# **Generative Model for Imputing Imaging Mass Spectrometry from Serial Two-Photon Tomography**

Kevin Yang

BC Cancer

675 W 10th Ave, Vancouver, BC

[kyang@bccrc.ca](mailto:kyang@bccrc.ca)

**Abstract**

*Serial Two-Photon Tomography (STPT) and Imaging Mass Spectrometry (IMC) are two popular imaging techniques in tumour analysis. STPT images describe tumour morphology, whereas IMC images describe protein abundance. Having both data modalities for the same tissue sample is often beneficial in clinical applications, as understanding the tumour landscape from both a morphological and proteomic perspective can inform treatment decisions. However, it is difficult to obtain both data modalities for a single tissue sample. To mitigate this issue, we present a Generative Model that, for the same tissue sample, only requires STPT images to impute corresponding IMC images. 18 aligned STPT and IMC physical sections have been identified and have been used for preliminary training; we are currently reaching out to relevant personnel to obtain more data. A naïve model based on convolutional layers has been trained to predict IMC from STPT images with poor results.*

# **Background**

Cancer is a complex disease driven by genetic mutations. Originating from a single mutated cell, subsequent rounds of proliferation and additional mutations ultimately give rise to the tumour (Nowell, 1976). This evolutionary process fosters tumour heterogeneity, producing genetically distinct populations of cells known as clones. Notably, these clonal populations differ from one another in their genetic makeup and have varied responses to treatment. Therefore, investigating tumour heterogeneity can direct treatment decisions to combat metastasis and recurrence, ultimately improving prognosis.

However, cancer is not exclusively made up of malignant cells, but also of surrounding normal cells [(Junttila and de Sauvage, 2013)](https://www.nature.com/articles/nature12626). This mixture of cell types is called the tumour microenvironment (TME). Interactions between malignant and non-malignant cells within the TME can provide insight into the mechanisms that promote cancer growth [(Egeblad et al., 2010)](https://pubmed.ncbi.nlm.nih.gov/20627072/); this in turn can inform and advance drug development. To fully understand the TME, a comprehensive picture of the morphology and the molecular profile of the tumour must be obtained [(Heindl et al., 2015)](https://www.nature.com/articles/labinvest2014155#Abs1). The introduction of powerful imaging techniques has enabled detailed illustrations of morphology and protein abundance profiles for tumors [(Bressan et al., 2021).](https://www.biorxiv.org/content/10.1101/2021.06.28.448342v1.full.pdf) In particular, serial two-photon tomography (STPT) uncovers tumour morphology, while imaging mass spectrometry reveals protein abundance inside the TME. These two imaging techniques are introduced in the following two subsections.

## **Serial Two-Photon Tomography (STPT)**

STPT is an automated imaging technique that produces high-throughput and high-fidelity imaging by integrating two-photon microscopy and tissue sectioning (Taranda and Turcan, 2021). The general workflow is as follows. First, place the specimen on the XYZ stage under the objective of a two-photon microscope; second, image an optical section as a mosaic of fields of view; third, cut off a slice of tissue using a vibrating blade; repeat. Importantly, the imaged portion is tens to hundreds of microns below the surface of the tissue (Amato et al., 2016). As a result, the imaged tissue will remain in a pristine state, unaffected by deformations from mechanical sectioning. Moreover, each serial section is imaged before being sliced, allowing for near perfect alignment of imaged tissue sections. Finally, STPT is an automatic procedure which does not require human intervention, greatly mitigating labor costs (Ragan et al., 2012). After imaging is completed, algorithms can reconstruct a 3D image from the 2D image slices (González-Solares et al., 2021). These 3D images uncover salient morphological features which can be used in drug development and TME analysis.

## **Imaging Mass Spectrometry (IMC)**

IMC is an imaging technique that can analyze the protein abundance of a specimen at single-cell resolution. The general workflow is as follows. First, an antibody-A picture containing text

Description automatically generatedstained tissue sample is ablated using a UV laser producing a plume of particles which are sent to a mass cytometer to undergo ionization where cell-specific signals are captured (Devine and Behbehani, 2021). The ability of IMC to resolve proteomic features of individual cells is crucial, since the TME is comprised of a mixture of cell-types, each varying in their protein abundance [(Spitzer and Nolan, 2016).](https://www.sciencedirect.com/science/article/pii/S009286741630410X) Additionally, IMC conveys spatial relationships between individual cells. This further increases its utility in TME analysis, since the spatial context in which cells operate is informative of their function in cancerous tumors [(Baharlou et al., 2019).](https://www.frontiersin.org/articles/10.3389/fimmu.2019.02657/full) In view of these capabilities, IMC can reveal biomarkers for monitoring therapy efficacy as well as targets for antibodies in immunotherapy, making it a valuable asset in TME analysis.

## **Difficulties in obtaining STPT and IMC**

However, it is costly to perform both STPT and IMC on the same solid tumor section. Although it is possible to first image a slice of a solid tumour using STPT and subsequently do IMC on that image slice, there are two caveats in doing this. The first is that it is difficult to align the image sets. Mechanical perturbations that take place when the vibratome slices the solid tumor during STPT and uncertainties in the exact spatial location of signals from IMC confounds the alignment process. The second is that performing both imaging techniques is expensive. Considering these, we aim to construct a generative model that reconstructs IMC images from STPT images. This model will learn how to reliably relate STPT and IMC images. It will take in STPT images as input and output IMC images as if IMC was done. Using our model, researchers and practitioners will be able to effectively obtain both STPT and IMC images while only incurring costs from STPT. That’s two for the price of one!

Figure 1: Images produced by STPT (left) and IMC (right)

caption. Captions are set in roman, 9 point. Use a Drawing area to make space for figures.

# **Methods**

## **Image Dataset Description**

18 physical sections of a solid tumour from breast cancer patient tissue were imaged using both STPT and IMC imaging techniques. Each physical section has 2 STPT images at two imaging depths (30 μm and 38 μm) (Bressan et al., 2021), and 40 IMC 2D grayscale images - each corresponding to a different protein marker. The 4 channels in STPT images correspond to far red, red, green, and blue. See Figure 1 for an example of STPT and IMC images. Refer to the Methods Section of González-Solares et al. (2021) for more details on the image data, image data acquisition process and image preprocessing steps.

## **Image Preprocessing**

While training, we found that our hardware was not able to handle STPT and IMC images because of their size. Therefore, we resorted to crop each image into smaller chunks so that our network could train on these smaller images.

First, we performed normalization and log scaling on all images. Next, for STPT images, we concatenated the two optical sections per physical section, giving an 8-channel tensor; for IMC images, we concatenated all 40 2D grayscale images, giving a 40-channel tensor. We cropped 16 pixels off from each side, as the border consists of solely black pixels. Finally, we used the function torch.split from the PyTorch library to crop original images into 256x256 chunks. We treated this collection of 256x256 images as our new dataset.

## **U-Net**

We used a U-Net architecture [(Ronneberger et al., 2015)](https://arxiv.org/pdf/1505.04597.pdf) as a benchmark to our point-cloud inspired architecture described in the next subsection. Our network architecture was nearly the same to that of the original U-Net paper. We only added a batch normalization layer before each ‘block’ of two 3x3 convolutions, followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation. We used mean-squared error (MSE) the AdamW [(Loshchilov and Hutter, 2019)](https://arxiv.org/abs/1711.05101) optimizer with learning rate 0.001 and weight decay 0.01 for training. 10% of all images were randomly selected to make up the validation set. We also ensured that the validation set contained images from each physical section. Minibatches of size 64 were used and the model was trained for 5 epochs.

## **Point Clouds**

3D data are becoming ubiquitous because of technological advancements in applications such as robotics, autonomous driving, and virtual reality (Wu et al., 2020). Point clouds, a collection of (x,y,z) points in Euclidean space, are one of many representations of 3D data (Qi et al., 2017). Special considerations must be taken account when working with point clouds. For example, algorithms typically employ a symmetric function to respect the permutation invariance of point clouds (Qi et al., 2017, Wu et al., 2020).

We trained the ‘hourglass’ neural network architecture found in Fan et al. (2016). The original aim in the paper was to reconstruct a 3D point cloud object from a single 2D image. In our application, we noted that IMC images of individual channels resembled point clouds (See Figure x) and STPT images were multi-channeled 2D images. Hence, the reconstruction task from STPT to IMC images was akin to the one described in Fan et al.

We followed the architecture described in Fan et al. ReLU activation functions and batch normalization were done after convolutional layers. We diverge from Fan et al. at the last layer where we perform a point cloud rendering technique based on the gaussian equation. The output according to the original architecture in Fan et al. is an (Nx3) tensor. However, since IMC images are not truly point clouds, but 2D grayscale images, we converted the (Nx3) tensor to a 2D image using a custom point cloud rendering technique. First, we first constructed two 256x256 mesh grids using torch.meshgrid and concatenated them to a (2x256x256) tensor. Next, we extracted the x and y coordinates from the (Nx3) tensor, to form an (Nx2) tensor. Using the gaussian equation, we

## **Illustrations, graphs, and photographs**

All graphics should be centered. Please ensure that any point you wish to make is resolvable in a printed copy of the paper. Resize fonts in figures to match the font in the body text and choose line widths which render effectively in print. Readers (and reviewers), even of an electronic copy, may choose to print your paper to read it. You cannot insist that they do otherwise, and therefore must not assume that they can zoom in to see tiny details on a graphic.

## **Color**

Please refer to the author guidelines on the CVPR 2022 web page for a discussion of the use of color in your document.

If you use color in your plots, please keep in mind that a significant subset of reviewers and readers may have a color vision deficiency; red-green blindness is the most frequent kind. Hence avoid relying only on color as the discriminative feature in plots (such as red *vs.* green lines) but add a second discriminative feature to ease disambiguation.

# **Results**

To evaluate performance of the U-Net-based and the point cloud-based model, we selected both cropped STPT images for reconstruction. We could not use full STPT images in our model due to hardware constraints. Cropped STPT images were chosen based on how informative they were of the model’s performance (e.g., we did not choose cropped images of the black background). We ran forward passes on selected STPT images to obtain reconstructed IMC images. Each forward pass generated 40 2D grayscale images packaged in a 40-channel tensor – each channel corresponding to a different protein marker. For convenience, we only compare reconstruction on 3 randomly chosen channels. We compared each model’s reconstructed IMC image by plotting both reconstructed and ground truth images using matplotlib, and by computing mean-squared error.

## **U-Net Reconstructions**

A picture containing dark

Description automatically generated

# **References**

1. FirstName Alpher, Frobnication. *IEEE TPAMI*, 12(1):234– 778, 2002.
2. FirstName Alpher and FirstName Fotheringham-Smythe. Frobnication revisited. *Journal of Foo*, 13(1):234–778, 2003.
3. FirstName Alpher, FirstName Fotheringham-Smythe, and FirstName Gamow. Can a machine frobnicate? *Journal of Foo*, 14(1):234–778, 2004.
4. FirstName Alpher and FirstName Gamow. Can a computer frobnicate? In *CVPR*, pages 234–778, 2005.
5. FirstName LastName. The frobnicatable foo filter, 2014. Face and Gesture submission ID 324. Supplied as supplemental material fg324.pdf.
6. FirstName LastName. Frobnication tutorial, 2014. Supplied as supplemental material tr.pdf.