GENERATIVE DEEP LEARNING IN DIGITAL PATHOLOGY WORKFLOW

A PREPRINT

David Morrison*
School of Medicine
University of St-Andrews
dm236@st-andrews.ac.uk

Peter Caie
School of Medicine
University of St-Andrews

David Harris-Birtill
School of Computer Science
University of St-Andrews

May 9, 2022

ABSTRACT

Many modern histopathology laboratories are in the process of digitising their workflows. Once images of the tissue exist as digital data, it becomes feasible to research the augmentation or automation of clinical reporting and diagnosis. The application of modern computer vision techniques, based on Deep Learning, promise systems that can identify pathologies in slide images with a high degree of accuracy. Generative modelling is an approach to machine learning and deep learning that can be used to transform and generate data. It can be applied to a broad range of tasks within digital pathology including the removal of color and intensity artefacts, the adaption of images in one domain into those of another, and the generation of synthetic digital tissue samples. This review provides an introduction to the topic, considers these applications, and discusses some future directions for generative models within histopathology.

1 Introduction

Clinical histopathology is at an exciting paradigm shift with many laboratories replacing traditional microscopy with high-resolution scanners and large digital displays. Unlike traditional slides, digital images can be shared electronically, marked up simultaneously by multiple pathologists, and assessed automatically[1]. It is hoped that the deployment into clinical practice of systems that automate and augment diagnostic reporting will lead to a significant increase in assessment capacity alongside quicker reporting times. This article will provide a brief introduction to Deep Generative Models (DGM), review their current use in digital pathology, and envision future applications for them within the field. To contextualise this work, DGMs are discussed in relation to the current state of the art deep learning techniques for pathology and the problems that generative techniques can solve within a conventional pipeline.

Before discussing the place that generative models could take in the field of automated histopathology, it is necessary to describe the current typical workflow of machine learning in digital pathology and some of the common issues that can hinder downstream reporting tasks. Mason and Wiggins[2] describe a taxonomy of data science tasks, independent of pathology, organising them into five categories, undertaken sequentially; obtain, scrub, explore, model, and interpret. This model can be used to understand the process of applying machine learning in digital pathology. Data is obtained through the fixing, staining, and scanning of tissue to transform it into a set of whole slide images. These images are then scrubbed, or pre-processed, to remove artefacts and make them ready to be used in the modelling phase. Tasks such as stain normalisation, data augmentation and patch generation fall into this category. In the exploration phase, resulting scrubbed data is analysed, either automatically or by a human, to determine an appropriate modelling technique, such

*Data Scientist iCAIRD project and Postgraduate Research Student in the School of Computer Science

as which neural network architecture to use. In digital pathology, there are a large number of different pathologies and tissue types that may be of interest and it's therefore impractical to iteratively try every possible modelling technique and, in the case of ensemble learning, every combination of technique. Modelling is where the machine learning system is trained and evaluated. In the interpretation phase, human pathologists are presented with the predictions of the model and can make use of it in the context of clinical or research work.

Figure-1 presents the Mason and Wiggins taxonomy[2] and although it may seem relatively simple, the automation of Whole Slide Image (WSI) analysis and diagnosis presents several significant challenges[3]. Foremost is the issue of data size; Whole Slide Images are multi-gigabyte images in the range of approximately 100,000 x 100,000 pixels. This makes the direct application of modern computer vision algorithms impractical on non-specialist computing hardware. Typical solutions to overcome this have been applied which include: downsampling the image and breaking the image up into smaller sub-images called patches.

Secondly, data availability is problematic for most researchers. Supervised machine learning requires labels for each sample. In WSI analysis, this may mean assigning a category to each slide as a whole, identifying a set of points-of-interest on the tissue, or drawing around areas to segment tissue types or pathologies. For each of these, a trained specialist in histopathology is required. The process is time-consuming, expensive, and there is often a lot of inter-observer and intra-observer variability between the labels provided by pathologists. As a result, data sets used to train automated digital pathology models tend to be small compared to those available in other computer vision subfields, such as ImageNet[4], where non-specialists can straight-forwardly provided labels, e.g. labelling a cat vs a dog. This situation, however, has been improved by the release of tissue annotated open data sets, such as Camelyon16[5] and 17[6], furthermore initiatives such as iCAIRD[7] and Pathlake[8] are being undertaken to provide large, well annotated and curated WSI data sets linked to clinicopathological data. These make rich digital pathology training material widely available, albeit within narrowly defined clinical reporting and specific tissue types.

Thirdly, WSI analysis sufferers from a number of domain-specific image artefacts caused by the process of surgical removal, fixing, cutting, staining, and scanning the tissue. These can include folds in the tissue, retraction artefact, variations in the application of chemicals in the staining process, small cracks and imperfections in the glass slide and coverslip, partial blurring of the image caused by focusing errors, and image resolution and compression differences between different scanners and file formats.

Despite these challenges, computer vision techniques based on supervised and weakly-supervised learning have been used to successfully automate some common assessment tasks in histopathology. For example, cell nucleus identification, pathology classification, and cancer segmentation[9]. Unsurprisingly, state of the art results on slide classification tasks, such as the work by Campanella, Hanna, Geneslaw, Miraflor, Werneck, Silva, Busam, Brogi, Reuter, Klimstra, and Fuchs[10] on prostate cancer, basal cell carcinoma and breast cancer nodal metastases, rely on large data sets.

2 Deep Generative Models

This section briefly introduces required terminology from computer vision, deep learning, and generative modelling before describing their uses in a digital pathology workflow. Firstly, an image filter or kernel is a rectangular matrix that can be applied to parts of a digital image to extract information, called features, from it. To apply a filter, a dot-product is performed (component-wise multiplication followed by a sum) between the filter and a section of the image with the same dimensions. In computer vision, this operation is referred to as a convolution. By sliding the filter across the image and performing the convolution at each point, this operation can produce a new matrix, known as a feature map. Filters that recognise primitive features, such as horizontal or vertical lines, can be hand-crafted, however more complex ones must be learned by the model. Neural networks are the most commonly used of machine-learning approaches[11]. A convolutional neural network[12] is a machine learning approach that enables image filters to be learned from data rather than programmed explicitly.

Generative models are an approach to machine learning in which systems attempt to estimate the probability of a specific sample being picked at random based on some training data[13]. Once there is an estimate for the probability density function over the training set, the model can be used to generate new examples. For example, a model can be trained that generates new images of cats by training it on a large number of images of cats. Generative models are contrasted with discriminative models, which estimate the probability of an output value given an input value (this includes classification and regression problems). Recently, generative models based on deep learning have shown to be good at generating novel data across a range of domains and tasks.

The most effective techniques, such as Generative Adversarial Networks (GANs)[13] and Variational Autoencoders (VAEs)[14], come from a class of models know as latent variable generative models. In such systems, a model is trained

that takes the lower-dimensional representation of data, called the latent space vector, and generates high-dimensional data from it. GANs and VAEs differ in the way they are trained, but both conceptualise generation as decoding. By changing what data is passed in, as the latent space vector, model parameters can be learned that enable the model to perform data translation tasks. In their recent review of GANs in Pathology[15], Tschuchnig, Oostingh, and Gadermayr[16] split the GANs up based on what kind of translation task the model is training for. This put the emphasis on task, for example image-to-image translation verses label-to-image translation. The rest of this review will describe how different generative models have been trained to perform different translation tasks and how these could be usefully applied to the automated reporting of a clinical task within a digital pathology pipeline.

Generative Adversarial Networks[13] are a class of generative model in which a network, known as the generator, is trained by having it attempt to trick a second model, known as the discriminator. The discriminator and the generator are trained simultaneously. During training, the generator is sampled from by having to translate noise into fake data. The discriminator is then trained on a combination of the fake data, labelled as fake, and the real data, labelled as real. The generator is then trained by having it generate fake data and asking the discriminator to predict labels for it. The training loss for the generator is based on how well the discriminator can tell them apart, i.e. how well the generator can fool it (generating fake data that the discriminator classes as real). This simultaneous training procedure can cause GANs to both be computationally expensive and suffer from difficulty in converging to an accurate solution.

3 Generative Models in the Digital Pathology Pipeline

Generative models have the potential to overcome a number of the issues that researchers are confronted with when developing computer vision systems for digital diagnosis and reporting. For example, when data sets stained at different institutions are used there can often be a lot of variation in color and intensity. It can be expensive and time-consuming to acquire high-quality labelled training data and generative models are being applied to the creation of synthetic datasets to overcome this. Generative models can also be used to stain tissue virtually, reducing the tissue preparation overhead.

3.1 Color and Intensity Normalisation

During tissue preparation, particularly staining, variations in color and intensity can be introduced between different whole slide images. These artefacts can complicate the interpretation of the slide by pathologists and computers. When this occurs, similar tissue features can present differently or different ones similarly. Such artefacts are introduced from several sources such as: differences between scanners, the thickness of the cut tissue samples, and the amounts and concentrations of chemicals used in varying staining protocols. These issues can be mitigated in three ways: color information can be ignored, models can be trained that learn features that are insensitive to the artefacts[17] or, images can be normalised to account for differences.

By converting the image to greyscale, much of the information that is provided by the staining process is lost. Analysis techniques for greyscale pathology images have to rely on other features, for example texture and morphology[18, 19], leading to lower performance on downstream tasks. In other situations, artefacts can be compensated for by applying a large number of color perturbations to the training data so that a wide range of variations were presented to the model during training[17]. This technique requires the perturbations to be statistically similar to the color and intensity variations across the data to be assessed, information that is not always available, and requires increased computational and memory overheads due to the large amount of data augmentation.

Ruifrok and Johnston[20] proposed a novel method based on color-deconvolution that depended on user-determined color information to reconstruct images for each stain. This method provided state of the art results for stain-normalisation but proved limited in its applicability to extensive studies due to the user needing to estimate the values used in the deconvolution manually. Whereas Magee, Treanor, Crellin, Shires, Smith, Mohee, and Quirke[21] presented a method for estimating the required color deconvolution parameters from the image data, eliminating the need for user input. This work was extended by Khan, Rajpoot, Treanor, and Magee[22] to account for image specific color variations and to improve the training data used to separate the different stains.

A limitation of color-deconvolution techniques is their failure to take into account information outside of the image color, for example tissue structure or texture. Generative models are able to address this limitation. Stain normalisation can be thought of as an image generation problem. Generative models have proved useful for image generation and recently have been applied to generate normalised pathology slides. Three different approaches have been applied to this task; stain-style tranfer[23], CycleGAN[24] based image-to-image translation, and Pix2Pix-based translation[25].

3.1.1 Stain-Style Transfer

Neural style transfer[26] is an image translation technique that transfers the style of one source image onto the content of another to generate a target image. The terms style and content can be a little misleading at first; content refers to aspects of the image, like the shape and arrangement of nuclei and cells and the tissue architecture that they comprise, and style refers to aspects such as color, like the haematoxylin and eosin shades, and texture, for example the nuclear chromatin. Style representations are derived from correlations in between the same location in different activation maps of the same layer of a neural network. For example, there might be a filter that recognises blue pixels and another that recognises a curve. If they consistently activate together, then this would represent that curves are generally blue. Stain normalisation can be thought of as a kind of style transfer from the source to the target, however it's important that only the color distribution is transformed, not other histopathological features.

Stain-style Transfer[23] uses a modification on GANs to perform color normalisation and its use was demonstrated in histopathology by its application on patches extracted from the Camelyon16 dataset[5]. The normalised patches were shown to improve tumor classification. In this technique, this input into the GANs generator is changed from noise to the unnormalised image. A Conditional GAN[27] is then used in which both the generator and discriminator are trained to generate and discriminate class labels for each patch, in this case tumor or non-tumor, in addition to the fake or real labels. On its own, this was shown to produce distortion in the patches' non-color histopathological features. In order to address these issues, two other loss functions were added to the system; reconstruction-loss, to minimise the difference between the source and generated images, and feature-preserving loss which derives a loss by comparing the activations of the final layer of the discriminator when the source and generated images are passed though the network. This approach was shown to improve the classification accuracy of a CNN-based model trained on image patches extracted from the Camelyon16 dataset. Bentaieb and Hamerneh[28] propose a similar approach in which the generator architecture is replaced with a U-Net encoder-decoder style network, called the Stain Transfer Network, and the discriminator is given an additional classification task. This approach was assessed on both classification and segmentation tasks, across three separate data-sets, showing it can be used to improve the identification of a wide range of tissue and pathology types.

3.1.2 Pix2Pix-base Image-to-Image Translation

Pix2Pix[25] is an extension of conditional GANs which, like other image-to-image translation models, learns the mapping from one image domain to another. The difference with Pix2Pix is that it also learns a loss function in order to train the translation model. This means that models based on Pix2Pix can be trained to translate between different domains without the need to specify a specific loss function for that translation, something that is hard to do. Like Conditional GANs for image-to-image translation, Pix2Pix requires image pairs, one from each domain, as example translations. Salehi and Chalechale[29] apply this approach successfully to the stain normalisation using five different H&E datasets. The method involves de-staining the patches by reducing them to greyscale, before synthetically re-staining them in a way that ensures that the color is consistent. This is similar to the artificial staining proposed by Rana, Yauney, Lowe, and Shah[30], discussed under data adaptation, and has been shown to perform well across a range of statistical measurements comparing ground-truth stained images against those re-stained using the GAN. This indicates that they will improve downstream assessment tasks, such as tumor classification and segmentation, in a similar way to the Stain-style Transfer techniques[23, 28].

3.1.3 CycleGAN-based Image-to-Image Translation

One of the key disadvantages of Pix2Pix is the need for paired images from the source and target domain, for example co-registered images before and after staining. CycleGAN [24] removes this requirement, allowing models to be trained to translate from a source to a target domain without the need for paired examples. This is done by training an inverse mapping from the source to target domain, at the same time as training the translation. By comparing the original image to one that has had the forward and inverse transformation applied to it, a loss called cycle-consistency loss is derived. When the Generator is trained, cycle-consistency loss is minimised, as is the conventional adversarial loss derived from trying to fool the discriminator.

Within digital patholgy, de Bel, Hermsen, Kers, van der Laak, and Litjens[31] showed that by modifying the original cycleGAN[24] to use a U-Net[32] style architecture, it made it more suitable for use with pathology images. They were able to use the system to artificially stain images to a high quality. The technique was applied to two datasets of renal tissue sections stained with Periodic Acid-Schiff from different staining centers. Models trained using the normalised data was shown to have increase accuracy when segmenting various objects of interest within the renal slides such as, arteries, tubuli, and glomeruli. However, the system was shown to be able to generate changes in texture, something that breaks the constraint that the transform should preserve non-color tissue features and potentially introduces unwanted bias into the generated datasets.

3.2 Data Adaptation

Data adaptation is the task of taking the data in one domain, say H&E WSIs, and translating them into images that resemble those in a different domain, say immunofluorescence WSIs. This can be useful as a data augmentation technique, allowing for images labelled in one domain to be used effectively for learning in another domain. Doing this relies on the image translation process retaining the correct labels. For example, if something is labelled as a cell nucleus, it has to still look like a cell nucleus once it has been translated.

One possible use of this data adaptation is to enrich patches with additional channels showing different fluorescence labels that highlight different kinds of information. This is called multiplexing and has traditionally been achieved though re-labelling the same tissue multiple times and scanning in each fluorophore separately. There are two issues with this; after multiple re-labelling the tissue quality begins to degrade, and scanning requires the slides to be precisely aligned in order to allow the tissue to be co-registered. By doing virtual staining, the tissue is not degraded and, because a single scan is used, there are no issues related to alignment.

A histopathological-to-immunofluorescence translation model that uses Pix2Pix[25] has been introduced by Burlingame, Margolin, Gray, and Chang[33]. They adapt Pix2Pix by adding an adaptive regularisation term during training that changes based on the prevalence of stained tissue in the patch. Patches with a low amount of stained tissue are penalised. This composites for the relative ease of translating patches with low amounts of tissue. The systems ability to generate realistic immunofluorescence stains from hematoxylin and eosin stains, opens up the possibility of quickly providing information about cellular complexity when only a standard H&E stain is available.

Another possible application of image translation is artificial staining, in which the source domain is an unstained image and the target is stained ones. If it's possible to do this in a consistent way, it can remove the need for lab based staining with its associated variations, requiring stain normalisation, and for the potential of human error. Rana, Yauney, Lowe, and Shah[30] apply a modified Pix2Pix[25] model in which the generator made use of a U-Net architecture[32] in order to translate between an unstained WSI taken from a prostrate core biopsy and virtual H&E stains of the same image. Examination of the virtually stained images by pathologists showed that the system correctly stained many different histological structures, including glands, stroma, nerve, and vascular spaces.

Data adaptation can also be used as a form of data augmentation. In DASGAN[34], Kapil, Wiestler, Lanzmich, Silva, Steele, Rebelatto, Schmidt, and Brieu use a cycleGAN[24] to generate virtual PD-L1 stains from existing cytokeratin stain that has been marked up with a costly segmentation label. This data was then used to train an image segmentation model for the tumor epithelium that outperformed the same model without the additional data.

3.3 Data Synthesis

Data synthesis is perhaps the most exciting prospect for generative models, especially in the field of AI-based reporting of histopathology. In digital pathology, creating a ground truth is expensive and time consuming. If accurately labelled synthetic samples could be generated, then this problem would be alleviated. The amount of data would only be limited by the resources available to run the generative model. However a conundrum exists here and data synthesis is challenging. If there is enough data to train a generative model to generate new labelled data, then it's likely there is enough data to train an accurate classifier. Useful data synthesis requires one or both of two things; that the generative model is able to learn different representations, more useful in generation, than a possible classifier, or that extra information is somehow added to the generative process (for example, using a guide image).

PathologyGAN[35] is a study that undertakes the first of these approaches. It uses BigGAN[36], a version of GANs that applies orthogonal regularisation to the generator allowing for more control over the generator's output. PathologyGAN uses a relativistic average discriminator[37], a modification to the discriminator which estimates the probability that a real sample is more realistic than a randomly sampled fake data point. This addition was shown to improve the quality of the generated images and converge faster. This system enables the creation of large datasets suitable for training.

The generation of new images can also be achieved by reformulating the problem as image translation, as shown by Wei, Suriawinata, Vaickus, Ren, Liu, Wei, and Hassanpour[38]. They take normal colonic mucosa images and generate synthetic colorectal polyp images on them. As with many other image-to-image translation models, their system is a variation on cycleGAN[24]. They train the system on a data set that is filtered to only include patches that can be unambiguously classified using a ResNet[39] classifier. Using this approach enabled them to augment their existing colorectal histology a classifier with a 10% improvement in AUC. The field of data synthesis is still wide open for follow-on work, as noted in both of these papers.

4 Future Directions

As has been shown, within specific pathology domains, high-quality synthetic data-sets with labels can be generated using GANs[35, 40] and these have been shown to improve the performance of discriminative models trained on their data. Currently these techniques have been applied to the synthesis of patches, rather than complete whole slide images. When diagnosing, a human pathologist will mostly work at a low magnification, for example 10x, and rely on architectural features that are lost when the image is broken down into patches. There is potential to train on similar low magnification images in order to exploit these features. A single WSI can be split into many thousands of patches, meaning that the training sets for patch classifiers are many times larger when the image is patched at high magnifications. At a lower magnification, the number of images available for training reduces dramatically, making such approaches less feasible. This is where generative approaches, such as those above, could be used to generate a large number of low magnification synthetic images, containing architectural features. Using other kinds of image synthesis, such as traditional computer graphics techniques, in combination with generative models [41], may provide a useful method in this domain and is an exciting future direction.

Generative models have the potential to enable medical data of all kinds, including pathology slides, to be used to train machine learning models without them needing access to the original patient-identifiable data-set. Training a model on non-anonymized data and then using the model to generate a new artificial anonymous data-set, may provide a way to overcome the clinical firewalls that, due to patient confidentiality, prohibit many researchers accessing the original data. This topic is the subject a large amount of research within the deep learning community[42, 43, 44]. Sufficiently deep generative models are capable of memorising their training data in a way that can cause potentially confidential information to leak into any synthetic data. This has important implications for data governance going forward. In order to make generative models public it's critical to ensure they are trained in such a way that removes the possibility of confidential data being leaked into any synthetic data-set and therefore guidelines for doing this while maintaining privacy are required and will need to be adhered to.

5 Conclusion

This paper has reviewed recent advances in the application of generative models to digital pathology. Work in this domain seeks to address issues of color and intensity artefacts, data adaptation, and data synthesis, and how generative models can address these. However, there remain a number of open challenges in the digital pathology workflow that generative models can assist with. Multi-resolution WSI synthesis may provide a way to train deep models that exploit architectural tissue features in a way that is currently unpractical due to a lack of data. Secondly, differential privacy for WSI datasets may allow for a much large amount of useful data to be released publicly. The application of generative models has proven useful in improving digital pathology workflows and this fast-developing technology holds much promise in this field.

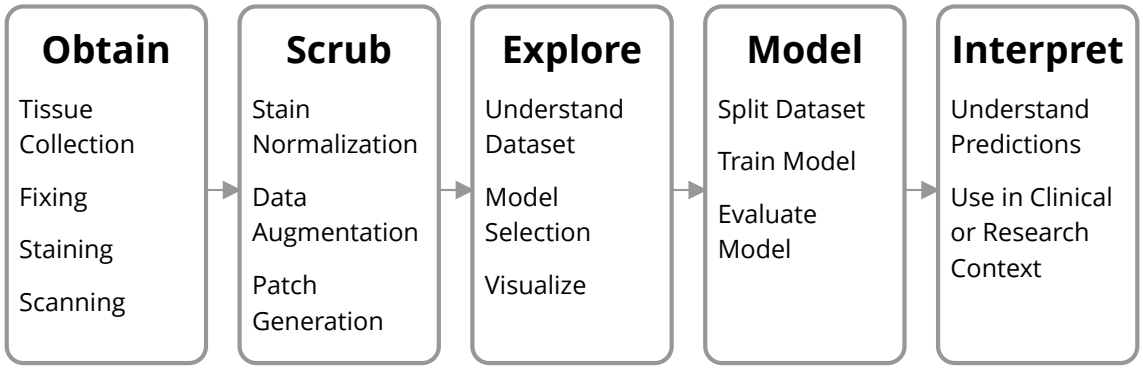

Figure 1: A taxonomy of data science tasks applied to automated Whole Slide Image Analysis

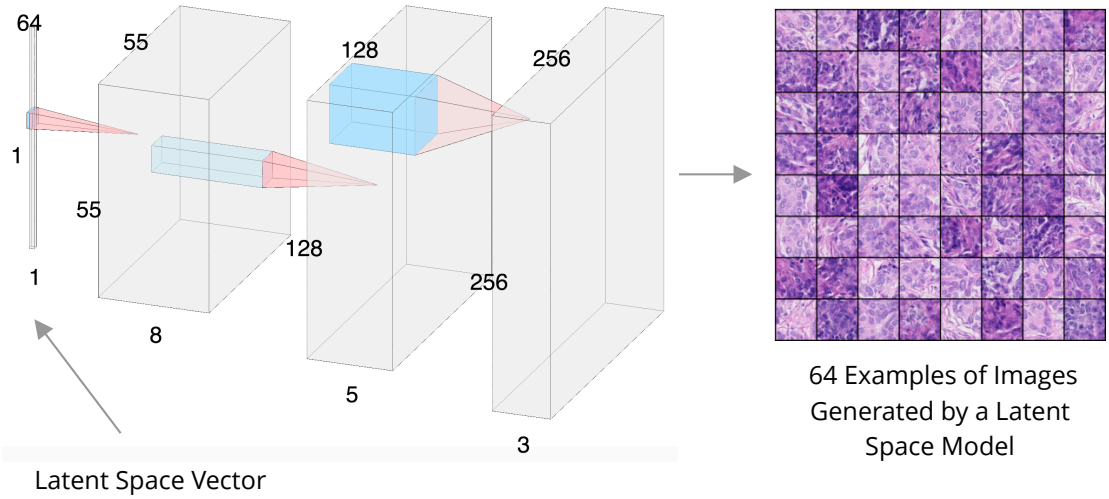

Figure 2: In latent space models, a low-dimensional vector is used to generate data. In this example a series of convolutions are used to achieve this transformation. The example output of the network shows 64 images, each 64 pixels wide and 64 pixels high.

References

[1] Neofytos Dimitriou, Ognjen Arandjelović, and Peter D Caie. Deep learning for whole slide image analysis: An overview. *Frontiers in Medicine*, 6, 2019.

[2] Wiggins C. Mason H. A taxonomy of data science. `http://www.dataists.com/2010/09/a-taxonomy-of-data-science/`. Accessed: 2020-09-18.

[3] Daisuke Komura and Shumpei Ishikawa. Machine learning methods for histopathological image analysis. *Computational and structural biotechnology journal*, 16:34–42, 2018.

[4] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.

[5] Babak Ehteshami Bejnordi, Mitko Veta, Paul Johannes Van Diest, Bram Van Ginneken, Nico Karssemeijer, Geert Litjens, Jeroen AWM Van Der Laak, Meyke Hermsen, Quirine F Manson, Maschenka Balkenhol, 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.

[6] Peter Bandi, Oscar Geessink, Quirine Manson, Marcory Van Dijk, Maschenka Balkenhol, Meyke Hermsen, Babak Ehteshami Bejnordi, Byungjae Lee, Kyunghyun Paeng, Aoxiao Zhong, et al. From detection of individual metastases to classification of lymph node status at the patient level: the camelyon17 challenge. *IEEE transactions on medical imaging*, 38(2):550–560, 2018.

[7] icaird: The industrial centre for artificial intelligence research in digital diagnostics. `https://icaird.com/`. Accessed: 2020-09-18.

[8] Pathlake: Pathology image data lake for analytics knowledge education. `https://www.pathlake.org/`. Accessed: 2020-09-18.

[9] Andrew Janowczyk and Anant Madabhushi. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of pathology informatics*, 7, 2016.

[10] Gabriele Campanella, Matthew G Hanna, Luke Geneslaw, Allen Miraflor, Vitor Werneck Krauss Silva, Klaus J Busam, Edi Brogi, Victor E Reuter, David S Klimstra, and Thomas J Fuchs. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature medicine*, 25(8):1301–1309, 2019.

[11] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[12] Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.

[13] Ian Goodfellow, 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*, pages 2672–2680, 2014.

[14] Geoffrey E Hinton and Richard S Zemel. Autoencoders, minimum description length and helmholtz free energy. In *Advances in neural information processing systems*, pages 3–10, 1994.

[15] Maximilian Ernst Tschuchnig, Gertie Janneke Oostingh, and Michael Gadermayr. Generative adversarial networks in digital pathology: A survey on trends and future potential. *arXiv*, pages arXiv–2004, 2020.

[16] Maximilian Ernst Tschuchnig, Gertie Janneke Oostingh, and Michael Gadermayr. Generative adversarial networks in digital pathology: A survey on trends and future potential. *arXiv preprint arXiv:2004.14936*, 2020.

[17] Yun Liu, Krishna Gadepalli, Mohammad Norouzi, George E Dahl, Timo Kohlberger, Aleksey Boyko, Subhashini Venugopalan, Aleksei Timofeev, Philip Q Nelson, Greg S Corrado, et al. Detecting cancer metastases on gigapixel pathology images. *arXiv preprint arXiv:1703.02442*, 2017.

[18] Clara Mosquera-Lopez, Sos Agaian, Alejandro Velez-Hoyos, and Ian Thompson. Computer-aided prostate cancer diagnosis from digitized histopathology: a review on texture-based systems. *IEEE reviews in biomedical engineering*, 8:98–113, 2014.

[19] Yu-Ying Liu, Mei Chen, Hiroshi Ishikawa, Gadi Wollstein, Joel S Schuman, and James M Rehg. Automated macular pathology diagnosis in retinal oct images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding. *Medical image analysis*, 15(5):748–759, 2011.

[20] Arnout C Ruifrok, Dennis A Johnston, et al. Quantification of histochemical staining by color deconvolution. *Analytical and quantitative cytology and histology*, 23(4):291–299, 2001.

[21] Derek Magee, Darren Treanor, Doreen Crellin, Mike Shires, Katherine Smith, Kevin Mohee, and Philip Quirke. Colour normalisation in digital histopathology images. In *Proc Optical Tissue Image analysis in Microscopy, Histopathology and Endoscopy (MICCAI Workshop)*, volume 100, pages 100–111. Citeseer, 2009.

[22] Adnan Mujahid Khan, Nasir Rajpoot, Darren Treanor, and Derek Magee. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. *IEEE Transactions on Biomedical Engineering*, 61(6):1729–1738, 2014.

[23] Hyungjoo Cho, Sungbin Lim, Gunho Choi, and Hyunseok Min. Neural stain-style transfer learning using gan for histopathological images. *arXiv preprint arXiv:1710.08543*, 2017.

[24] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

[25] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

[26] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576*, 2015.

[27] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.

[28] Aïcha BenTaieb and Ghassan Hamarneh. Adversarial stain transfer for histopathology image analysis. *IEEE transactions on medical imaging*, 37(3):792–802, 2017.

[29] Pegah Salehi and Abdolah Chalechale. Pix2pix-based stain-to-stain translation: A solution for robust stain normalization in histopathology images analysis. *arXiv preprint arXiv:2002.00647*, 2020.

[30] Aman Rana, Gregory Yauney, Alarice Lowe, and Pratik Shah. Computational histological staining and destaining of prostate core biopsy rgb images with generative adversarial neural networks. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 828–834. IEEE, 2018.

[31] Thomas de Bel, Meyke Hermsen, Jesper Kers, Jeroen van der Laak, Geert JS Litjens, et al. Stain-transforming cycle-consistent generative adversarial networks for improved segmentation of renal histopathology. In *MIDL*, pages 151–163, 2019.

[32] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[33] Erik A Burlingame, Adam A Margolin, Joe W Gray, and Young Hwan Chang. Shift: speedy histopathological-to-immunofluorescent translation of whole slide images using conditional generative adversarial networks. In *Medical Imaging 2018: Digital Pathology*, volume 10581, page 1058105. International Society for Optics and Photonics, 2018.

[34] Ansh Kapil, Tobias Wiestler, Simon Lanzmich, Abraham Silva, Keith Steele, Marlon Rebelatto, Guenter Schmidt, and Nicolas Brieu. Dasgan–joint domain adaptation and segmentation for the analysis of epithelial regions in histopathology pd-l1 images. *arXiv preprint arXiv:1906.11118*, 2019.

[35] Adalberto Claudio Quiros, Roderick Murray-Smith, and Ke Yuan. Pathology gan: Learning deep representations of cancer tissue. *arXiv preprint arXiv:1907.02644*, 2019.

[36] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale gan training for high fidelity natural image synthesis. *arXiv preprint arXiv:1809.11096*, 2018.

[37] Alexia Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard gan. *arXiv preprint arXiv:1807.00734*, 2018.

[38] Jerry Wei, Arief Suriawinata, Louis Vaickus, Bing Ren, Xiaoying Liu, Jason Wei, and Saeed Hassanpour. Generative image translation for data augmentation in colorectal histopathology images. *arXiv preprint arXiv:1910.05827*, 2019.

[39] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[40] Le Hou, Ayush Agarwal, Dimitris Samaras, Tahsin M Kurc, Rajarsi R Gupta, and Joel H Saltz. Unsupervised histopathology image synthesis. *arXiv preprint arXiv:1712.05021*, 2017.

[41] Ashish Shrivastava, Tomas Pfister, Oncel Tuzel, Joshua Susskind, Wenda Wang, and Russell Webb. Learning from simulated and unsupervised images through adversarial training. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2107–2116, 2017.

[42] Brett K Beaulieu-Jones, Zhiwei Steven Wu, Chris Williams, Ran Lee, Sanjeev P Bhavnani, James Brian Byrd, and Casey S Greene. Privacy-preserving generative deep neural networks support clinical data sharing. *Circulation: Cardiovascular Quality and Outcomes*, 12(7):e005122, 2019.

[43] Liyue Fan. A survey of differentially private generative adversarial networks. In *The AAAI Workshop on Privacy-Preserving Artificial Intelligence*, 2020.

[44] Chugui Xu, Ju Ren, Deyu Zhang, Yaoxue Zhang, Zhan Qin, and Kui Ren. Ganobfuscator: Mitigating information leakage under gan via differential privacy. *IEEE Transactions on Information Forensics and Security*, 14(9):2358–2371, 2019.