



## **Fluorescence Microscopy: Generation of Synthetic Images using Generative AI**

**MSc Artificial Intelligence and Machine Learning**

**Anagha Ramadas Mulloth**

**Student ID: 2583584**

**Supervisor: Dr Alexander Krull**

School of Computer Science

College of Engineering and Physical Sciences

University of Birmingham

2023-24

---

## Abstract

---

This study explores the generation of synthetic fluorescence images using two distinct types of input images to evaluate the efficiency of an existing architecture and the potential of a novel input type. Utilizing the Pix2Pix model, a conditional Generative Adversarial Network (GAN) architecture, the study conducts training and validation processes, achieving satisfactory results. The research focuses on enhancing the quantity and quality of microscopic image datasets by generating synthetic images from real images. This approach addresses the limitations of data availability in the medical domain, where confidentiality and the high cost of creating extensive datasets pose challenges. The study validates the reliability of a newly published dataset of white blood cell microscopic images using a popular supervised model, confirming the results with two different types of inputs. The findings demonstrate the potential of generative AI in improving image quality and expanding the scope of medical imaging applications.

---

## Acknowledgements

---

I would like to thank my supervisor Dr Alexander Krull for his immense support and guidance throughout this project. I thank the valuable feedback from my inspector Dr Sharu Jose, that helped improve the quality of the project. Last but not least, I would like to thank my family, and friends for all the support. Thanks to God.

---

## Contents

---

|   |            |
|---|------------|
| <b>Abstract</b>   | <b>ii</b>  |
| <b>Acknowledgements</b>                                     | <b>iii</b> |
| <b>List of Figures</b>                                      | <b>v</b>   |
| <b>List of Tables</b>                                       | <b>vi</b>  |
| <b>1 Introduction</b>                                       | <b>1</b>   |
| <b>2 Background Reading</b>                                 | <b>3</b>   |
| <b>3 Related Work</b>                                       | <b>6</b>   |
| <b>4 Methodology</b>  | <b>10</b>  |
| 4.1 Dataset . . . . .                                       | 10         |
| 4.2 Data Preparation . . . . .                              | 12         |
| 4.3 Proposed Architecture . . . . .                         | 12         |
| 4.3.1 Generator: U-Net . . . . .                            | 12         |
| 4.3.2 Discriminator: PatchGAN . . . . .                     | 13         |
| 4.4 Input to Models . . . . .                               | 14         |
| 4.4.1 Case 1: Brightfield Images as Input . . . . .         | 14         |
| 4.4.2 Case 2: Masked Fluorescence Images as Input . . . . . | 16         |
| <b>5 Results</b>  | <b>18</b>  |
| <b>6 Discussion and Future Work</b>                         | <b>21</b>  |
| <b>7 Conclusion</b>   | <b>25</b>  |
| <b>A Appendix</b>   | <b>31</b>  |

---

## List of Figures

---

|     |   |    |
|-----|---|----|
| 2.1 | Conditional Generative Adversarial Network architecture . . . . . | 5  |
| 4.1 | Brightfield Images of WBCs . . . . .                              | 11 |
| 4.2 | Fluorescence Images of WBCs . . . . .                             | 11 |
| 4.3 | U-Net Encoder Decoder Architecture . . . . .                      | 13 |
| 4.4 | PatchGAN Discriminator . . . . .                                  | 15 |
| 4.5 | Input with 1 channel visible and 5 channels masked . . . . .      | 16 |
| 4.6 | Input with 1 channel visible and 5 channels masked . . . . .      | 16 |
| 4.7 | Case 1: Brightfield Images as Input . . . . .                     | 17 |
| 4.8 | Case 2: Masked Fluorescent Images as Input . . . . .              | 17 |
| 5.1 | Generated Images - Training . . . . .                             | 19 |
| 5.2 | Generated Images - Validation . . . . .                           | 20 |

---

## List of Tables

---

|     |           |    |
|-----|-----------|----|
| 4.1 | Dataset   | 10 |
| 5.1 | SSIM Loss | 19 |

# CHAPTER 1

---

## Introduction

---

In the twenty-first century, artificial intelligence (AI) has grown in popularity and is revolutionising nearly every industry, including marketing, telecommunications, aviation, defence, and healthcare. AI has drawn interest from experts and the general public alike by generating innovative ideas and creating new possibilities. Artificial intelligence (AI) has been increasingly popular in recent years, largely due to improvements in machine learning methods, computing capacity, and data availability. Chatbots, virtual assistants, and other tools and systems are becoming commonplace in many facets of daily life due to the rapid growth of artificial intelligence.

The healthcare industry has shown to be a fruitful field for innovative research, notably in the areas of medical imaging and diagnostics, given its inherent obstacles and low tolerance for risk. Artificial intelligence (AI) has significantly changed this business through the application of image processing and computer vision techniques [27].

AI improves patient care, diagnosis, and treatment in a variety of ways in the healthcare industry. By 2030, artificial intelligence (AI) will have radically changed the way healthcare is delivered and how patients are treated, making it an invaluable tool for patients, researchers, and medical professionals [8]. AI is transforming the healthcare industry in several domains, such as clinical decision support systems, predictive analysis, drug discovery, and precision medicine. AI algorithms, for example, might be able to swiftly and reasonably find possible drug concepts by analysing large databases. This approach speeds up drug research.

AI is also crucial for quickening the recognition of medical trends, which are required to enhance patient outcomes and lower healthcare expenses by facilitating the early identification of illness patterns and prompt action. Significant improvements in the precision and effectiveness of image and pattern recognition algorithms have facilitated the early detection of conditions like cancer and heart problems [3]. Artificial intelligence (AI) systems can accurately assess genetic sequences, pathology slides, and medical pictures. This accuracy

allows for the crucial insight of a correct diagnosis and treatment approach. Chatbots and virtual assistants driven by artificial intelligence (AI) are revolutionising patient care through monitoring chronic illnesses, delivering timely medical information, and developing personalised health recommendations. The involvement of patients and their access to healthcare services are enhanced by these developments.

Within the field of artificial intelligence, computer vision (CV) focuses on teaching robots how to perceive, process, and evaluate visual information from their environment. Computer vision is transforming the analysis and interpretation of medical images, including MRIs, CT scans, and X-rays, in the healthcare industry. Using computer vision algorithms to quickly process and analyse images lowers the risk of human error and speeds up the diagnosis process. In an emergency situation, where every second counts, this agility is essential [13].

By seeing patterns and anomalies that the human eye might miss, CV improves diagnosis accuracy. Reducing false positives and guaranteeing that patients receive prompt, appropriate care depend heavily on this improved accuracy. Additionally, better patient outcomes are achieved by early disease detection, which is made feasible by CV's predictive capabilities, particularly for disorders like cancer [31].

Generative artificial intelligence (AI) is a recent development in computer vision that is transforming the field by creating synthetic data for computer vision system training [2]. Thus, there is less expense and risk of privacy infringement when using real data. Additionally, generative AI techniques learn data labelling more quickly and precisely than human techniques [33].

Applications of computer vision in healthcare include automated health monitoring, early diagnosis, cancer identification, and surgical help. More precise diagnosis is made possible by CV algorithms' ability to distinguish between cancerous and healthy tissue. With CV's real-time support, surgeon performance is improved and problems are reduced during surgery [19]. For instance, the real-time guiding of the Triton system calculates blood loss during surgical procedures [23].

CV is used by automated health monitoring systems to analyse vast amounts of medical images and pinpoint areas of interest that could need more investigation. By categorising patients according to risk profiles, this automation speeds up the diagnosing process.

Clinical decision-making in medical research and trials is guided by CV, which also facilitates the understanding of complex imaging data. By facilitating easy access to patient data, the integration of CV with electronic health records improves interdisciplinary collaboration among healthcare teams.

In this study, we aim to improve the area of cell microscopy in medical AI using the advancements in generative AI. More specifically, we focus on enhancing the quantity and quality of microscopic image datasets by generating synthetic images from real images using generative adversarial network. The availability of data is highly limited especially in medical domain due to various reasons including confidentiality. Furthermore, creating huge datasets is expensive and time consuming. We try to address these issues and conduct a study on the reliability of a newly published dataset of white blood cell microscopic images on a popular supervised model, and verify the same using two different types of inputs.

# CHAPTER 2

---

## Background Reading

---

A key component of medical imaging, cell microscopy offers vital information about the composition and operation of cells. It is essential for identifying illnesses, comprehending biological functions, and creating novel therapeutic approaches. Varied microscopy methods have varied benefits and drawbacks, which makes them appropriate for various medical imaging applications.

Among the most basic and popular methods is brightfield microscopy. It entails shining white light on a sample and looking down on it. Brightfield photons exhibit contrast because of the sample's ability to absorb light, producing a dark image against a brilliant background. This is an easy-to-use, economical solution that needs very little equipment. When examining materials that have inherent colour, like certain organelles or pigmented tissues, it is especially helpful. But brightfield microscopy is not without its limitations. For translucent and colourless materials, it offers poor contrast, frequently requiring staining, which can be intrusive and inappropriate for living cells. Furthermore, the visible light wavelength limits the resolution, which can make it difficult to see minute cellular details [1].

One potent non-invasive method with a lot of applications in medicine is fluorescence imaging, which is especially useful for studying and visualising biological processes in living things [30]. Fluorescence imaging makes use of fluorophores, which are molecules that release light when excited by electromagnetic radiation. This technique enables a thorough analysis of molecular structures and mechanisms [9]. Understanding the kinetics of gene and protein expression as well as molecular interactions inside cells depends on this skill.

The high sensitivity of fluorescence imaging makes it possible to detect biological molecules at extremely low concentrations, which is one of its main benefits. In medical diagnostics, where early disease diagnosis can greatly improve patient outcomes, this sensitivity is especially helpful. Fluorescence imaging is a vital tool in clinical settings because it may be used to track the course of diseases, evaluate the effectiveness of medicines, and guide surgical procedures[15].

One particular kind of fluorescence imaging that has become essential for

researchers examining biological processes at the cellular level is fluorescence microscopy. It makes it possible to see subcellular elements like molecules and organelles, which sheds light on how cells work and how diseases are caused. This discipline has undergone a revolution thanks to the advent of fluorescent markers like green fluorescent proteins (GFP), which enable precise live-cell imaging and the monitoring of complex intracellular movements.

Fluorescence imaging is widely employed in the medical field, especially in oncology, for diagnostic purposes. Compared to conventional white light endoscopy, it frequently has higher sensitivity and specificity for the identification of precancerous and cancerous tissues. Fluorescence imaging is also used in a number of other medical specialities, such as gastroenterology, pulmonology, and dermatology, where it helps with a broad variety of diagnoses [28].

The therapeutic applications of fluorescence imaging technology have been further broadened by recent developments, including the creation of near-infrared (NIR) fluorescent probes. With their enhanced tissue penetration and decreased background interference, these probes increase the precision and dependability of fluorescence imaging in clinical settings. Furthermore, the combination of fluorescence imaging and machine learning approaches is opening the door to more advanced diagnostic instruments that can analyse intricate imaging data and offer insightful information about the mechanisms behind disease.

Many cells are opaque to the unaided eye, making microscopic examination of their architecture impossible. It could be difficult to discern their cellular components from the background because they seem colourless or lack contrast. One often used method in medical imaging for visualising in cell microscopy is cell staining. One technique for staining cells with fluorescent dyes is fluorescence staining. When certain wavelengths excite these fluorescent dyes, they release light, producing high contrast images that set the stained elements against a darker background. To see various cell components or proteins, different fluorescent dyes are employed [32]. Nevertheless, this method is frequently costly and time-consuming. Stains can also be harmful to cells, lowering their quality and decreasing their usefulness for live-cell imaging.

Using deep learning and computational techniques, virtual staining is a novel technology that creates the illusion of stained cells in photographs of unstained cells. This method is less expensive and time-consuming because it doesn't use dyes and is non-invasive. It also makes it possible to examine the same cells under different staining circumstances, which makes a more thorough and structural investigation possible. Consequently, virtual staining is widely employed as a substitute technique for cell visualisation.[29].

To digitally produce stained images from conventional cell photos, a variety of deep learning approaches have been investigated [14]. Convolutional Neural Networks (CNNs) are one type of these that have been used to virtually stain photos. An further method uses cascaded neural networks (C-DNNs), which are a sequence of neural networks structured such that the output of one network becomes the input for the following. The quality and accuracy of the virtually stained images can be improved by using this technique to first build a virtual stain and then refine it through subsequent networks [7]. Furthermore, preserving the tissue's structural integrity is crucial in applications where Structurally Conditioned GANs are most helpful [22]. By compressing the input data into a low-dimensional space and recreating it, autoencoders can make the complicated transformations needed for virtual staining easier to understand. These

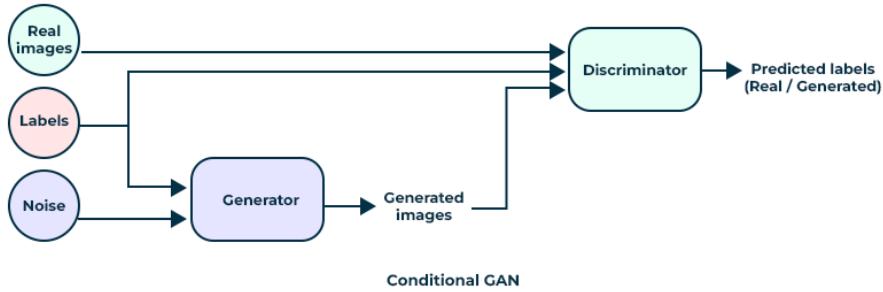


Figure 2.1: Conditional Generative Adversarial Network architecture

methods make use of deep learning to offer a more effective way to produce high-quality stained images without causing damage to the original cell structure or requiring repetitive dye application.

Conditional Generative Adversarial Networks (cGANs) are considered one of the best deep learning methods for virtual staining for paired pictures. Their ability to handle paired data and produce high-fidelity, virtually stained images that closely resemble traditional stains is partly responsible for this. Images produced by this method have a better resolution and seem more realistic than those produced by traditional CNNs, which might not have the same level of realism and detail [24]. When paired images—that is, both unstained and stained images—are available, cGANs perform especially well. To ensure that the generated images are both extremely accurate, they use this paired data to learn the mapping from the input (unstained) image to the target (stained) image. For applications in pathology where an exact representation of tissue characteristics is required, this fidelity is essential. Moreover, different stains can be freely generated from the same input image by conditioning cGANs on the desired stain information. This feature is quite helpful in creating different kinds of stains without actually changing the sample. Adversarial loss, which can be troublesome in CNNs, is used in cGANs to reduce typical artefacts like blurring. cGANs produce clearer and more detailed images by optimising the generator to trick the discriminator [20]. In Figure 2.1, the cGAN architecture is displayed.

# CHAPTER 3

---

## Related Work

---

Image-to-Image translation is a technique in computer vision and machine learning that involves transforming an input image from one domain to a corresponding output image in another domain while preserving essential features and structures. This process makes use of deep learning models like Generative Adversarial Networks (GANs), conditional GANs (cGANs), and Convolutional Neural Networks (CNNs), to learn complex mappings between different image domains [10]. These domains could represent different styles, resolutions, or semantic categories.

Several popular models have been developed for image-to-image translation, each with unique features and capabilities. Pix2pix model is a notable conditional GAN model. It uses a U-Net architecture, which combines an encoder and decoder network to capture detailed pixel-to-pixel features, enabling high quality image generation. This model is particularly effective for tasks requiring precise control over the output. However, this model requires paired images for training, which means the dataset should contain corresponding output image in the target domain for each input image [18]. CycleGAN is an unsupervised image-to-image translation model which overcomes this drawback of Pix2pix model. It does not require paired training data. This model uses ResNet architecture for the generator, which makes it well-suited for learning complex transformations. This model incorporates cycle consistency loss which enables mapping of translated image back to original image [36]. In 2022, [16] was published providing a detailed review of the research that has happened in this domain of Image-to-Image translation in the past couple of years, explaining the drawbacks of models that required paired image data, and unsupervised models out there that were introduced to solve this issue. The paper also highlights other drawbacks and research gaps in this area and also suggests potential areas to explore in the future for improvement. The Pix2pix and CycleGAN models find application in a variety of domains including style transfer, image augmentation, medical imaging, and so on. The choice of model depends on the nature of the dataset, and hence neither of them can be supremely classified as better

than the other.

Image-to-Image translation has a wide range of applications across various fields. They are used for transforming images to adopt the style of another image, such as converting a photo into a painting style. In 2022, [12] paper investigated generation of synthetic paintings that look similar to realistic paintings created by humans. The authors use an approach based on StyleGAN model which is a conditional GAN model that trains on large amounts of paintings actually created by human artisans. Their results showed that the computer generated images were quite similar to actual paintings and emulate most characteristics that were checked for manually with regard to evaluation of the quality and authenticity of art. Though such synthetic images cannot be considered as a substantial replacement of human created artwork which is a symbol of unique talent, this study could pave the way for future research regarding generating synthetic images and the influence of inherent defects or biases in the training data on the output images.

Image-to-image translation can be used for adding color to grayscale images, which becomes useful in restoring old photographs or enhancing medical images. [5] paper published in 2023 used the conditional GAN method for colorization of images. The authors use a model based on U-Net architecture and train the model using the COCO dataset. The intent of the research was to demonstrate the use of the technique of image colorization to solve real-world problems in different domains like photography, coloring black and white videos and even in medical imaging. Another application of image-to-image translation is Super-Resolution, which becomes valuable in fields like satellite imagery and medical diagnostics. Many publications in the past focus on super resolution and cross modality image transformations to improve the visualization of cellular components and finer details. In 2022, [25] introduced a technique based on the unsupervised learning model of CycleGAN with regards to super resolution. The paper tries to address the limitation of volumetric imaging due to the ill-posed problems in optical microscopy and the requirement of paired input data for training. They try to find a work around for multiple image degradation factors that affect the quality of microscopy image data. In 2023, [17] introduced a new model TCAN, that can be used to improve the resolution of confocal fluorescence microscopy images. The model takes inspiration from U-Net model and DFCAN (deep fourier channel attention network) in deep learning and employs the architecture of a conditional GAN model using the same. Though TCAN primarily focuses on confocal microscopic images, the model also shows superior performance in general on different types of microscopic images.

Image-to-image translation is a valuable technique for semantic segmentation, as it modifies images to highlight specific objects or regions, particularly useful in fields like autonomous driving. This method is also beneficial for generating additional datasets by altering existing images, thereby improving model training performance. The paper [21], presents an unsupervised approach to image translation aimed at domain adaptation in semantic segmentation tasks. It addresses the challenge of adapting models to new domains using image translation methods. The research proposes integrating real images with synthetic ones created through image translation. Although the paper mainly concentrates on using augmented datasets for object recognition in underwater environments, the approach can be adapted for similar applications in other areas. The paper [34] explores the use of conditional Generative Adversarial Networks (GANs) to

enhance autonomous vehicle navigation by translating images from one viewpoint to another. The authors propose a framework based on the Pix2Pix model, incorporating a perceptual loss to improve image quality. This approach is fine-tuned alongside a pre-trained imitation learning model that outputs driving controls, ensuring the synthesized images are not only visually accurate but also functional for driving tasks. The results indicate that conditional GANs can successfully handle cross-view image translation without relying on auxiliary information from the target domain, highlighting the potential for using alternate camera feeds to compensate for missing data in autonomous vehicles. This work suggests broader applications, such as addressing wider camera angle discrepancies or generating ground views from aerial perspectives, to enhance autonomous systems.

Pix2Pix model has been a recent choice for studies and experiments involving image-to-image tasks across different domains. The paper [35], presents enhancements to the traditional Pix2Pix model to address its limitations in image generation tasks. The authors introduce two significant modifications: the integration of a U-Net++ network as the generator and the addition of a differential image discriminator. These enhancements aim to reduce information loss during the encoding and decoding processes and to impose stronger constraints during image generation. The U-Net++ network, with its denser skip connections, helps retain more detailed information, resulting in clearer and more detailed images. The differential image discriminator is designed to focus on the differences between generated and real images, thereby improving the quality of the generated outputs. The paper demonstrates the effectiveness of these improvements using datasets such as facades and sketch portraits. The experimental results show that the modified Pix2Pix model surpasses the original model and the CycleGAN in terms of image quality and clarity. Specifically, the model with the U-Net++ generator shows significant improvement in the Fréchet Inception Distance (FID) metric, while the model with the differential image discriminator excels in the Inception Score (IS) and Structural Similarity Index (SSIM) metrics. The joint model, which combines both enhancements, provides balanced improvements across all metrics. Overall, this study highlights the potential of using advanced network architectures and specialized discriminators to enhance the performance of image-to-image translation models, making them more effective for tasks requiring high-quality image generation. The paper [6] explores the application of the Pix2Pix model in generating synthetic corneal tomography images. This research is particularly focused on addressing challenges associated with small datasets and class imbalance, which are common issues in medical imaging. The study utilizes the Pix2Pix conditional GAN (cGAN) to create color-coded Scheimpflug camera corneal tomography images. The dataset comprises images that were labeled and preprocessed for training the Pix2Pix cGAN. The quality of the generated images was evaluated using metrics such as the Fréchet Inception Distance (FID), mean square error, structural similarity index, and peak signal-to-noise ratio. The results demonstrated that the synthesized images were of high quality, both subjectively and objectively. When these synthetic images were used alongside real images to train a deep convolutional neural network (DCNN) for classifying keratoconus, the performance surpassed that of models trained solely on original images or those augmented through traditional methods. This study underscores the potential of Pix2Pix cGANs in generating high-quality synthetic images for medical applications, thereby

---

### 3. Related Work

overcoming limitations posed by small datasets and imbalanced classes. The approach provides a scalable solution for augmenting datasets, which can be beneficial for both experimental and clinical applications in ophthalmology and potentially other medical fields.

# CHAPTER 4

---

## Methodology

---

### 4.1 Dataset

We use The Berkeley Single Cell Computational Microscopy Dataset (BSCCM) for our research [26]. This dataset contains images of white blood cells categorized based on the surface proteins present in their cell membranes. The BSCCM dataset contains approximately 400,000 images of white blood cells which are organized based on their type of illumination such as LED array and fluorescence light. The LED array based illumination patterns enable label-free imaging of cells while the fluorescence light was used to capture different features of cells under 6 different fluorescent channels. The different illumination patterns based on LED array are: brightfield, darkfield, and differential phase contrast. The fluorescence lamp illuminated cell images contain one of the following three properties: single antibody at a time, all antibodies together, and no antibodies. For our research purpose, we consider only fluorescence images with no antibodies. The WBCs in this dataset are also classified into two based on their labels. The first is the classification of WBCs into lymphocytes, monocytes, and granulocytes. The second where the cells are further classified into 10 subtypes of the previous 3 classes, plus red blood cells, and 2 cell types which are unclear. Figure 4.1 shows brightfield images and Figure 4.2 shows target fluorescence images of white blood cells. The details of the sub-dataset that we prepare by extracting images from the BSCCM Dataset are given in Table 4.1.

| Input                             | Target                             | Train Dataset | Test Dataset |
|-----------------------------------|------------------------------------|---------------|--------------|
| Brightfield                       | Fluorescent Images with 6 channels | 2000          | 500          |
| Fluorescent Images with 1 channel | Fluorescent Images with 6 channels | 2000          | 500          |

Table 4.1: Dataset

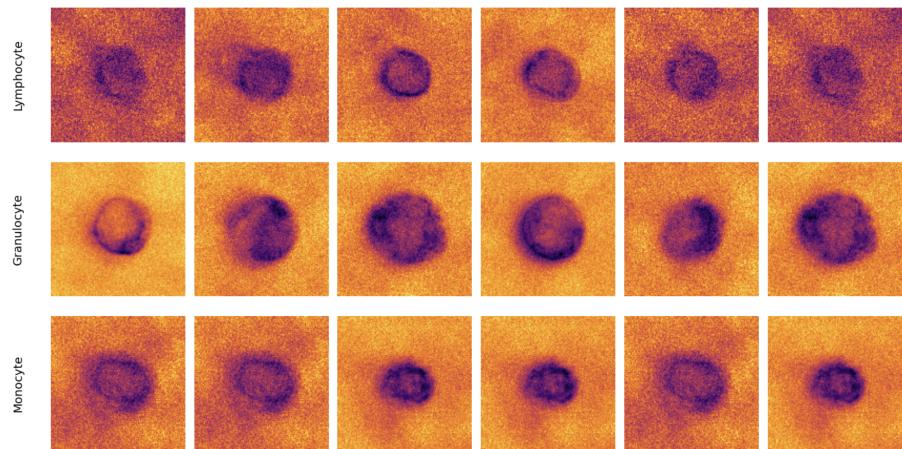


Figure 4.1: Brightfield Images of WBCs

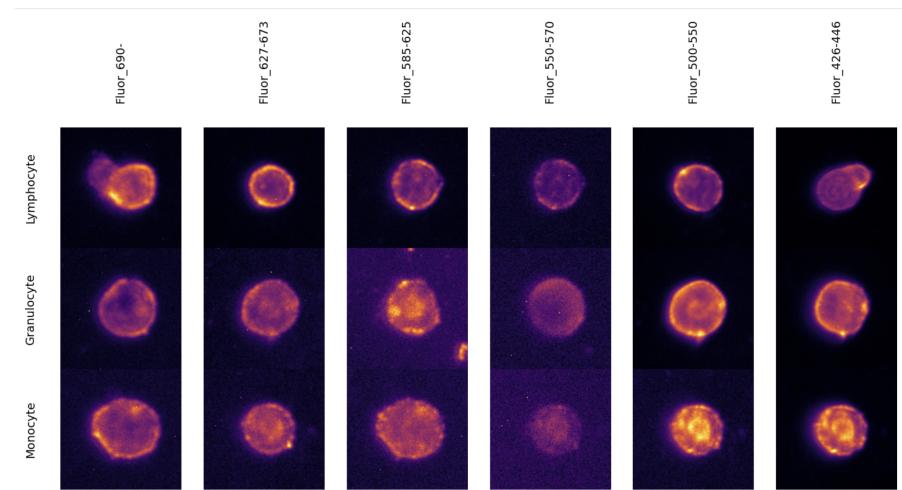


Figure 4.2: Fluorescence Images of WBCs

## 4.2 Data Preparation

The images in the dataset were originally of ‘uint16’ datatype. Also, they follow the image format ‘NCHW’, where N, C, H, and W represents the batch size, channels, height, and width respectively. This data is converted and stored in a new file of ‘.npz’ format (numpy arrays) for future convenience as data stored as numpy arrays are stored in a compressed format and hence optimises storage. Further more, images stored as numpy arrays can be easily retrieved without any change to the images’ original properties. We use tensorflow and keras frameworks, which by default use ‘NHWC’ (channels last) format for images. Hence we change the format of our data from ‘NCHW’ to ‘NHWC’ to avoid discrepancies. The pixel values in different channels have a huge range difference. Hence, the images are also normalized using min-max normalization to scale pixel values between 0 and 1 while preserving their relative difference in pixel values.

## 4.3 Proposed Architecture

We use a conditional generative adversarial network (cGAN) architecture for the prediction of fluorescent channels. The BSCCM dataset we use, provide us with paired images, that is, the target fluorescent images for the brightfield images are also available. The Pix2Pix model is the most popular cGAN model for image-to-image translation tasks when paired input images are available (CycleGAN was suggested as the best model for unsupervised learning purposes). We try to adapt the existing Pix2Pix model for our use case. The U-Net model is used as the Generator and PatchGAN model is used as the discriminator. We used the Pix2Pix model to generate fluorescent images with 6 channels as target for 2 different input use cases which are explained in section 4.4.1 and 4.4.2 respectively. The pipeline of the model for both the use cases are given in Figure 4.7 and Figure 4.8 respectively. Each model is trained from scratch without any pre-trained parameters, for 50 epochs on 1000 paired input images. A fresh set of 500 images are used for validation. The training was done using GPU resources allotted by the University.

### 4.3.1 Generator: U-Net

The U-Net architecture serves as an effective generator, leveraging its strengths in image segmentation to produce high-quality output images conditioned on input data. The U-Net’s encoder-decoder structure, combined with skip connections, is particularly well-suited for generating images that maintain both global context and fine details, crucial for tasks requiring high fidelity and accuracy.

The encoder path captures essential features from the input image, which serves as the condition for the cGAN. It consists of convolutional layers followed by max-pooling operations, which systematically reduce spatial dimensions while increasing feature depth. This path is critical for extracting hierarchical features that inform the generator about the input image’s structure and content. At the core of the U-Net, the bottleneck layer acts as a bridge between the encoder and decoder paths. It processes the condensed feature maps, preparing them for reconstruction in the decoder path. The decoder

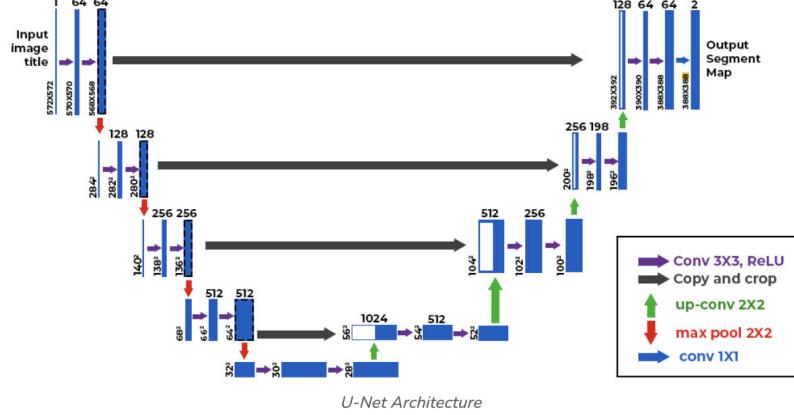


Figure 4.3: U-Net Encoder Decoder Architecture

path reconstructs the image from the encoded features, using upsampling layers to increase spatial dimensions. This path mirrors the encoder, with transposed convolutions facilitating the up-sampling process. The decoder path's role is to generate an output image that is conditioned on the input, ensuring that the generated image aligns with the given condition.

Skip connections are integral to the U-Net's effectiveness as a generator in cGANs. They connect corresponding layers in the encoder and decoder, allowing high resolution features to be directly transferred to the decoder. These connections help preserve spatial details and ensure that the generated images maintain structural integrity and detail, reducing likelihood of artifacts. The final layer of the U-Net generator is a one-by-one convolution that outputs the generated image. This layer ensures that the output matches with the desired dimensions and characteristics of the target image.

Using U-Net as a generator in cGANs is advantageous because it can produce images that are not only visually realistic but also adhere closely to the conditions specified by the input data. This capability is particularly valuable in applications like medical imaging, where maintaining the accuracy and detail of anatomical structures is crucial. The U-Net's architecture allows the GAN to generate high-quality images that are both contextually relevant and detailed, making it an ideal choice for tasks requiring precise image-to-image translation. Figure 4.3 shows the general architecture of U-Net model [4].

#### 4.3.2 Discriminator: PatchGAN

The PatchGAN discriminator is an integral component of the conditional GAN, particularly in models like Pix2Pix, where it plays a crucial role in distinguishing between real and generated images. Unlike traditional discriminators that evaluate an entire image to determine its authenticity, the PatchGAN discriminator focuses on smaller patches within the image. This approach allows it to effectively capture local texture details, making it particularly adept at tasks where preserving fine-grained features is essential.

The PatchGAN discriminator divides the input image into smaller patches,

of size  $N \times N$  (in our case  $8 \times 8$ ), and evaluates each patch independently. This design choice enables the discriminator to focus on local patterns and textures rather than the global structure of the image. Each patch is assessed to determine whether it is real or fake, and the results are averaged to produce an overall authenticity score for the entire image.

The architecture consists of several convolutional layers that process each patch. These layers are responsible for extracting features that help in distinguishing real patches from fake ones. The convolutional layers are designed to have a receptive field that matches the size of the patches, ensuring that each output of the network corresponds to a specific patch in the input image. By focusing on local patches, PatchGAN can effectively model the image as a Markov random field, where each patch is considered independently. This approach is particularly useful for capturing high-frequency details such as edges and textures, which are often critical in image-to-image translation tasks. The smaller patch size reduces the number of parameters in the discriminator, making it computationally efficient and faster to train compared to a full-image discriminator.

In the context of cGANs, the PatchGAN discriminator is conditioned on both the input and the generated images. It evaluates whether the patches in the generated image align with the expected patterns based on the input image. This configuration is particularly effective in scenarios where the generator is tasked with transforming an input image into a corresponding output image, such as in style transfer or image super-resolution.

The PatchGAN discriminator is implemented as a deep convolutional neural network, with layers configured to ensure that the effective receptive field corresponds to the desired patch size. During training, the discriminator is optimized to correctly classify patches as real or fake, while the generator is trained to produce patches that deceive the discriminator. This adversarial training process helps improve the quality of the generated images. The PatchGAN discriminator's focus on local patches makes it a powerful tool for tasks that require detailed texture and pattern recognition. Its ability to efficiently evaluate smaller sections of an image allows it to maintain high performance while reducing computational complexity. This makes it particularly well-suited for use in conditional GANs, where maintaining fidelity of local image features is crucial for generating realistic and high-quality images. Figure 4.4 shows PatchGAN discriminator [11].

## 4.4 Input to Models

### 4.4.1 Case 1: Brightfield Images as Input

We feed our model with simple brightfield images as input to predict the 6 target fluorescent channels. Figure 4.1 shows the brightfield images used as input. This is the simplest use case of virtual staining where we give images of cells as they are visible to the naked eye without any staining and try to get target images in the form of fluorescent cell images.

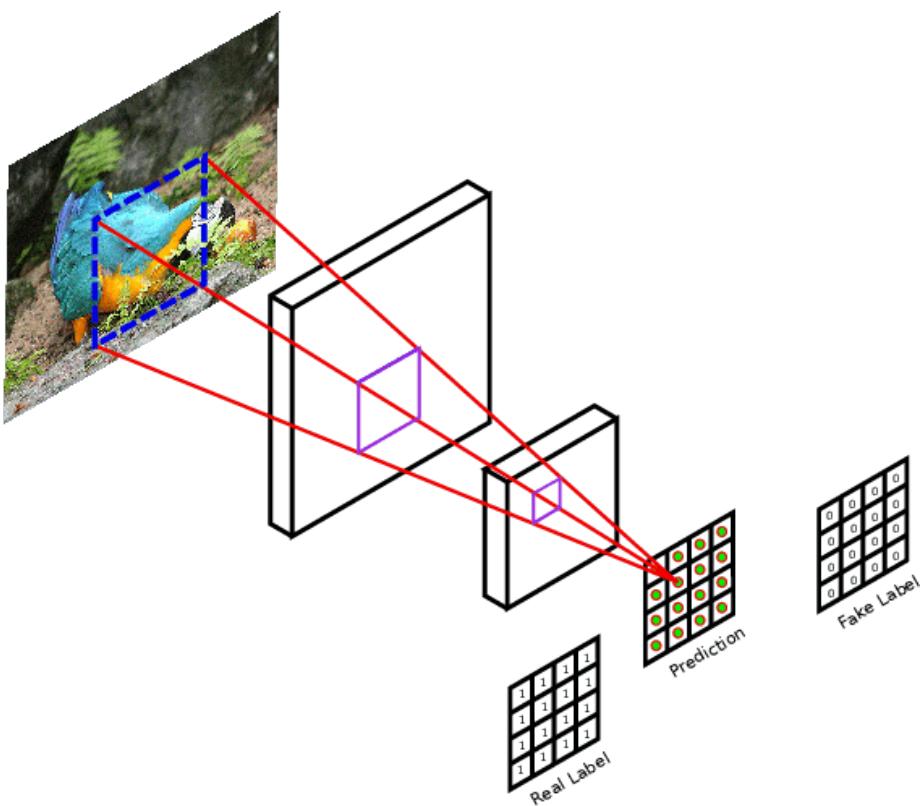


Figure 4.4: PatchGAN Discriminator

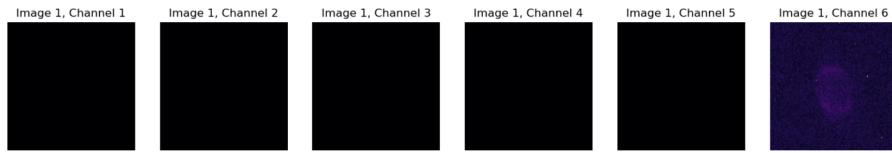


Figure 4.5: Input with 1 channel visible and 5 channels masked

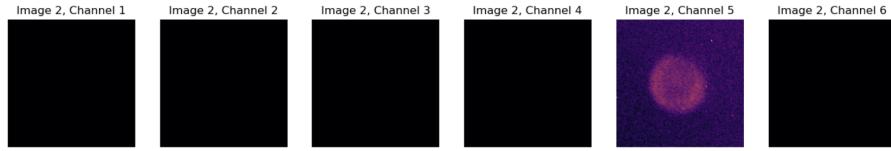


Figure 4.6: Input with 1 channel visible and 5 channels masked

#### 4.4.2 Case 2: Masked Fluorescence Images as Input

We give fluorescent images as input, where only 1 channel is visible in an image at a time, while other 5 fluorescent channels are kept masked. The choice of channels to be masked is random, meaning different combination of channels would be masked in different input images. In this way, the model gets to see all channels independently as part of different images while training and tries to generate targets where all 6 channels are predicted for the input images. Figure 4.5 and 4.6 give 2 examples of input in the second use case. Figure 4.5 shows an input image where only channel 6 is visible to the model while the other 5 fluorescence channels remain masked. Figure 4.6 shows another input image to the model where channel 5 is visible and the other 5 channels remain masked. This use case becomes beneficial in scenarios where we stain different subsets of cells with different fluorescent stains to visualize different combinations of proteins and other structural features. Preparing a dataset with images of all the cells would result in a similar condition where different input images would be stained with different fluorescent dyes and the model can learn to predict all fluorescent channels of each cell irrespective of the input fluorescent channel visible for the cell.

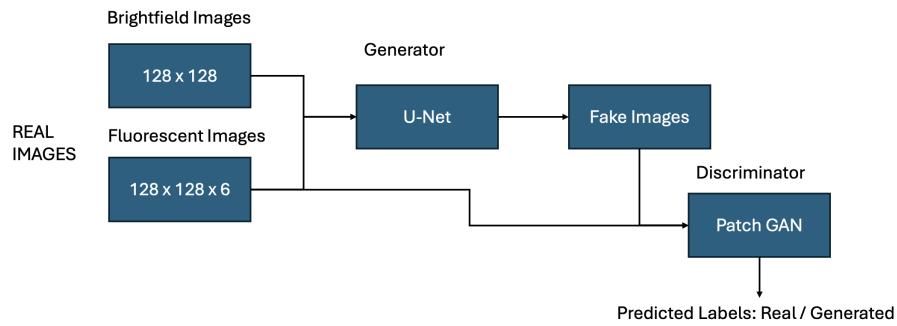


Figure 4.7: Case 1: Brightfield Images as Input

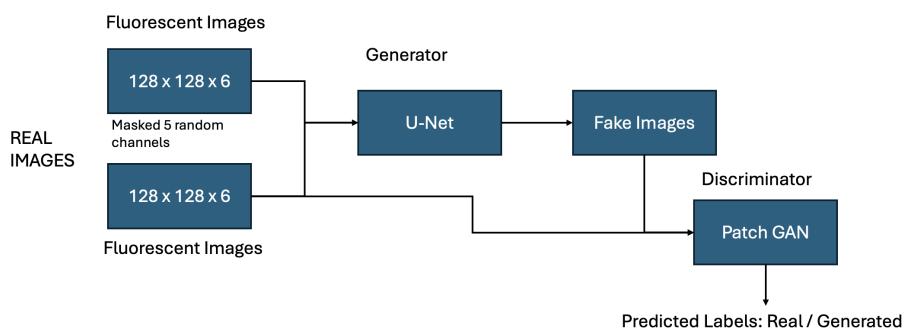


Figure 4.8: Case 2: Masked Fluorescent Images as Input

# CHAPTER 5

---

## Results

---

A popular technique for assessing the perceived quality of digital photos and movies is the Structural Similarity Index Metric (SSIM), especially when comparing two images. It is particularly pertinent for assessing conditional Generative Adversarial Networks (cGANs), such as Pix2Pix, since fidelity to the reference image and image quality are crucial. SSIM is a perception-based model that incorporates key perceptual phenomena including brightness and contrast masking and views image degradation as perceived changes in structural information. This makes it a better option than more conventional measures that only evaluate absolute mistakes, such as Mean Squared Error (MSE) or Peak Signal-to-Noise Ratio (PSNR).

Three primary factors are evaluated by SSIM in order to determine how similar two images are: brightness, contrast, and structure. Slide-across local windows ( $X^*X$ ) are used to calculate these components. The index value is a decimal value between -1 and 1, where 0 denotes no similarity, 1 denotes perfect structural similarity, and -1 denotes perfect anti-correlation.

As a cGAN model created to convert an input image into an equivalent output image, Pix2Pix needs to be evaluated by comparing the generated images' realism and quality to the ground truth. Because it gives an indication of how well the generated image retains the structural information of the reference image, SSIM is very helpful in this situation.

SSIM is a trustworthy statistic for assessing image quality from a human viewpoint since it closely matches human visual perception. It takes into account structural information, which is essential for comprehending the texture and content of images. SSIM is less sensitive to uniform changes in brightness or contrast than Mean Squared Error or Peak Signal-to-Noise Ratio. Instead, it focuses on structural changes that are more perceptually relevant. SSIM can produce a comprehensive quality map of the image using local windows, indicating locations in which the generated image deviates from the reference.

Although SSIM can withstand some distortions, it might not work well on photos with a lot of noise or blur since these can distort structural similarity.

| Input                             | Training | Validation |
|-----------------------------------|----------|------------|
| Brightfield Images                | 0.1012   | 0.2895     |
| Fluorescent Images with 1 channel | 0.1066   | 0.2841     |

Table 5.1: SSIM Loss

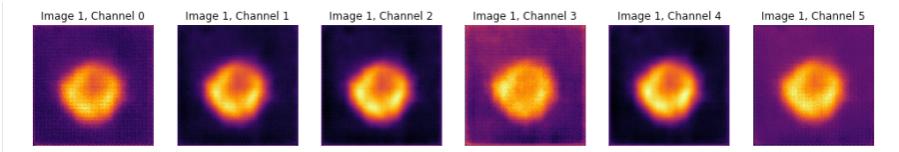


Figure 5.1: Generated Images - Training

Because of the sliding window method, calculating SSIM can be computationally demanding, particularly for big photos or video sequences.

SSIM can be used as a loss function in Pix2Pix model training to direct the generator to produce images that are structurally similar to the target image. To promote pixel-wise similarity, the SSIM loss is commonly used with other loss functions, such as the L1 loss and the adversarial loss from the GAN framework. The definition of the SSIM loss is as follows:

$$SSIMLoss = 1 - SSIM(x, y) \quad (5.1)$$

By encouraging the generator to maximise the SSIM index, this formulation results in images that have a high degree of structural resemblance to the reference. The model can concentrate on maintaining significant structural elements by integrating SSIM loss, producing more realistic and aesthetically pleasing outcomes.

A series of created photos and the matching ground truth images are used to compute SSIM scores. Better performance is indicated by higher SSIM scores, which imply that the generated images' structural content is quite similar to that of the reference images. Examining the generated images visually is also essential. The realism and authenticity of the images can be evaluated by humans, who might offer insights that quantitative measurements would overlook. Table 5.1 provides the SSIM loss of our Pix2Pix model for both training and validation. Figures 5.1 and 5.2, respectively, show the first output fluorescence image produced by the generator for the training and validation datasets.

From the Figures 5.1 and 5.2, note that Channels 1, 2, and 4 show more similar traits while Channels 0, 3, and 5 seem more similar to each other. This property also holds true for the target images that were given to the generator. The visual examination of results combined with SSIM loss for training and validation indicate that the Pix2Pix model is an ideal choice for this dataset. The number of images used for training the model were limited by the storage capacity. cGANs in general are observed to give better results with more data. Therefore, it can be safely assumed that this model would give better results with a bigger dataset and training capacity.

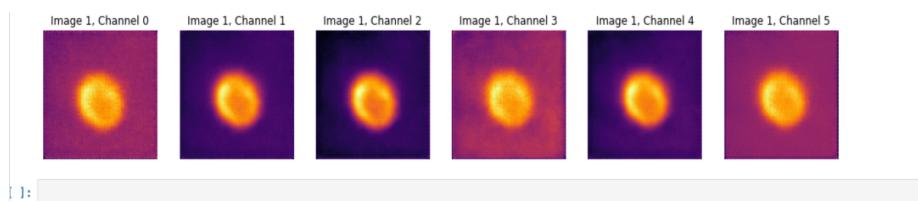


Figure 5.2: Generated Images - Validation

# CHAPTER 6

---

## Discussion and Future Work

---

Artificial intelligence (AI) has rapidly become one of the technologies of the twenty-first century that will have the largest influence on a wide range of industries, including marketing and healthcare. The integration of medical imaging and diagnostics, particularly with computer vision (CV) and machine learning algorithms, has revolutionised the field of healthcare. The aforementioned advancements have presented novel prospects for enhancing patient care, treatment effectiveness, and diagnostic precision. The current work emphasises on staining virtual cell images using conditional Generative Adversarial Networks (cGANs) inside this framework. Traditional staining methods, while effective, often need chemical reactions that can be costly, time-consuming, and even dangerous for cell samples. Conversely, cGAN-based virtual staining offers a non-invasive alternative that preserves living cells' integrity while producing thorough visual insights. This approach reduces the risks associated with chemical stains and allows for repeated studies, making it particularly helpful for longitudinal study. The project's objective is to demonstrate how cGANs can generate virtually stained, high-quality images from inputs from brightfield microscopy, providing a practical and affordable means of diagnosing medical issues. By employing artificial intelligence, this research contributes to the ongoing advancement of medical imaging technologies by highlighting their potential to improve patient outcomes and streamline healthcare processes.

The aforementioned study has significantly advanced the field of medical imaging by demonstrating the capacity of conditional Generative Adversarial Networks (cGANs), and specifically the Pix2Pix model, to produce high-quality fluorescent images from brightfield shots. This achievement is closely linked to the primary goal of the study, which is to enhance image quality without the need for messy staining techniques. By achieving this objective, the study has opened up new possibilities for the application of artificial intelligence in medical imaging, especially in pathology.

The primary objective of the study was to examine the potential of cGANs to improve the quality of medical images. Pix2Pix, a type of cGAN, was utilised

to generate high-fidelity fluorescent images from brightfield photos. This conversion process is crucial because it eliminates the need for traditional staining methods, which can be costly and time-consuming. The study's outcomes lend credence to the research themes by demonstrating how AI may effectively and practically improve image quality. This alignment not only validates the research concept but also shows how AI technology can revolutionise medical imaging methods.

This finding broadens our understanding of the use of AI in medical imaging and has significant theoretical implications. The study supports the idea that by employing deep learning techniques, AI can increase the accuracy and efficiency of diagnostic processes. This advancement is particularly significant in the field of pathology, where precise image interpretation is essential. The findings suggest that by providing more accurate and efficient picture analysis, the incorporation of AI technologies—like cGANs—into traditional diagnostic workflows has the potential to completely transform current practices. Better diagnostic outcomes and a deeper understanding of many medical conditions could come from this modification.

The practical applications of the research are equally important. One of the most notable benefits is the potential for virtual staining, which reduces the costs and risks associated with chemical staining procedures. Chemicals are used in conventional staining, which can be hazardous and requires careful handling and disposal. By eliminating these chemicals, virtual staining offers a safer and more affordable alternative. This method also allows samples to be examined again without losing quality, which is very helpful for long-term study. The ability to do several tests on the same sample without compromising its quality is very advantageous in clinical settings, where sample integrity is very important. Furthermore, the use of AI-driven virtual staining techniques may lead to more productive workflows in clinical laboratories. These techniques reduce the amount of time and materials required for traditional staining treatments, which can increase laboratory productivity. This efficiency is essential in high-volume applications where fast turnaround times are essential. Moreover, the reduced reliance on chemical stains should yield economic advantages for healthcare facilities, making these technologies a financially viable option for widespread adoption.

To improve image quality without physical staining, the research shows how cGANs, specifically the Pix2Pix model, can efficiently synthesise high-quality fluorescent images from brightfield inputs. The theoretical implications show how AI can increase diagnostic efficiency and accuracy, while the practical implications show significant cost savings and operational advantages in clinical settings. As the field advances, the use of AI in medical imaging operations offers the potential to improve patient outcomes and advance diagnostic approaches.

The study explores the application of conditional Generative Adversarial Networks (cGANs) in medical imaging, with an emphasis on virtual staining techniques. This result is in line with earlier studies that demonstrate the effectiveness of AI models, such cGANs, for translating image domains. Previous research has primarily focused on style transfer and picture improvement; this study extends these findings to medical imaging. The work uses cGANs to virtual staining, challenging traditional staining methods that often require extensive manual processing and can be costly and time-consuming. This opens up new possibilities for non-invasive imaging.

This paper constitutes a significant contribution by providing an in-depth analysis of cGANs in the context of virtual staining, which closes a gap in the literature. Although they are thought to be the best, traditional staining procedures have limitations, such as the need for specific tools and trained personnel, which makes them less useful in settings with little funding. With the potential to revolutionise medical diagnostics, deep learning-powered virtual staining offers a rapid and cost-effective replacement that can generate histological stains from label-free images.

This work not only establishes the viability of using cGANs for virtual staining, but also shows how artificial intelligence (AI) may progress and improve medical diagnoses. By offering a non-invasive, efficient technique for tissue analysis, our work paves the way for future advancements in medical imaging, particularly with regard to enhancing diagnostic capabilities in remote and resource-constrained settings.

It appears that the model constructed and evaluated in this study makes use of the BSCCM dataset. It has some flaws and potential for improvement, just like any other scientific study. Below, we look at these restrictions, how they impacted the study’s conclusions, and potential directions for future investigation.

One of the primary limitations of the study is its reliance on the BSCCM dataset. This dataset may not contain all conceivable cell types that physicians may see in practice, despite its value. The model might not have a solid enough foundation if the dataset doesn’t accurately represent the large range of cell types found in real-world scenarios. Because of this, the model’s applicability to cell types not represented in the dataset may be restricted. This limitation highlights the significance of utilising diverse datasets to ensure the broad applicability of the model.

An further limitation concerns potential variations in the model’s performance when applied to different datasets or imaging circumstances. The model may perform best under the specific conditions and dataset that it was trained on, but it may lose its usefulness if these things change. This variation casts question on the model’s generalisability, or its ability to function effectively in a range of settings. A model should exhibit robustness across several datasets and imaging conditions in order to be truly valuable in clinical settings. The aforementioned limitations have a substantial impact on the study’s conclusions. The model might yield promising results in controlled settings, but it’s not obvious if it can be used with other datasets or in real-world scenarios. Because the model depends on a single dataset and specific imaging settings, its performance could not translate well to other scenarios. This limitation suggests that although the current results are encouraging, caution should be exercised in interpreting them. The true test of the model’s efficacy will be to apply it to a range of datasets and scenarios. Future work should focus on a few key areas to address these limitations and enhance the model’s usefulness. Applying the model to a wider variety of datasets is one crucial area that needs more investigation. By testing the model with various cell types and imaging modalities, researchers may more precisely assess the model’s resilience and generalisability. This approach will help identify any areas in which the model exhibits subpar performance and provide recommendations for improving its ability to handle a greater range of scenarios. Research on other state-of-the-art AI models, such CycleGANs, can be effectively pursued by integrating them. These

## 6. Discussion and Future Work

models have the potential to enhance the model's ability to learn from a range of unlabelled data sets and offer valuable insights into unsupervised learning techniques. By combining various models, researchers can look at new ways to improve the model's usefulness and functionality. For more developments, the model's architecture must be improved. Two aspects of the model that need to be addressed for practical application are noise handling and image quality. By strengthening the model's architecture, researchers can increase the model's robustness and ensure that it produces high-quality results even in challenging situations. Finally, increasing the number of loss functions that give structural integrity priority might further enhance the quality of the images that are generated. These loss functions would contribute to improving the accuracy and reliability of the outputs by guaranteeing that the model maintains the essential structural properties of the images it processes. In conclusion, even though the study presents a workable model, it is important to acknowledge its limitations and how they could impact the results. Future research should aim to overcome these limitations by combining advanced AI models, examining a range of datasets, and improving the model's architecture and loss functions. Scientists may ensure that the model can be applied in a range of clinical scenarios and increase its resilience by doing this.

# CHAPTER 7

---

## Conclusion

---

The use of conditional Generative Adversarial Networks (cGANs) for virtual staining of microscopic images in this research effort has advanced the field of medical imaging significantly. We have shown that it is feasible to produce high-quality fluorescent images from brightfield microscopy images without the use of conventional staining techniques by utilising the capabilities of deep learning, especially the Pix2Pix model.

Our study validates the performance of Pix2Pix model on the newly published BSCCM dataset and also experiments with a new type of input, fluorescent images with random channels that generate images with all possible fluorescent channels that were given as input to the model. Even though the conventional method of staining using chemical dyes work well, it has a number of disadvantages, such as the need for potentially hazardous chemicals, labour-intensive steps, and the inability to reuse samples for different investigations. The development of a cGAN-based virtual staining technique, creates new avenue for safe, reliable, and quick cellular imaging.

The model used, Pix2Pix seems to give encouraging results with respect to image quality and structural similarity. An an assessment indicator, the Structural Similarity Index (SSIM) offered insightful information on our model's functionality. The generated images retain a high degree of structural resemblance to the ground truth, as indicated by the SSIM scores obtained for both training and validation datasets. This suggests that our approach can yield accurate and dependable virtual stains.

The adaptability of our method is one of its main advantages. Creating fluorescent images from brightfield inputs and predicting all fluorescent channels from partially masked fluorescent images are two different use cases that we investigated. This dual strategy shows how our model may be tailored to diverse imaging conditions, which could increase its usefulness in a range of clinical and research contexts.

The ramifications of this study go beyond the field of microscopy itself. Our approach has the potential to greatly improve laboratory operations, lower

expenses, and have a less negative impact on the environment by minimising the need for chemical stains. Virtual Staining's non-destructive nature also makes it possible to analyse the same sample again, creating new opportunities for longitudinal research and in-depth cellular analyses.

It is imperative to recognise the constraints of our research, nevertheless. While our research has a strong foundation thanks to its reliance on a single dataset, the Berkeley Single Cell Computational Microscopy Dataset (BSCCM), there is a chance that our findings won't be as applicable to different cell types or image settings. To ensure model's robustness across a variety of cellular samples and imaging modalities, future work should concentrate on broadening the range of datasets utilised for training and validation.

Furthermore, even if our findings seem encouraging, more research is required to completely comprehend how well the model functions in actual clinical settings. The shift from controlled research settings to real-world applications frequently reveals unanticipated difficulties, thus comprehensive testing in a variety of clinical settings will be necessary to confirm the accuracy and dependability of our virtual staining method.

There are a number of fascinating directions this field has to go in terms of study and development. Investigating more complex GAN designs, such CycleGANs, is one possible path that can enhance the model's capacity to manage unpaired datasets. When coupled brightfield and fluorescent photos are not easily accessible, this might be especially helpful. Future work should also focus on improving the model's architecture to better withstand noise and maintain intricate structural elements. This can entail integrating attention mechanisms to concentrate on important cellular properties or experimenting with various loss functions that prioritise structural integrity.

There are also exciting prospects for integrating our virtual staining process with other cutting-edge medical imaging and diagnostic technologies. For example, our method in conjunction with automated cell categorisation algorithms may yield potent diagnostic instruments that can quickly detect and describe cellular anomalies. Moreover, this method has potential uses outside of conventional microscopy. Our virtual staining technique may be essential to improving the quality and interpretability of digital tissue samples as digital pathology gains more popularity. Wide-ranging effects may result from this for telemedicine and remote diagnosis, especially in places with low resources where access to sophisticated staining methods may be restricted.

It is important to think about how this technology may affect ethics. It will be essential to develop policies and standards for the use of AI-generated images in clinical practice and research as they grow more and more similar to images created using conventional techniques. To preserve confidence in scientific results and medical diagnoses, it will be crucial to maintain strict validation procedures and to guarantee transparency regarding the source of pictures. To sum up, our findings mark a substantial advancement in the use of artificial intelligence in medical imaging. We have aided in the continuing technological transformation in healthcare by proving that cGANs may be used for virtual staining. Our research creates new avenues for cellular investigation and medical diagnostics in addition to providing a workable answer to the drawbacks of conventional staining techniques.

This technique has a wide range of possible applications, from increased productivity in research labs to better diagnostics in medical settings. We are

getting closer to a day when sophisticated imaging methods will be more widely available, reasonably priced, and ecologically benign as we keep improving and developing this strategy. To fully utilise this technology, biologists, computer scientists, and medical specialists will need to continue working together. It will be imperative to carry out more model optimisation, dataset expansion, and thorough testing of the technology in various real-world circumstances. To guarantee the successful integration of AI-generated images into clinical practice, efforts should also be undertaken to educate and train healthcare personnel in the use and interpretation of these images. It's obvious that innovation will continue to flourish at the nexus of medical imaging and artificial intelligence in the years to come. Our work on virtual staining using cGANs is just one aspect of this fascinating field. We can strive towards a future where sophisticated diagnostic technologies are more widely available, practical, and efficient and a future that will eventually improve patient outcomes and further our understanding of cellular biology by continuing to push the envelope of what artificial intelligence in healthcare is capable of.

Our research advances the larger objective of personalised medicine in the great picture of medical advancement. More advanced virtual staining methods may make it possible to study cellular structures in a more detailed and personalised manner, which could result in more focused interventions and therapies. In the end, how well this technology can expand scientific understanding and enhance patient care will determine how successful it is. Even though there is still more to be done, our study's findings offer a solid basis for upcoming advancements in this fascinating and quickly developing sector. We are getting closer to a time when cutting-edge imaging technologies are a standard component of medical care, helping both patients and healthcare professionals. This is because we are still improving and developing these techniques.

---

## Bibliography

---

- [1] Bright-field microscopy - Wikipedia.
- [2] Generative AI Synthetic Data Techniques: Leverage the Power.
- [3] Transforming healthcare with AI: The impact on the workforce and organizations — McKinsey.
- [4] U-Net Architecture Explained - GeeksforGeeks.
- [5] COLORIZING BLACK AND WHITE IMAGES USING CONDITIONAL GAN. *International Research Journal of Modernization in Engineering Technology and Science*, 4 2023.
- [6] H. Abdelmotaal, A. A. Abdou, A. F. Omar, D. M. El-Sebaity, and K. Abdelazeem. Pix2pix Conditional Generative Adversarial Networks for Scheimpflug Camera Color-Coded Corneal Tomography Image Generation. *Translational Vision Science & Technology*, 10(7), 6 2021.
- [7] B. Bai, X. Yang, Y. Li, Y. Zhang, N. Pillar, and A. Ozcan. Deep learning-enabled virtual histological staining of biological samples, 12 2023.
- [8] J. Bajwa, U. Munir, A. Nori, and B. Williams. Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthcare Journal*, 8(2):e188, 7 2021.
- [9] C. A. Combs. Fluorescence Microscopy: A Concise Guide to Current Imaging Methods. *Current protocols in neuroscience / editorial board, Jacqueline N. Crawley ... [et al.]*, 0 2(SUPPL.50):Unit2.1, 2010.
- [10] P. Coscia, A. Genovese, F. Scotti, and V. Piuri. Applications and Limits of Image-to-Image Translation Models. Technical report, 2023.
- [11] U. Demir, G. Unal, and U. Demir. Patch-Based Image Inpainting with Generative Adversarial Networks Vision and Imaging View project Shape Representations View project Patch-Based Image Inpainting with Generative Adversarial Networks. 2018.

- [12] K. Dobler, F. H. Ubscher, J. Westphal, A. Sierra-M ´ Unera, G. De Melo, and R. Krestel. Art Creation with Multi-Conditional StyleGANs. Technical report, 2022.
- [13] A. Esteva, K. Chou, S. Yeung, N. Naik, A. Madani, A. Mottaghi, Y. Liu, E. Topol, J. Dean, and R. Socher. Deep learning-enabled medical computer vision. *npj Digital Medicine* 2021 4:1, 4(1):1–9, 1 2021.
- [14] M. J. Fanous, N. Pillar, and A. Ozcan. Digital staining facilitates biomedical microscopy. *Frontiers in Bioengineering*, 3, 2023.
- [15] S. M. Hickey, B. Ung, C. Bader, R. Brooks, J. Lazniewska, I. R. Johnson, A. Sorvina, J. Logan, C. Martini, C. R. Moore, L. Karageorgos, M. J. Sweetman, and D. A. Brooks. Fluorescence microscopy—an outline of hardware, biological handling, and fluorophore considerations. *Cells*, 11(1), 1 2022.
- [16] H. Hoyez, C. Schockaert, J. Rambach, B. Mirbach, and D. Stricker. Unsupervised Image-to-Image Translation: A Review, 11 2022.
- [17] B. Huang, J. Li, B. Yao, Z. Yang, E. Y. Lam, J. Zhang, W. Yan, and J. Qu. Enhancing image resolution of confocal fluorescence microscopy with deep learning. *PhotoniX*, 4(1), 12 2023.
- [18] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, and B. A. Research. Image-to-Image Translation with Conditional Adversarial Networks. Technical report.
- [19] L. R. Kennedy-Metz, P. Mascagni, A. Torralba, R. D. Dias, P. Perona, J. A. Shah, N. Padov, and M. A. Zenati. Computer Vision in the Operating Room: Opportunities and Caveats. *IEEE transactions on medical robotics and bionics*, 3(1):2, 2 2021.
- [20] L. Kreiss, S. Jiang, X. Li, S. Xu, K. C. Zhou, K. C. Lee, A. Mühlberg, K. Kim, A. Chaware, M. Ando, L. Barisoni, S. A. Lee, G. Zheng, K. J. Lafata, O. Friedrich, and R. Horstmeyer. Digital staining in optical microscopy using deep learning - a review, 12 2023.
- [21] E. H. Lee, B. Park, M. H. Jeon, H. Jang, A. Kim, and S. Lee. Data augmentation using image translation for underwater sonar image segmentation. *PLoS ONE*, 17(8), 8 2022.
- [22] Y. Li, N. Pillar, J. Li, T. Liu, D. Wu, S. Sun, G. Ma, K. de Haan, L. Huang, Y. Zhang, S. Hamidi, A. Urisman, T. Keidar Haran, W. D. Wallace, J. E. Zuckerman, and A. Ozcan. Virtual histological staining of unlabeled autopsy tissue. *Nature Communications*, 15(1), 12 2024.
- [23] Y.-J. Li, L.-G. Zhang, H.-Y. Zhi, K.-H. Zhong, W.-Q. He, Y. Chen, Z.-Y. Yang, L. Chen, X.-H. Bai, X.-L. Qin, D.-F. Li, D.-D. Wang, J.-T. Gu, J.-L. Ning, K.-Z. Lu, J. Zhang, Z.-Y. Xia, Y.-W. Chen, and B. Yi. A better method for the dynamic, precise estimating of blood/haemoglobin loss based on deep learning of artificial intelligence. *Annals of Translational Medicine*, 8(19):1219–1219, 10 2020.

---

## 7. BIBLIOGRAPHY

- [24] A. Osokin, A. Chessel, R. E. C. Salas, and F. Vaggi. GANs for Biological Image Synthesis. 8 2017.
- [25] H. Park, M. Na, B. Kim, S. Park, K. H. Kim, S. Chang, and J. C. Ye. Deep learning enables reference-free isotropic super-resolution for volumetric fluorescence microscopy. *Nature Communications*, 13(1), 12 2022.
- [26] H. Pinkard, C. Liu, F. Nyatigo, D. A. Fletcher, and L. Waller. The Berkeley Single Cell Computational Microscopy (BSCCM) Dataset. 2 2024.
- [27] L. Pinto-Coelho. How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications. *Bioengineering*, 10(12), 12 2023.
- [28] A. Refaat, M. L. Yap, G. Pietersz, A. P. G. Walsh, J. Zeller, B. del Rosal, X. Wang, and K. Peter. In vivo fluorescence imaging: success in preclinical imaging paves the way for clinical applications. *Journal of Nanobiotechnology* 2022 20:1, 20(1):1–22, 10 2022.
- [29] Y. Rivenson, K. de Haan, W. D. Wallace, and A. Ozcan. Emerging Advances to Transform Histopathology Using Virtual Staining, 1 2020.
- [30] A. Sieroń, K. Sieroń-Stołtny, A. Kawczyk-Krupka, W. Latos, S. Kwiatek, D. Straszak, and A. M. Bugaj. The role of fluorescence diagnosis in clinical practice. *OncoTargets and therapy*, 6:977, 2013.
- [31] X. Tang. The role of artificial intelligence in medical imaging research. *BJR Open*, 2(1):20190031, 11 2020.
- [32] H. Wieslander, A. Gupta, E. Bergman, E. Hallström, and P. J. Harrison. Learning to see colours: Biologically relevant virtual staining for adipocyte cell images. *PLoS ONE*, 16(10 October), 10 2021.
- [33] X. Yi, E. Walia, and P. Babyn. Generative adversarial network in medical imaging: A review. *Medical Image Analysis*, 58:101552, 12 2019.
- [34] D. Zhang, B. Sanders, G. Byrd, F. Luo, V. Krovi, D. Gorsich, J. M. Smereka, and M. Brudnak. CROSS-VIEW IMAGE TRANSLATION USING CONDITIONAL GAN FOR AUTONOMOUS DRIVING. Technical report.
- [35] X. Zhao, H. Yu, and H. Bian. Image to Image Translation Based on Differential Image Pix2Pix Model. *Computers, Materials & Continua*, 77(1):181–198, 10 2023.
- [36] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. 3 2017.

# APPENDIX A

---

## Appendix

---

The code for the entire project can be found at : <https://git.cs.bham.ac.uk/projects-2023-24/axm2075>

Dataset : <https://github.com/Waller-Lab/BSCCMs>

Inorder to run the code you need not download the dataset. Data required have been downloaded and stored as a compressed numpy array.

The path to the files are as follows: BSCCM-main/bsccm/..

Case 1 : case1\_brightfield\_input\_images.ipynb

Case 2 : case2\_masked\_fluorescence\_input\_images.ipynb

the brightfield images are stored in : fluor\_images\_transposed\_asnumpy.npz  
the masked fluorescence images are stored in : masked\_dataset.npz

In case 1: Code for splitting images into input and target for generator are given.

Note: If you want to train the model from scratch, the code for cleaning the checkpoints have been added as a comment. Just uncomment and execute the same. Make sure to comment it back after execution.

In case 2: The input and target for generator are stored as modified\_data1.npz and modified\_data2.npz which gets called in the Data Generator. The code just needs to be executed from the start.

The comments are given in the jupyter notebook itself. While training the model, the weights of generator and discriminator are stored as checkpoints after each epoch to avoid training the model from scratch each time. Also, due to RAM limitations the training and validation dataset cannot be loaded at the same time. Hence for validating the model, the kernel need to be restarted after training the model and validation should be done by skipping the loading and executing of training data.