











Article

Automated Cellular-Level Dual Global Fusion of Whole-Slide Imaging for Lung Adenocarcinoma Prognosis

Songhui Diao ^{1,2,3}, Pingjun Chen ³, Eman Showkatian ³, Rukhmini Bandyopadhyay ³, Frank R. Rojas ⁴, Bo Zhu ⁵, Lingzhi Hong ^{3,5}, Muhammad Aminu ³, Maliazorina B. Saad ³, Morteza Salehjehromi ³, Amgad Muneer ³, Sheeba J. Sujit ³, Carmen Behrens ⁵, Don L. Gibbons ⁵, John V. Heymach ⁵, Neda Kalhor ⁶, Ignacio I. Wistuba ⁴, Luisa M. Solis Soto ⁴, Jianjun Zhang ^{5,7}, Wenjian Qin ^{1,*} and Jia Wu ^{3,5,*}

- ¹ Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
 - ² Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen 518055, China
 - ³ Department of Imaging Physics, Division of Diagnostic Imaging, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
 - ⁴ Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
 - ⁵ Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
 - ⁶ Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
 - ⁷ Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
- * Correspondence: wj.qin@siat.ac.cn (W.Q.); jwu11@mdanderson.org (J.W.)

Simple Summary: Lung cancer is the leading cause of cancer death in the United States and worldwide. Currently, deep learning-based methods show significant advances and potential in pathology and can guide lung cancer diagnosis and prognosis prediction. In this study, we present a fully automated cellular-level survival prediction pipeline that uses histopathologic images of lung adenocarcinoma to predict survival risk based on dual global feature fusion. The results show meaningful, convincing, and comprehensible survival prediction ability and manifest the potential of our proposed pipeline for application to other malignancies.

Abstract: Histopathologic whole-slide images (WSI) are generally considered the gold standard for cancer diagnosis and prognosis. Survival prediction based on WSI has recently attracted substantial attention. Nevertheless, it remains a central challenge owing to the inherent difficulties of predicting patient prognosis and effectively extracting informative survival-specific representations from WSI with highly compounded gigapixels. In this study, we present a fully automated cellular-level dual global fusion pipeline for survival prediction. Specifically, the proposed method first describes the composition of different cell populations on WSI. Then, it generates dimension-reduced WSI-embedded maps, allowing for efficient investigation of the tumor microenvironment. In addition, we introduce a novel dual global fusion network to incorporate global and inter-patch features of cell distribution, which enables the sufficient fusion of different types and locations of cells. We further validate the proposed pipeline using The Cancer Genome Atlas lung adenocarcinoma dataset. Our model achieves a C-index of 0.675 (± 0.05) in the five-fold cross-validation setting and surpasses comparable methods. Further, we extensively analyze embedded map features and survival probabilities. These experimental results manifest the potential of our proposed pipeline for applications using WSI in lung adenocarcinoma and other malignancies.

Keywords: embedded features; global fusion; cellular architecture; whole-slide image; survival prediction; lung adenocarcinoma



Citation: Diao, S.; Chen, P.; Showkatian, E.; Bandyopadhyay, R.; Rojas, F.R.; Zhu, B.; Hong, L.; Aminu, M.; Saad, M.B.; Salehjehromi, M.; et al. Automated Cellular-Level Dual Global Fusion of Whole-Slide Imaging for Lung Adenocarcinoma Prognosis. *Cancers* **2023**, *15*, 4824. <https://doi.org/10.3390/cancers15194824>

Academic Editors: Filippo Pesapane and Michaela Cellina

Received: 8 September 2023

Revised: 24 September 2023

Accepted: 27 September 2023

Published: 1 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cancer mortality and morbidity remain fatal global burdens, among which lung cancer, the leading cause of cancer death in the United States and worldwide, has a higher risk of death than other cancers [1]. Lung adenocarcinoma (LUAD) is a common histological subtype of lung cancer and accounts for about 40% of all lung cancers [2]. However, LUAD is often diagnosed at an advanced stage. Despite significant efforts to improve LUAD treatment, survival rates are still suboptimal [3,4]. Precise survival prediction contributes to the mitigation of overtreatment and a reduction in economic costs [5].

Pathological assessments have been the gold standard for cancer diagnosis, and their role also extends to prognostication [6,7]. With the application of the whole-slide imaging scanner, digital pathologic images from tissue glass slides have markedly aided the development of computational pathology, including the prediction of the survival risk of patients [8]. This facilitates the study of more precise interconnections between pathologic image structures and the LUAD prognosis.

With the development of computing technology, numerous artificial intelligence studies have sought to automate tissue-level diagnosis [9–11], with some reports of deep learning matching pathologist-level performance [12,13]. Compared to cancer diagnosis, predicting treatment response or survival has thus far received less attention. Essentially, a cancer diagnosis is a classical classification task for a given glass or digital slide based on the examination of its morphology [14]. In contrast, patient prognosis is a more complex predictive problem that can be driven by both intrinsic (e.g., host anti-tumor immunity and performance status) and extrinsic (e.g., treatment and lifestyle) factors [15]. Pathological images provide a method for quantifying the intrinsic properties of the tumor and its host. However, the large scale of pathological slides and intratumor heterogeneity pose a significant challenge in distilling whole-slide image (WSI) patterns to predict survival. Existing WSI feature extraction methods can be categorized into two main paradigms: the patch-based approach [16] and the nuclei-based approach [17]. The patch-based approach first divides the WSI into several regular-sized patches, then utilizes a pre-trained convolutional neural network (CNN) to extract patch-level features, and finally adopts voting [18] or multiple instance learning (MIL) [19] to fuse different patches to establish a survival prediction model. A major limitation of the patch-based method is the lack of appropriate biological explanations of results when extracting patch features and merging patches. Some approaches seek to address this challenge by employing attention mechanisms [10,20]. On the other hand, there is more transparency and direct interpretability for nuclei-based approaches that aim to describe different cell populations and their spatial interactions. Such frameworks usually start with nuclei segmentation of WSI and then characterize nuclear morphology [21] or construct a multi-nuclear-type model to quantify intratumor heterogeneity [17]. With this approach, the tumor prognostic analysis is more dependent on contextual information and information about various cell types in tumor microenvironments. However, these nuclei-based models do not adequately consider the relationships between different cell types as well as within the same cell type, which means that these methods lack sufficient consideration of contextual information. Meanwhile, these models are mainly applied to predefined regions of interest rather than WSI owing to the need for greater computing resources.

Recently, several new studies on cancer prognosis using feature extraction have focused on segmented nuclei. For example, Lu et al. [22] extracted 615 features relating to nuclear shape, texture, and orientation disorder from nuclei obtained using watershed segmentation from each tissue microarray spot. Then, the 15 most prognostic quantitative histomorphometric features were identified and used to predict survival risk. In parallel, Alsubaie et al. [23] characterized the heterogeneity of LUAD by using morphological features of segmented tumor cell nuclei, then generated a heat map of the entire WSI and used Cox proportional risk regression models to find the most important discriminatory features to implement prognostic analyses. Another study by Chen et al. [7] was based on calculating partial eigenvalues for the cell nuclei to represent certain textural features,

and ultimately selecting a subset of features for prognostic analysis. These models have some important limitations. First, while splitting the nucleus or extracting features from the segmented nucleus, the method of extracting features manually is not conducive to an automated pipeline. Second, there is no explicit classification of different types of nuclei, and a large amount of data is needed to effectively learn the characteristics of different cell types. Third, contextual information about the same or different types of nuclei is not fully utilized. Therefore, it is essential to develop a fully automatic pipeline to effectually utilize the contextual information of multiple cell types.

In this paper, we present a fully automated cellular-level dual-branch global fusion pipeline to predict survival based on WSI. Compared to nuclei-based prognosis methods, our pipeline can better utilize global contextual information, including different cell types and locations. Specifically, we first compute an embedded WSI map based on the segmentation and classification of all cell nuclei on the WSI, which maintains their spatial information. This approach has better interpretability compared to methods that compute patches or nucleus features extracted from some black-box mechanisms. Next, the features of the overall WSI and the features derived from relationships between different regions within the WSI are calculated separately for the WSI-embedded maps. The former features focus on global cell type and distribution, while the latter features focus more on contextual relationships between different or within the same cell types. Finally, these two types of global features are further fused to assess their implication for prognosis. We evaluated our pipeline using The Cancer Genome Atlas (TCGA) lung adenocarcinoma dataset and demonstrated promising performance. The proposed method is more robust compared to individual global features. Meanwhile, this overall framework has the potential to predict survival and treatment responses for other cancer malignancies.

2. Materials and Methods

2.1. Data Curation and Pre-Processing

We applied the proposed methodology to the TCGA LUAD dataset, which was accessed from the TCGA data portal. Survival and phenotype information were obtained from UCSC Xena. The investigation focused on frozen section pathology images captured at a magnification of 20×, with corresponding overall survival (OS) data. Following data curation and rigorous quality control procedures, a cohort of 210 patients was meticulously assembled, with one WSI randomly selected from each patient. Figure 1 shows the OS distribution of the cases in the dataset, presented by year. Patients' OS was concentrated between 0 and 4 years (Figure 1A). To randomly select training and testing cohorts, we then stratified patients according to whether their survival duration ended in the first half or the last half of a year. When stratified into half-year increments over the span of a year, the distribution of patients across these different OS intervals was relatively balanced (Figure 1B,C). Figure 1D underscores the equitable distribution of patients between the two survival events, where event 1 represents death and 0 represents a censoring event, except for those falling within the 1- to 2-year OS interval.

To establish a balanced partitioning, wherein the training and testing cohorts exhibited comparability across key characteristics such as gender, age, race/ethnicity, TNM categories, overall stage, and OS [24], the utilization of the propensity score-matching algorithm was instrumental in splitting the five-fold dataset. This rigorous approach ensured that our subsequent analyses were conducted on a well-matched and representative subset of the initial patient cohort. When the five-fold cross-validation was applied to divide the current dataset, the OS distributions for each fold were similar.

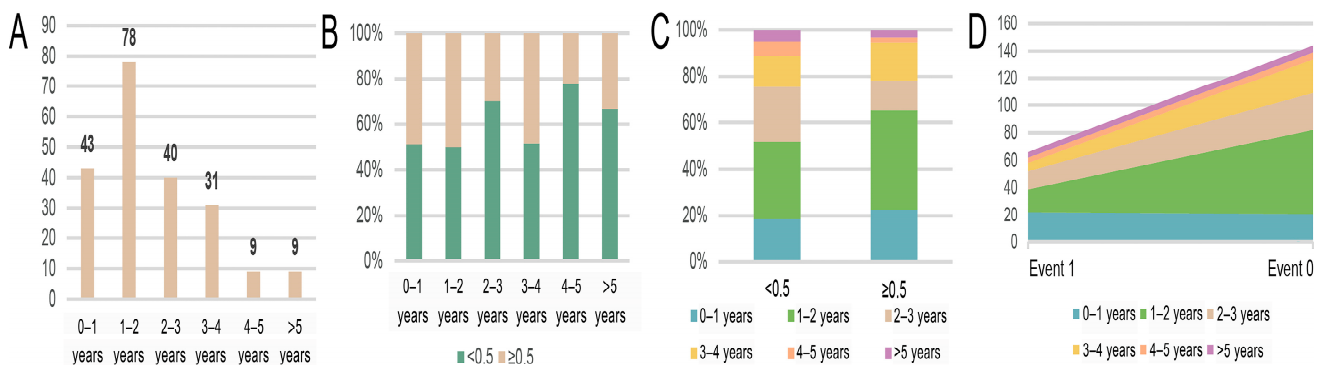


Figure 1. The distribution of dataset statistics. (A) represents the number of cases in different survival years; (B) shows the percentage of patients whose survival duration ended in the first half or the second half of the indicated year based on (A); (C) represents the number of cases per survival year based on (B); (D) shows the number of cases in each survival year for different survival events. Event is the observed state of survival, i.e., event 1 represents death and 0 represents a censoring event.

2.2. Overview of the Proposed Pipeline

Figure 2 shows the proposed automatic segmentation of nuclei and the overall process of its application in survival prediction that mainly includes block-based nuclei segmentation and classification on WSI, aggregating the composition and structure of different cells, embedding those locations and class categories in the embedded WSI map, and extracting features from the embedded WSI map based on a dual-branch feature network to construct a survival prediction model. These steps are detailed in the following sections.

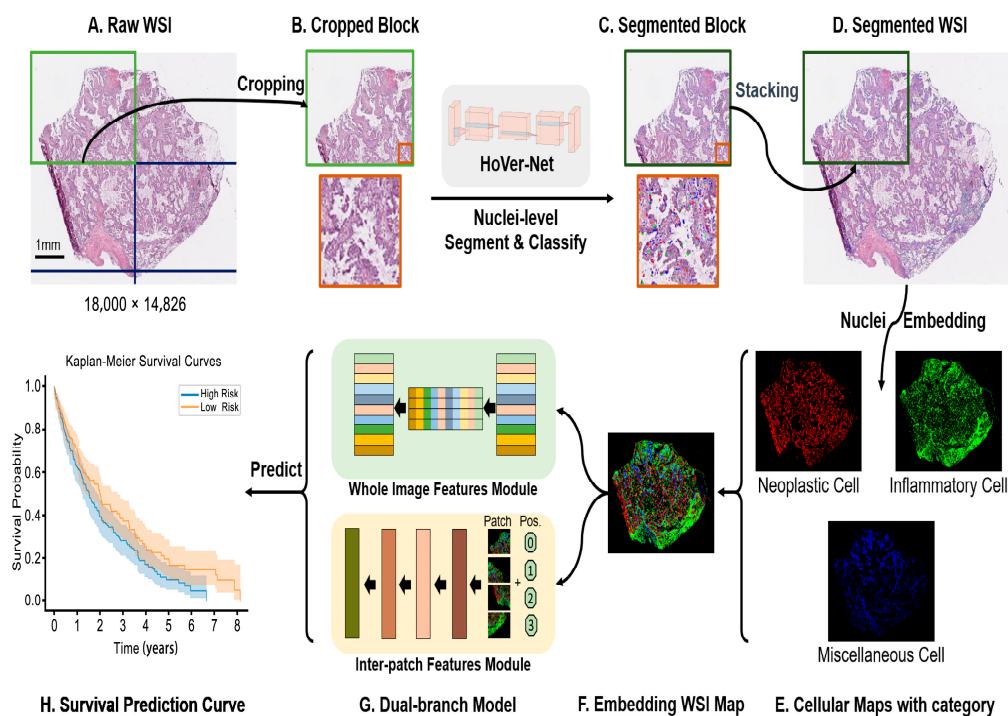


Figure 2. Overview of the proposed pipeline for survival prediction. (A,B): Cropping a whole-slide image (WSI) into multiple blocks. (B,C): Nuclei-based segmentation and classification. (C,D): Stacking segmented blocks as a segmented WSI. (D,E): Generating different types of nuclear maps. E-F: Embedding cell categories on the WSI map. (F,G): Extracting features based on the dual-branch model. (G,H): Prediction of survival risk from final features.

2.3. Block-Based Nuclei Segmentation

Transfer learning has shown outstanding performance in the field of deep learning [25,26]. We harnessed the power of transfer learning by employing the HoVer-Net architecture [27], which has been pre-trained on the PanNuke dataset [28] for both nuclei segmentation and cell type classification tasks. Multiple works have demonstrated the effectiveness of segmentation on the TCGA LUAD dataset directly using the pre-trained model without fine-tuning [29–32]. For example, Chen et al. [32] qualitatively assessed the results of segmentation by three pathologists, in which an accuracy of 0.902 was achieved for cell classification. Traditional methodologies for WSI analysis conventionally involve an initial step of identifying tissue regions, followed by the application of machine learning models to these designated areas. However, this conventional approach has several practical challenges. Many existing tissue-detection algorithms struggle to achieve robust generalization on WSI. Consequently, recourse to manual hyperparameter adjustments becomes necessary to attain satisfactory outcomes. To deal with this predicament, we introduced a novel block-based strategy tailored for the effective handling of gigapixel WSI. Meanwhile, the model-cropped WSIs into multiple blocks enables parallel computation to speed up inference due to computational efficiency and memory constraints. Moreover, we cropped the blocks without overlap in large sizes to minimize the effect of the border. Our approach entails the segmentation and labeling of cells within discrete sections of the WSI. To implement this, we partition each given WSI into smaller blocks (Figure 2B), with a predefined block size of 8000×8000 pixels. Subsequently, we employ the HoVer-Net on these individual blocks separately, enabling the precise segmentation and labeling of cellular entities (Figure 2C).

Upon attaining outputs from the block-wise processing, we stack the blocks to generate WSI maps with comprehensive cell classification (Figure 2D). Altogether, the categories of cells present on the lung adenocarcinoma slides encompass six distinct classes based on the PanNuke dataset, specifically neoplastic, non-neoplastic epithelial, inflammatory, connective, dead, and non-nuclei categories. In the context of our study, we excluded the non-nuclei category, retained the neoplastic and inflammatory categories, and merged the non-neoplastic epithelial, connective, and dead categories into a consolidated “miscellaneous” category. Subsequently, the segmented nuclei were allocated to three distinct categories, denoted as neoplastic (tumor), inflammatory (immune), and miscellaneous (Figure 2E). It should be noted that our results are calculated using direct inference from HoVer-Net and do not require additional annotations for re-training the model.

Thereby, we can compute the WSI with the nuclear annotation from the raw WSI, as shown in Figure 3. Among them, columns (a) and (b) are sampled from the general position and zoomed in on the local detail map, which shows that the segmentation achieves excellent performance and most of the cells are annotated. The sample of magnified area in column (c) is taken from the folded position and shows that the annotations predominantly appear outside of this artifact region. Similarly, in column (d), the model avoids holes in the tissue. This shows that the algorithm has strong robustness to image variations. These outcomes lay the basis for subsequent accurate prognostic analysis.

2.4. Cellular Embedding of Segmented WSI

A WSI, scanned at $20\times$ ($0.50 \mu\text{m}/\text{pixel}$), typically has billions of pixels, which usually contain over a million cells. Extracting cell features from this huge-scale WSI and visualizing them is still challenging. Since we are more focused on the global architecture and distribution of different cell populations than on the contours and textures of individual nuclei, our approach is to compress the local nuclear information to embed cell types in the WSI. First, we encode the abundances of the three cell types (neoplastic, inflammatory, and miscellaneous) for each 64×64 pixel window in three separate channels following Chen et al. [7]. Then, an embedded map of different cellular types (Figure 2F) with a 64-fold dimensional reduction from the raw WSI is obtained. We chose a down-sampling factor under 64 because cell images are commonly cropped to 256×256 pixels [33–35].

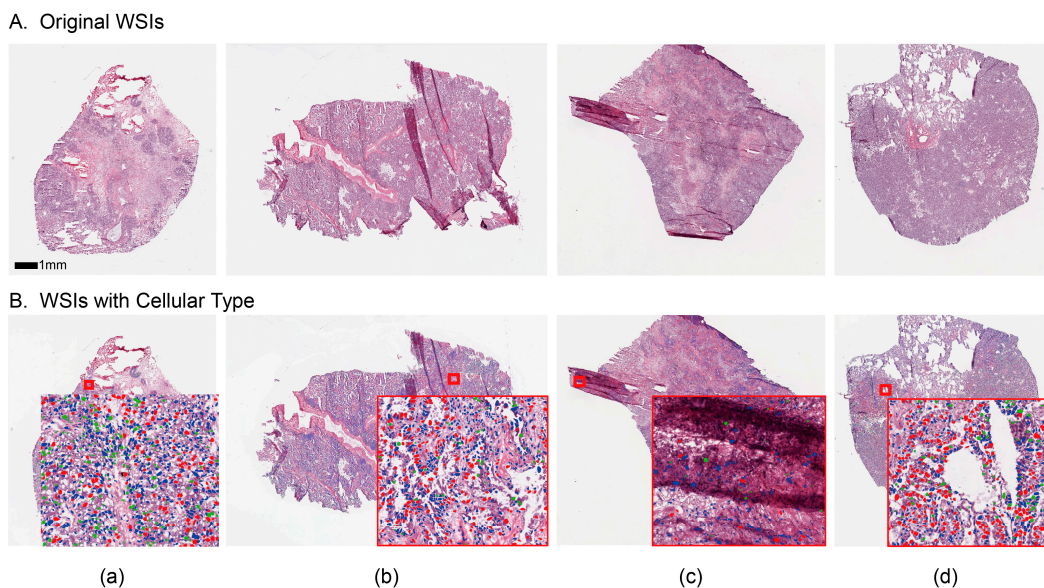


Figure 3. Examples of whole-slide images (WSI) without and with cellular type annotations, shown by magnification. (A) shows the original WSIs, and (B) shows the segmented WSIs with cellular type, in which (a), (b), (c) and (d) represent different samples.

2.5. Dual Global Fusion for Survival Prediction

Figure 2G shows the process of feature extraction and prognostic analysis for embedded WSI maps. This model mainly consists of dual branches; one of the modules focuses on extracting global features from the maps, and the other focuses on extracting contextual features between inter-patches cropped from embedded maps. These two types of features are then combined to predict the patient prognosis.

Specifically, to compute the global features of an embedded WSI map, a network, N_{whole} , is built based on the DeepMixer [36]. This tool is a newly proposed simple network architecture based on multi-layer perceptron (MLP), which achieves competitive performance in image classification tasks. Since its architecture comprises mainly MLP, it is more conducive to learning global features [37,38]. The N_{whole} handpicks only the feature extraction module from the MLP-Mixer, removing the classification projection headers according to our task. Each MLP block contains two fully connected layers and a nonlinearity Gaussian error linear unit [39] applied to each row independently of its input data tensor. One Mixer layer can be written as follows:

$$M = FC(GELU(FC(x))), \quad (1)$$

$$MC = M(LN(X^{S \times C})), \quad (2)$$

$$MP = M(LN(X^{C \times S})), \quad (3)$$

$$Z_l^1 = (Z_{l-1}^1 + MP(Z_{l-1}^1)) + MC(Z_{l-1}^1 + MP(Z_{l-1}^1)), \quad l \in (1, \dots, L) \quad (4)$$

where x is the input, LN is layer normalization [40], L is the total number of Mixer blocks in N_{whole} , and M denotes an MLP block, which returns the output of the same dimension as the input by specifying the hidden layer dimension. MC denotes a channel-mixing MLP, which accepts input with a resolution of $S \times C$, and similarly, MP is a patch-mixing MLP that accepts input images of $C \times S$ resolution. The MC allows communication between different channels, and the MP allows communication between different spatial locations. The feature reuse mechanism aims to reuse the upper layer Z_l^1 . The final global feature is obtained after multiple Mixer layers in branch one.

The second branch applies ViT [41], a visual framework based on the self-attention mechanism with outstanding performance for handling contextual information. Similarly, we extract only the feature coding module as the branch N_{inter} , removing the classification headers. Specifically, the input X is first tokenized by reshaping as a sequence, which consists of all flattened 2D patches x_p , where p is the patch number. To define the spatial position of different patches and learn the contextual information from inter-patches, the position embeddings are added to the patch embedding as follows:

$$Z_0^2 = [x_1E; x_2E; \dots; x_pE] + E_{pos.} \quad (5)$$

where E is the linear projection and $E_{pos.}$ denotes the position embedding. Then, the initial embedding Z_0^2 is fed into the transformer layer to learn the contextual information from inter-patches. The transformer layer consists of multi-head self-attention (MSA) and MLP , which can be formulated as follows:

$$Z_l^2 = MSA\left(LN\left(Z_{l-1}^2\right)\right) + Z_{l-1}^2 + MLP\left(MSA\left(LN\left(Z_{l-1}^2\right)\right) + Z_{l-1}^2\right), \quad l \in (1, \dots, L) \quad (6)$$

where the calculation of MSA is as previously described [42]. After the L layer calculation, the final context feature with attention weight Z_l^2 is obtained in branch two. Finally, we concatenate two features with different focuses and add patient survival events as the final embedded WSI map features $Z = [Z_l^1, Z_l^2, \mu]$.

In a general survival analysis, patients are usually grouped into two categories: censored and non-censored. We use a binary $\mu \in (0, 1)$ to represent the survival events under observation. A censored case means the patient did not experience death during the follow-up period, which denotes they had survived longer than the recorded time. We predicted survival risk based on continuous survival data, differing from classification-based prognostic models. The proposed model directly generates survival risks r corresponding to patients i based on the embedded WSI map features. Hence, our loss function is defined as negative Cox log partial likelihood [43] based on the survival time t as follows:

$$Loss(r_i) = \sum_i \mu_i \left(-r_i + \log \sum_{j:t, j \geq t_j} \exp(r_j) \right). \quad (7)$$

3. Results

3.1. Experiment Implementation Details

For the final prediction of the extracted features, we added a fully connected layer in which the size of the input vector was 257. For each branch of the proposed method as well as the comparative methods, we used the same output length, i.e., 128. The depth of each branch was 2. The feature map as an input of the survival prediction model needs to be of a fixed size. We defined the input resolution as 256. Meanwhile, routine data augmentations were applied to our input data in our experiments, such as right-angle rotation, scaling, flipping, and transformations in brightness, contrast, and saturation. The patch size of the dual branch was set as 8. For the training, the optimizer was Adam, with a 1×10^{-4} learning rate, 5×10^{-3} weight decay, and 0.99 momentum. The batch size and epoch were 64 and 100. We initialized the weights with a fixed random number for initialization. We trained models until the last epoch as a stopping criterion. All models were implemented using PyTorch (version 1.12.1), and all training processes were trained on the two NVIDIA A100 SXM4 40GB graphics processing units in Linux (version 5.4.0-153-generic).

3.2. Evaluation of Prognosis Performance

We first evaluated the performance of the proposed pipeline. Based on the quantitative risk output from the prediction model, we divided patients into low-risk and high-risk groups using the optimal cut-off derived from the quantile classification scheme [44]. The cut-off values were calculated based on all samples in the training or validation

datasets. The Kaplan–Meier estimator and two-sided log-rank test were used to evaluate the significance of patient stratification, and Antolini’s concordance index (C-index) was used to measure the goodness-of-fit of the model discriminative ability [45]. Figure 4A–E illustrates the Kaplan–Meier survival analysis results for each fold model. It can be seen that the p -values between our high-risk group and low-risk group are close to or less than 0.0001 in each fold, which was statistically significant. This shows not only that the data split is reasonable but also that the proposed method is robust.

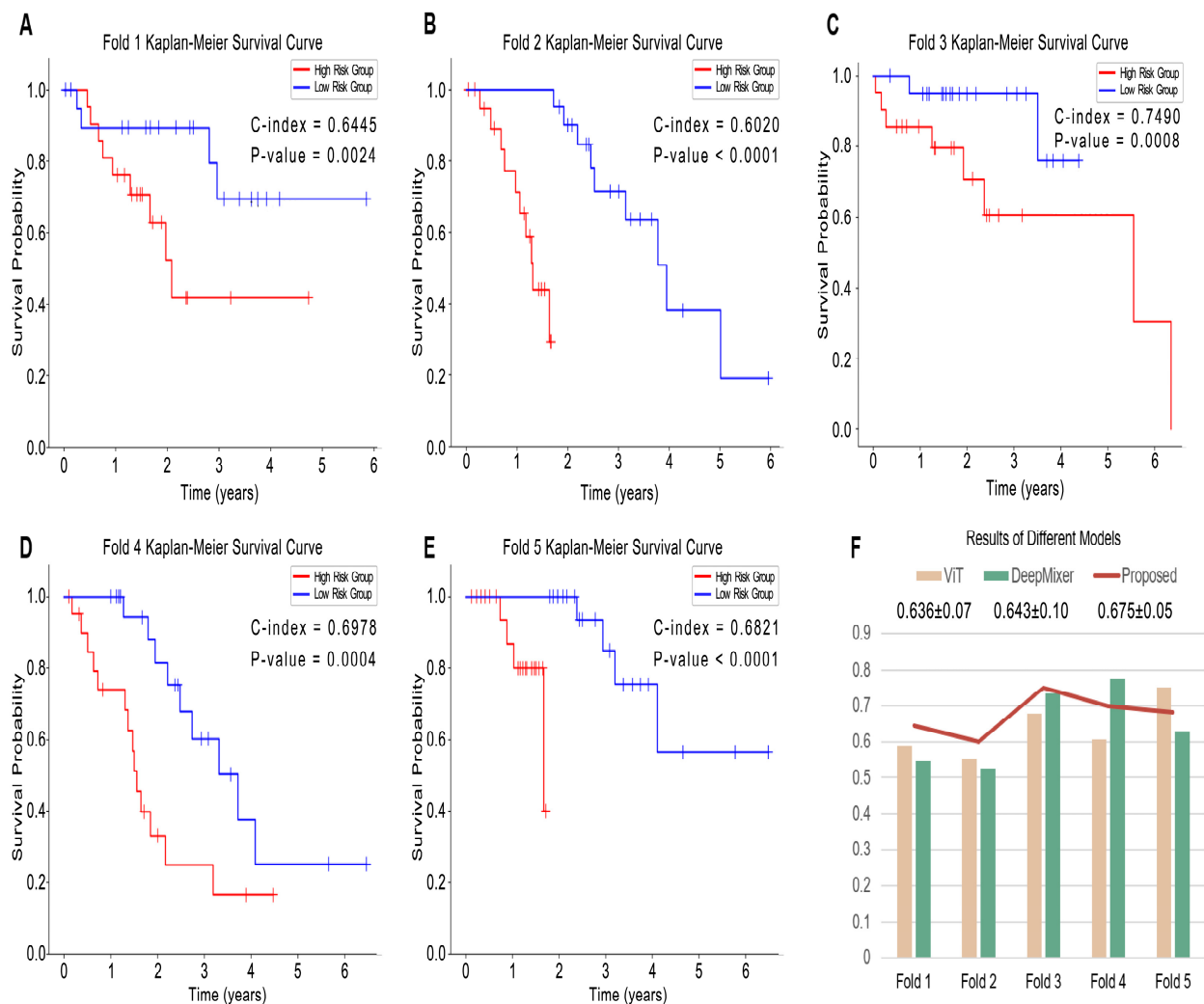


Figure 4. Survival analysis using the proposed methods and its constituent methods. (A–E) shows the results of our proposed method in every fold. (F) compares the C-indexes of different methods and includes the mean and standard deviation for each method.

Among the five folds, the third fold showed the highest C-index value (0.749), and the second and fifth folds had the lowest p -values (both below 0.0001). In the third fold, the results showed that both the high-risk and low-risk groups had a lower survival risk as well as survival times compared to the other folds. Of note, this indicates that our model is more effective in predicting consistent outcomes for patients who have longer survival times. This same finding was validated in the distribution of survival risk in the fourth and fifth folds. In the first fold, there was better differentiation between the low-risk and high-risk groups of results in some patients with longer survival times, while the stratification was not consistent for some patients with shorter survival times. Overall, the performance of all models showed significance in stratifying high-risk and low-risk groups.

Next, our proposed method comprised two constituent models, so we used two different models, ViT and DeepMixer, for our ablation and comparison experiments. As shown in Table 1, the average C-index of our proposed method (0.675) was higher than that of the others. Our method also achieved a higher C-index in the first through third folds. In the fourth fold, our method's C-index was lower than DeepMixer's but higher than ViT's; in the fifth fold, it was lower than ViT's but higher than DeepMixer's, as shown in Figure 4F. The proposed method's C-index also had a lower standard deviation. The results show that our dual-branch structure has better robustness and can compensate for the performance deficiencies of the two branches separately. Furthermore, this gives us an enlightening direction to improve our work in the future. We anticipate making full use of the advantages of the different branches to improve the final performance.

Table 1. Performance comparison with other methods.

Methods	C-Index	p-Value
ViT	0.636 ± 0.07	1.8×10^{-6}
DeepMixer	0.643 ± 0.10	3.7×10^{-5}
Our	0.675 ± 0.05	5.6×10^{-5}

3.3. Feature Visualization Analysis

To further evaluate the performance of the proposed method, we performed a visual study based on the characteristics of each patient. The following analysis unifies the visualization technique using the Uniform Manifold Approximation and Projection (UMAP) [46] tool to understand the underlying data structure. The neighbor number is the most critical parameter of UMAP, which effectively controls how UMAP balances local versus global structure. Low numbers of neighbors will push UMAP to focus more on the local structure by constraining the number of neighboring points considered when analyzing the data in high dimensions. Hence, we conducted two group experiments where the neighbor numbers were equal to 4 and 5 for each visualization. Each column represents the result of a different fold. The corresponding label of the color bar is the survival years.

First, we analyzed the survival risk as shown in Figure 5. We observed that both groups had good separation where no complete overlap was noticed between the samples. Samples with similar survival years were all in the same cluster, and each cluster kept a distance from other clusters. This distinct clustering is crucial for model performance for predicting survival probabilities. When the neighbor number was 4, the number of samples in each cluster was closer to the survival years and the boundaries between different clusters were sharper compared to when the neighbor number was 5, but the latter had a more substantial clustering effect. These results indicate that the proposed method predicts survival probability relatively accurately.

Next, we analyzed the final embedded WSI map features Z and those without survival events. Figures 6 and 7 show the results of the two groups' visualizations, respectively. The results for both groups show that features with events had better differentiation and more pronounced clustering. A linear relationship between samples in the features with the events was perceived which we hypothesized to be particularly beneficial for optimization based on the loss function of survival ranking. This linearity also demonstrates that survival events are deeply integrated into the model and are instrumental in predicting the final probability of survival. Additionally, we also found that better clustering results were obtained and that the cluster number was closer to the survival years when the neighbor number was 4, which was consistent with the previous analysis of survival probability visualization. These results show that the features learned by the model are effective and are better predictors of patient survival probability compared to other models.

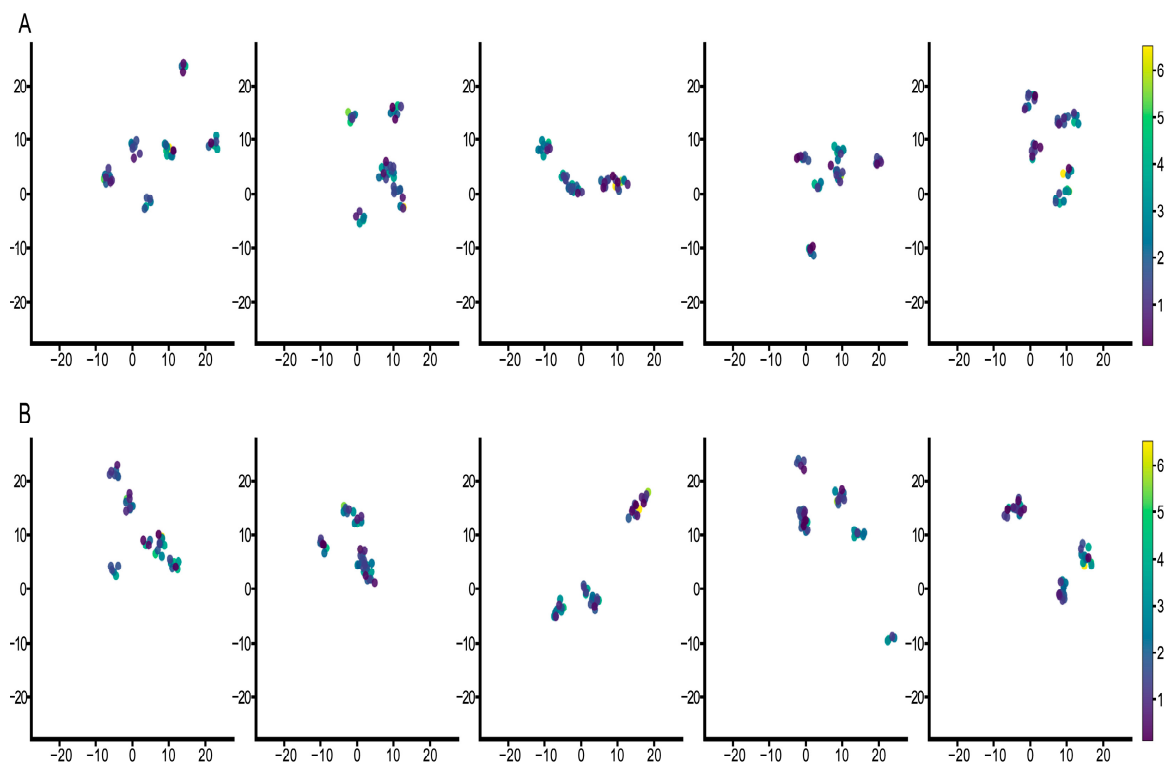


Figure 5. UMAP visualization for analysis of survival probability. (A) shows the results with the number of neighbors equal to 4, and (B) shows results with the number of neighbors equal to 5. Each column represents the result of each fold. Colors indicate survival time.

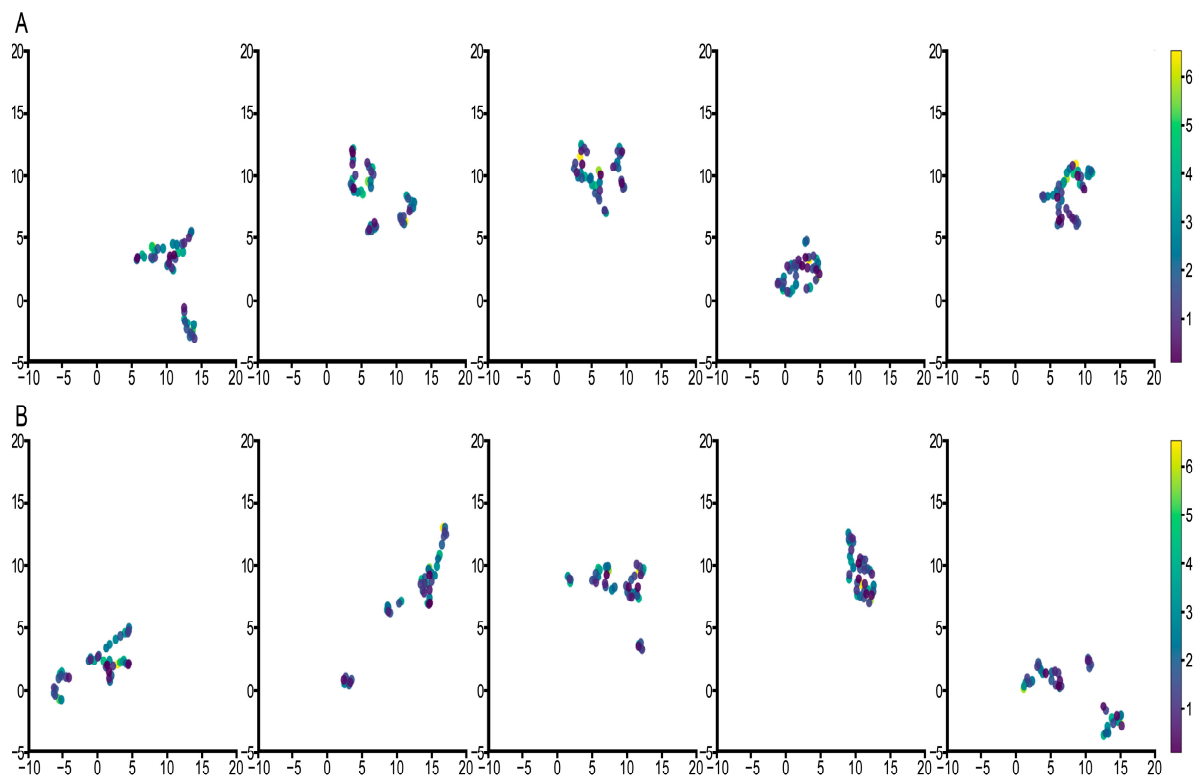


Figure 6. UMAP visualization for analysis of embedded WSI map features when the neighbors' number is 4. (A) represents the result for the feature without survival event, and (B) represents the feature with survival event. Each column represents the result of each fold.

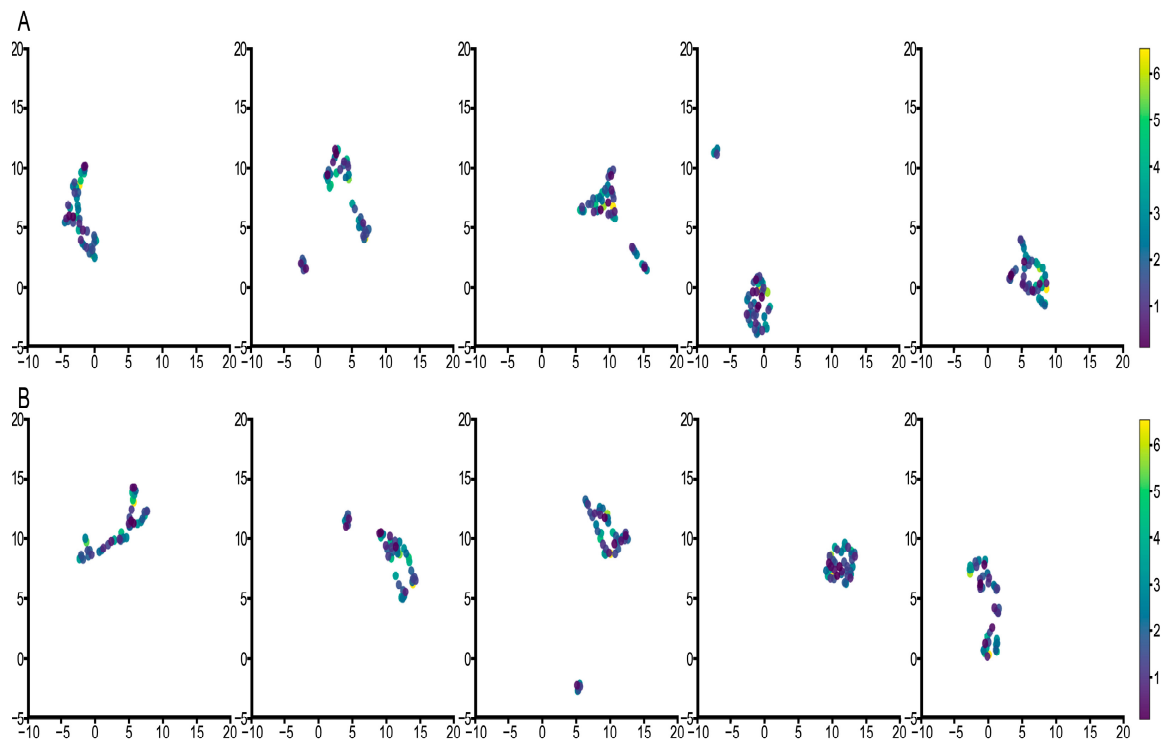


Figure 7. UMAP visualization for analysis of embedded WSI map features when the neighbors' number is 5. (A) represents the result for the feature without survival event, and (B) represents the feature with survival event. Each column represents the result of each fold.

4. Discussion

Here, we present a fully automated cellular-level dual-branch survival prediction pipeline for lung adenocarcinoma pathologic images, which achieved promising results using the LUAD dataset of TCGA. Compared to black-box methods that directly use pathologic images, our proposed method uses more interpretable images at the nucleus level for survival analysis, which makes the results more convincing and comprehensible. Some kernel-based methods partly use manual feature extraction, which means that full automation has not been achieved. These methods extract features directly with convolution, which is more concerned with local regions and exhibits general limitations on modeling explicit long-range relations [47]. These methods not only are unfavorable for learning global features but also have difficulty learning features between cells that are far away from each other.

To address these challenges, we first segment and classify the nuclei from raw WSI. After performing category screening, we generate an embedded WSI map from the segmented images of cell nuclei. The embedded map more directly demonstrates the distribution of various types of cells in the tissue, especially neoplastic cells and inflammatory cells, which are indispensable for the analysis of the tumor microenvironment. The three categories of cells are aggregated into the final embedded WSI map. We also perform $64 \times$ down-sampling to make subsequent survival prediction more efficient. The generated embedded maps are then fed into our dual-branch network. One branch learns global features and the other learns inter-patch features, which only consider the global distribution of various types of cells, but also allow for the computation of long-distance cellular correlations. Subsequently, we fuse these two features from different branches for the final survival prediction, which enables complete automation of survival analysis.

To demonstrate the proposed method, we performed several experiments using the LUAD dataset. First, the model was cross-validated with five folds, and survival Kaplan-Meier curves were plotted separately for each predicted outcome in two groups, high and low risk, and the C-index was calculated. The models all achieved good performance,

and the results of the analysis revealed that our results in each fold were statistically significant. Ablation and comparison studies with other methods showed that the proposed method had a higher average consistency, and less variance in the results of different folds, demonstrating the robustness of the proposed method. Furthermore, visualization of the survival probabilities and the final embedded WSI map features revealed a clustered relationship between the probability predicted by our method and the survival years, as well as a great predictive performance of the embedded map features extracted using our method. Our method provides a fully automated dual-branch-based pipeline for survival prediction using cellular-level information only, which will be valuable in clinical applications.

Our study has several limitations. First, the computational cost of WSI nuclei segmentation and classification is relatively high, and further investigation on how to quickly infer the segmentation results at larger batch sizes is needed. Second, the current prognostic model is validated on a smaller-scale dataset; future validation is planned with other, larger LUAD datasets. Third, this study focuses on predicting prognosis in patients with pathologic images; how to apply our approach to multimodal data such as genes for survival prediction is well worth exploring. Finally, in the proposed dual-branch network, we use a concatenate strategy to fuse different features, and how to merge such features dynamically and adaptively is an important future research direction.

5. Conclusions

This paper presents a fully automated dual-branch-based cellular-level survival prediction pipeline for lung adenocarcinoma pathologic images. This method generates embedded WSI maps for intelligent automated survival analysis by flexibly describing the composition and structure of different cell populations on WSI. This allows us to investigate different cell populations as well as the tumor microenvironment within gigapixel WSI. Furthermore, we use a dual-branch model to learn whole-slide and inter-patch features of cell distribution, allowing the network to sufficiently learn different types and different locations of cells. We demonstrate the effectiveness and robustness of the proposed method with superior experimental results on the LUAD dataset compared to other models. The proposed method is a general WSI-based learning pipeline, which allows researchers to explore the clinical prognosis for predicting different cancer types. In the future, we will expand techniques based on the proposed model to integrate embedded maps with genetic data and improve survival prediction.

Author Contributions: Conceptualization, S.D., P.C. and E.S.; methodology, S.D., P.C. and M.B.S.; software, S.D. and R.B.; validation, S.D., P.C. and J.W.; formal analysis, S.D., F.R.R., B.Z. and L.H.; investigation, S.D., M.A., M.S. and A.M.; data curation, F.R.R., I.I.W., L.M.S.S. and S.J.S.; writing—original draft preparation, S.D. and P.C.; writing—review and editing, S.D., P.C., J.V.H., N.K., I.I.W., L.M.S.S., J.Z., W.Q. and J.W.; visualization, S.D., C.B., D.L.G. and J.V.H.; supervision, W.Q. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Shenzhen Science and Technology Program of China grant JCYJ20200109115420720, and the National Natural Science Foundation of China (No. 62271475); and the Youth Innovation Promotion Association CAS (2022365).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Patient consent was waived as the TCGA data is public.

Data Availability Statement: All the data are available from the public databases, including the TCGA database (<https://portal.gdc.cancer.gov/>, last accessed on 30 January 2022) and UCSC Xena (<https://tcga.xenahubs.net/>, last accessed on 30 January 2022).

Acknowledgments: The manuscript was edited by Sarah Bronson of the Research Medical Library at The University of Texas MD Anderson Cancer Center.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
2. Duma, N.; Santana-Davila, R.; Molina, J.R. Non-small cell lung cancer: Epidemiology, screening, diagnosis, and treatment. *Mayo Clin. Proc.* **2019**, *94*, 1623–1640. [[CrossRef](#)] [[PubMed](#)]
3. Zengin, T.; Önal-Süzek, T. Analysis of genomic and transcriptomic variations as prognostic signature for lung adenocarcinoma. *BMC Bioinform.* **2020**, *21*, 368. [[CrossRef](#)] [[PubMed](#)]
4. Wang, Q.; Li, M.; Yang, M.; Yang, Y.; Song, F.; Zhang, W.; Li, X.; Chen, K. Analysis of immune-related signatures of lung adenocarcinoma identified two distinct subtypes: Implications for immune checkpoint blockade therapy. *Aging* **2020**, *12*, 3312. [[CrossRef](#)]
5. Altman, D.G. Prognostic models: A methodological framework and review of models for breast cancer. *Cancer Investig.* **2009**, *27*, 235–243. [[CrossRef](#)]
6. Rorke, L.B. Pathologic diagnosis as the gold standard. *Cancer* **1997**, *79*, 665–667. [[CrossRef](#)]
7. Chen, P.; Saad, M.B.; Rojas, F.R.; Salehjahreni, M.; Aminu, M.; Bandyopadhyay, R.; Hong, L.; Ebare, K.; Behrens, C.; Gibbons, D.L. Cellular Architecture on Whole Slide Images Allows the Prediction of Survival in Lung Adenocarcinoma. In Proceedings of the International Workshop on Computational Mathematics Modeling in Cancer Analysis, Singapore, 18 September 2022; pp. 1–10.
8. Xu, Z.; Lim, S.; Shin, H.-K.; Uhm, K.-H.; Lu, Y.; Jung, S.-W.; Ko, S.-J. Risk-aware survival time prediction from whole slide pathological images. *Sci. Rep.* **2022**, *12*, 21948. [[CrossRef](#)]
9. Diao, S.; Hou, J.; Yu, H.; Zhao, X.; Sun, Y.; Lambo, R.L.; Xie, Y.; Liu, L.; Qin, W.; Luo, W. Computer-aided pathologic diagnosis of nasopharyngeal carcinoma based on deep learning. *Am. J. Pathol.* **2020**, *190*, 1691–1700. [[CrossRef](#)]
10. Singh, P.; Cirrone, J. A data-efficient deep learning framework for segmentation and classification of histopathology images. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; pp. 385–405.
11. Lu, M.Y.; Chen, T.Y.; Williamson, D.F.; Zhao, M.; Shady, M.; Lipkova, J.; Mahmood, F. AI-based pathology predicts origins for cancers of unknown primary. *Nature* **2021**, *594*, 106–110. [[CrossRef](#)]
12. Diao, S.; Tian, Y.; Hu, W.; Hou, J.; Lambo, R.; Zhang, Z.; Xie, Y.; Nie, X.; Zhang, F.; Racoceanu, D. Weakly supervised framework for cancer region detection of hepatocellular carcinoma in whole-slide pathologic images based on multiscale attention convolutional neural network. *Am. J. Pathol.* **2022**, *192*, 553–563. [[CrossRef](#)]
13. Hekler, A.; Utikal, J.S.; Enk, A.H.; Berking, C.; Klode, J.; Schadendorf, D.; Jansen, P.; Franklin, C.; Holland-Letz, T.; Krahl, D. Pathologist-level classification of histopathological melanoma images with deep neural networks. *Eur. J. Cancer* **2019**, *115*, 79–83. [[CrossRef](#)] [[PubMed](#)]
14. Luo, X.; Zang, X.; Yang, L.; Huang, J.; Liang, F.; Rodriguez-Canales, J.; Wistuba, I.I.; Gazdar, A.; Xie, Y.; Xiao, G. Comprehensive computational pathological image analysis predicts lung cancer prognosis. *J. Thorac. Oncol.* **2017**, *12*, 501–509. [[CrossRef](#)] [[PubMed](#)]
15. Wang, S.; Yang, D.M.; Rong, R.; Zhan, X.; Fujimoto, J.; Liu, H.; Minna, J.; Wistuba, I.I.; Xie, Y.; Xiao, G. Artificial intelligence in lung cancer pathology image analysis. *Cancers* **2019**, *11*, 1673. [[CrossRef](#)] [[PubMed](#)]
16. Yao, J.; Zhu, X.; Jonnagaddala, J.; Hawkins, N.; Huang, J. Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. *Med. Image Anal.* **2020**, *65*, 101789. [[CrossRef](#)]
17. Lu, C.; Koyuncu, C.; Corredor, G.; Prasanna, P.; Leo, P.; Wang, X.; Janowczyk, A.; Bera, K.; Lewis, J., Jr.; Velcheti, V. Feature-driven local cell graph (FLoCK): New computational pathology-based descriptors for prognosis of lung cancer and HPV status of oropharyngeal cancers. *Med. Image Anal.* **2021**, *68*, 101903. [[CrossRef](#)]
18. Othman, N.A.; Abdel-Fattah, M.A.; Ali, A.T. A Hybrid Deep Learning Framework with Decision-Level Fusion for Breast Cancer Survival Prediction. *Big Data Cogn. Comput.* **2023**, *7*, 50. [[CrossRef](#)]
19. Sandarenu, P.; Millar, E.K.; Song, Y.; Browne, L.; Beretov, J.; Lynch, J.; Graham, P.H.; Jonnagaddala, J.; Hawkins, N.; Huang, J. Survival prediction in triple negative breast cancer using multiple instance learning of histopathological images. *Sci. Rep.* **2022**, *12*, 14527. [[CrossRef](#)]
20. Diao, S.; Tang, L.; He, J.; Zhao, H.; Luo, W.; Xie, Y.; Qin, W. Automatic Computer-Aided Histopathologic Segmentation for Nasopharyngeal Carcinoma Using Transformer Framework. In Proceedings of the International Workshop on Computational Mathematics Modeling in Cancer Analysis, Singapore, 18 September 2022; pp. 141–149.
21. Lu, C.; Lewis, J.S., Jr.; Dupont, W.D.; Plummer, W.D., Jr.; Janowczyk, A.; Madabhushi, A. An oral cavity squamous cell carcinoma quantitative histomorphometric-based image classifier of nuclear morphology can risk stratify patients for disease-specific survival. *Mod. Pathol.* **2017**, *30*, 1655–1665. [[CrossRef](#)]
22. Lu, C.; Romo-Bucheli, D.; Wang, X.; Janowczyk, A.; Ganesan, S.; Gilmore, H.; Rimm, D.; Madabhushi, A. Nuclear shape and orientation features from H&E images predict survival in early-stage estrogen receptor-positive breast cancers. *Lab. Invest.* **2018**, *98*, 1438–1448.
23. Alsubaie, N.M.; Snead, D.; Rajpoot, N.M. Tumour nuclear morphometrics predict survival in lung adenocarcinoma. *IEEE Access* **2021**, *9*, 12322–12331. [[CrossRef](#)]
24. Rosenbaum, P.R.; Rubin, D.B. The central role of the propensity score in observational studies for causal effects. *Biometrika* **1983**, *70*, 41–55. [[CrossRef](#)]

25. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A comprehensive survey on transfer learning. *Proc. IEEE* **2020**, *109*, 43–76. [[CrossRef](#)]
26. Tsuneki, M.; Abe, M.; Kanavati, F. A deep learning model for prostate adenocarcinoma classification in needle biopsy whole-slide images using transfer learning. *Diagnostics* **2022**, *12*, 768. [[CrossRef](#)]
27. Graham, S.; Vu, Q.D.; Raza, S.E.A.; Azam, A.; Tsang, Y.W.; Kwak, J.T.; Rajpoot, N. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med. Image Anal.* **2019**, *58*, 101563. [[CrossRef](#)] [[PubMed](#)]
28. Gamper, J.; Alemi Koochbanani, N.; Benet, K.; Khuram, A.; Rajpoot, N. Pannuke: An open pan-cancer histology dataset for nuclei instance segmentation and classification. In Proceedings of the Digital Pathology: 15th European Congress, ECDP 2019, Proceedings 15, Warwick, UK, 10–13 April 2019; pp. 11–19.
29. Gupta, R.; Le, H.; Van Arnam, J.; Belinsky, D.; Hasan, M.; Samaras, D.; Kurc, T.; Saltz, J.H. Characterizing immune responses in whole slide images of cancer with digital pathology and pathomics. *Curr. Pathobiol. Rep.* **2020**, *8*, 133–148. [[CrossRef](#)]
30. Wei, Z.; Zhao, X.; Chen, J.; Sun, Q.; Wang, Z.; Wang, Y.; Ye, Z.; Yuan, Y.; Sun, L.; Jing, J. Deep learning-based stratification of gastric cancer patients from hematoxylin and eosin-stained whole slide images by predicting molecular features for immunotherapy response. *Am. J. Pathol.* **2023**, *193*, 1517–1527. [[CrossRef](#)] [[PubMed](#)]
31. Ding, R.; Rodriguez, E.; Da Silva, A.C.A.L.; Hsu, W. Using Graph Neural Networks to Capture Tumor Spatial Relationships for Lung Adenocarcinoma Recurrence Prediction. In Proceedings of the 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), Cartagena de Indias, Colombia, 18–21 April 2023; pp. 1–5.
32. Chen, P.; Rojas, F.R.; Hu, X.; Serrano, A.; Zhu, B.; Chen, H.; Hong, L.; Bandyopadhyay, R.; Aminu, M.; Kalhor, N. Pathomic Features Reveal Immune and Molecular Evolution from Lung Preneoplasia to Invasive Adenocarcinoma. *Mod. Pathol.* **2023**, *36*, 100326. [[CrossRef](#)]
33. Majurski, M.; Manescu, P.; Padi, S.; Schaub, N.; Hotaling, N.; Simon, C., Jr.; Bajcsy, P. Cell image segmentation using generative adversarial networks, transfer learning, and augmentations. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019.
34. Riana, D.; Jamil, M.; Hadiani, S.; Na'am, J.; Sutanto, H.; Sukwadi, R. Model of Watershed Segmentation in Deep Learning Method to Improve Identification of Cervical Cancer at Overlay Cells. *TEM J.* **2023**, *12*, 813–819. [[CrossRef](#)]
35. Herbsthofer, L.; Ehall, B.; Tomberger, M.; Prietl, B.; Pieber, T.R.; López-García, P. Procedural generation of synthetic multiplex immunohistochemistry images using cell-based image compression and conditional generative adversarial networks. In Proceedings of the Medical Imaging 2022: Digital and Computational Pathology, San Diego, CA, USA, 4 April 2022; pp. 140–148.
36. Tolstikhin, I.O.; Houlsby, N.; Kolesnikov, A.; Beyer, L.; Zhai, X.; Unterthiner, T.; Yung, J.; Steiner, A.; Keysers, D.; Uszkoreit, J. Mlp-mixer: An all-mlp architecture for vision. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 24261–24272.
37. Lai, Z.; Deng, H. Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron. *Comput. Intell. Neurosci.* **2018**, *2018*, 1–13. [[CrossRef](#)]
38. Gao, J.; Lan, J.; Wang, B.; Li, F. SDANet: Spatial deep attention-based for point cloud classification and segmentation. *Mach. Learn.* **2022**, *111*, 1327–1348. [[CrossRef](#)]
39. Hendrycks, D.; Gimpel, K. Gaussian error linear units (gelus). *arXiv* **2016**, arXiv:1606.08415.
40. Ba, J.L.; Kiros, J.R.; Hinton, G.E. Layer normalization. *arXiv* **2016**, arXiv:1607.06450.
41. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv* **2020**, arXiv:2010.11929.
42. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.
43. Zhu, X.; Yao, J.; Huang, J. Deep convolutional neural network for survival analysis with pathological images. In Proceedings of the 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Shenzhen, China, 15–18 December 2016; pp. 544–547.
44. Budczies, J.; Klauschen, F.; Sinn, B.V.; Györfy, B.; Schmitt, W.D.; Darb-Esfahani, S.; Denkert, C. Cutoff Finder: A comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. *PLoS ONE* **2012**, *7*, e51862. [[CrossRef](#)]
45. Antolini, L.; Boracchi, P.; Biganzoli, E. A time-dependent discrimination index for survival data. *Stat. Med.* **2005**, *24*, 3927–3944. [[CrossRef](#)]
46. McInnes, L.; Healy, J.; Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv* **2018**, arXiv:1802.03426.
47. Zhang, Y.; Liu, H.; Hu, Q. Transfuse: Fusing transformers and cnns for medical image segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2021: 24th International Conference, Proceedings, Part I 24, Strasbourg, France, 27 September–1 October 2021; pp. 14–24.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.