PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Anonymising pathology data using generative adversarial networks

Morrison, David, Harris-Birtill, David

David Morrison, David Harris-Birtill, "Anonymising pathology data using generative adversarial networks," Proc. SPIE 12039, Medical Imaging 2022: Digital and Computational Pathology, 1203917 (4 April 2022); doi: 10.1117/12.2611803

SPIE.

Event: SPIE Medical Imaging, 2022, San Diego, California, United States

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 09 May 2022 Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Anonymising Pathology Data using Generative Adversarial Networks

David Morrison[a] and David Harris-Birtill[a]

[a]School of Computer Science, University of St Andrews, Jack Cole Building, North Haugh, St Andrews, KY16 9SX

ABSTRACT

Anonymising medical data for use in machine learning is important to preserve patient privacy and, in many circumstances, is a requirement before data can be made available. One approach to anonymising image data is to train a generative model to produce data that is statistically similar to the input data and then use the output of the model for downstream tasks, such as image classification, instead of the original sensitive data. In digital pathology, it's not yet well understood how using generative models to anonymise histology slide data impacts the performance of downstream tasks. To begin addressing this, we present an evaluation of a histology image classifier trained using patches extracted from the Camelyon 16 dataset and compare it to a classifier trained on the same number of synthetic images generated with a Deep Convolutional Generative Adversarial Network (DCGAN), from the same data. When predicting the class of an image patch as either cancer or normal it's shown that the accuracy reduces from 0.78 for original alone to 0.59 for synthetic alone, and the recall is significantly reduced from 0.70 to 0.44 when training exclusively on the same amount of synthetic data. If retaining a similar accuracy is required for the downstream task, then either the original data must be used or an improved anonymisation strategy must be devised. We conclude that using this DCGAN to anonymise the dataset, degrades the accuracy of the classifier which implies that it has failed to capture the required variation in the original data to generalise and act as a sufficient anonymisation strategy.

Keywords: GANs, Generative Adversarial Networks, Anonymisation, Histopathology, Digital Pathology, Medical Anonymisation

1. DESCRIPTION OF PURPOSE

Data anonymisation is often required in medical domains to preserve patient privacy and to comply with the terms under which the data has been collected.[1] Generative models, such as Generative Adversarial Networks (GANs)[2] have performed well at generating and anonymising data across a range of domains and tasks, such as brain MRI data[3] and human face datasets.[4] Currently, their use in digital pathology has been limited, though some recent work has used GANs in digital pathology for data augmentation to provide more samples when data is limited. For example, Wei et al.,[5] make use of a cycleGAN-based[6] model to show that GANs can be used to enrich datasets by translating from one tissue type to another in order, increasing the amount of rarer tissue samples. PathologyGAN[7] shows that GANs can be trained that generate data suitable for effective data augmentation without the need for additional input images. See Tschuchnig et al.[8] or Morrison et al.[9] for two recent reviews. Currently, anonymisation has not been the specific subject of investigation. Often clinical data providers, such as hospitals, are unable to release patient data for institutional or legal reasons. However, synthetically generated data with similar statistical properties do not suffer from the same ownership and data governance restrictions. Therefore, using a GAN to synthetically generate data from the original images provides an opportunity to exploit a large number of previously inaccessible medical imaging datasets for use in tasks, such as detecting cancer.

Our purpose with this study is to assess the appropriateness of Deep Convolutional Generative Adversarial Networks (DCGAN) as an anonymisation technique for histopathology datasets.

Further author information: (Send correspondence to David Morrison)
David Morrison: E-mail: dm236@st-andrews.ac.uk

Medical Imaging 2022: Digital and Computational Pathology, edited by John E. Tomaszewski, Aaron D. Ward, Proc. of SPIE Vol. 12039, 1203917 · © 2022 SPIE
1605-7422 · doi: 10.1117/12.2611803

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 09 May 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

2. METHODS

To assess the effect of using only synthetic (anonymised) generated data instead of the original clinical data we trained three deep neural networks: one on original data, one on synthetic (a.k.a synthetic) data, and one on 50% original and 50% synthetic data. The model trained on the original data provides an accuracy that can be used to measure the relative performance of the other models and thus the effect on the accuracy of using all synthetic data or a mix of the two data sources.

Each of the three networks has the same architecture and hyper-parameters. They were all simple six-layer convolutional neural networks acting as binary classifiers, classifying the patches into normal or tumour. They were trained using a Stochastic Gradient Descent (SGD) optimiser with a learning rate of 0.001, a momentum of 0.9, and a weight decay of 0.0005. A learning rate step scheduler was used with a step size of 2000 and a gamma of 0.5. The original data was provided by the Camelyon16 whole slide image dataset[10] which is scans of biopsies taken from breast lymph nodes. These images were then preprocessed into a set of 256x256 sized patches at a x40 magnification and then downscaled, for reasons of efficiency, to 64x64. Each patch was labelled either normal or tumour based on the provided annotations.

The synthetic data was generated using two separately trained DCGANs,[11] one synthesising normal patches and another synthesising tumour patches. Each GAN was trained for 400 epochs with a batch size of 256 patches per batch. Pytorch[12] was used to accelerate the training process for the GANs and the classifiers and the program was run on an NVidia DGX-1[13] (eight Tesla V100 GPUs interconnected with NVLink). Training the GAN to generate normal labelled patches took 3h 49min 12s and training the GAN to generate tumour labelled patches took 1h 59min 44s.

The DCGANs were trained using datasets drawn from the same larger set of patches as the original classification sets. All sampling was without replacement, however, the same test set, composed of original data, was used for each classifier so they could be reliably compared against each other. Table 1 show the compositions of each dataset. Figure 2 shows examples of the synthetic tumour patches that were generated.

Model	Train	Validate	Test
GAN (Normal)	35k original	-	-
GAN (Tumour)	35k original	-	-
Original	70k original	30k original	15k original
Synthetic	70016 synthetic	30k original	15k original
Mixed	35k original, 35k synthetic	15k original, 15k synthetic	20k original

Table 1. Composition of training, validate, and testing sets for different classifiers and the GANs.

The source code for the project can be found at https://www.github.com/davemor/pathgen.

3. RESULTS

Table 2 show the precision, recall, and f1-scores for each of the models when tested using the same test set.

Model	Accuracy	Precision	Recall	F1-score
Original	0.78	0.83	0.70	0.76
Mixed	0.75	0.82	0.65	0.73
Synthetic	0.59	0.64	0.44	0.52

Table 2. Results of predicting on the same test set with each model.

This preliminary study shows three things. Firstly it shows that a DCGAN when trained on 35,008 image patches of breast cancer or healthy tissue (from the Camylon dataset) is able to produce images which to the non-expert eye look cell-like in nature with reasonable appearance, as shown in figures 1 and 2. However, when these synthetic patch images are used in place of the original patch images to train a basic CNN classifier the classification performance is shown to degrade across a range of metrics (including accuracy, precision, recall and F1-score) as shown in table 2. We also note that the performance of the basic CNN model on the original

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 09 May 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

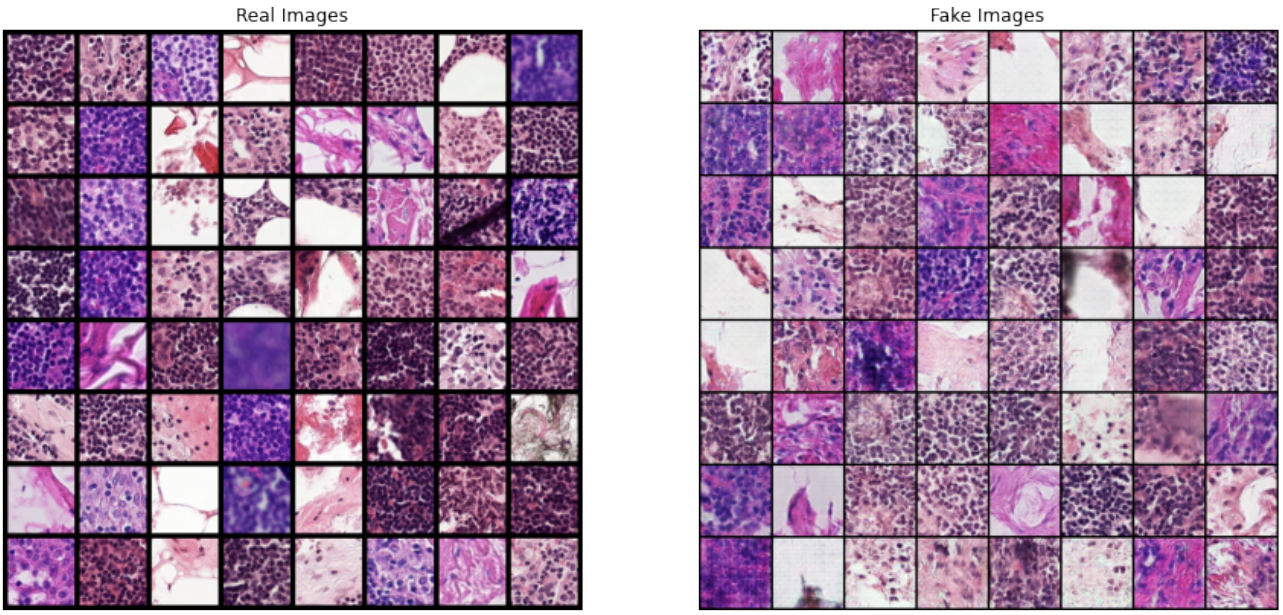

Figure 1. 64 examples of normal patches extracted from Camelyon 16 (left) and 64 examples of normal patches generated by the GAN (right).

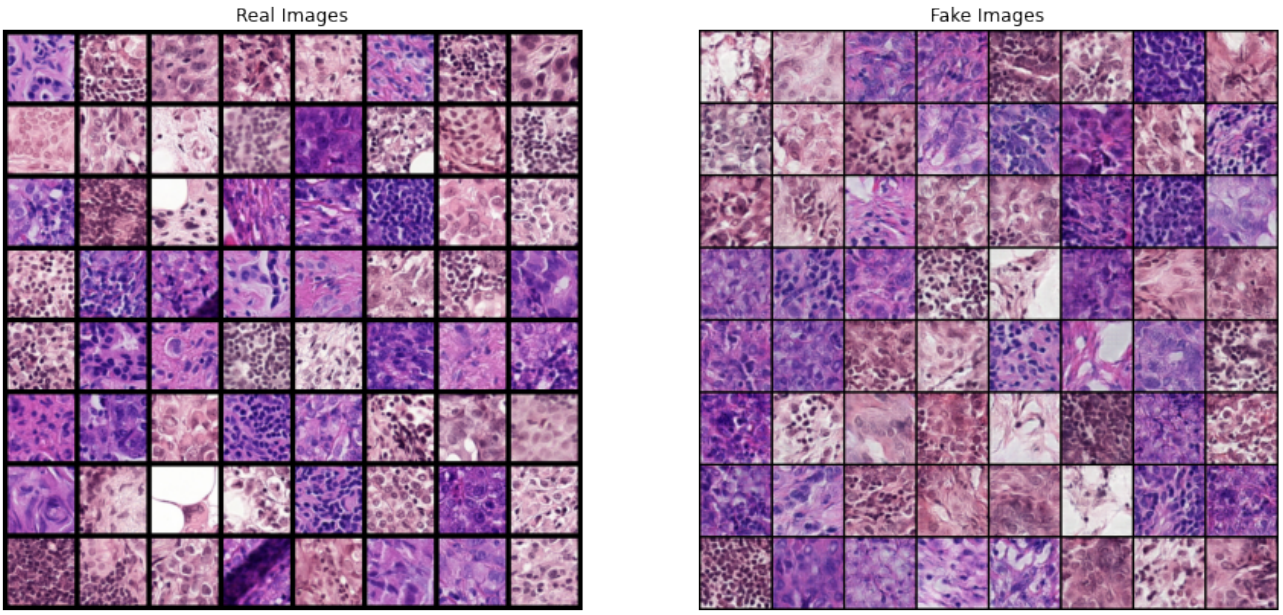

Figure 2. 64 examples of tumor patches extracted from Camelyon 16 (left) and 64 examples of tumor patches generated by the GAN (right).

data is not state-of-the-art, current state-of-the-art performance for Cameylon 16 dataset is 98.4%[14] for patch classification which uses a much deeper network based on GoogLeNet.[15] To enable training times to allow for iteration this preliminary study does not focus on obtaining the top performance for the patch classification task, this preliminary study is to assess the potential for using GANs as a synthetic data source for use in anonymisation of pathology patches assuming all meta-data is removed.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 09 May 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

4. CONCLUSIONS

From the results shown in table 2, it can be seen that the performance decreased across all four metrics when the synthetic data was used for training. This implies that the DCGAN has failed to capture the required variation in the original data to generalise and act as a sufficient anonymisation strategy. Though it is unlikely that a DCGAN-based synthetic dataset using our specific training regime could be used to replace non-anonymised data of the same number of samples, the power of a generated model is that it can be used to create a dataset of unlimited size. An interesting next step of this work will be to generate a series of datasets of increasing size, to assess the performance as the amount of synthetic data is increased. In addition, varying the models used for both the GANs and the classifiers may have a significant effect on the outcome of the experiment particularly if they are replaced with state-of-the-art networks such as GoogLeNet[14,15] for the classifiers and StyleGAN2[16] for image generation. The third line of enquiry is to attempt to measure the degree of anonymisation provided by the synthetic dataset perhaps by looking at the cross-correlations of all the images in the real and synthetic sets. Such a measure is important to ensure that the dataset is being correctly anonymised and that the generative models are not memorising the input data and outputting that same data into the synthetic dataset.

REFERENCES

[1] Patel, M., Looney, P., Young, K., and Halling-Brown, M., "Automated collection of medical images for research from heterogeneous systems: trials and tribulations," in [*Medical Imaging 2014: PACS and Imaging Informatics: Next Generation and Innovations*], **9039**, 90390C, International Society for Optics and Photonics (2014).

[2] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., "Generative adversarial nets," in [*Advances in neural information processing systems*], 2672–2680 (2014).

[3] Shin, H.-C., Tenenholtz, N. A., Rogers, J. K., Schwarz, C. G., Senjem, M. L., Gunter, J. L., Andriole, K. P., and Michalski, M., "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," in [*International workshop on simulation and synthesis in medical imaging*], 1–11, Springer (2018).

[4] Maximov, M., Elezi, I., and Leal-Taixé, L., "Ciagan: Conditional identity anonymization generative adversarial networks," in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*], 5447–5456 (2020).

[5] Wei, J., Suriawinata, A., Vaickus, L., Ren, B., Liu, X., Wei, J., and Hassanpour, S., "Generative image translation for data augmentation in colorectal histopathology images," *Proceedings of machine learning research* **116**, 10 (2019).

[6] Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A., "Unpaired image-to-image translation using cycle-consistent adversarial networks," in [*Proceedings of the IEEE international conference on computer vision*], 2223–2232 (2017).

[7] Quiros, A. C., Murray-Smith, R., and Yuan, K., "Pathologygan: Learning deep representations of cancer tissue," *arXiv preprint arXiv:1907.02644* (2019).

[8] Tschuchnig, M. E., Oostingh, G. J., and Gadermayr, M., "Generative adversarial networks in digital pathology: A survey on trends and future potential," *arXiv preprint arXiv:2004.14936* (2020).

[9] Morrison, D., Harris-Birtill, D., and Caie, P. D., "Generative deep learning in digital pathology workflows," *The American Journal of Pathology* (2021).

[10] Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., Van Der Laak, J. A., Hermsen, M., Manson, Q. F., Balkenhol, M., et al., "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer," *Jama* **318**(22), 2199–2210 (2017).

[11] Radford, A., Metz, L., and Chintala, S., "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434* (2015).

[12] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al., "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems* **32**, 8026–8037 (2019).

[13] "Nvidia dgx-1: Deep learning server for ai research."

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 09 May 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

[14] Wang, D., Khosla, A., Gargeya, R., Irshad, H., and Beck, A. H., "Deep learning for identifying metastatic breast cancer," *arXiv preprint arXiv:1606.05718* (2016).

[15] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A., "Going deeper with convolutions," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 1–9 (2015).

[16] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., and Aila, T., "Analyzing and improving the image quality of stylegan," in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*], 8110–8119 (2020).

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 09 May 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use