Histopathological Image Classification with Cell Morphology Aware Deep Neural Networks

Andrey Ignatov [1,2]
andrey@vision.ee.ethz.ch

Josephine Yates [2]
josephine.yates@inf.ethz.ch

Valentina Boeva [2]
valentina.boeva@inf.ethz.ch

ETH Zurich, Computer Vision Lab [1] & Computational Cancer Genomics Group [2]

Abstract

Histopathological images are widely used for the analysis of diseased (tumor) tissues and patient treatment selection. While the majority of microscopy image processing was previously done manually by pathologists, recent advances in computer vision allow for accurate recognition of lesion regions with deep learning-based solutions. Such models, however, usually require extensive annotated datasets for training, which is often not the case in the considered task, where the number of available patient data samples is very limited. To deal with this problem, we propose a novel DeepCMorph model pre-trained to learn cell morphology and identify a large number of different cancer types. The model consists of two modules: the first one performs cell nuclei segmentation and annotates each cell type, and is trained on a combination of 8 publicly available datasets to ensure its high generalizability and robustness. The second module combines the obtained segmentation map with the original microscopy image and is trained for the downstream task. We pre-trained this module on the Pan-Cancer TCGA dataset consisting of over 270K tissue patches extracted from 8736 diagnostic slides from 7175 patients. The proposed solution achieved a new state-of-the-art performance on the dataset under consideration, detecting 32 cancer types with over 82% accuracy and outperforming all previously proposed solutions by more than 4%. We demonstrate that the resulting pre-trained model can be easily fine-tuned on smaller microscopy datasets, yielding superior results compared to the current top solutions and models initialized with ImageNet weights. The codes and pre-trained models presented in this paper are available at: https://github.com/aiff22/DeepCMorph

1. Introduction

Whole slide images (WSI), which are microscopy photos of stained tissue regions, are an important source of data for digital pathology. They contain rich visual information about tissue cell type composition, cell states, immune system activity, vasculature, cell abnormalities, *etc*., and thus are widely used for the diagnosis and treatment selection for patients. Over the past years, histopathology image analysis has proven to be useful for such tasks as cancer screening and classification [6, 34, 44, 72, 78, 97], tumor grading [8, 9, 45, 49, 61, 93], patient survival prediction [2, 21, 68, 77, 91, 96], mitosis detection [5, 55, 85], gene mutant prediction [13, 60, 95], tumor immune infiltration quantification [1, 79, 94], gene expression [15, 66] and biomarker prediction [51, 84], therefore making it crucial to develop performant WSI analysis tools.

While initially whole slide images were analyzed manually by pathologists, during the past years machine learning based approaches were gradually adopted for this task as they allowed for more accurate and efficient data processing. The first attempts were based on traditional computer vision methods [41, 70, 80] using handcrafted feature descriptors such as histograms of oriented gradients, Gabor filters, SIFT features, *etc*. These approaches were later significantly outperformed by convolutional neural network (CNN) based models [7, 14, 73, 81, 90] that were trained in an end-to-end fashion and did not require any manual feature engineering. Recently, vision transformer (ViT) architectures were also proposed for WSI classification and retrieval problems, often demonstrating superior performance on these tasks [29, 36, 48, 92, 98]. A typical approach to adapt these models for WSI analysis has been to use an existing neural network architecture and train it from scratch or initialize the model with weights obtained on the ImageNet dataset [31, 39, 44, 60, 66, 89]. Since numerous WSI datasets contain only tens or hundreds of samples, a significant limitation arises: while larger network architectures with millions of parameters may have the potential to learn more robust features and attain greater accuracy when trained on histopathology data, their practical performance is often constrained by insufficient training data and overfitting problems [39]. One recently presented approach to deal with this problem is to pre-train models in an unsupervised

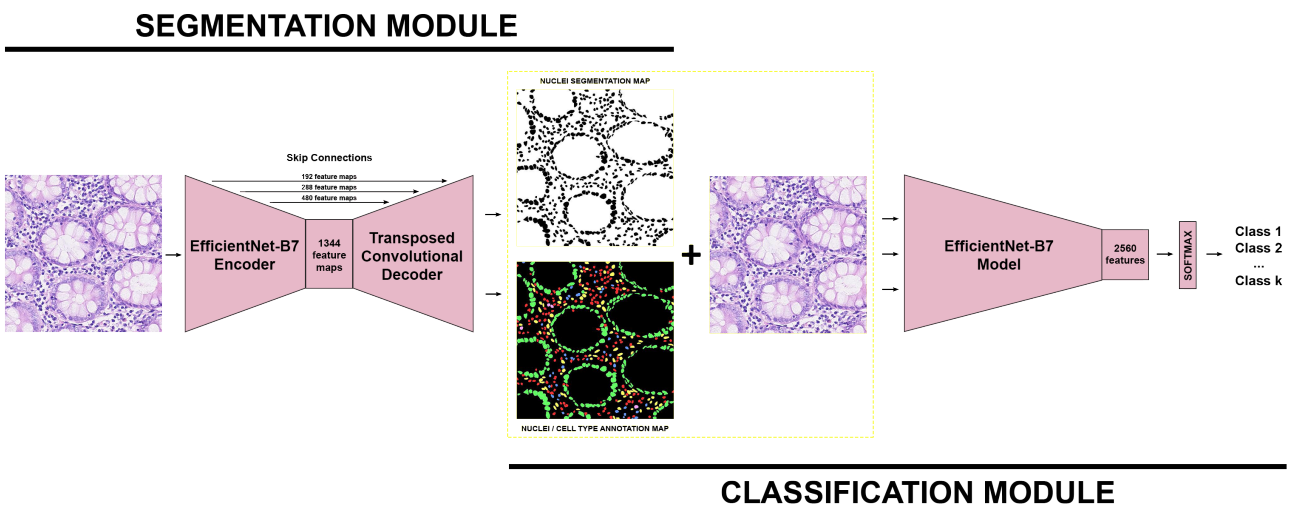

Figure 1. Overview of the proposed DeepCMorph network architecture. The model consists of two separate modules: the first one performs nuclei segmentation and cell type annotation. Its outputs are then stacked together with the original histopathology image and are passed to the second module performing the final classification task.

manner using contrastive learning on a large corpus of diverse whole slide images. This led to a significant accuracy boost when tuning the obtained networks on WSI classification tasks [50, 87, 88]. In this paper, we follow a different approach — rather than optimizing the network to generate representative features for each WSI patch, we suggest pre-training the model to learn cell morphology. Specifically, the model is trained to segment nuclei and classify cell types on histopathological slides. This information is directly integrated into the model and combined with the corresponding H&E stained image to make predictions on the final downstream task.

Nuclei morphology plays an important role in understanding cell development and underlying cellular processes. Deviations from normal nucleus shape are often good indicators of external stress and thus are informative markers for a number of diseases, *e.g.* breast cancers [11, 56], carcinomas [20] or prostate cancers [10]. Nucleus morphology also changes during cell cycle progression and allows to identify actively dividing cells, which is a key characteristic of mutated cancer cells. Besides that, nuclei segmentation is additionally used for cell identification and counting: as accurate segmentation of the entire cell is very difficult in general due to invisibility of cell membranes (defining cell boundaries) on microscopy images, nuclei are usually used for recognition of single cells and their positions. Due to its vital role for histopathology image analysis, a large number of different nuclei segmentation datasets have been proposed over the past years [22, 25–28, 47, 52, 57, 58, 71, 82, 82, 83]. Various efficient deep learning-based solutions were also developed for this task [27, 30, 32, 33, 37, 67, 74], making it possible to achieve precise nuclei segmentation results for various cell types. In this work, we build the segmentation part of our model on top of the previous solutions, further enhancing its robustness by combining multiple datasets and applying extreme data augmentations.

Additionally, the cell type composition differs between healthy and diseased tissues and can serve to analyze disease progression. For instance, having large amounts of immune cells in a specific tissue region is generally associated with a strong inflammatory response [69]. In tumors, immune cell infiltration is strongly correlated with positive disease progression and improved patient survival [53]. Therefore, an ability to distinguish between non-immune and immune cells and their subtypes might significantly improve the predictive power of the model. As several nuclei segmentation datasets provide additional cell type annotations [22, 26, 27, 82], we explicitly use this information for training our segmentation module that generates both nuclei segmentation and cell type annotation maps. In this paper, we demonstrate that such information allows to achieve a significant accuracy boost, outperforming the latest transformer-based models trained on millions of histopathology images while also being more interpretable.

Our main contributions are:

- We propose a novel histopathological image analysis DeepCMorph model that explicitly learns cell morphology: its segmentation module is trained to identify different cell types and nuclei morphological features.
- The segmentation module is independent of the classification block; this allows to train it on all previously published nuclei classification and cell type annotation datasets and leverage this information when training the classification model part.
- Unlike the recently proposed transformer-based foun-

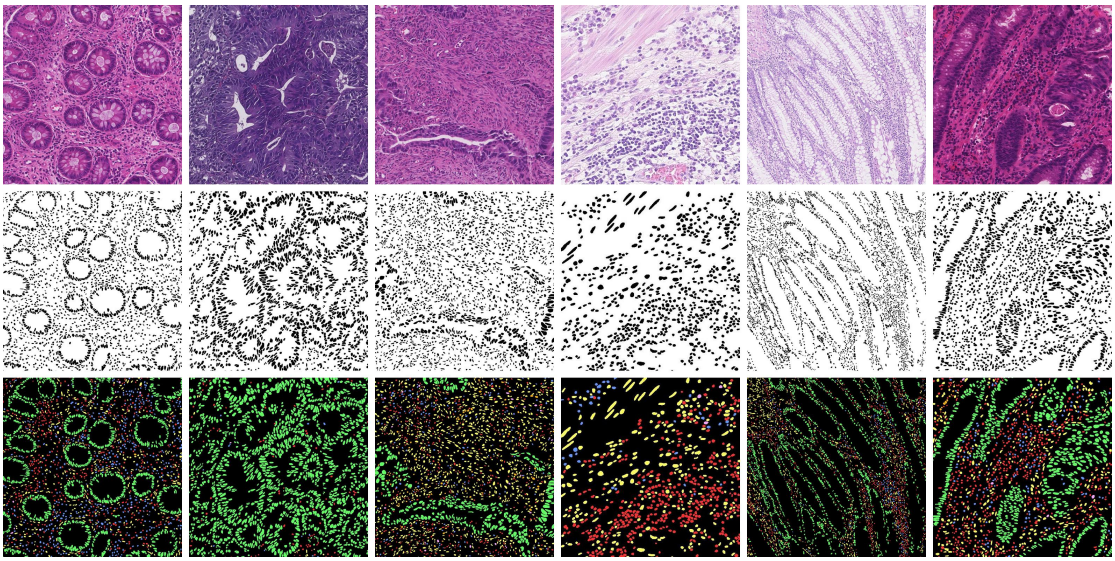

Figure 2. Sample images used for training the segmentation DeepCMorph module. Top row – original H&E stained image patches, middle row – target nuclei segmentation maps, bottom row – cell annotation maps. For latter, red color encodes *lymphocytes*, green: *epithelial cells*, blue: *plasma cells*, orange: *neutrophils*, magenta: *eosinophils*, yellow: *connective tissue*.

dation models [50, 87, 88], this solution is trained in a fully-supervised fashion, which lowers the training time and computational resources by an order of magnitude: for example, compared to the *CTransPath* [88] model that requires 250 hours of training on 48 Nvidia V100 GPUs, our network can be trained on only one GPU in less than one week.

- We apply extreme input data regularization to ensure that the model is not prone to learning any potential batch effect present in the data.
- The model is pre-trained on 8 publicly available segmentation datasets and a large-scale Pan-Cancer TCGA dataset containing over 270K histopathological image patches. On this dataset, DeepCMorph established a new state-of-the-art result, bypassing all previous solutions by over 4% of accuracy.
- We show that that pre-trained model can be easily tuned on smaller datasets for tissue classification with excellent accuracy.
- The proposed model has a fully convolutional architecture and can therefore be applied to images of arbitrary resolutions and aspect ratios.
- We publicly release the codes and pre-trained models to facilitate the development of new performant histopathological image analysis tools.

2. Datasets

This section describes nuclei segmentation, cell type annotation and tissue classification datasets that are used for training the segmentation and classification model blocks.

2.1. Nuclei Segmentation and Annotation Datasets

As our goal is to obtain a robust nuclei segmentation / cell type annotation model module that can be applied to tissues from various organs, we used the majority of previously published datasets proposed for these tasks. In particular, we combined data from 8 segmentation datasets described below to ensure a large diversity of tissue types, patients and imaging conditions:

• **Lizard dataset** [26] contains 291 H&E stained histopathological images of colon tissue with 495179 labeled nuclei. Besides providing target nuclei segmentation maps, this dataset also annotates the type of the cell corresponding to each nucleus: *epithelial cells, connective tissue cells, lymphocytes, plasma cells, neutrophils* and *eosinophils*. The Lizard dataset is itself a combination of six different databases: **DigestPath**, **CRAG** [25], **GlaS** [71], **CoNSeP** [27], **PanNuke** [22] and **TCGA** [28] that were collected in total from 16 different centers and three countries, which ensures the diversity of the data from both biological and technical perspectives. The images are provided at 20× magnification.

• **CryoNuSeg dataset** [52] contains 30 annotated H&E stained images of resolution 512×512 pixels with 7596 nuclei downloaded from the TCGA [28] database. Histopathological images from this dataset correspond to ten different human organs: *adrenal gland, larynx, lymph node, mediastinum, pancreas, pleura, skin, testis, thymus* and *thyroid gland*, and are provided at a magnification of 40×.

• **MoNuSAC dataset** [82] contains 209 annotated H&E stained images with 31411 nuclei downloaded from the TCGA [28] database. This data corresponds to 46 patients from 32 hospitals and four organs: *breast, kidney, lung* and

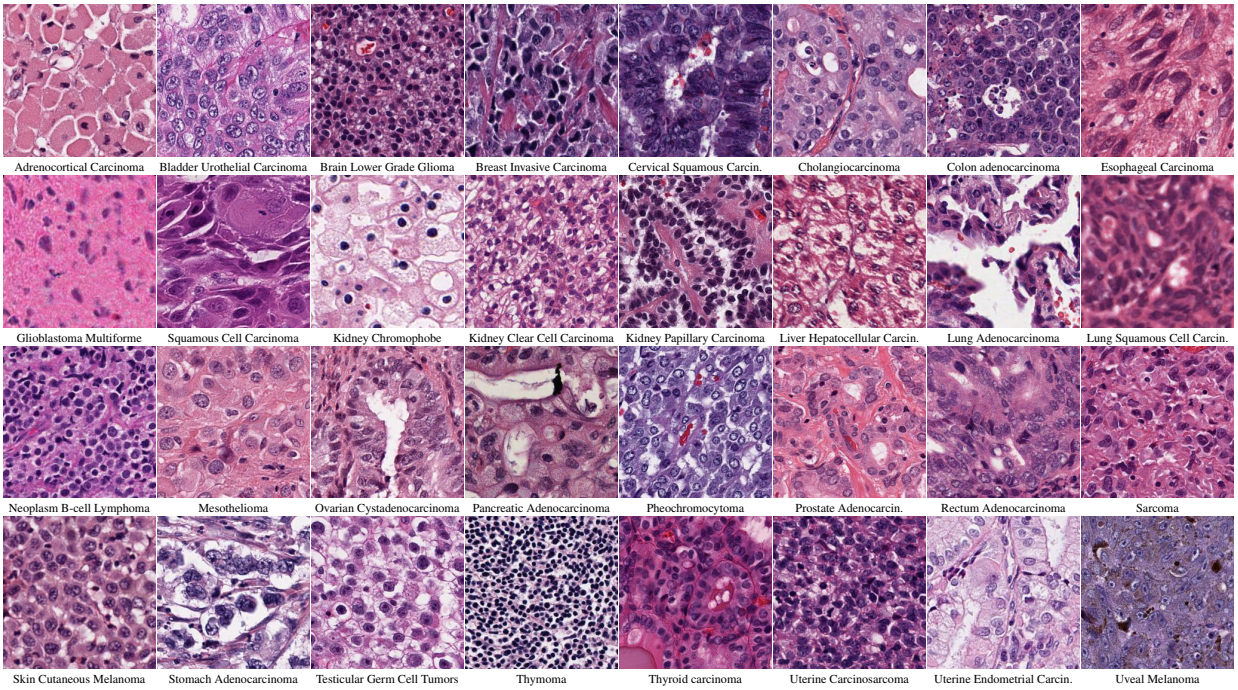

Figure 3. Sample H&E stained image patches for 32 different cancer types from the Pan Cancer TCGA dataset [44].

prostate. The images are provided at 40× magnification.

• **BNS dataset** [57] contains 33 annotated H&E stained images of resolution 512×512 pixels with 2754 nuclei. The data corresponds to breast cancer, the images are provided at 40× magnification.

• **TNBC dataset** [58] contains 50 annotated H&E stained images of resolution 512×512 pixels with 4022 nuclei. The data corresponds to 11 patients and breast organs, the images are provided at 40× magnification.

• **KUMAR dataset** [47] contains 30 annotated H&E stained images of resolution 1000×1000 pixels with 21623 nuclei. The data corresponds to 30 patients and seven organs: *breast, kidney, liver, prostate, bladder, colon* and *stomach*. The images are provided at 40× magnification.

• **MICCAI (CPM)–15/17 datasets** [83] contain 79 annotated H&E stained images with 10475 nuclei downloaded from the TCGA [28] database. The images correspond to four different cancer types: *non-small cell lung cancer, head and neck squamous cell carcinoma, glioblastoma multiforme* and *lower grade glioma*, and are provided at both 20× and 40× magnification.

• **PanNuke dataset** [22] contains 7901 annotated H&E stained images of resolution 256×256 pixels with 216400 nuclei from 19 different organs. Segmentation maps provided in this dataset were generated automatically and then revised by humans. Because of this, the targets for difficult or ambiguous cases are not very accurate [26] as they are limited by the performance of the used FCNN neural network. In this work, PanNuke dataset is used only for initial model pre-training.

Since the datasets use different approaches for storing images and annotations, they were first converted to the same data representation format: all data samples were transformed into 4-channel PNG images, where the first 3 channels encoded RGB image values, and the last one – the target segmentation and / or cell type annotation maps. This allowed to significantly simplify data pre-processing and reduce the size of the entire dataset from 40 GB to less than 2 GB. Overall, the combined dataset contains over eight thousand histopathological images, sample H&E stained input data and target segmentation maps are shown in Fig. 2. The instructions for downloading this data are provided on the official project webpage [1].

2.2. Tissue Classification Datasets

We consider four different histopathological datasets for training the classification DeepCMorph module. The large-scale Pan Cancer TCGA dataset is used for initial model pre-training and benchmarking its predictive capacity. The other three datasets are mainly used to assess DeepCMorph generalization abilities and its performance on other tissue subtypes. A detailed description of the considered datasets is provided below:

• **Pan Cancer TCGA dataset** [44] was obtained by processing 8736 diagnostic slides downloaded from the

[1] https://github.com/aiff22/DeepCMorph

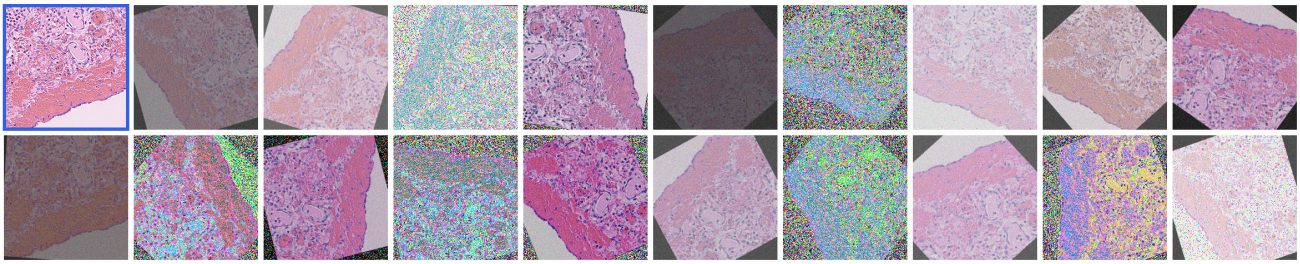

Figure 4. The original image patch (top left, denoted by blue frame) and training patches generated by the proposed data augmentations.

TCGA [28] database and belonging to 7175 patients. For each WSI image, pathologists selected a number of representative patches corresponding to tumor regions. Overall, the dataset contains over 1.6 million patches extracted at 6 different magnification factors for 32 different cancer types, which list is provided in Fig. 3. The authors performed color normalization using [63] to reduce the potential batch effect. Due to a large diversity, this dataset is perfectly suited for pre-training complex deep learning models for various histopathological tissue classification tasks. In this work, we consider the largest magnification factor of 20× (0.5 μm/pixel) and split the dataset randomly into training, validation and test parts using 70:15:15 ratio, which resulted in 188210, 41750 and 41750 patches, respectively. Notably, the splits were stratified per patient, *i.e.*, images from a single patient were present only in the train or test set.

• **NCT-CRC-HE dataset** [40] contains 100K 224×224 px image patches extracted from 136 colorectal adenocarcinoma samples from the National Center for Tumor Diseases (NCT). Nine tissue classes are present in the dataset: *adipose, background, debris, lymphocyte, mucus, smooth muscle, normal colon mucosa, cancer-associated stroma* and *colorectal adenocarcinoma epithelium*. All images are color normalized. A separate set of 7180 image patches from 50 patients with colorectal adenocarcinoma is used for testing.

• **Colorectal cancer (CRC) dataset** [41] contains 5000 150×150 px image patches corresponding to eight tissue types: *epithelium, simple stroma, complex stroma, lymphoid follicles, debris, mucosal glands, adipose* and *background*. Because of its small size, this dataset serves as a good benchmark for model generalization abilities. We split the dataset randomly into training, validation and test parts using 8:1:1 ratio.

• **UniToPatho dataset** [8] contains 8699 patches of resolution 1812×1812 px extracted from 292 WSIs. It was designed for colorectal polyp classification and adenomas grading task and has six tissue classes: *normal tissue, hyperplastic polyp, tubular adenoma (low-grade and high-grade dysplasia)* and *tubulo-villous adenoma (low-grade and high-grade dysplasia)*. Training / test data splits are explicitly provided by the authors.

2.3. Data Augmentations

Since the majority of our datasets contain heterogeneous data where samples corresponding to different classes (diseases) are often collected by different institutions, a pronounced *batch effect* might be present in the data. Batch effect refers to the consequence of variations in tissue processing techniques across different labs or pathologists that may introduce distinct signatures specific to each site. These signatures can then be utilized to uniquely identify the corresponding WSIs. [35]. Recently, Fang et al. demonstrated that neural networks can even recognize camera sensor models directly from images [17], illustrating how microscopy equipment and image post-processing software might also introduce unique and identifiable signatures. Following the results from [17] indicating that such signatures are greatly destroyed by applying data augmentations, we adopted a similar strategy to mitigate any batch effect that might be present in the data. We used the following data augmentations:

- Random image rescaling (by 0–20%),
- Random change of image aspect ratio (by 0–10%),
- Random image rotation (by 0–360 degrees),
- Random image sharpness adjustment (by 2×),
- Random image brightness adjustment (by 0–50%),
- Random image hue adjustment (by 0–10%),
- Random image contrast adjustment (by 0–70%),
- Random image saturation adjustment (by 0–30%),
- Addition of random Gaussian noise to images.

These augmentations are applied to all nuclei segmentation, cell type annotation and the Pan Cancer TCGA datasets. Sample results of such augmentations are demonstrated in Fig. 4, indicating that color-, scale- and texture-related batch effect should be generally eliminated due to the considered extreme scales.

3. Proposed Method

This section provides the architectural and training details of the proposed DeepCMorph model (Fig. 1) that consists of two independent segmentation and classification modules described below.

3.1. Segmentation Module

We aimed to design a solution capable of processing arbitrary resolution images, thus we chose a fully convolutional segmentation model with a U-Net [65] like structure. Its high-level architecture is shown in Fig. 1: a normal 3-channel RGB image is first processed by the EfficientNet-B7 [76] encoder. Its outputs are then passed to a decoder module consisting of 4 upscaling blocks, each one with two subsequent convolution and one transposed convolution layers. Three skip connections are used to pass features obtained at the beginning of EfficientNet blocks 2, 3 and 4 with 192, 288, and 480 feature maps, respectively, to the decoder module. The decoder produces two outputs of the same resolution as the input image: nuclei segmentation map (1 channel) and cell type annotation map (7 channels corresponding to 7 cell annotation classes). The model is trained with a combination of the Binary Cross Entropy (BCE) loss applied to nuclei segmentation maps and the Categorical Cross Entropy (CCE) loss function applied to cell type annotation maps:

$$BCE(\mathbf{y}, \mathbf{p}) = -\frac{1}{N}\sum_{i}^{N}(y_i log(p_i) + (1 - y_i) log(1 - p_i)),$$

$$CCE(\mathbf{y}, \mathbf{p}) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c} \cdot log(p_{i,c}),$$

where $\mathbf{y}$ is a matrix of the ground truth labels and $\mathbf{p}$ is a matrix of predicted values for each pixel and sample. The segmentation module was trained in multiple stages:

1. First, the PanNuke dataset was used for initial model pre-training and then discarded during all later steps.
2. Next, the combined nuclei segmentation datasets were used for model training. At this stage, the cell type annotation data was ignored and only BCE loss used.
3. After convergence, the model was subjected to a further round of training with the addition of cell type annotation data. We used an equal amount of BCE and CCE loss at this step.
4. Pre-trained EfficientNet-B7 encoder model was further tuned on the Pan Cancer TCGA dataset [44] independently of the decoder model to distinguish between 32 cancer types. Experimental results revealed that this procedure significantly improved the quality of the learned feature maps of the EfficientNet-B7 encoder.
5. Steps 2-3 were repeated again, this time using the EfficientNet-B7 encoder already trained on both nuclei segmentation and cancer tissue classification tasks.

After segmentation module training, the module's weights were frozen and its predicted nuclei segmentation and cell type annotation maps were used as additional inputs to the classification DeepCMorph module.

3.2. Classification Module

The classification module of the DeepCMorph model is also based on the EfficientNet-B7 architecture. This model stacks the original RGB histopathology image with the segmentation and annotation maps from the segmentation module, thus its input has 11 feature maps (3 RGB + 1 nuclei segmentation + 7 cell type annotation). The outputs of the last EfficientNet-B7 global average pooling layer are passed to a dense layer with a softmax activation function that produces the final model predictions. We should note that this pooling layer allows the network to handle images of arbitrary sizes as it averages each of 2560 feature maps from the last model layer, thus producing 2560 features irrespective of the input image resolution. Dropout with a rate of 0.2 is additionally applied here to achieve higher feature robustness. The model is trained to minimize the Categorical Cross Entropy (CCE) loss function on classification tasks. The training process of this module also involved multiple stages, we provide a summary of all steps below:

Step 1: As random / ImageNet weights are not perfectly suited for model initialization for the considered tasks, we first pre-trained the classification module to perform the same nuclei segmentation and cell type annotation task as described in the previous section. We added an additional decoder block and trained the model using the same procedure (steps 1–5) as was applied to the segmentation module. The main difference here is that no skip connections between the encoder and decoder were used, given we targeted feature accumulation in the last (bottleneck) EfficientNet-B7 layer. At this stage, instead of the segmentation and annotation maps from the segmentation module, random noise was introduced. This allowed the EfficientNet-B7 model to learn to analyze the input histopathological rather than solely functioning as an autoencoder that compressed the input features and propagates them to the output layer.

Step 2: Once model pre-training on the segmentation task was done, random noise used in the previous step was replaced with real feature maps from the segmentation module, and the model was tuned for additional 5 epochs on the same segmentation task. The goal of this step is to let the model perform integration of the learned histopathological image processing features with the feature maps from the segmentation module, while not completely repurposing its filters to only encode the input segmentation / annotation maps. To determine the appropriate stopping point for training, we monitored the accuracy on the validation set. Following a small initial decrease, we noticed a rapid increase, indicating successful integration.

Step 3: During the last step, the decoder was removed from the classification module and the pre-trained EfficientNet-B7 model was finally tuned on the final downstream tissue classification task.

Method	Dice Score	Binary PQ	Multi PQ
DeepCMorph, Extreme Data Augmentations	0.8365	0.714	0.340
DeepCMorph, Moderate Data Augmentations	0.8406	0.728	0.368

Table 1. Nuclei segmentation (Dice Score) and annotation (Panoptic Quality Score) accuracy on combined segmentation datasets.

3.3. Implementation Details

The model is implemented in PyTorch [59] and trained on four *Nvidia 2080 Ti* GPUs with 12 GB of RAM. At each step, model parameters were optimized using the Adam [43] algorithm with a learning rate of 1e–4 till model convergence (if not stated otherwise) and then additionally tuned with a learning rate of 2e–5. The segmentation module was trained on 224×224 px image crops with a batch size of 36. The resolution of the input images received by the classification module depended on the task and dataset, for images of size 224×224 px the batch size was set to 64. The entire DeepCMorph model has 89M parameters.

4. Experimental Results

This section provides and analyzes experimental results obtained with the DeepCMorph model. First, we assess the performance of the segmentation module on nuclei segmentation / cell type annotation data. Next, we check the results obtained on the Pan Cancer TCGA dataset and compare DeepCMorph performance to the previously proposed solutions. Finally, we check the generalization ability of the proposed model on three other tissue classification datasets.

4.1. Nuclei Segmentation and Annotation Results

After the segmentation module was trained on a combination of the previously described datasets, we first briefly analyzed its performance to ensure a high quality of the generated segmentation and annotation maps. Numerical results obtained on the hold-out test set showing both the segmentation (Dice Score) and annotation (PQ Score) quality are presented in Table 1. They indicate that the model is able to accurately segment nuclei and annotate cell types for different human organs and tissue types, which was also confirmed by visual observations of segmentation results. One can also notice that the proposed excessive histopathological image augmentation approach leads to only minor accuracy degradations while making the model considerably more resistant towards potential batch effects.

We additionally benchmarked DeepCMorph segmentation performance separately on the Lizard [26] dataset. Numerical results shown in Table 2 indicate that it performs

Method	Dice Score
U-Net [26, 65]	0.735
Micro-Net [26, 62]	0.786
HoVer-Net [26, 27]	0.828
DeepCMorph [Segmentation Module]	**0.832**

Table 2. Segmentation accuracy results on the Lizard [26] dataset.

Method	BA, %	Accuracy, %
VGG16 based solution [44]		30.7
ResNet-18 based solution [23]		33.5
ResNet-18 based solution [3]		54.1
CTransPath [88] features + SVM		73.38
UNI (ViT-Large foundation model) [12]	65.7	–
EfficientNet-B7 (ImageNet initialized weights)		75.89
EfficientNet-B7 (Pre-trained on nuclei segmentation)		78.72
CTransPath [88] tuned on the TCGA dataset		78.77
DeepCMorph, Extreme Augmentations	71.81	82.00
DeepCMorph, Moderate Augmentations	**72.79**	**82.73**

Table 3. Accuracy results on the Pan Cancer TCGA dataset [44]. BA here stands for Balanced Accuracy score.

comparably or slightly better than other conventional deep learning based segmentation methods. As the initial experiments demonstrated that the DeepCMorph tissue classification accuracy is almost unsusceptible to small nuclei segmentation and annotation errors as they can be tolerated by the classification module, this performance was considered as sufficient for our further experiments.

4.2. Pan Cancer TCGA Data Classification

We first trained the model on the Pan Cancer TCGA dataset. As it contains 32 different cancer types and data from thousands of patients, it works as a severe benchmark for histopathological tissue classification models, being able to reveal the differences in their predictive capacity due to high task complexity. We compared the performance of our model against CNN-based solutions previously proposed for this dataset [3, 23, 44]; a recent *UNI* [12] foundation vision transformer model pre-trained on 100K diagnostic H&E-stained WSIs; *CTransPath* [88] foundation transformer model pre-trained with contrastive learning on 15M unlabeled patches cropped from WSIs; two EfficientNet-B7 models: one with ImageNet initialized weights and one pre-trained on our combined segmentation datasets.

The results for all models are presented in Table 3. DeepCMorph achieved an accuracy of 82.7%, outperforming the second best CTransPath foundation model that was additionally fine-tuned on the considered dataset by almost 4%. An advantage of 7% was obtained over the UNI model when considering the balanced accuracy reported by the authors. The importance of learning cell morphology was demonstrated by EfficientNet-B7 results: the first model trained starting from ImageNet weights demonstrated an accuracy of only 75.9%. Pre-training EfficientNet-B7 on nuclei segmentation and cell type annotation task improved the accuracy to 78.7%, still leaving a gap of 4% compared to the DeepCMorph that is using explicit annotations.

Same as with the DeepCMorph segmentation module, we can notice that the proposed extreme data augmentation leads to an accuracy drop of less than 1%. The resulting model still outperforms all other solutions by over 3% while being insensitive to variations in staining, noise, sharpness and lighting conditions, thus significantly reducing the impact of any potential batch effect present in the data.

Method	BA, %	Accuracy, %
DenseNet based solution [42]	90.3	92.9
VGG19 based solution [40]		94.3
Inception-v3 based solution [86]		94.8
ResNet-50 based solution [75]		94.8
VGG16 based solution [4]		95.3
CONCH (ViT-Base foundation transformer model) [50]	93.0	–
iBOT (ViT-Large transformer model) [19]	94.4	95.8
DINO (ViT transformer model) [38]	94.5	95.9
Ensemble of 4 models (DenseNet, IncResNetV2, Xception and custom) [24]		96.16
EfficientNet-B7 (ImageNet initialized weights)	94.76	96.18
Ensemble of 5 models (Same as [24] + VGG16) [46]		96.26
CTransPath [88]		96.52
DeepCMorph	**95.59**	**96.99**

Table 4. Accuracy results on the NCT-CRC-HE-7K validation dataset [40]. BA stands for Balanced Accuracy score.

4.3. Results on the NCT-CRC-HE Dataset

NCT-CRC-HE is a popular dataset with a large number of previously proposed deep learning solutions. We tuned the DeepCMorph model on this dataset, initializing it with weights obtained on the TCGA classification task, and assessed its performance on the conventional NCT-CRC-HE-7K test split. Table 4 shows the results for different methods obtained on this task. With an accuracy of 96.99%, DeepCMorph outperformed all other solutions, including visual transformer based CONCH [50], iBOT [19], DINO [38] and CTransPath [88] models that were pre-trained on a large cohort of histopathological data. We should additionally highlight a relatively good performance of the baseline ImageNet-initialized EfficientNet-B7 model that achieved an accuracy of 96.18%, which is comparable to the results of more complex solutions. This also justifies the choice of the DeepCMorph backbone as the underlying EfficientNet-B7 model has a powerful architecture allowing to identify and learn complex patterns from histopathological data.

4.4. Results on the NCT-CRC-HE Dataset

DeepCMorph results on the Colorectal Cancer (CRC) dataset are shown in Table 5. As this dataset contains only 625 image patches per class, we used it to check the generalization ability of the proposed model. Same as in the previous section, we initialized DeepCMorph with weights obtained on the TCGA classification task and tuned it for a few epochs on the CRC data. Despite the small number of training samples, the model was able to achieve top results on this task with an accuracy of 98.33%. This shows that the filters and features learned by the model on the TCGA data are transferable to other histopathological tasks, making it a powerful tool for WSI data analysis and classification.

Method	Accuracy, %
Conventional CV feature descriptors [41]	87.40
Ensemble of 4 models (DenseNet, IncResNetV2, Xception and custom) [24]	92.83
VGG19 based solution [18]	93.58
KimiaNet (DenseNet based model) [64]	96.80
Ensemble of 6 models (AlexNet, GoogleNet, VGG, ResNet, IncV3 and IncResV2) [54]	97.60
EfficientNet-B7 (ImageNet initialized weights)	96.46
CTransPath [88]	98.20
DeepCMorph	**98.33**

Table 5. Accuracy on the Colorectal Cancer (CRC) dataset [41].

Method	BA, %	Accuracy, %
ResNet-18 based solution [8] (σ = 800)	40.0	–
DeepCMorph (TCGA pre-trained) features + SVM	42.51	46.31
DeepCMorph tuned on UniToPatho	**47.35**	**55.81**

Table 6. Accuracy results on the UniToPatho dataset [8]. Here, DeepCMorph is applied to WSI patches of resolution 1812×1812 pixels. BA stands for Balanced Accuracy score.

4.5. Results on the UniToPatho Dataset

UniToPatho dataset is used in different contexts including WSI retrieval [87] or hierarchical multi-scale WSI scan processing [16]. In this work, we focus only on working with the provided 1812×1812 px WSI patches to demonstrate that the DeepCMorph model can be used for high-resolution histopathological image data analysis, unlike transformer-based models restricted to a specific (usually small) input image size. Table 6 shows the results of the proposed solution obtained using two setups: when DeepCMorph model pre-trained on TCGA data is used to produce features for the considered patches that are later classified with SVM, and when it is additionally fine-tuned on smaller 256×256 px crops and then applied to the original patches to provide direct predictions. Even the first setup managed to achieve a higher balanced accuracy (42.51%) compared to the ResNet-18 model trained on UniToPatho in [8], which again demonstrates the versatility of the features learned by the DeepCMorph on TCGA data. After a short additional fine-tuning, the balanced accuracy improves further, reaching 47.35%. Thus, DeepCMorph can be used to classify or annotate large-resolution patches or even entire WSI scans without a need for tile-based image processing and aggregation of the obtained crop-level results.

5. Conclusion

In this work, we considered the problem of histopathological image analysis and proposed a novel DeepCMorph model leveraging the understanding of cell morphology for more accurate tissue classification. DeepCMorph's segmentation module provides additional nuclei segmentation and cell type annotation predictions to the classification module, and is trained on a combination of 8 publicly available segmentation datasets. The classification module was first pre-trained on a large-scale TCGA dataset containing over 270K tissue patches for 32 different cancer types. The proposed DeepCMorph solution achieved the state-of-the-art results on four different tissue classification tasks, outperforming foundation transformer models like CTransPath pre-trained on millions of WSI patches. Due to a fully-convolutional architecture, the model can classify images of arbitrary sizes, avoiding the need for tile-base WSI processing. Finally, we open source the proposed solution and provide pre-trained models to facilitate the development of efficient histopathological image processing methods.

References

[1] Shahira Abousamra, Rajarsi Gupta, Le Hou, Rebecca Batiste, Tianhao Zhao, Anand Shankar, Arvind Rao, Chao Chen, Dimitris Samaras, Tahsin Kurc, et al. Deep learning-based mapping of tumor infiltrating lymphocytes in whole slide images of 23 types of cancer. *Frontiers in oncology*, 11:806603, 2022. 1

[2] Saloni Agarwal, Mohamedelfatih Eltigani Osman Abaker, and Ovidiu Daescu. Survival prediction based on histopathology imaging and clinical data: A novel, whole slide cnn approach. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24*, pages 762–771. Springer, 2021. 1

[3] Mario Alejandro García, Martín Nicolas Gramática, and Juan Pablo Ricapito. Intermediate task fine-tuning in cancer classification. *Journal of Computer Science & Technology (JCS&T)*, 23(2), 2023. 7

[4] TE Anju and S Vimala. Finetuned-vgg16 cnn model for tissue classification of colorectal cancer. In *International Conference on Intelligent Sustainable Systems*, pages 73–84. Springer, 2023. 8

[5] Maschenka CA Balkenhol, David Tellez, Willem Vreuls, Pieter C Clahsen, Hans Pinckaers, Francesco Ciompi, Peter Bult, and Jeroen AWM van der Laak. Deep learning assisted mitotic counting for breast cancer. *Laboratory investigation*, 99(11):1596–1606, 2019. 1

[6] Peter Bandi, Oscar Geessink, Quirine Manson, Marcory Van Dijk, Maschenka Balkenhol, Meyke Hermsen, Babak Ehteshami Bejnordi, Byungjae Lee, Kyunghyun Paeng, Aoxiao Zhong, et al. From detection of individual metastases to classification of lymph node status at the patient level: the camelyon17 challenge. *IEEE transactions on medical imaging*, 38(2):550–560, 2018. 1

[7] Sugata Banerji and Sushmita Mitra. Deep learning in histopathology: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 12(1):e1439, 2022. 1

[8] Carlo Alberto Barbano, Daniele Perlo, Enzo Tartaglione, Attilio Fiandrotti, Luca Bertero, Paola Cassoni, and Marco Grangetto. Unitopatho, a labeled histopathological dataset for colorectal polyps classification and adenoma dysplasia grading. In *2021 IEEE International Conference on Image Processing (ICIP)*, pages 76–80. IEEE, 2021. 1, 5, 8

[9] Wouter Bulten, Kimmo Kartasalo, Po-Hsuan Cameron Chen, Peter Ström, Hans Pinckaers, Kunal Nagpal, Yuannan Cai, David F Steiner, Hester Van Boven, Robert Vink, et al. Artificial intelligence for diagnosis and gleason grading of prostate cancer: the panda challenge. *Nature medicine*, 28(1):154–163, 2022. 1

[10] Neil M Carleton, George Lee, Anant Madabhushi, and Robert W Veltri. Advances in the computational and molecular understanding of the prostate cancer cell nucleus. *Journal of cellular biochemistry*, 119(9):7127–7142, 2018. 2

[11] Jia-Mei Chen, Ai-Ping Qu, Lin-Wei Wang, Jing-Ping Yuan, Fang Yang, Qing-Ming Xiang, Ninu Maskey, Gui-Fang Yang, Juan Liu, and Yan Li. New breast cancer prognostic factors identified by computer-aided image analysis of he stained histopathology images. *Scientific reports*, 5(1):10690, 2015. 2

[12] Richard J Chen, Tong Ding, Ming Y Lu, Drew FK Williamson, Guillaume Jaume, Andrew H Song, Bowen Chen, Andrew Zhang, Daniel Shao, Muhammad Shaban, et al. Towards a general-purpose foundation model for computational pathology. *Nature Medicine*, pages 1–13, 2024. 7

[13] Nicolas Coudray, Paolo Santiago Ocampo, Theodore Sakellaropoulos, Navneet Narula, Matija Snuderl, David Fenyö, Andre L Moreira, Narges Razavian, and Aristotelis Tsirigos. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature medicine*, 24(10):1559–1567, 2018. 1

[14] Athena Davri, Effrosyni Birbas, Theofilos Kanavos, Georgios Ntritsos, Nikolaos Giannakeas, Alexandros T Tzallas, and Anna Batistatou. Deep learning on histopathological images for colorectal cancer diagnosis: a systematic review. *Diagnostics*, 12(4):837, 2022. 1

[15] Muhammad Dawood, Kim Branson, Nasir M Rajpoot, and Fayyaz ul Amir Afsar Minhas. All you need is color: image based spatial gene expression prediction using neural stain learning. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 437–450. Springer, 2021. 1

[16] Ruining Deng, Can Cui, Lucas W Remedios, Shunxing Bao, R Michael Womick, Sophie Chiron, Jia Li, Joseph T Roland, Ken S Lau, Qi Liu, et al. Cross-scale multi-instance learning for pathological image diagnosis. *Medical Image Analysis*, page 103124, 2024. 8

[17] Zilin Fang, Andrey Ignatov, Eduard Zamfir, and Radu Timofte. Sqad: Automatic smartphone camera quality assessment and benchmarking. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 20532–20542, 2023. 5

[18] Kevin Faust, Quin Xie, Dominick Han, Kartikay Goyle, Zoya Volynskaya, Ugljesa Djuric, and Phedias Diamandis. Visualizing histopathologic deep learning classification and anomaly detection using nonlinear feature space dimensionality reduction. *BMC bioinformatics*, 19:1–15, 2018. 8

[19] Alexandre Filiot, Ridouane Ghermi, Antoine Olivier, Paul Jacob, Lucas Fidon, Alice Mac Kain, Charlie Saillard, and Jean-Baptiste Schiratti. Scaling self-supervised learning for histopathology with masked image modeling. *medRxiv*, pages 2023–07, 2023. 8

[20] Edgar G Fischer. Nuclear morphology and the biology of cancer cells. *Acta cytologica*, 64(6):511–519, 2020. 2

[21] Yu Fu, Alexander W Jung, Ramon Viñas Torne, Santiago Gonzalez, Harald Vöhringer, Artem Shmatko, Lucy R Yates, Mercedes Jimenez-Linan, Luiza Moore, and Moritz Gerstung. Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *Nature cancer*, 1(8):800–810, 2020. 1

[22] Jevgenij Gamper, Navid Alemi Koohbanani, Ksenija Benes, Simon Graham, Mostafa Jahanifar, Syed Ali Khurram,

Ayesha Azam, Katherine Hewitt, and Nasir Rajpoot. Pannuke dataset extension, insights and baselines. *arXiv preprint arXiv:2003.10778*, 2020. 2, 3, 4

[23] Mario Alejandro García, Martín Nicolás Gramática, Leandro Ariel Rostagno, and Juan Pablo Ricapito. Clasificación de tipos de cáncer en imágenes de histopatología. 7

[24] Sourodip Ghosh, Ahana Bandyopadhyay, Shreya Sahay, Richik Ghosh, Ishita Kundu, and KC Santosh. Colorectal histology tumor detection using ensemble deep neural network. *Engineering Applications of Artificial Intelligence*, 100:104202, 2021. 8

[25] Simon Graham, Hao Chen, Jevgenij Gamper, Qi Dou, Pheng-Ann Heng, David Snead, Yee Wah Tsang, and Nasir Rajpoot. Mild-net: Minimal information loss dilated network for gland instance segmentation in colon histology images. *Medical image analysis*, 52:199–211, 2019. 2, 3

[26] Simon Graham, Mostafa Jahanifar, Ayesha Azam, Mohammed Nimir, Yee-Wah Tsang, Katherine Dodd, Emily Hero, Harvir Sahota, Atisha Tank, Ksenija Benes, et al. Lizard: a large-scale dataset for colonic nuclear instance segmentation and classification. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 684–693, 2021. 2, 3, 4, 7

[27] Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical image analysis*, 58:101563, 2019. 2, 3, 7

[28] Robert L Grossman, Allison P Heath, Vincent Ferretti, Harold E Varmus, Douglas R Lowy, Warren A Kibbe, and Louis M Staudt. Toward a shared vision for cancer genomic data. *New England Journal of Medicine*, 375(12):1109–1112, 2016. 2, 3, 4, 5

[29] Ahmet Gokberk Gul, Oezdemir Cetin, Christoph Reich, Nadine Flinner, Tim Prangemeier, and Heinz Koeppl. Histopathological image classification based on self-supervised vision transformer and weak labels. In *Medical Imaging 2022: Digital and Computational Pathology*, volume 12039, pages 366–373. SPIE, 2022. 1

[30] Yue Han, Yang Lei, Viktor Shkolnikov, Daisy Xin, Alicia Auduong, Steven Barcelo, Jan Allebach, and Edward J Delp. An ensemble method with edge awareness for abnormally shaped nuclei segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4314–4324, 2023. 2

[31] Bryan He, Ludvig Bergenstråhle, Linnea Stenbeck, Abubakar Abid, Alma Andersson, Åke Borg, Jonas Maaskola, Joakim Lundeberg, and James Zou. Integrating spatial gene expression and breast tumour morphology via deep learning. *Nature biomedical engineering*, 4(8):827–834, 2020. 1

[32] Reka Hollandi, Nikita Moshkov, Lassi Paavolainen, Ervin Tasnadi, Filippo Piccinini, and Peter Horvath. Nucleus segmentation: towards automated solutions. *Trends in Cell Biology*, 32(4):295–310, 2022. 2

[33] Fabian Hörst, Moritz Rempe, Lukas Heine, Constantin Seibold, Julius Keyl, Giulia Baldini, Selma Ugurel, Jens Siveke, Barbara Grünwald, Jan Egger, et al. Cellvit: Vision transformers for precise cell segmentation and classification. *Medical Image Analysis*, page 103143, 2024. 2

[34] Le Hou, Dimitris Samaras, Tahsin M Kurc, Yi Gao, James E Davis, and Joel H Saltz. Patch-based convolutional neural network for whole slide tissue image classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2424–2433, 2016. 1

[35] Frederick M Howard, James Dolezal, Sara Kochanny, Jefree Schulte, Heather Chen, Lara Heij, Dezheng Huo, Rita Nanda, Olufunmilayo I Olopade, Jakob N Kather, et al. The impact of site-specific digital histology signatures on deep learning model accuracy and bias. *Nature communications*, 12(1):4423, 2021. 5

[36] Kobiljon Ikromjanov, Subrata Bhattacharjee, Yeong-Byn Hwang, Rashadul Islam Sumon, Hee-Cheol Kim, and Heung-Kook Choi. Whole slide image analysis and detection of prostate cancer using vision transformers. In *2022 international conference on artificial intelligence in information and communication (ICAIIC)*, pages 399–402. IEEE, 2022. 1

[37] Jeremiah W Johnson. Automatic nucleus segmentation with mask-rcnn. In *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 2 1*, pages 399–407. Springer, 2020. 2

[38] Mingu Kang, Heon Song, Seonwook Park, Donggeun Yoo, and Sérgio Pereira. Benchmarking self-supervised learning on diverse pathology datasets. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3344–3354, 2023. 8

[39] Jakob Nikolas Kather, Lara R Heij, Heike I Grabsch, Chiara Loeffler, Amelie Echle, Hannah Sophie Muti, Jeremias Krause, Jan M Niehues, Kai AJ Sommer, Peter Bankhead, et al. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nature cancer*, 1(8):789–799, 2020. 1

[40] Jakob Nikolas Kather, Johannes Krisam, Pornpimol Charoentong, Tom Luedde, Esther Herpel, Cleo-Aron Weis, Timo Gaiser, Alexander Marx, Nektarios A Valous, Dyke Ferber, et al. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS medicine*, 16(1):e1002730, 2019. 5, 8

[41] Jakob Nikolas Kather, Cleo-Aron Weis, Francesco Bianconi, Susanne M Melchers, Lothar R Schad, Timo Gaiser, Alexander Marx, and Frank Gerrit Zöllner. Multi-class texture analysis in colorectal cancer histology. *Scientific reports*, 6(1):1–11, 2016. 1, 5, 8

[42] Alexander Khvostikov, Andrey Krylov, Ilya Mikhailov, Pavel Malkov, and Natalya Danilova. Tissue type recognition in whole slide histological images. In *CEUR Workshop Proc. 3027*, volume 50, 2021. 8

[43] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 7

[44] Daisuke Komura, Akihiro Kawabe, Keisuke Fukuta, Kyohei Sano, Toshikazu Umezaki, Hirotomo Koda, Ryohei Suzuki, Ken Tominaga, Mieko Ochi, Hiroki Konishi, et al. Universal

encoding of pan-cancer histology by deep texture representations. *Cell Reports*, 38(9), 2022. 1, 4, 6, 7

[45] Michał Koziarski, Bogusław Cyganek, Bogusław Olborski, Zbigniew Antosz, Marcin Żydak, Bogdan Kwolek, Paweł Wkasowicz, Andrzej Bukała, Jakub Swadźba, and Piotr Sitkowski. Diagset: a dataset for prostate cancer histopathological image classification. *arXiv preprint arXiv:2105.04014*, 2021. 1

[46] Anurodh Kumar, Amit Vishwakarma, and Varun Bajaj. Crccn-net: Automated framework for classification of colorectal tissue using histopathological images. *Biomedical Signal Processing and Control*, 79:104172, 2023. 8

[47] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE transactions on medical imaging*, 36(7):1550–1560, 2017. 2, 4

[48] Tristan Lazard, Marvin Lerousseau, Etienne Decencière, and Thomas Walter. Giga-ssl: Self-supervised learning for gigapixel images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4304–4313, 2023. 1

[49] Nicolas Loménie, Capucine Bertrand, Rutger HJ Fick, Saima Ben Hadj, Brice Tayart, Cyprien Tilmant, Isabelle Farré, Soufiane Z Azdad, Samy Dahmani, Gilles Dequen, et al. Can ai predict epithelial lesion categories via automated analysis of cervical biopsies: The tissuenet challenge? *Journal of Pathology Informatics*, 13:100149, 2022. 1

[50] Ming Y Lu, Bowen Chen, Drew FK Williamson, Richard J Chen, Ivy Liang, Tong Ding, Guillaume Jaume, Igor Odintsov, Andrew Zhang, Long Phi Le, et al. Towards a visual-language foundation model for computational pathology. *arXiv preprint arXiv:2307.12914*, 2023. 2, 3, 8

[51] Wenqi Lu, Michael Toss, Muhammad Dawood, Emad Rakha, Nasir Rajpoot, and Fayyaz Minhas. Slidegraph+: Whole slide image level graphs to predict her2 status in breast cancer. *Medical Image Analysis*, 80:102486, 2022. 1

[52] Amirreza Mahbod, Gerald Schaefer, Benjamin Bancher, Christine Löw, Georg Dorffner, Rupert Ecker, and Isabella Ellinger. Cryonuseg: A dataset for nuclei instance segmentation of cryosectioned h&e-stained histological images. *Computers in biology and medicine*, 132:104349, 2021. 2, 3

[53] Marit M Melssen, Natasha D Sheybani, Katie M Leick, and Craig L Slingluff Jr. Barriers to immune cell infiltration in tumors. *Journal for Immunotherapy of Cancer*, 11(4), 2023. 2

[54] Loris Nanni, Stefano Ghidoni, and Sheryl Brahnam. Ensemble of convolutional neural networks for bioimage classification. *Applied Computing and Informatics*, 17(1):19–35, 2021. 8

[55] Ramin Nateghi, Habibollah Danyali, and Mohammad Sadegh Helfroush. A deep learning approach for mitosis detection: application in tumor proliferation prediction from whole slide images. *Artificial intelligence in medicine*, 114:102048, 2021. 1

[56] Sidra Nawaz and Yinyin Yuan. Computational pathology: Exploring the spatial dimension of tumor ecology. *Cancer letters*, 380(1):296–303, 2016. 2

[57] Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Nuclei segmentation in histopathology images using deep neural networks. In *2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)*, pages 933–936. IEEE, 2017. 2, 4

[58] Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE transactions on medical imaging*, 38(2):448–459, 2018. 2, 4

[59] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019. 7

[60] Hui Qu, Mu Zhou, Zhennan Yan, He Wang, Vinod K Rustgi, Shaoting Zhang, Olivier Gevaert, and Dimitris N Metaxas. Genetic mutation and biological pathway prediction based on whole slide images in breast carcinoma using deep learning. *NPJ precision oncology*, 5(1):87, 2021. 1

[61] Ashwin Raju, Jiawen Yao, Mohammad MinHazul Haq, Jitendra Jonnagaddala, and Junzhou Huang. Graph attention multi-instance learning for accurate colorectal cancer staging. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part V 23*, pages 529–539. Springer, 2020. 1

[62] Shan E Ahmed Raza, Linda Cheung, Muhammad Shaban, Simon Graham, David Epstein, Stella Pelengaris, Michael Khan, and Nasir M Rajpoot. Micro-net: A unified model for segmentation of various objects in microscopy images. *Medical image analysis*, 52:160–173, 2019. 7

[63] Erik Reinhard, Michael Adhikhmin, Bruce Gooch, and Peter Shirley. Color transfer between images. *IEEE Computer graphics and applications*, 21(5):34–41, 2001. 5

[64] Abtin Riasatian, Morteza Babaie, Danial Maleki, Shivam Kalra, Mojtaba Valipour, Sobhan Hemati, Manit Zaveri, Amir Safarpoor, Sobhan Shafiei, Mehdi Afshari, et al. Fine-tuning and training of densenet for histopathology image representation using tcga diagnostic slides. *Medical image analysis*, 70:102032, 2021. 8

[65] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015. 6, 7

[66] Benoît Schmauch, Alberto Romagnoni, Elodie Pronier, Charlie Saillard, Pascale Maillé, Julien Calderaro, Aurélie Kamoun, Meriem Sefta, Sylvain Toldo, Mikhail Zaslavskiy, et al. A deep learning model to predict rna-seq expression of tumours from whole slide images. *Nature communications*, 11(1):3877, 2020. 1

[67] Uwe Schmidt, Martin Weigert, Coleman Broaddus, and Gene Myers. Cell detection with star-convex poly-

gons. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II 11*, pages 265–273. Springer, 2018. 2

[68] Wei Shao, Tongxin Wang, Zhi Huang, Zhi Han, Jie Zhang, and Kun Huang. Weakly supervised deep ordinal cox model for survival prediction from whole-slide pathological images. *IEEE Transactions on Medical Imaging*, 40(12):3739–3747, 2021. 1

[69] Edward R Sherwood and Tracy Toliver-Kinsky. Mechanisms of the inflammatory response. *Best Practice & Research Clinical Anaesthesiology*, 18(3):385–405, 2004. 2

[70] Prashant Shukla, Shekhar Verma, et al. A compact fuzzy rule interpretation of svm classifier for medical whole slide images. In *TENCON 2017-2017 IEEE Region 10 Conference*, pages 1588–1592. IEEE, 2017. 1

[71] Korsuk Sirinukunwattana, Josien PW Pluim, Hao Chen, Xiaojuan Qi, Pheng-Ann Heng, Yun Bo Guo, Li Yang Wang, Bogdan J Matuszewski, Elia Bruni, Urko Sanchez, et al. Gland segmentation in colon histology images: The glas challenge contest. *Medical image analysis*, 35:489–502, 2017. 2, 3

[72] Zhigang Song, Shuangmei Zou, Weixun Zhou, Yong Huang, Liwei Shao, Jing Yuan, Xiangnan Gou, Wei Jin, Zhanbo Wang, Xin Chen, et al. Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nature communications*, 11(1):4294, 2020. 1

[73] Chetan L Srinidhi, Ozan Ciga, and Anne L Martel. Deep neural network models for computational histopathology: A survey. *Medical image analysis*, 67:101813, 2021. 1

[74] C Stringer, T Wang, M Michaelos, and M Pachitariu Cellpose. A generalist algorithm for cellular segmentation., 2021, 18. *DOI: https://doi. org/10.1038/s41592-020-01018-x. PMID: https://www. ncbi. nlm. nih. gov/pubmed/33318659*, pages 100–106. 2

[75] Kai Sun, Yushi Chen, Bingqian Bai, Yanhua Gao, Jiaying Xiao, and Gang Yu. Automatic classification of histopathology images across multiple cancers based on heterogeneous transfer learning. *Diagnostics*, 13(7):1277, 2023. 8

[76] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019. 6

[77] Pei-Chen Tsai, Tsung-Hua Lee, Kun-Chi Kuo, Fang-Yi Su, Tsung-Lu Michael Lee, Eliana Marostica, Tomotaka Ugai, Melissa Zhao, Mai Chan Lau, Juha P Väyrynen, et al. Histopathology images predict multi-omics aberrations and prognoses in colorectal cancer patients. *Nature communications*, 14(1):2102, 2023. 1

[78] Nelson Zange Tsaku, Sai Chandra Kosaraju, Tasmia Aqila, Mohammad Masum, Dae Hyun Song, Ananda M Mondal, Hyun Min Koh, and Mingon Kang. Texture-based deep learning for effective histopathological cancer image classification. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 973–977. IEEE, 2019. 1

[79] Riku Turkki, Nina Linder, Panu E Kovanen, Teijo Pellinen, and Johan Lundin. Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. *Journal of pathology informatics*, 7(1):38, 2016. 1

[80] Mira Valkonen, Kimmo Kartasalo, Kaisa Liimatainen, Matti Nykter, Leena Latonen, and Pekka Ruusuvuori. Metastasis detection from whole slide images using local features and random forests. *Cytometry Part A*, 91(6):555–565, 2017. 1

[81] Jeroen Van der Laak, Geert Litjens, and Francesco Ciompi. Deep learning in histopathology: the path to the clinic. *Nature medicine*, 27(5):775–784, 2021. 1

[82] Ruchika Verma, Neeraj Kumar, Abhijeet Patil, Nikhil Cherian Kurian, Swapnil Rane, Simon Graham, Quoc Dang Vu, Mieke Zwager, Shan E Ahmed Raza, Nasir Rajpoot, et al. Monusac2020: A multi-organ nuclei segmentation and classification challenge. *IEEE Transactions on Medical Imaging*, 40(12):3413–3423, 2021. 2, 3

[83] Quoc Dang Vu, Simon Graham, Tahsin Kurc, Minh Nguyen Nhat To, Muhammad Shaban, Talha Qaiser, Navid Alemi Koohbanani, Syed Ali Khurram, Jayashree Kalpathy-Cramer, Tianhao Zhao, et al. Methods for segmentation and classification of digital microscopy tissue images. *Frontiers in bioengineering and biotechnology*, 7:433738, 2019. 2, 4

[84] Sophia J Wagner, Daniel Reisenbüchler, Nicholas P West, Jan Moritz Niehues, Jiefu Zhu, Sebastian Foersch, Gregory Patrick Veldhuizen, Philip Quirke, Heike I Grabsch, Piet A van den Brandt, et al. Transformer-based biomarker prediction from colorectal cancer histology: A large-scale multicentric study. *Cancer Cell*, 41(9):1650–1661, 2023. 1

[85] Haibo Wang, Angel Cruz-Roa, Ajay Basavanhally, Hannah Gilmore, Natalie Shih, Mike Feldman, John Tomaszewski, Fabio Gonzalez, and Anant Madabhushi. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *Journal of Medical Imaging*, 1(3):034003–034003, 2014. 1

[86] Kuan-Song Wang, Gang Yu, Chao Xu, Xiang-He Meng, Jianhua Zhou, Changli Zheng, Zhenghao Deng, Li Shang, Ruijie Liu, Shitong Su, et al. Accurate diagnosis of colorectal cancer based on histopathology images using artificial intelligence. *BMC medicine*, 19:1–12, 2021. 8

[87] Xiyue Wang, Yuexi Du, Sen Yang, Jun Zhang, Minghui Wang, Jing Zhang, Wei Yang, Junzhou Huang, and Xiao Han. Retccl: Clustering-guided contrastive learning for whole-slide image retrieval. *Medical image analysis*, 83:102645, 2023. 2, 3, 8

[88] Xiyue Wang, Sen Yang, Jun Zhang, Minghui Wang, Jing Zhang, Wei Yang, Junzhou Huang, and Xiao Han. Transformer-based unsupervised contrastive learning for histopathological image classification. *Medical image analysis*, 81:102559, 2022. 2, 3, 7, 8

[89] Philippe Weitz, Yinxi Wang, Kimmo Kartasalo, Lars Egevad, Johan Lindberg, Henrik Grönberg, Martin Eklund, and Mattias Rantalainen. Transcriptome-wide prediction of prostate cancer gene expression from histopathology images using co-expression-based convolutional neural networks. *Bioinformatics*, 38(13):3462–3469, 2022. 1

[90] Yawen Wu, Michael Cheng, Shuo Huang, Zongxiang Pei, Yingli Zuo, Jianxin Liu, Kai Yang, Qi Zhu, Jie Zhang, Hong-

hai Hong, et al. Recent advances of deep learning for computational histopathology: principles and applications. *Cancers*, 14(5):1199, 2022. 1

[91] Ellery Wulczyn, David F Steiner, Zhaoyang Xu, Apaar Sadhwani, Hongwu Wang, Isabelle Flament-Auvigne, Craig H Mermel, Po-Hsuan Cameron Chen, Yun Liu, and Martin C Stumpe. Deep learning-based survival prediction for multiple cancer types using histopathology images. *PloS one*, 15(6):e0233678, 2020. 1

[92] Hongming Xu, Qi Xu, Fengyu Cong, Jeonghyun Kang, Chu Han, Zaiyi Liu, Anant Madabhushi, and Cheng Lu. Vision transformers for computational histopathology. *IEEE Reviews in Biomedical Engineering*, 2023. 1

[93] Jun Xu, Haoda Lu, Haixin Li, Chaoyang Yan, Xiangxue Wang, Min Zang, Dirk G de Rooij, Anant Madabhushi, and Eugene Yujun Xu. Computerized spermatogenesis staging (css) of mouse testis sections via quantitative histomorphological analysis. *Medical image analysis*, 70:101835, 2021. 1

[94] Zeyan Xu, Yong Li, Yingyi Wang, Shenyan Zhang, Yanqi Huang, Su Yao, Chu Han, Xipeng Pan, Zhenwei Shi, Yun Mao, et al. A deep learning quantified stroma-immune score to predict survival of patients with stage ii–iii colorectal cancer. *Cancer cell international*, 21:1–12, 2021. 1

[95] Rikiya Yamashita, Jin Long, Teri Longacre, Lan Peng, Gerald Berry, Brock Martin, John Higgins, Daniel L Rubin, and Jeanne Shen. Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *The Lancet Oncology*, 22(1):132–141, 2021. 1

[96] Jiawen Yao, Xinliang Zhu, Jitendra Jonnagaddala, Nicholas Hawkins, and Junzhou Huang. Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. *Medical Image Analysis*, 65:101789, 2020. 1

[97] Hong Yu, Xiaofan Zhang, Lingjun Song, Liren Jiang, Xiaodi Huang, Wen Chen, Chenbin Zhang, Jiahui Li, Jiji Yang, Zhiqiang Hu, et al. Large-scale gastric cancer screening and localization using multi-task deep neural network. *Neurocomputing*, 448:290–300, 2021. 1

[98] Huaqi Zhang, Huang Chen, Jin Qin, Bei Wang, Guolin Ma, Pengyu Wang, Dingrong Zhong, and Jie Liu. Mcvit: Multi-path cross-scale vision transformer for thymoma histopathology whole slide image typing. *Frontiers in Oncology*, 12:925903, 2022. 1