2022 44th Annual International Conference of
the IEEE Engineering in Medicine & Biology Society (EMBC)
Scottish Event Campus, Glasgow, UK, July 11-15, 2022

2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 978-1-7281-2782-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/EMBC48229.2022.9871251

A Graph Based Neural Network Approach to Immune Profiling of Multiplexed Tissue Samples

Natalia Garcia Martin [1,2], Stefano Malacrino [1,3], Marta Wojciechowska [1,4], Leticia Campo [4], Helen Jones [5], David C. Wedge [6], Chris Holmes [1,2], Korsuk Sirinukunwattana [1,4], Heba Sailem [1,4], Clare Verrill [3,7], and Jens Rittscher [1,4,7]

***Abstract*— Multiplexed immunofluorescence provides an unprecedented opportunity for studying specific cell-to-cell and cell microenvironment interactions. We employ graph neural networks to combine features obtained from tissue morphology with measurements of protein expression to profile the tumour microenvironment associated with different tumour stages. Our framework presents a new approach to analysing and processing these complex multi-dimensional datasets that overcomes some of the key challenges in analysing these data and opens up the opportunity to abstract biologically meaningful interactions.**

I. Introduction

Novel tissue multiplexing imaging platforms [1], [2] allow the analysis of a broad range of cell types in the tissue architecture context. These approaches open up new opportunities for improving our understanding of disease, monitoring therapeutic response, and the development of high-dimensional clinical tests. Here, we are interested in profiling the complex interaction between the tumour and the immune system within the tumour microenvironment (TME), which dictates the tumour progression. While current cancer classification highly relies on the extent of the primary tumour (T), lymph node involvement (N) and metastatic presence (M), visualising multiple protein targets in the same tissue allows us to interrogate the role of adaptive immune cell infiltration in colorectal cancer (CRC) prognosis.

The analysis of multiplexing data requires the combination of spatial information that captures the changes in tissue architecture with measurements of protein expression. When compared to standard digital pathology, multiplexing datasets are typically much smaller and contain imaging artifacts and strong variations in protein expression, making this a particularly challenging problem. Moreover, interpretability is a key aspect when working with multiplexed data to be able to link the analysis to any underlying biological hypothesis.

Building on recent success of applying graph neural networks (GNNs) to histopathology, we introduce a novel framework for analysing multiplexed immunofluorescence (IF) images using GNNs. Constructing graphs from multiplexed IF data is non-trivial due to the stated challenges. Our approach overcomes these challenges by: (1) including a selection of network metrics that capture the interactions between the immune cells and the tumour; (2) a hierarchical structure that considers both the cell-level and the spatial tissue arrangement; (3) implicit denoising from the use of message-passing; (4) data augmentation on the graphs to account for the limited amount of training data; and (5) the opportunity to interpret results in order to identify the tissue areas contributing the most to the predictions. In summary, we propose a GNN model to profile the tumour microenvironment associated with different tumour stages in an explainable setting.

II. Methods

Figure 1 provides a summary of the overall approach. Prior to the analysis, we carefully pre-process the data by applying fluorescent image correction algorithms and subsequently identify cell nuclei using a segmentation approach previously validated on a different dataset. Rather than performing a global analysis of the slide, we perform a local analysis in selected regions of interest (RoIs)(see Fig. 1A). The two-layer graph described in Section III-B is constructed to abstract the key biological interactions of the underlying tissue. It first captures the location of cells, certain morphological measurements and protein expressions to form a cell-graph (Fig.

[1]Big Data Institute, University of Oxford, Li Ka Shing Centre for Health Information and Discovery, Oxford, UK natalia.garciamartin@spc.ox.ac.uk
[2]Department of Statistics, University of Oxford, Oxford, UK
[3]Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK
[4]Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK,
[5]Department of Colorectal Surgery, Oxford University Hospitals NHS Trust, Oxford, UK
[6]Manchester Cancer Research Centre, University of Manchester, Manchester, UK
[7]NIHR Oxford Biomedical Research Centre, Oxford, UK

*NGM is supported by Cancer Research UK (CRUK), through a CRUK Oxford Centre Prize DPhil Studentship (C2195/A27450). SM and KS are supported by the PathLAKE Centre of Excellence for digital pathology and artificial intelligence (Grant ref: File Ref 104689/application number 18181). MW is funded by the UK Engineering and Physical Sciences Research Council and Medical Research Council (EP/L016052/1) and in part by Perspectum Ltd. HS is funded by a Sir Henry Wellcome Fellowship (204724/Z/16/Z). This work was supported by the Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Science (CIFMS), China (grant number: 2018-I2M-2-002). This research was further supported by Bowel Research UK and by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. Computation used the Oxford Biomedical Research Computing (BMRC) facility, a joint development between the Wellcome Centre for Human Genetics and the Big Data Institute supported by Health Data Research UK and the NIHR Oxford Biomedical Research Centre.

1D). After message-passing, the updated cell embeddings are aggregated at the tile level and concatenated with the set of hand-crafted immune-interaction features that we describe in Section II-D. The second set of graphs are constructed at the RoI level, with nodes representing tile centroids to form a tile-graph (Fig. 1B). These RoI-level graphs are then fed into the model for pT stage prediction (Fig. 1C). Finally, post-hoc explainability methods, presented in Section II-F, are utilised to visualise the relationship between immune interaction profiles and prediction of tumour stage.

A. Multiplexed IF data

The Perkin-Elmer Vectra platform features an immune panel consisting of six fluorescent markers. DAPI is used for nuclei segmentation. Cytokeratin is used to delineate epithelial cells. A further four markers are included to depict immune cells: CD4 for helper T-cells, CD8 for cytotoxic T-cells, CD20 for B-cells, and Foxp3 for regulatory T-cells. The system also provides a seventh channel corresponding to the imaging system's autofluorescence isolation capacity which improves signal-to-noise ratio [3].

B. Segmentation of nuclei and cell phenotyping

Multiplexed image data requires careful preprocessing as tiles are scanned independently. Hence before segmentation, a background and shading correction [16] was performed to improve the stitching of individual image tiles. Moreover, contrast limited adaptive histogram equalisation (CLAHE) [6] was used to improve contrast in the DAPI channel, which is used for segmentation. The segmentation network employed to identify cell nuclei consists of a 3-class (cell inside, cell boundary, background) modified U-Net architecture comprised of the original U-Net [17] decoder and a ResNet [18] encoder. The model was pre-trained on the fluorescence samples from the publicly available BBBC038v1 dataset [19] to ensure the correct identification of cells of different sizes and to ignore DAPI positive fragments not corresponding to actual cells. Segmentation masks were projected onto the remaining channels to measure the average nuclei protein expression. Cells were then assigned to the cell type corresponding to the marker whose expression was located in the highest percentile rank.

C. Two-layer graph representation

After image segmentation we define an undirected graph $G = (V, E)$, with vertices V and edges E. Similar to Pati *et al.* [4] we employ a two-layer graph representation, with (1) cell-graphs [5] constructed on small randomly sampled tiles, where nodes represent nuclei centroids in order to quantify local patterns of immune interaction, and (2) a tile-level graph able to aggregate information from the multiple tissue regions. The graph topology is represented by an adjacency matrix $A \in \mathbb{R}^{N\times N}$, and node features are represented by the matrix $X \in \mathbb{R}^{N\times D}$, with $N = |V|$ and feature dimension D. We construct A based on a distance threshold as follows:

$$A_{ij} = \begin{cases} 1 & \text{if } d(i,j) < k \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where k determines which nodes in the graph are connected. The choice of k is described in Section III-B.

D. Cell-graph feature extraction

We calculate a total of 68 handcrafted network metrics at the cell-graph level to acquire information about the distribution of each cell population of interest.

These include the average clustering and square clustering coefficients, the assortativity, radius, density, transitivity, and the closeness of each cell type population, as defined by Schult *et al.* [8]. The ratios between each pair of immune cell densities (e.g. CD4-CD8 ratio), a known prognostic factor for cancer progression [9] are also computed. To measure the degree of mixing between tumour and immune cells, we additionally compute the ratio of immune-tumour to immune-immune interactions [10].

E. Cell-graph and tile-graph neural network

We employ Graph Neural Networks (GNNs) to obtain a graph representation $H \in \mathbb{R}^{N\times P}$ from our initial cell embeddings $H^0 = X \in \mathbb{R}^{N\times D}$, where P is the number of output features. Using the notation from [11] and [12], we first perform a number of message passing steps to obtain the node embeddings h_v for each cell v in the cell-graph CG, which we then combine into a global cell-graph embedding h_{CG} for each tile. The message passing consists of an aggregation and combination of the neighbouring nodes features. For the k^{th} GNN layer:

$$a_v^{(k)} = \text{AGG}^{(k)}\left(\left\{h_u^{(k-1)} : u \in \mathcal{N}(v)\right\}\right) \tag{2}$$

$$h_v^{(k)} = \text{COMBINE}^{(k)}\left(h_v^{(k-1)}, a_v^{(k)}\right) \tag{3}$$

$$h_{CG} = \text{READOUT}\left(\left\{h_v^{(k)} \mid v \in CG\right\}\right), \tag{4}$$

where $\mathcal{N}(v)$ denotes the set of neighbours of v. We use the graph convolutional network (GCN) operator defined in [13]:

$$h_v^{(k)} = W_1^{(k)} h_v^{(k-1)} + W_2^{(k)} \sum_{u\in\mathcal{N}(v)} h_u^{(k-1)}. \tag{5}$$

The updated cell-graph embeddings h_{CG} are then combined with the selected network metrics m_{CG} listed in the previous sub-section to define the tile-graph embeddings for each tile t:

$$h_{\text{t}}^{(0)} = \text{CONCAT}\left(m_{\text{CG}}, h_{CG}\right). \tag{6}$$

These tile embeddings h_t for each tile t in the tile-graph TG are then updated by applying Eqs. 2-3 again, where nodes now correspond to tiles. The readout layer then combines the information from the multiple tiles to obtain the final embedding h_{TG} for the RoI:

$$h_{TG} = \text{READOUT}\left(\left\{h_t^{(k)} \mid t \in TG\right\}\right). \tag{7}$$

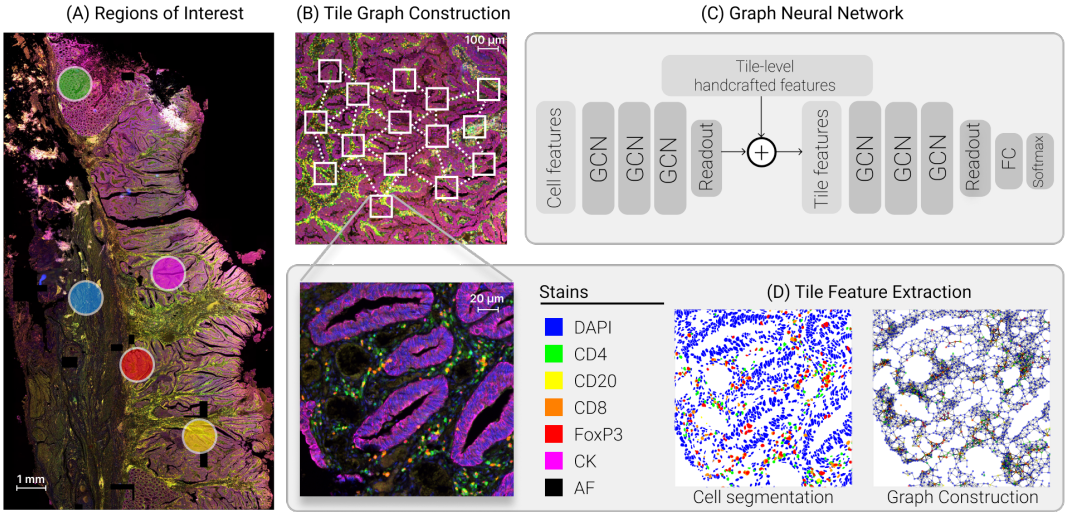

Fig. 1. Overview of the proposed method described in Section II. The regions of interest correspond to the tumour centre (yellow and magenta), invasive tumour front (red), background mucosa (green), and peritumoural stroma (blue).

F. Post-hoc explainability

We employ Integrated Gradients (IG) [14] in the tile-graph to understand the significance of each tile node in predicting tumour stage. We do so by computing the IG attribution of each edge and aggregating the attributions of the edges connecting each node. The IG edge attribution is computed by comparing each edge mask with a baseline of edge weights set to zero. Since we use unweighted graphs, the initial edge weights are all one. The IG for each edge e_i is computed as follows:

$$\mathrm{IG}_{e_i} = \int_{\alpha=0}^{1} \frac{\partial F(x_\alpha)}{\partial w_{e_i}} d\alpha, \tag{8}$$

where x_α corresponds to the original input graph but with all edge weights set to α, w_{e_i} denotes the current edge weight, and $F(x)$ is the output of the model for an input x. The integral is approximated using a Gauss-Legendre quadrature.

In order to identify the key features impacting the prediction, we further run the GNN Explainer model [15], which maximises the mutual information MI between the prediction of the trained model and that of the simplified explainer model given by a subgraph G_S and a subset of features T: $\max_{G_S,T} \mathrm{MI}(Y,(G_S,T))$.

III. Experiments

A. Dataset

Paraffin-embedded tissue samples of 41 rectal primary tumours were used to investigate the risk of disease progression and recurrence. Specialist GI pathologists reported tumour stage on matching H&E slides: 25 of these samples were assigned a pT1 tumour stage, while 16 samples were considered to be more advanced (13 pT2, 3 pT3). Specific regions of interest such as those shown in Fig. 1A were provided by a pathologist for the tumour centre, invasive tumour front, background mucosa, and peritumoural stroma, guided by the matched H&E image. Annotation areas correspond to the standard 1mm diameter disk size used for biopsies and tissue microarrays (TMAs), allowing for a future integrative analysis with TMA cohorts.

B. Graph construction

RoIs of the size of 2048x2048 pixels corresponding to the bounding box of the disk annotations are selected to investigate immune-cell interactions across samples and regions. From each RoI, 200 256x256 tiles are randomly chosen to construct cell-graphs using NetworkX [8], with nodes positioned at the centroid of each nucleus. We set k from Eq. 1 to be 30 pixels. This results in a small node degree as well as a small number of disconnected nodes in order to reduce graph complexity and facilitate metric computation. Given the resolution of 20x, the 30 pixel threshold corresponds to about 15μm—between two to three times the size of the average diameter of the observed nuclei, 6μm—and captures the immediate neighbourhood of each cell. Moreover, the tile of 256x256 pixels captures a length of about 120μm hence the size of the tile is within a distance within which the cell can communicate effectively [7].

For each node, we record the average expression for the five markers of interest (CD4, CD8, CD20, FoxP3, CK), the area occupied by the cell, and the cell solidity. These 7 features are inserted as node features. We subsequently perform three message-passing steps to update the node features by encompassing information from nearby cells, which are aggregated using mean pooling and transformed into a vector of length 16. Additionally, for each tile, we compute the set of 68 hand-crafted immune-interaction features enumerated in Section II-D. Nuclei and cell-interaction features are then concatenated into a vector of length 84 per tile. The second set of graphs are constructed at the RoI level, with nodes representing the 200 sampled tiles positioned at their tile centroids and node features corresponding to the selected 84 attributes. These RoI-level graphs are then fed into the model

TABLE I

MEAN AND STANDARD DEVIATION OF ROI-LEVEL CLASS-WEIGHTED F1-SCORES MEASURED ON THE TEST SET AND AVERAGED OVER THREE DISTINCT TRAIN-TEST SPLITS.

Region	GCN Mean pool	GCN Add pool	GCN Max pool	MIL Attention pool	MIL Mean pool	MLP
All	$58.4_{\pm 1.7}$	$61.6_{\pm 4.2}$	$61.6_{\pm 3.2}$	$55.1_{\pm 1.4}$	$49.5_{\pm 1.6}$	$49.0_{\pm 0.7}$
Centre	$53.8_{\pm 6.4}$	$60.7_{\pm 8.9}$	$58.0_{\pm 0.1}$	$50.0_{\pm 2.1}$	$49.7_{\pm 3.4}$	$49.2_{\pm 0.8}$
Front	$63.6_{\pm 8.5}$	$60.4_{\pm 9.9}$	**72.9** $_{\pm 7.8}$	$54.4_{\pm 5.8}$	$40.0_{\pm 7.1}$	$48.9_{\pm 0.7}$
Mucosa	$47.5_{\pm 12.6}$	$50.2_{\pm 8.1}$	$60.8_{\pm 9.9}$	$63.5_{\pm 11.3}$	$46.1_{\pm 9.1}$	$48.4_{\pm 1.9}$
Stroma	$56.6_{\pm 7.4}$	$58.9_{\pm 8.5}$	$61.6_{\pm 6.5}$	$57.7_{\pm 3.9}$	$53.0_{\pm 4.5}$	$48.4_{\pm 1.2}$

for pT stage prediction.

C. Data augmentation

The augmented set is obtained by constructing the networks using a subset of 80% of the nodes at each step (160 tiles) and by varying the threshold k that needs to be surpassed for an edge to be included between two neighbours by sampling a value in the pixel range {150, 175, 200, 225, 250}, resulting in a variety of tighter and sparser graphs. The node subsampling and edge modifications ensure that networks in the training set are sufficiently different to avoid over-fitting. For the test set, only a single network is constructed per RoI using the default distance threshold of 200 pixels for adjacency construction and the full set of tile nodes in the RoI (200 tiles).

D. Implementation

The model consists of three GraphConv [13], [20] layers with ReLu activation and global pooling aggregation. Experiments are conducted in PyTorch 1.7.0 using PyTorch Geometric [20].

Data are split into training and testing at the patient level. We use 70% (134) of the RoIs for training and 30% (59) for testing. Due to the limited sample size, a pseudo-validation set is constructed by randomly sampling (with pT stratification) 10% of the pre-augmented training data, and used for hyperparameter grid-search. Model performance is measured according to their weighted F1-scores on the test set. The model is trained using an L2-regularised Adam optimiser and a weighted cross entropy loss. We employed early stopping based on the pseudo-validation set. We tune the hyperparameters using a grid search. The values that provide the best performance in terms of class-weighted F1 score correspond to a dropout ratio of 0.5, a learning rate of 10e-5, a weight decay of 10e-5, 32 hidden layers, and a batch size of 64.

Two baseline models are considered: the first one consists of a multi-layer perceptron (MLP) which takes as input the average individual nuclei features (size, shape and marker expression) without taking the cell topology into consideration. The second consists of a multi-instance learning (MIL) approach which takes as input both the cell-level features and the cell-graph features, computed as in section II-D. We consider this model with both attention and average pooling.

E. Post-hoc explainability

We compute Integrated Gradients [14] using the model interpretability library for PyTorch Captum [21] to obtain an importance score of individual edges and nodes for the pT stage prediction of each instance in the test set. We can then compare areas of predictive importance across the different selected RoI regions. The GNN Explainer model (implemented using PyTorch Geometric) is used to obtain feature importances across all tiles in the test set.

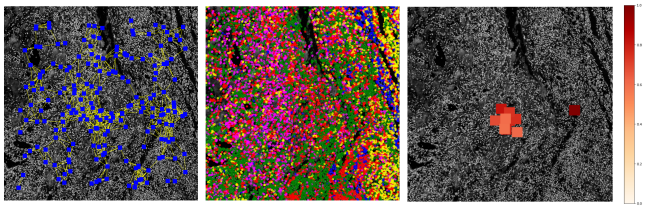

Fig. 2. An example of an invasive front RoI for a pT0 sample classified correctly. (Left) Tile-graph of 200 256x256 tiles overlaid on DAPI. (Centre) Cell-graph corresponding to the 2048x2048 RoI: blue - epithelial, green - T-helper, red: cytotoxic T-cell, magenta: T-reg, yellow: B-cell. (Right) Top ten tiles classified as important using integrated gradients for predicting tumour stage from immune interaction features.

F. Results and discussion

As shown in Table I, in the majority of the graph-based experiments the invasive front was the region with the highest predictive power, followed by the peritumoural stroma, known to have a high prognostic impact. Moreover, all the graph-based models present an improvement over the baseline models: this result suggests that the network topology plays an important role in tumour stage classification. Among the graph-based models, global max-pooling performed better than average pooling. Due to the limited number of samples with pT3, the classification of these RoIs was challenging. However, the majority of these RoIs were predicted to have pT2 stage, demonstrating that the model has learned to identify immune features related to an advanced cancer state. The proportion of interactions between CD4+ and CD8+ cells, the interactions between FoxP3 positive and epithelial cells, and the average expression of CD20 were determined by GNN Explainer as the top three features affecting tumour stage classification. Fig. 2 shows an example of tiles selected by IG as important in an invasive front RoI: it can be observed that the network considers a large cluster of regulatory T-cells as the most significant area for the prediction.

IV. CONCLUSIONS

Our experiments demonstrate that the proposed two-layer GNN opens up new possibilities for interrogating multi-

plexed immuno-fluorescence images. As the model is capable of predicting tumour stage with an F1-score above 60%, we conclude that the model captures disease relevant information at a local level. The improvement over the baseline observed with the models that use GCNs, which are able to capture more complex spatial interactions, suggests that effectively modeling the cell topology plays an important role in the tumour stage classification. However, the improvement in performance is not the only advantage gained through the use of the proposed method. First, we were able to naturally denoise the marker expressions of the cells by means of the message-passing steps. Second, our hierarchical graph structure generated biologically meaningful entities which would not have been otherwise acquired through the use of convolutional neural networks. Third, by applying post-hoc explainability methods on the tile-graph we were able to identify the regions that contributed the most to the classification hence profiling local interaction patterns. This will enable a follow up analysis to identify explicit entities at the cellular and tissue level that are of biological and clinical interest.

ACKNOWLEDGMENT

The authors thank Avelino Javer for the development and training of the cell segmentation algorithm and Joel Lefebvre for the development of the Python implementation of the BaSiC shading correction method (PyBaSiC).

REFERENCES

[1] Stack, E., Wang, C., Roman, K. & Hoyt, C. Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods*. **70**, 46-58 (2014)

[2] Bodenmiller, B. Multiplexed epitope-based tissue imaging for discovery and healthcare applications. *Cell Systems*. **2**, 225-238 (2016)

[3] inForm software. (https://www.akoyabio.com/phenoptics/software/inform-tissue-finder/), (Accessed on 23/01/2022)

[4] Pati, P., Jaume, G., Fernandes, L., Foncubierta-Rodriguez, A., Feroce, F., Anniciello, A. & Others HACT-Net: A Hierarchical Cell-to-Tissue Graph Neural Network for Histopathological Image Classification. *Uncertainty For Safe Utilization Of Machine Learning In Medical Imaging, And Graphs In Biomedical Image Analysis*. 208-219 (2020)

[5] Gunduz, C., Yener, B. & Gultekin, S. The cell graphs of cancer. *Bioinformatics*. **20**, i145-i151 (2004)

[6] Pizer, Stephen M and Amburn, E Philip and Austin, John D and Cromartie, Robert and Geselowitz, Ari and Greer, Trey and ter Haar Romeny, Bart and Zimmerman, John B & Zuiderveld, Karel. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*. **39.3**, 355-368 (1987)

[7] Francis, Karl & Palsson, Bernhard O. Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion. *Proceedings of the National Academy of Sciences*. **94.23**, 12258-12262 (1997)

[8] Schult, D. & Swart, P. Exploring network structure, dynamics, and function using NetworkX. *Proceedings Of The 7th Python In Science Conferences (SciPy 2008)*. 11-16 (2008)

[9] Hadrup, S., Donia, M. & Straten, P. Effector CD4 and CD8 T cells and their role in the tumor microenvironment. *Cancer Microenvironment*. **6**, 123-133 (2013)

[10] Keren, L., Bosse, M., Marquez, D., Angoshtari, R., Jain, S., Varma, S., Yang, S., Kurian, A., Valen, D., West, R. & Others A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell*. **174**, 1373-1387 (2018)

[11] Kipf, T. & Welling, M. Semi-supervised classification with graph convolutional networks. *ArXiv Preprint ArXiv:1609.02907*. (2016)

[12] Xu, K., Hu, W., Leskovec, J. & Jegelka, S. How powerful are graph neural networks?. *ArXiv Preprint ArXiv:1810.00826*. (2018)

[13] Morris, C., Ritzert, M., Fey, M., Hamilton, W. & Others Weisfeiler and leman go neural: Higher-order graph neural networks. *Proceedings Of The AAAI Conference On Artificial Intelligence*. **33**, 4602-4609 (2019)

[14] Sundararajan, M., Taly, A. & Yan, Q. Axiomatic attribution for deep networks. *International Conference On Machine Learning*. 3319-3328 (2017)

[15] Ying, R., Bourgeois, D., You, J., Zitnik, M. & Leskovec, J. GNN Explainer: A tool for post-hoc explanation of graph neural networks. *ArXiv Preprint ArXiv:1903.03894*. (2019)

[16] Peng, T., Thorn, K., Schroeder, T., Wang, L., Theis, F., Marr, C. & Navab, N. A BaSiC tool for background and shading correction of optical microscopy images. *Nature Communications*. **8**, 1-7 (2017)

[17] Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. *International Conference On Medical Image Computing And Computer-assisted Intervention*. 234-241 (2015)

[18] He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. *Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition*. 770-778 (2016)

[19] Ljosa, V., Sokolnicki, K. & Carpenter, A. Annotated high-throughput microscopy image sets for validation.. *Nature Methods*. **9**, 637-637 (2012)

[20] Fe, M. & Lenssen, J. Fast Graph Representation Learning with PyTorch Geometric. *ICLR Workshop On Representation Learning On Graphs And Manifolds*. (2019)

[21] Kokhlikyan, N., Miglani, V., Martin, M., Wang, E., Alsallakh, B., Reynolds, J., Melnikov, A., Kliushkina, N., Araya, C., Yan, S. & Others Captum: A unified and generic model interpretability library for PyTorch.