DeepPathway: Predicting Pathway Expression from Histopathology Images

Muhammad Ahtazaz Ahsan[1], Karen Piper Hanley[2], Martin Fergie[1], Claire O'leary[3], Gerben Borst[3], Federico Roncaroli[4], Magnus Ratrray[1], Mudassar Iqbal[1*], and Syed Murutza Baker[1*]

[1]Division of Informatics, Imaging and Data Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, UK
[2]Division of Diabetes, Endocrinology & Gastroenterology, Faculty of Medical Sciences, The University of Manchester, UK
[3]Division of Cancer Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, UK
[4]Division of Neuroscience, Faculty of Biology, Medicine and Health, The University of Manchester, UK

***Abstract*—Spatial transcriptomics (ST) technologies provide spatially resolved gene expression along with image data, allowing the integrated analysis of complex tissue microenvironments. Despite their potential, the widespread adoption of ST remains limited due to high costs and technical challenges in data acquisition. Thus, there have been recent efforts to develop deep learning (DL)-based computational methods capable of inferring spatial gene expression from H&E images which are much cheaper and easily available. These methods demonstrate promising results in reconstructing transcriptomic landscapes within tissue sections. While existing approaches predominantly focus on gene-level prediction, biological processes are often regulated at the pathway level through coordinated activity among functionally related genes. Here we present DeepPathway, a contrastive learning-based approach trained on ST data to predict pathway expression from H&E images, with input pathway expression computed by summarizing the expression of constituent genes using established gene set definitions. We evaluate the performance of our method on two prostate cancer datasets and validate our approach on the H&E images acquired from The Cancer Genome Atlas (TCGA) clearly differentiating between normal and cancer tissues. Finally, we apply our method to predict hypoxia signatures using H&Es of brain tumour patients where ground truth hypoxia staining was available. Implementation code for DeepPathway is available at https://github.com/aahsan045/DeepPathway.**

I. Introduction

Tissue in multi-cellular organisms consist of diverse cell populations that coordinate their functions by spatially interacting with each other and with the microenvironment [1]. Single-cell RNA sequencing (scRNAseq) technologies allow high-throughput gene expression profiling to study individual cellular activity but lack the spatial context. To address this limitation, spatial transcriptomics (ST) [2]) technologies, such as the Visium platform by 10x Genomics and Slide-seq [3] have been developed. These methods retain the spatial context of cells while capturing gene expression data without disrupting the complexity of tissue organisation. ST platforms often include co-registered Haematoxylin and Eosin (H&E)-stained images providing a view of tissue architecture and heterogeneity at microscopic level. The integration of pathological, ST and molecular data enables researchers to accurately analyse tissue regions that can be relevant to diagnosis, disease progression or are driven by molecular pathways that can be targeted with effective treatment modalities [4].

Although ST provides detailed molecular insights, it is an expensive and time-consuming technique that requires unique expertise. In contrast, whole scanned H&E-stained images are inexpensive and readily available, e.g., from routine diagnostic provision. Recent advances in deep learning (DL)-based methods have enabled the prediction of spatial gene expression directly from histology images [5]–[19]. These models are trained on ST datasets and then used to infer gene expression patterns from H&E images alone. Thus, H&E-stained sections have become a cornerstone that offers a fast and practical alternative to multi-omics technologies to investigate cellular functions across diverse tissue types.

For example, ST-Net [5] utilises DenseNet-121 (pre-trained on ImageNet) for extracting features from H&E

* Joint corresponding authors

images and predicts expression levels for 250 genes. Similarly, BLEEP (Bimodal Embedding for Expression Prediction), proposed by Xie et al. [6], employs a bimodal contrastive learning framework [20] to learn a shared embedding space between H&E patches and corresponding gene expression profiles. Despite being trained on thousands of genes, these methods tend to predict only a limited subset of disease-relevant genes with reasonable accuracy.

Normal cell functions result from the coordination of complex, hierarchical gene networks[1]. The analysis of such networks and pathways rather than individual genes reflects more closely tissue homeostasis in normal and pathological conditions. To translate a set of individual gene expression profiles into a pathway expression profile at cellular or spot level resolution, methods such as AUCell [22] or Ucell [21] can be utilised. This provides a way to transform thousands of genes into a handful of relevant pathways which could be more accurately and efficiently learned and could have more translational impact in identifying biomarkers for diagnosing certain disease phenotypes.

With the aim to predict the pathway expression from H&E images, we develop a deep learning (DL)-based approach called DeepPathway, adapted from an existing contrastive learning framework, i.e., BLEEP [6]. Starting with a spot-by-gene expression matrix from ST experiment, we employ UCell [21] to quantify pathway activity, transforming the original spot-by-gene matrix into a spot-by-pathway matrix. To benchmark our approach and demonstrate its utility, we first compare gene-level (predicted gene expression subsequently summarized into pathways) and pathway-level prediction performance by adapting the original BLEEP framework across multiple ST datasets. BLEEP requires access to ST training data to infer the expression of test histology image, somewhat limiting its practical deployment in real-world clinical settings. To address this limitation, we introduce a novel two-stage learning framework. In the first stage, we train a BLEEP-based model using contrastive learning to align H&E image patches with their corresponding pathway expression profiles. In the second stage, we replace BLEEP's inference mechanism with a deep neural network (DNN) that directly predicts pathway expression from H&E features extracted by the trained first-stage model.

The key contributions of this work are outlined below.

1) We present a new approach adapted from an existing bimodal contrastive learning framework to predict the pathway expression from H&E images. We demonstrate the effectiveness of pathway prediction over gene prediction using prostate cancer datasets.
2) We propose a two-stage learning framework to predict pathway expression directly from H&E image.
3) We validate our approach on a larger prostate cancer dataset to show the effectiveness of pathway prediction. Additionally, we show the application of the proposed method on the H&Es of the prostate tumour samples from The Cancer Genome Atlas (TCGA) data. Finally, we use the proposed method to identify hypoxic regions in brain tumour patient samples using only H&E images which were validated using hypoxia specific staining.

[1]https://www.genome.gov/about-genomics/fact-sheets/Biological-Pathways-Fact-Sheet

II. Materials and Methods

A. Approach overview

A brief overview of the proposed method is present in Fig. 1 which shows as input paired H&E images and gene expression data coming from ST experiments. After initial gene expression and H&E image processing, we summarize the gene expression into pathway expression with UCell [21] method using MSigDB hallmark pathway definitions.

We organise the work reported here into three interconnected approaches, which are briefly described as follows. In the first approach, we use BLEEP as the baseline method to perform and compare gene prediction (then converted into predicted pathway expression) versus direct pathway prediction highlighting the utility of our approach. In the second approach, we improve image feature representations by integrating outputs from a trainable BLEEP image encoder with features from the non-trainable foundation model, *H-optimus-0* [23]. Finally, we replace the inference part of BLEEP with training a deep neural network (DNN) model for improved deployability. This results in a proposed two-stage learning framework in which, first a model is trained using BLEEP with Optimus, and then we train a DNN on image features extracted from the trained model (obtained in the first stage) to predict the pathway expression. Throughout all stages, we leverage the bimodal contrastive learning framework introduced in BLEEP [6], originally developed for predicting spatial gene expression from H&E images. The model takes randomly shuffled batch of samples for training. The main objective is to learn the same-sized embedding space between matching image-expression pairs. As model training continues, the model improves its performance

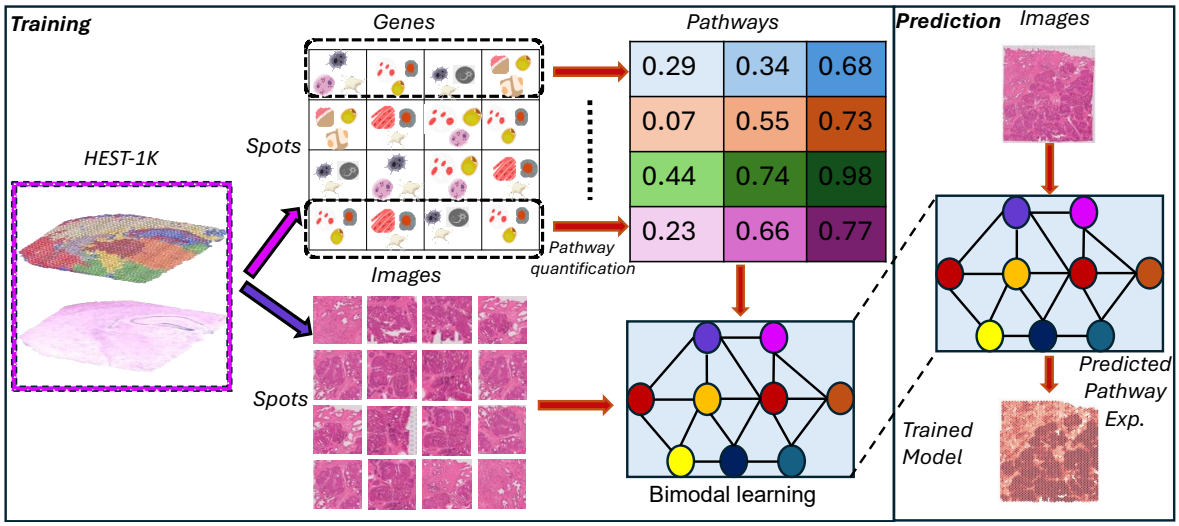

Fig. 1: Overview of the proposed method (**DeepPathway**); Spot-level gene expression and H&E image patches are acquired from HEST-1k database. Gene expression data is summarized into spots × pathway expression matrix using pathway quantification with UCell [21]. In the training part, a deep learning model is trained using bimodal contrastive learning on spot-level image patches and pathway expression data. At the inference time, only H&E image is passed to the trained model for predicting pathway expression.

by capturing meaningful cross-modal similarities while pushing apart dissimilar pairs, thereby enhancing the modality alignment. In the original BLEEP inference setup, test image embeddings are compared to those from the training set, and the pathway expression is estimated by averaging the outputs from the top k closest matches. In contrast, our two-stage framework eliminates the need for training data at inference time and pathway expressions are predicted directly from the image features using the trained DNN.

B. Bimodal contrastive learning for pathway expression prediction

Contrastive learning (CL) [24], [25] is applied to learn meaningful semantic representation by pulling similar samples closer to each other while pushing dissimilar samples apart. This idea has been extended to bimodal data such as image-text, to learn a joint feature representation across modalities. Additionally, it is widely used for developing DL-based applications such as image-caption generation [26] or text-to-image generation [27]. Adapting the idea of contrastive learning from BLEEP, given a spot (a spatially barcoded area where gene expression is profiled) $S = \{X_{img}, X_{expr}\}$, that consisted of paired image-expression profile, a learnable image encoder ($f_\theta(X_{img})$) and an expression encoder ($f_\phi(X_{expr})$) are used to produce embeddings Z_{img} and Z_{expr}, respectively. We then add the positional encoding of the corresponding spot location coordinates (x, y) into both embeddings by adapting idea of sinusoidal positional encoding (f_{PE} (x,y)) as proposed for transformers [28]. Finally, we calculate the contrastive loss using Equation 1.

$$
\begin{aligned}
Z_{img} &= f_\theta(X_{img}),\\
Z_{expr} &= f_\phi(X_{expr}),\\
Z_{img} &= Z_{img} + f_{PE}(x, y),\\
Z_{expr} &= Z_{expr} + f_{PE}(x, y),\\
logits &= Z_{img} \cdot Z_{expr}^T,\\
targets &= Softmax\left(\frac{Z_{img} \cdot Z_{img}^T + Z_{expr} \cdot Z_{expr}^T}{2}\right),\\
loss_{img} &= cross_entropy(logits, targets),\\
loss_{expr} &= cross_entropy(logits^T, targets^T),\\
loss_{CL} &= \frac{loss_{img} + loss_{expr}}{2}
\end{aligned}
\tag{1}
$$

where $logits$ is the paired similarity vector between both modalities treated as a ground truth and $targets$ is the averaged internal similarity between both modalities followed by *softmax*. Then cross-entropy is applied between them to calculate the image and expression loss followed by taking their mean for producing contrastive loss. The final loss (Ł) is calculated by adding L_1 regularization to both image and expression embeddings,

$$
Ł = loss_{CL} + \lambda(\|Z_{img}\|_1 + \|Z_{expr}\|_1) \ . \tag{2}
$$

C. Dataset

We obtained publicly available 10X Visium datasets from the HEST-1k database [29], which provides uni-

formly processed high quality images, gene expression profile, metadata for each ST sample, covering multiple organs and diseases. Specifically in HEST-1k, all whole-slide images (WSIs) are normalized and transformed to generic Pyramidal-based TIFF image objects that can be integrated with computational tools such as OpenSlide [30] or image viewers like QUPath [31]. In this work, we used two different datasets from HEST-1K, i.e., human prostate cancer (one we refer to as small prostate cancer dataset [29] and other as large human prostate cancer dataset [32] hereafter). The small prostate cancer dataset contains five ST samples, WSIs with their paired spatial gene expression profiles where resolution of each WSI is around 0.61 μm/pixel. The total number of spots from the five samples after pre-processing (described in the next section) is $19,245$. The large data from prostate adenocarcinoma consisted of 15 specimens, all diagnosed as Gleason 4+3 score group 3, covering a total of $34,845$ spots with the image resolution of approximately 0.34 μm/pixel.

D. Data processing

For each dataset, we performed gene filtering to remove genes that are expressed in less than 10 spots and spot filtering to remove spots that have less than 500 genes expressed. Then we selected the top 10,000 highly variable genes from each sample. The reason for selecting a large number of such genes is to keep the maximum overlap between each sample's genes with the pathway-specific genes (see below). For pathways definitions, we used the Molecular Signature Database (MSigDB) hallmark collection [33], which is commonly used for cancer studies. It consists of 50 pathway definitions with the unique number of genes equal to 4383. After finding the unique common genes, i.e., intersection of pathway genes with each dataset's unique genes, the number of genes for small prostate cancer dataset and large prostate cancer dataset are found to be 3900 and 3917, respectively. After acquiring the spot-level gene count data, we performed the total normalization followed by the *log1p* transformation using *scannpy*. Normalizing the data deals with different sequencing depths and helps in comparison across spots while the *log1p* transformation reduces the skewness in the data.

Then we performed the H&E image processing on the spots left after gene expression processing. For each dataset, we cropped the image patch of 224 $\times$ 224 pixels centered at each spot's location using the given resolution of microns per pixel through the TIA toolbox [34]. We removed spots where more than 75% of the image represented the background, i.e., having pixel value in all three color channels more than 200.

E. Pathway expression quantification

We used UCell [21] for quantifying the pathway expression given the gene expression measurements and pathway definitions from MSigDB hallmark. In this method, genes in each spot are ranked (from high to low) based on their expression values. Then, for a jth spot, the UCell score (U_j) for a pathway composed of n genes is calculated using Eq. 3.

$$U_j = 1 - \frac{(\sum_{i=0}^{n} r_{ij} - \frac{n(n+1)}{2})}{n.R_{max} - \frac{n(n+1)}{2}} \tag{3}$$

where r_{ij} is the rank value of ith the gene in the jth spot. The R_{max} is the maximum rank value used to drop the uninformative tail of genes with low expression. We selected the R_{max} value for each dataset independently using the following method. For a given spot within a sample, we sorted the gene expression from highest to lowest values and obtained the rank index of the first zero value. We repeated this for all spots and took the median. Then we took the median across all samples within the dataset to obtain the final R_{max} threshold. By following this procedure, we selected the R_{max} value of 1000 for small and large prostate data. Also, we selected only those pathways that have at least 70% of the genes overlapping with the ST expression dataset genes (that are union of all samples' genes in the dataset) and contain minimum 20 genes. By using this threshold, we found the number of pathways for small and large prostate cancer datasets to be 50 and 47, respectively.

F. Model training and validation

Using a dataset of ST samples that consist of H&E image patches and their corresponding expression profiles, we assessed the performance of each model using leave-one-out cross validation, where one sample from the data is held out for testing while remaining data (comprising multiple tissue samples) is divided into spot-level training and validation data with a split ratio of 80% and 20%, respectively. We used ResNet50 for gene prediction experiments as it gives 2048-dimensional feature vector to train and predict thousands of genes while we used ResNet18 (which provides 512-dimensional feature representation) for pathway expression prediction due to its lower computational complexity as well as suitable choice to train and predict a few pathways. Models were trained using the Adam optimizer with a learning rate of 0.001 and training batch size of 256. We kept the same embedding size of 256 for input image projection, gene/pathway expression projection, and positional encoding data. We also integrated pre-trained *H-optimus-0* (a vision-transformer model, trained

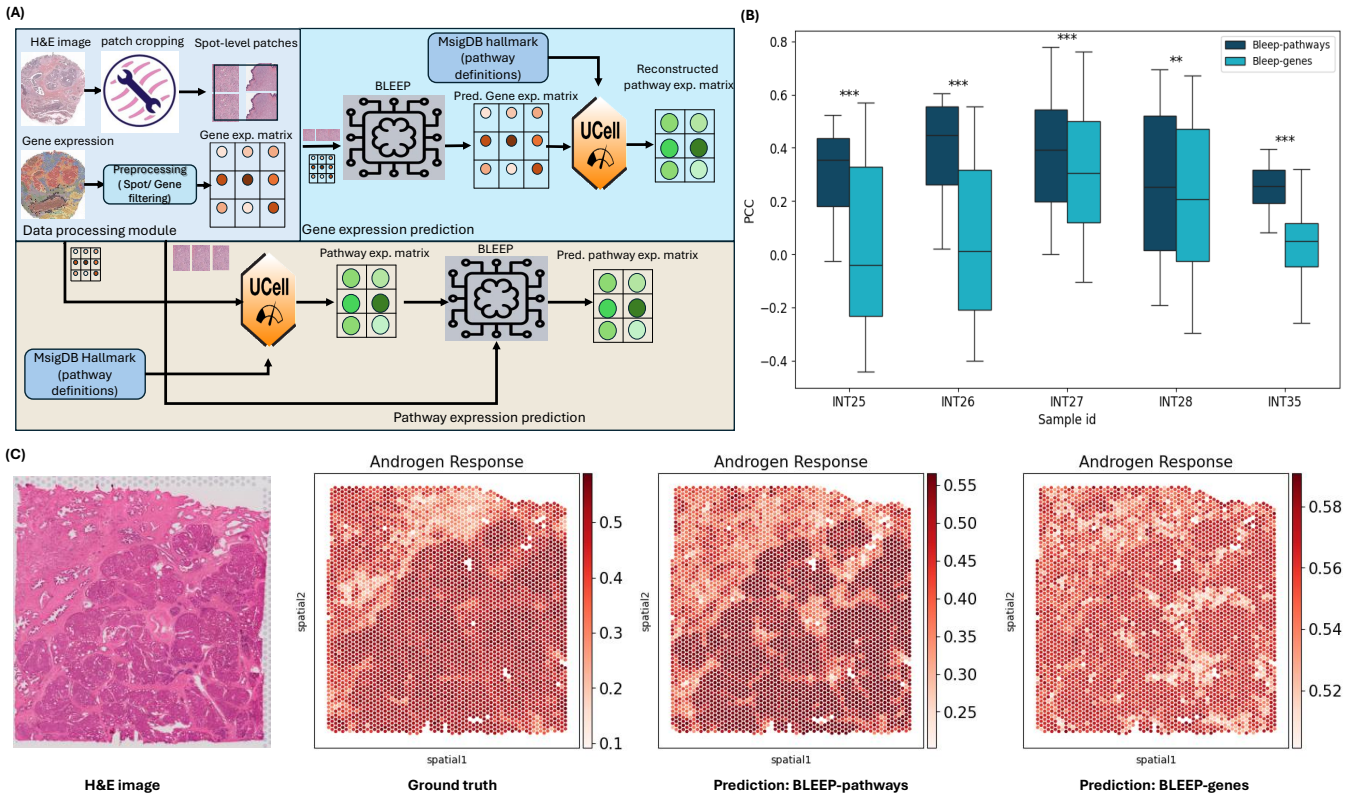

Fig. 2: **(A)** Overview of the baseline method where spot-level data is initially processed (data processing module). Then, BLEEP is first trained on paired H&E images and gene expression data to predict the gene expression. Then predicted gene expression is passed to the UCell for obtaining the pathway expression. Whereas in pathway expression prediction (bottom half), gene expression matrix is first converted into pathway expression matrix followed by training BLEEP to predict the pathway expression. Finally, the pathway expression matrix from gene expression prediction (BLEEP-genes) is compared with the predicted pathway expression (BLEEP-pathways) obtained from pathway expression prediction. The "*" symbol shows the magnitude of p-value calculated from paired t-test of two PCC distributions ("***" for p-value <0.001, "**" for p-value <0.01). **(B)** PCC distribution plot of BlEEP (pathways) versus BLEEP (genes) for all samples. **(C)** Spatial visualization of ground truth pathway expression from UCell, predicted *Androgen response* pathway expression from BLEEP(pathways) and BLEEP (genes) having PCC value of 0.67 and 0.37, respectively, with H&E image (**INT28** sample.)

on more than $500,000$ H&E stained whole slide histology images), in our experiments which provides an embedding of 1538-dimensional vector for each image patch. All performance metrics are evaluated on test data. For pathway prediction experiments, we used the criterion of early stopping with a patience level of 5 to monitor the improvements in validation loss to prevent overfitting. To access the quality of predicted expression, we used the Pearson correlation coefficient (PCC), as well as spatial visualization of ground truth and predicted expression.

III. Results

A. Bimodal contrastive learning for gene versus pathway prediction on prostate cancer data

Initially, we applied BLEEP to perform gene prediction and pathway prediction on the small prostate cancer dataset to establish the utility of pathway-level prediction while using the same underlying contrastive learning approach. At first, spot-level paired gene and H&E image data is processed and BLEEP having ResNet50 as an image encoder is trained on paired data to predict gene expression. Then, predicted gene expression for the test sample is converted into pathway expression using MSigDB hallmark pathway definitions using Ucell (we call this approach "BLEEP-genes"). In parallel, we passed the processed gene expression data along with the MSigDB pathway definitions to UCell to obtain pathway expression data. Then, we trained BLEEP having ResNet18 as an image encoder with paired pathway expression data to predict the pathway expression (we call this approach "BLEEP-pathways"). Finally, reconstructed pathway expression obtained from BLEEP-genes and predicted pathway expression from BLEEP-pathways are compared (Fig. 2A). To compare the performance of BLEEP-pathways with BLEEP-

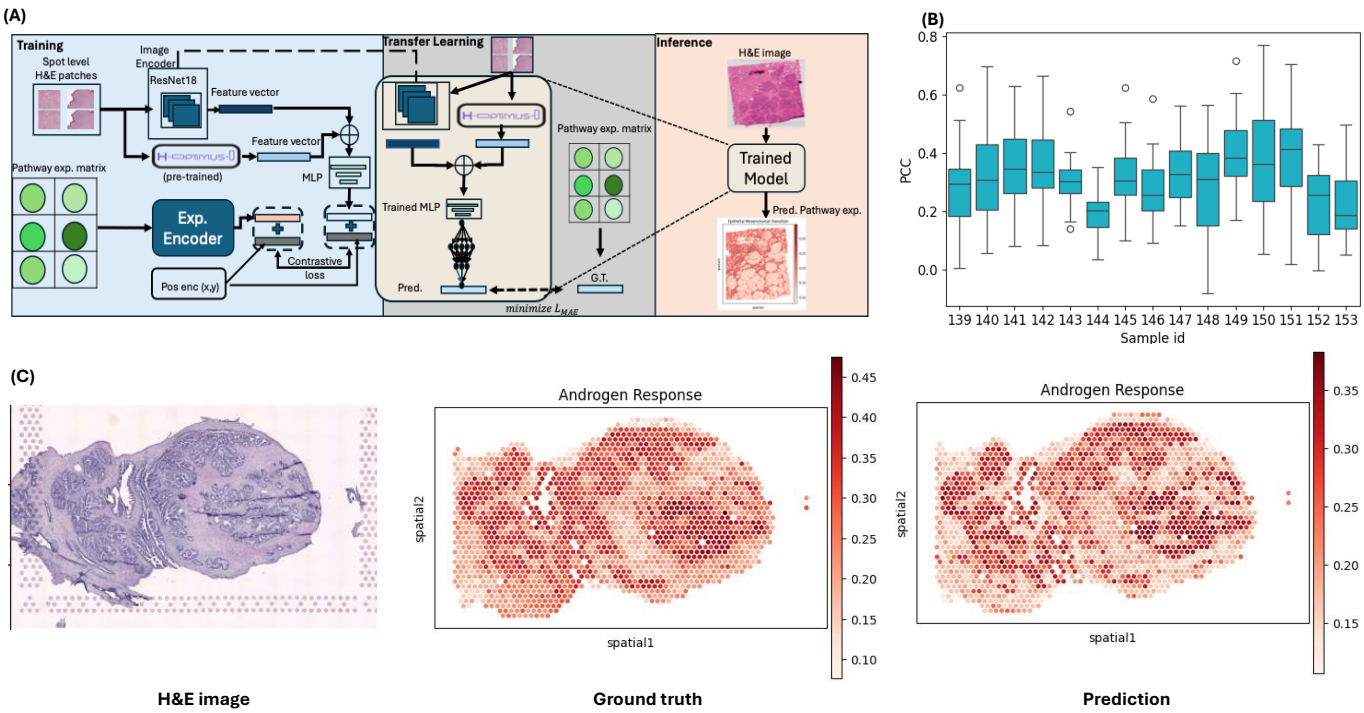

Fig. 3: **(A)** Overview of the proposed two-stage learning framework. At first, ResNet18 is trained on combined image features that are obtained from ResNet18 and *H-optimus-0* foundation model using bimodal contrastive learning. In second stage, a DNN is trained on image features extracted from trained stage 1 model to predict pathway expression by minimizing MAE. During inference time, only H&E image is used to predict pathway expression**(B)** PCC distribution plot of all samples evaluated using leave-one-out cross validation. **(C)** test H&E image and spatial visualization of the Androgen Response pathway shows the ground truth pathway expression—obtained from UCell—alongside the predicted pathway expression from the trained two stage model. The predicted expression achieved a PCC value of 0.76 (sample ID=150)

genes, we first computed the PCC between the ground truth pathway expression of the test sample—and both the reconstructed and predicted pathway expressions. Overall, BLEEP-pathways clearly outperformed BLEEP-genes (Fig. 2B). For example, in the case of Androgen Response (Fig. 2C), BLEEP-pathways has PCC of 0.67 while BLEEP-genes with 0.37. On the other hand, BLEEP-pathways and BLEEP-genes achieved a max PCC value of 0.80 and 0.76 for the Epithelial Mesenchymal Transition (EMT) pathway in "INT27" sample, respectively, which plays crucial role in tumour metastasis [35]. This depicts that genes in the EMT pathway are well correlated with the histology image features, which helped in predicting the genes and its constituent pathway very well in both scenarios. Additionally, we computed the PCC between top 5 highly expressed and highly variable pathways (Supplementary Fig. 1). This consistency in pathway prediction depict that it is more meaningful to perform pathway prediction over gene prediction as it leverages the advantage of focusing on aggregate biological signal rather than individual gene fluctuations.

B. Improved pathway prediction model deployability via two-stage learning framework

One of the limitations of the BLEEP model is that it requires the training data (latent projections) to infer the predicted pathway expression for the test data. BLEEP predicts expression by projecting test image into the embedding space and retrieving the average of top k matched candidates from the reference (training) dataset. However, this might not always feasible considering future deployability of these models in translational settings. Thus, we propose our own method by replacing the inference part of BLEEP with a two-stage learning framework (Fig. 3A), which is explained as follows. First, we trained BLEEP having ResNet18 as an image encoder with *H-optimus-0* using contrastive learning. In the second stage, we trained a DNN model by giving these features obtained from stage-1 models as an input to predict a vector of pathway expression (multivariate regression). This approach minimizes the mean absolute error (MAE) between the ground truth and prediction, which is calculated as $MAE(y, y') = \frac{\sum_{i=1}^{N} \|y - y'\|_1}{N}$, where y is ground truth pathway expression obtained from UCell and y' is the predicted pathway expression

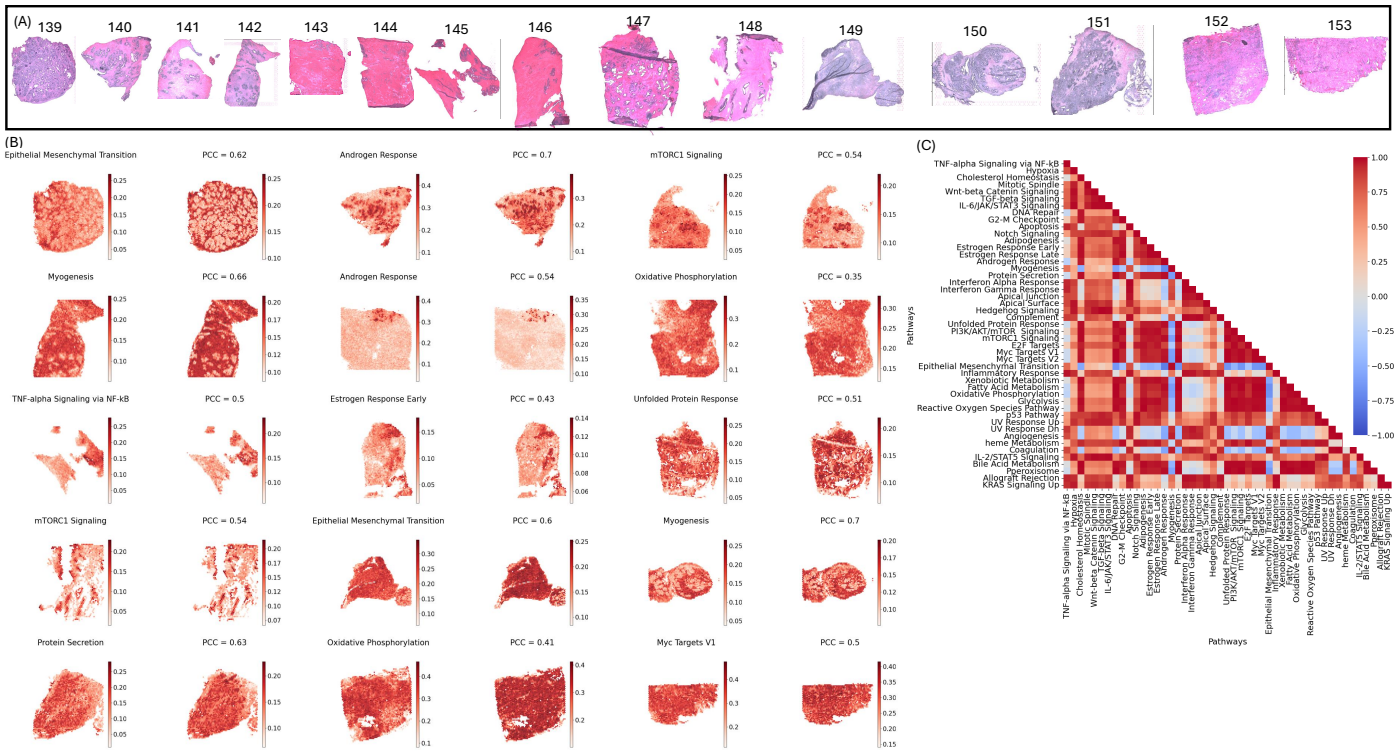

Fig. 4: **(A)** Thumbnails of H&E images. **(B)** Examples of selected pathways (from top 10 highly predicted based on PCC values) within tissues from all samples of large prostate cancer dataset. **(C)** Heatmap of inter-pathway correlation calculated using predicted pathway expression obtained from the proposed two-stage model to identify positively correlated pathways (shown for sample ID=150).

for a batch of N samples. Additionally, here we have not added the positional encodings, as we are obtaining the image features from the trained model which has already learned the spatial similarities between those modalities. During inference, we used the trained ResNet18, *H-optimus-0* (stage 1), and the trained DNN model (stage 2) to predict the patch-level pathway expression. We evaluated this approach on large data from 15 samples of human prostate adenocarcinoma, classified as Gleason 4+3, grade group 3 [32]. Using leave-one-out cross validation, we evaluated the proposed model stage-wise, i.e., stage-1 followed by stage-2. The results for stage-1 model (Supplementary Fig. 2A) in terms of PCC between ground truth pathway expression and predicted pathway expression of all pathways for each test sample are shown in Supplementary Fig. 2B. H&E image and spatial visualization of the best predicted pathway in stage-1 (Myogenesis with PCC value of 0.77) is shown in (Supplementary Fig. 2C). Also, high and low variation of PCC distribution in different samples depicts substantial differences in model performance across samples, even though all samples originate from the same patient. This also shows that intrinsic heterogeneity within the tissue or local variations in image characteristics could influence the model's performance. Additionally, we evaluated stage-1 only approach on small prostate cancer data where EMT pathway (in INT27 sample) achieved a maximum PCC of 0.81 (Supplementary Fig. 3). We also conducted an ablation study of the proposed two-stage learning framework by performing the experiments on BLEEP with and without *H-optimus-0* foundation model, keeping rest of the architecture same (small prostate cancer and large prostate cancer dataset are shown in Supplementary Fig. 4 and Supplementary Fig. 5, respectively). These results show that our proposed approach achieved largely similar results in terms of pathway prediction without any notable performance drop irrespective of inclusion of the foundation model.

For the second stage, i.e., training an MLP for direct prediction, PCC distribution between ground truth and predicted pathway expression is shown in Fig. 3B. A similar trend of PCC distribution is observed when compared to Supplementary Fig. 2B, as the bimodal contrastive loss model remained same in both approaches. However, the maximum PCC value achieved in this approach is observed for Androgen Response pathway with the value of 0.76 on sample (ID=150). H&E image with spatial visualization of ground truth and predicted

pathway expression is shown in Fig. 3C. This shows that our proposed approach not only aligns the image-expression modality using contrastive learning but also learned a meaningful representation of image features with pathway expression through a simple DNN model.

We also analysed the pathways predicted well within each tissue sample and across all tissue samples. Thumbnails of H&E images used for evaluating the proposed two-stage approach are shown in Fig. 4(A). Pathways (both ground truth and prediction)along with their PCC which are predicted well within tissue samples is shown in Fig. 4(B). It can be seen that some slides have low PCC even for their well predicted pathways but most of the slides have exhibited a good correlation and highlighted the active regions of that pathway on the tissue. Only two pathways, i.e., Androgen Response and mTORC1 signaling are predicted well in all tissue samples (shown in Supplementary Fig. 6). We also computed the inter-pathway correlation (PCC between each pathway with other pathway across all spots) using the predicted pathway expression obtained from the proposed two-stage learning framework (heatmap for the best predicted sample (ID=150) is shown in Fig. 4(C)). For this test sample, we first observed positively correlated pathways using ground truth pathway expression data as shown in Supplementary Fig. 7. Almost similar pattern of positive correlation for predictions can be seen in Fig. 4(C). For example, for both ground truth and prediction, p53 pathway and mTORC1 Signaling pathway are positively correlated. The p53 is a tumour suppressor pathway and maintains the cell integrity by activating cell cycle arrest, DNA repair, or apoptosis in case of DNA damage. The mTORC1 is the regulatory pathway for cell growth and metabolism. Thus, in Gleason 4+3 prostate cancers, both pathways can co-exist to promote tumour growth as they contribute in cell survival and proliferation. Furthermore, EMT and Myogenesis pathways are positively correlated in both ground truth (Supplementary Fig. 7) and predictions (Fig. 4(C)), co-exist in prostate cancers and contribute to tumour diversity, specifically, with respect to metasttaic potential and differentiation state. This co-existence is visually validated further when we test sample (id=150) in the proposed approach, these two pathways marked the same tissue areas with the PCC of 0.71 and 0.69, respectively.

C. Pathway expression mapping in prostate cancer using TCGA whole slide images

We extended the application of our model to the cancer genome atlas (TCGA) prostate cancer dataset, which is most widely used public cancer data resource. We acquired whole slide images (WSIs) of 8 patients which have both normal and prostate tumour tissue slides from TCGA data portal. Then, we used the best performing weights of our proposed model to predict the patch-wise pathway expression using only WSIs. We selected three representative pathways, i.e, Hypoxia, Monogenesis, and Androgen Response, for patch-wise visualisation. Hypoxia is known to promote tumour progression and metastatic spread [36]–[39] while activation of androgen receptor signalling drives proliferation and survival of prostate cancer cells [40]. Finally, stromal changes in prostate carcinoma including the activation of the myogenesis programme in smooth muscle cells are now recognised to correlate with poor outcomes [41]. Fig. 5 shows the results of the predictions obtained for these three pathways on two TCGA tumour patients (row-wise). The results for remaining tumour patients and their normal tissue sections are shown in Supplementary Fig. 8 and Supplementary Fig. 9, respectively. It can be seen that there is strong expression of hypoxia in tumour samples when compared with normal samples. Also, Androgen response and Myogenesis shows an inverse behavior with respect to their expression values in both normal or tumour samples. However, the expression intensity of Myogenesis is higher in tumour slides as compared with normal slides.

D. Prediction of hypoxia signatures from H&Es of brain tumour patient samples

Hypoxia is one of the hallmarks of cancer and it is caused by the unbalanced oxygen availability due to insufficient blood supply and the excessive metabolic demand of tumour cells. Hypoxia causes substantial pressure on the tumour and its microenvironment leading to immune evasion and metabolic adaptation of neoplastic cells, maintenance of tumour cell stemness and resistance to treatment. Hypoxia can be traced in tissues using Pimonidazole (PIMO) which binds selectively thiol-containing proteins or forms metabolites after reduction of its nitro group macromolecules when oxygen concentration falls below 10 mmHg [42]

Here we applied DeepPathway to samples from patients with glioblastoma (GBM) who received PIMO as part of a clinical trial [43]. Consecutive sections were stained for H&E and PIMO as ground truth and analysed with DeepPathway. We first acquired the 3 ST-based glioblastoma samples from Ravi et al. [44] that have the H&E image resolution of around 0.5 MPP. Then we performed preprocessing of the H&E images after applying the Reinhard H&E image stain normalization to obtain

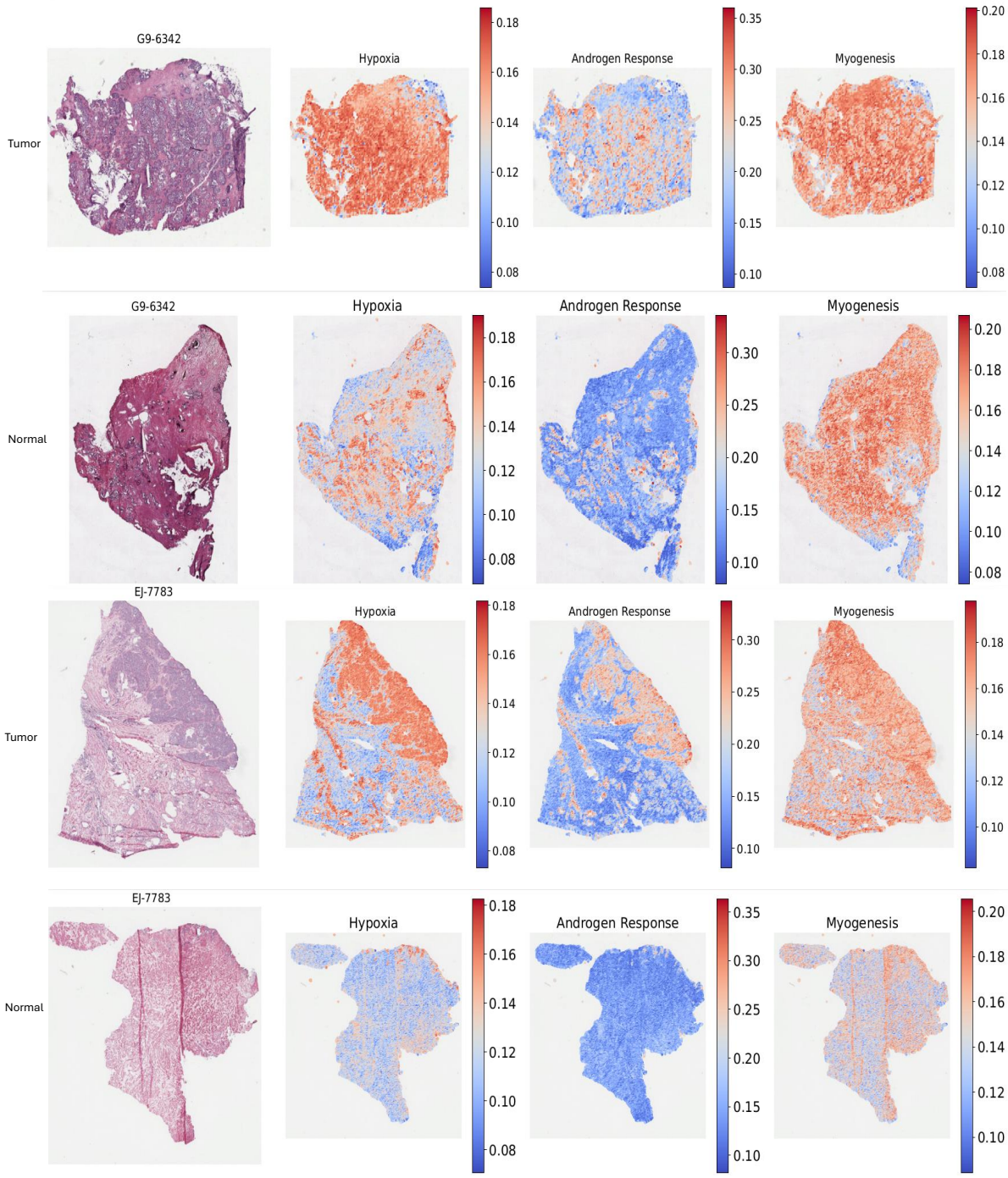

Fig. 5: Spatial visualization of predictions obtained from the proposed model for Hypoxia, Androgen Response, and Monogenesis pathways for TCGA patients having both tumour and normal tissue sections.

the same histology colors distribution and for smooth model training. Gene expression processing and pathway quantification, was performed as described in Section II. We trained our proposed two-stage model on those three GBM samples by taking the best performing model's weights after training on the large prostate cancer dataset. After training the first-stage model, i.e., BLEEP with *H-optimus-0*, we used the same GBM samples, to train the second stage model of the proposed method. After training both models, we used in-house brain-tumour patients H&Es to predict the hypoxia pathway expression. During inference, H&E image patches are cropped as follow. Starting from the (0,0) location of the image, non-overlapping patches of 224×224 were taken and passed to the image processing pipeline to filter out non-informative patches. Only informative patches were then

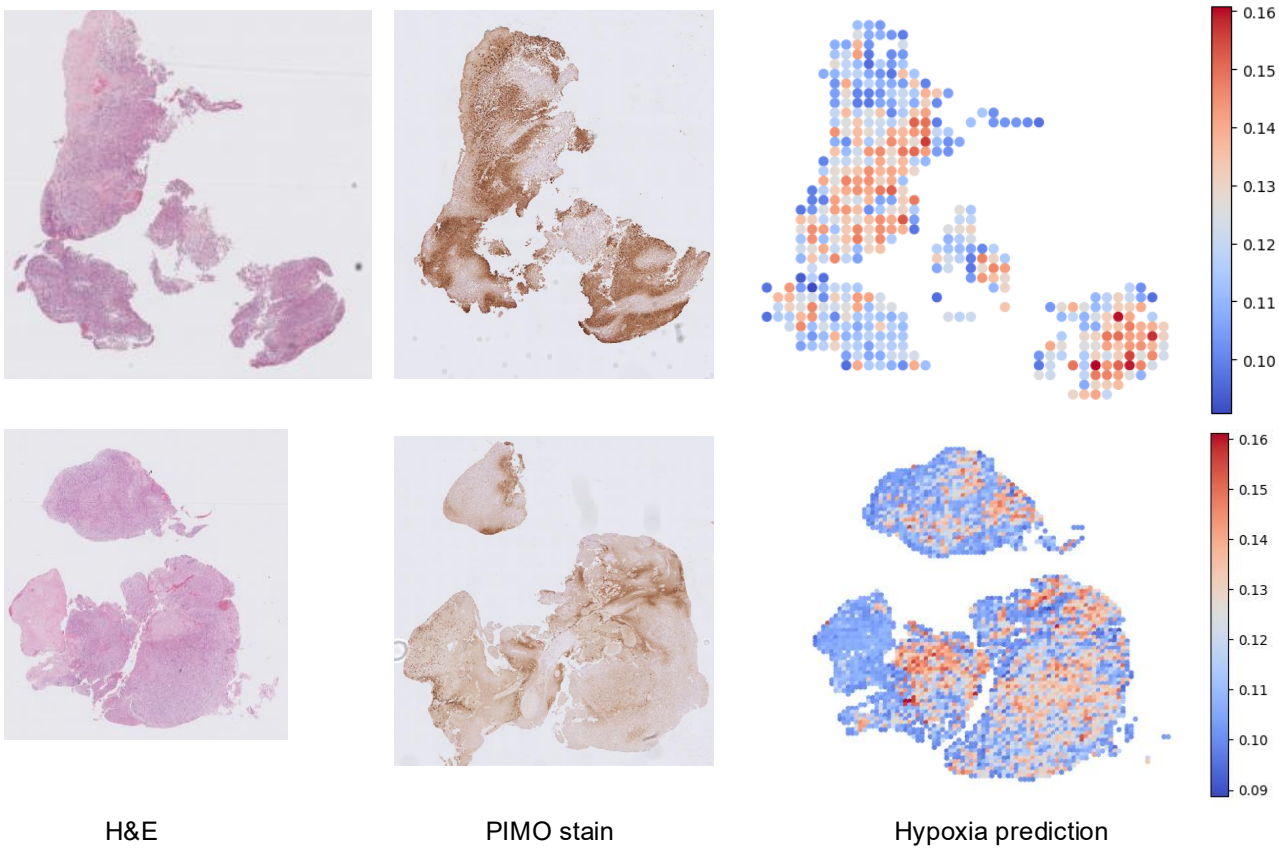

Fig. 6: Visualization of H&E image, PIMO staining validation and predicted Hypoxia signatures from H&Es of brain tumour patients using the proposed model.

passed to the trained model to infer pathway expression and their locations were preserved for visualization. Results for two brain tumour patients having hypoxia clearly marked with PIMO stain are shown in Fig. 6. This approach provided a robust validation of DeepPathway and of its translational potential, providing a good visual concordance with the PIMO stain despite being trained on limited and low-quality ST brain tumour specific data.

IV. Discussion and Conclusion

In this paper, we proposed DeepPathway, a novel approach for the prediction of pathway expression from H&E images which could have potential translational applications for diagnostics and treatment monitoring in clinical settings. Our method builds on existing work for the prediction of gene expression from H&E images, utilising ST data (gene expression with co-registered images). For both gene and pathway expression prediction tasks, quality of images and their preprocessing could have significant impact on the model development and performance. DL-based expression prediction models mainly use spot-level H&E image patch, which is cropped around the center of the spot. However, in ST data, tissue sections may have variable image resolution due to the technical differences between experiments or variable spot diameters even using the same technology like Visium 10X. This leads to H&E patch sizes with inconsistent morphological features, i.e., cell shapes, density, and their surroundings. Additionally, in clinical settings, tissue H&Es without ST data can be stored in standard imaging formats like portable network graphics (PNG), where the image resolution in microns per pixel (MPP) is unavailable. To achieve the constant-sized image patch which is not affected by the spot diameter or the type of image, we have used TIA toolbox which provides a function called *VirtualWSIReader* that can take any type of H&E image and provide H&E patches of same size (say 224×224) with same MPP resolution (say 0.5). Additionally, large-scale ST datasets, e.g. HEST-1k provides normalized and transformed WSIs in a pyramidal objects (TIFF) along with the metadata to obtain smooth image. However, in future, a multi-resolution DL-based framework for leveraging different magnifications can be developed to achieve more generalizable results by incorporating the images at any resolution.

Besides images, preprocessing and batch correction of multi-sample gene expression data might be needed before training the DL models for gene or pathway expression prediction. There are some statistical methods like Harmony [45] or probabilistic methods like single-cell variational inference [46] that uses the autoencoders based approach to remove batch effects. These methods transforms the gene counts data into another embeddings space which may contain negative gene expression values. However, in BLEEP-genes pipeline of our method, we have not applied any batch correction method for the sake of obtaining positive valued gene expression prediction that are used to obtain reconstructed pathway expression. We also believe that transforming the gene expression data into pathway expression data using UCell reduces the batch effects because UCell is a rank-based method, which can handle the global shifts in gene expression levels caused by batch effects. Thus, our results show that pathway prediction achieved more smooth and robust prediction results when compared to the gene prediction across all samples.

Also, recently H&E foundation models are widely applied to extract rich histology features as they are trained on enormous amount of WSIs. We have used the first version of optimus foundation model, i.e., *H-optimus-0*, however, there are other foundation models [47], [48] along with *H-optimus-1*[2] (the model is trained on billions of histology images sampled from over 1 million slides of more than $800,000$ patients) can be used as a future work. These models could lead to development of a pathway prediction foundation model that can be used on different tissues, i.e., brain, liver, or lungs, etc.

Furthermore, we show the utility of our model on TCGA data where there are only H&E images available besides bulk RNA data. Our method can show spatially localized expression of different pathways using H&Es, differentiating the behavior of different pathways on tumour vs normal tissue slides for same patients. This can add further insight into the analysis of those samples, and could be explored further in conjunction with the bulk RNA data and other metadata available with those samples as a future application.

Lastly, we focused on hypoxia in glioblastoma as a paradigm of hypoxic tumour. Any implementation of computational methods to identify hypoxia in H&E image will improve the ability to stratify patients at risk of a worse outcome. Noteworthy, hypoxia can only be identified using the stain for HIF-1α as a surrogate biomarker of low oxygen concentrations. At the same time, this was an excellent opportunity for validation and potential clinical utility of our method, as we have access to these precious human brain tumour samples which were collected in a way that they could be stained with PIMO as a ground truth. Our method makes good visual concordance with the PIMO stain despite being trained on limited and low-quality ST brain tumour specific data. These analyses can be further extended and validated to make the case for potential clinical application.

Overall, to best of our knowledge, this is the first approach to predict pathways from H&E images, and we believe with more training data, better quality images, refined pathway definitions, it will lead to further improvement and wider translational impact.

Acknowledgment: Muhammad Ahtazaz Ahsan was funded by MRC DTP award (MR/W007428/1).

We would like to thank all authors of the public data that is used in this study. We also thank Gabriella Forte for her support in the preparation of the in-house brain tumour histology slides and PIMO staining data.

REFERENCES

[1] X. Zhuang, "Spatially resolved single-cell genomics and transcriptomics by imaging," *Nature methods*, vol. 18, no. 1, pp. 18–22, 2021.

[2] V. Marx, "Method of the year: spatially resolved transcriptomics," *Nature methods*, vol. 18, no. 1, pp. 9–14, 2021.

[3] S. G. Rodriques, R. R. Stickels, A. Goeva, C. A. Martin, E. Murray, C. R. Vanderburg, J. Welch, L. M. Chen, F. Chen, and E. Z. Macosko, "Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution," *Science*, vol. 363, no. 6434, pp. 1463–1467, 2019.

[4] B. Zheng and L. Fang, "Spatially resolved transcriptomics provide a new method for cancer research," *Journal of Experimental & Clinical Cancer Research*, vol. 41, no. 1, p. 179, 2022.

[5] B. He, L. Bergenstråhle, L. Stenbeck, A. Abid, A. Andersson, Å. Borg, J. Maaskola, J. Lundeberg, and J. Zou, "Integrating spatial gene expression and breast tumour morphology via deep learning," *Nature biomedical engineering*, vol. 4, no. 8, pp. 827–834, 2020.

[6] R. Xie, K. Pang, S. Chung, C. Perciani, S. MacParland, B. Wang, and G. Bader, "Spatially resolved gene expression prediction from histology images via bi-modal contrastive learning," *Advances in Neural Information Processing Systems*, vol. 36, pp. 70626–70637, 2023.

[7] Q. Wang, W.-j. Chen, J. Su, G. Wang, and Q. Song, "Heclip: histology-enhanced contrastive learning for imputation of transcriptomics profiles," *Bioinformatics*, p. btaf363, 2025.

[8] T. Monjo, M. Koido, S. Nagasawa, Y. Suzuki, and Y. Kamatani, "Efficient prediction of a spatial transcriptomics profile better characterizes breast cancer tissue sections without costly experimentation," *Scientific reports*, vol. 12, no. 1, p. 4133, 2022.

[9] S. Bae, H. Choi, and D. S. Lee, "Discovery of molecular features underlying the morphological landscape by integrating spatial transcriptomic data with deep features of tissue images," *Nucleic acids research*, vol. 49, no. 10, pp. e55–e55, 2021.

[2]https://huggingface.co/bioptimus/H-optimus-0-1

[10] Y. Jia, J. Liu, L. Chen, T. Zhao, and Y. Wang, "Thitogene: a deep learning method for predicting spatial transcriptomics from histological images," *Briefings in Bioinformatics*, vol. 25, no. 1, p. bbad464, 2024.

[11] X. Xiao, Y. Kong, R. Li, Z. Wang, and H. Lu, "Transformer with convolution and graph-node co-embedding: an accurate and interpretable vision backbone for predicting gene expressions from local histopathological image," *Medical Image Analysis*, vol. 91, p. 103040, 2024.

[12] Y. Zeng, Z. Wei, W. Yu, R. Yin, Y. Yuan, B. Li, Z. Tang, Y. Lu, and Y. Yang, "Spatial transcriptomics prediction from histology jointly through transformer and graph neural networks," *Briefings in Bioinformatics*, vol. 23, no. 5, p. bbac297, 2022.

[13] M. Pang, K. Su, and M. Li, "Leveraging information in spatial transcriptomics to predict super-resolution gene expression from histology images in tumors," *bioRxiv*, pp. 2021–11, 2021.

[14] C.-H. Huang, Y. Park, J. Pang, and J. R. Bienkowska, "Single-cell gene expression prediction using h and e images based on spatial transcriptomics," in *Medical Imaging 2023: Digital and Computational Pathology*, vol. 12471, pp. 17–25, SPIE, 2023.

[15] M. Stevens, A. Nanou, L. W. Terstappen, C. Driemel, N. H. Stoecklein, and F. A. Coumans, "Stardist image segmentation improves circulating tumor cell detection," *Cancers*, vol. 14, no. 12, p. 2916, 2022.

[16] J. T. Ash, G. Darnell, D. Munro, and B. E. Engelhardt, "Joint analysis of expression levels and histological images identifies genes associated with tissue morphology," *Nature communications*, vol. 12, no. 1, p. 1609, 2021.

[17] W. Min, Z. Shi, J. Zhang, J. Wan, and C. Wang, "Multimodal contrastive learning for spatial gene expression prediction using histology images," *Briefings in Bioinformatics*, vol. 25, no. 6, p. bbae551, 2024.

[18] M. Pizurica, Y. Zheng, F. Carrillo-Perez, H. Noor, W. Yao, C. Wohlfart, A. Vladimirova, K. Marchal, and O. Gevaert, "Digital profiling of gene expression from histology images with linearized attention," *Nature Communications*, vol. 15, no. 1, p. 9886, 2024.

[19] Z. Shi, F. Zhu, and W. Min, "Inferring multi-slice spatially resolved gene expression from h&e-stained histology images with stmcl," *Methods*, vol. 234, pp. 187–195, 2025.

[20] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, *et al.*, "Learning transferable visual models from natural language supervision," in *International conference on machine learning*, pp. 8748–8763, PmLR, 2021.

[21] M. Andreatta and S. J. Carmona, "Ucell: Robust and scalable single-cell gene signature scoring," *Computational and structural biotechnology journal*, vol. 19, pp. 3796–3798, 2021.

[22] S. Aibar, S. Aerts, *et al.*, "Aucell: Analysis of 'gene set'activity in single-cell rna-seq data," 2016.

[23] C. Saillard, R. Jenatton, F. Llinares-López, Z. Mariet, D. Cahané, E. Durand, and J.-P. Vert, "H-optimus-0," 2024.

[24] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*, pp. 1597–1607, PmLR, 2020.

[25] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," *Advances in neural information processing systems*, vol. 33, pp. 18661–18673, 2020.

[26] B. Dai and D. Lin, "Contrastive learning for image captioning," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[27] H. Zhang, J. Y. Koh, J. Baldridge, H. Lee, and Y. Yang, "Cross-modal contrastive learning for text-to-image generation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 833–842, 2021.

[28] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[29] G. Jaume, P. Doucet, A. Song, M. Y. Lu, C. Almagro Pérez, S. Wagner, A. Vaidya, R. Chen, D. Williamson, A. Kim, *et al.*, "Hest-1k: A dataset for spatial transcriptomics and histology image analysis," *Advances in Neural Information Processing Systems*, vol. 37, pp. 53798–53833, 2025.

[30] A. Goode, B. Gilbert, J. Harkes, D. Jukic, and M. Satyanarayanan, "Openslide: A vendor-neutral software foundation for digital pathology," *Journal of pathology informatics*, vol. 4, no. 1, p. 27, 2013.

[31] M. Humphries, P. Maxwell, and M. Salto-Tellez, "Qupath: The global impact of an open source digital pathology system," *Computational and structural biotechnology journal*, vol. 19, pp. 852–859, 2021.

[32] A. Erickson, M. He, E. Berglund, M. Marklund, R. Mirzazadeh, N. Schultz, L. Kvastad, A. Andersson, L. Bergenstråhle, J. Bergenstråhle, *et al.*, "Spatially resolved clonal copy number alterations in benign and malignant tissue," *Nature*, vol. 608, no. 7922, pp. 360–367, 2022.

[33] A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J. P. Mesirov, and P. Tamayo, "The molecular signatures database hallmark gene set collection," *Cell systems*, vol. 1, no. 6, pp. 417–425, 2015.

[34] J. Pocock, S. Graham, Q. D. Vu, M. Jahanifar, S. Deshpande, G. Hadjigeorghiou, A. Shephard, R. M. S. Bashir, M. Bilal, W. Lu, *et al.*, "Tiatoolbox as an end-to-end library for advanced tissue image analytics," *Communications medicine*, vol. 2, no. 1, p. 120, 2022.

[35] V. Mittal, "Epithelial mesenchymal transition in tumor metastasis," *Annual Review of Pathology: Mechanisms of Disease*, vol. 13, no. 1, pp. 395–412, 2018.

[36] P. Manescu, J. Geradts, and D. Fernandez-Reyes, "Computational pathology detection of hypoxia-induced morphologic changes in breast cancer," *The American Journal of Pathology*, vol. 195, no. 4, pp. 663–670, 2025.

[37] L. Feldman, "Hypoxia within the glioblastoma tumor microenvironment: a master saboteur of novel treatments," *Frontiers in immunology*, vol. 15, p. 1384249, 2024.

[38] Z. Chen, F. Han, Y. Du, H. Shi, and W. Zhou, "Hypoxic microenvironment in cancer: molecular mechanisms and therapeutic interventions," *Signal transduction and targeted therapy*, vol. 8, no. 1, p. 70, 2023.

[39] B. Muz, P. De La Puente, F. Azab, and A. Kareem Azab, "The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy," *Hypoxia*, pp. 83–92, 2015.

[40] K. Fujita and N. Nonomura, "Role of androgen receptor in prostate cancer: a review.," *World Journal of Men's Health*, vol. 37, no. 3, 2019.

[41] M. Rasmussen, J. Fredsøe, P. V. Salachan, M. P. L. Blanke, S. H. Larsen, B. P. Ulhøi, J. B. Jensen, M. Borre, and K. D. Sørensen, "Stroma-specific gene expression signature identifies prostate cancer subtype with high recurrence risk," *NPJ Precision Oncology*, vol. 8, no. 1, p. 48, 2024.

[42] S. Kizaka-Kondoh and H. Konse-Nagasawa, "Significance of nitroimidazole compounds and hypoxia-inducible factor-1 for imaging tumor hypoxia," *Cancer science*, vol. 100, no. 8, pp. 1366–1373, 2009.

[43] M. Waqar, F. Roncaroli, I. Djoukhadar, L. Akkari, C. O'Leary, L. Hewitt, G. Forte, R. Jackson, E. Hessen, L. Withington, *et al.*, "Study protocol: preoperative brain irradiation in glioblastoma

(pobig)–a phase i trial," *Clinical and translational radiation oncology*, vol. 39, p. 100585, 2023.

[44] V. M. Ravi, P. Will, J. Kueckelhaus, N. Sun, K. Joseph, H. Salié, L. Vollmer, U. Kuliesiute, J. von Ehr, J. K. Benotmane, *et al.*, "Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma," *Cancer cell*, vol. 40, no. 6, pp. 639–655, 2022.

[45] I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P.-r. Loh, and S. Raychaudhuri, "Fast, sensitive and accurate integration of single-cell data with harmony," *Nature methods*, vol. 16, no. 12, pp. 1289–1296, 2019.

[46] R. Lopez, J. Regier, M. B. Cole, M. I. Jordan, and N. Yosef, "Deep generative modeling for single-cell transcriptomics," *Nature methods*, vol. 15, no. 12, pp. 1053–1058, 2018.

[47] R. J. Chen, T. Ding, M. Y. Lu, D. F. Williamson, G. Jaume, A. H. Song, B. Chen, A. Zhang, D. Shao, M. Shaban, *et al.*, "Towards a general-purpose foundation model for computational pathology," *Nature Medicine*, vol. 30, no. 3, pp. 850–862, 2024.

[48] E. Vorontsov, A. Bozkurt, A. Casson, G. Shaikovski, M. Zelechowski, K. Severson, E. Zimmermann, J. Hall, N. Tenenholtz, N. Fusi, *et al.*, "A foundation model for clinical-grade computational pathology and rare cancers detection," *Nature medicine*, vol. 30, no. 10, pp. 2924–2935, 2024.