pretrained model -> train a GAN with the label-1 data -> generated data -> send all data to train the pretrained D
 - 3.

Week 3

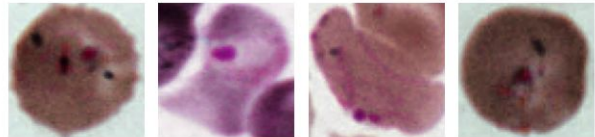
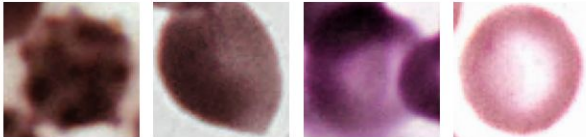
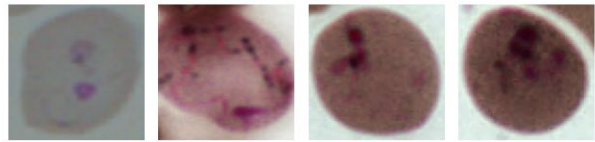
- Construct the first GAN model, DCGAN. Generate samples.
- Inception Score is not stable (by literature, not my experiment). Change the evaluation metric to Frechet Inception Distance (FID). Define the FID function to evaluate the generated samples. (Still work on it)

True Dataset

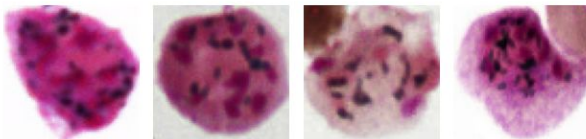
Type 0 Cells



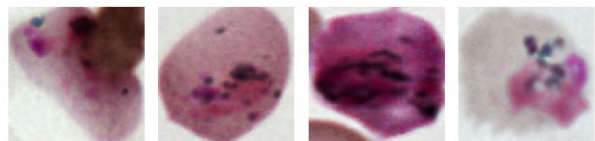
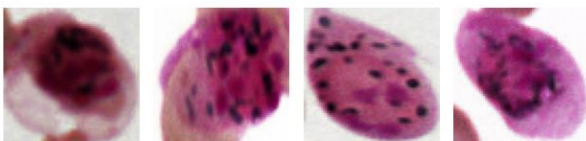
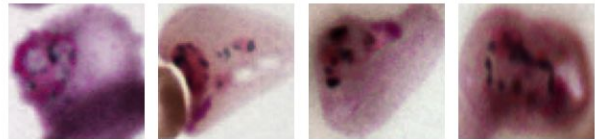
Type 1 Cells



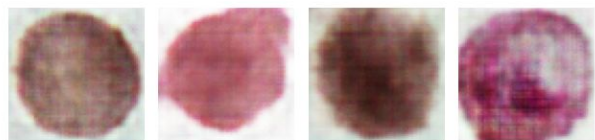
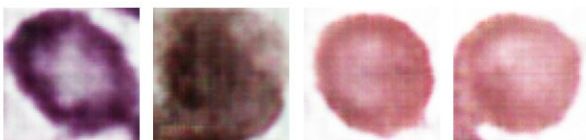
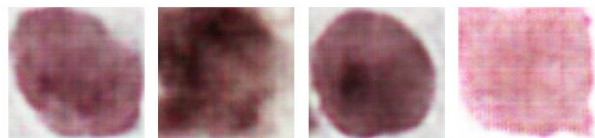
Type 2 Cells



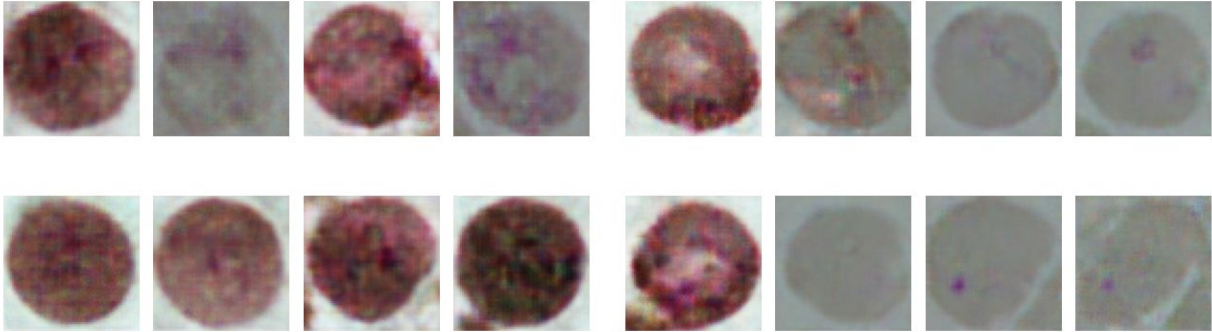
Type 3 Cells



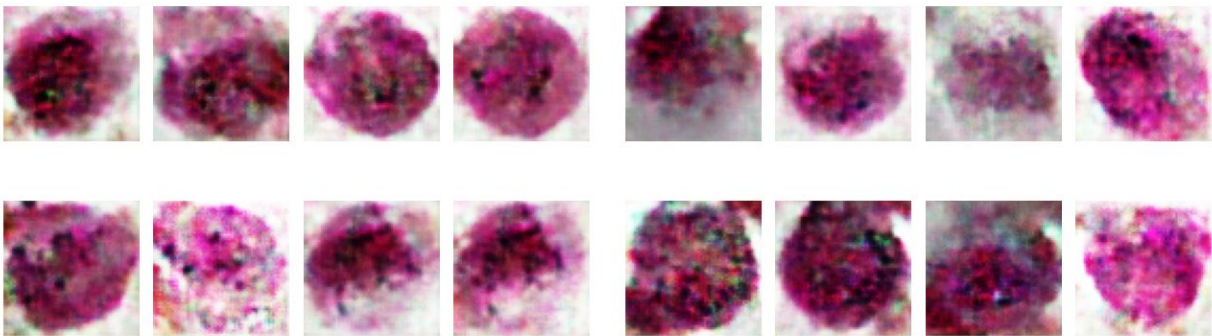
Generated samples (Input: all types of cells)



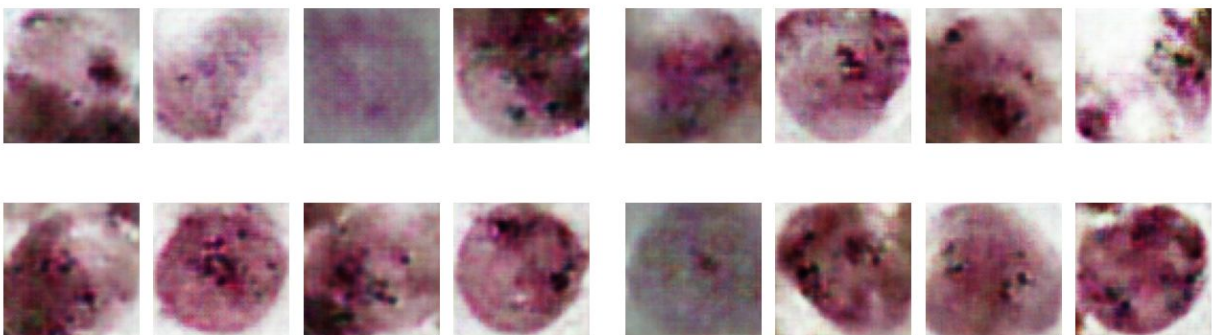
Generated samples (Input: one type of cells)
Type 1



Type 2



Type 3



Week 4

- Run a baseline CNN model for cell classification, then try the original/augmented dataset.

Without augmentation. - 30 epochs

```
Epoch 30/30
5508/5508 [=====] - 1s 208us/step - loss: 0.1125 - accuracy: 0.9609 - val_loss: 0.1768
1722/1722 [=====] - 0s 141us/step
Final accuracy on validations set: 93.90243887901306 %
Cohen Kappa 0.46529508789821983
F1 score 0.5056512774882392
Confusion Matrix:
[[1400   0   0   0]
 [  67   1   0   5]
 [   6   0  10  11]
 [ 126   0   3  93]]
```

With 3000 augmented samples. 1000 for each type 1 & 2 & 3. - 30 epochs

```
Epoch 30/30
7908/7908 [=====] - 1s 137us/step - loss: 0.1610 - accuracy: 0.9311 - val_loss: 0.1441
1722/1722 [=====] - 0s 138us/step
Final accuracy on validations set: 97.16898798942566 %
Cohen Kappa 0.8182436328608045
F1 score 0.6472630194084553
Confusion Matrix:
[[1395   0   0   5]
 [  27  20   0  26]
 [   0   0   6  21]
 [  11   0   5 206]]
```