



Article AI-Based Risk Score from Tumour-Infiltrating Lymphocyte Predicts Locoregional-Free Survival in Nasopharyngeal Carcinoma

Made Satria Wibawa ^{1,†}, Jia-Yu Zhou ^{2,3,†}, Ruoyu Wang ¹, Ying-Ying Huang ^{2,3}, Zejiang Zhan ^{2,3}, Xi Chen ^{2,3}, Xing Lv ^{2,3,‡}, Lawrence S. Young ⁴ and Nasir Rajpoot ^{1,5,*,‡}

- ¹ Tissue Image Analytics Centre, Department of Computer Science, University of Warwick,
- ² Coventry CV4 7AL, UK; made-satria.wibawa@warwick.ac.uk (M.S.W.); ruoyu.wang.2@warwick.ac.uk (R.W.)
 ² State Key Laboratory of Oncology in South China, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-Sen University Cancer Center, Guangzhou 510060, China; zhoujy@sysucc.org.cn (J.-Y.Z.); huangyy4@sysucc.org.cn (Y.-Y.H.); zhanzj@sysucc.org.cn (Z.Z.); chenxi15@sysucc.org.cn (X.C.); lvxing@sysucc.org.cn (X.L.)
- ³ Department of Nasopharyngeal Carcinoma, Sun Yat-Sen University Cancer Center, Guangzhou 510060, China
- ⁴ Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK; l.s.young@warwick.ac.uk
- ⁵ The Alan Turing Institute, London NW1 2DB, UK
- * Correspondence: n.m.rajpoot@warwick.ac.uk
- [†] These authors share joint first authorship.
- [‡] These authors share joint senior authorship.

Simple Summary: Plasma Epstein–Barr virus (EBV) DNA is an important prognostic marker for nasopharyngeal carcinoma (NPC). However, EBV DNA is less sensitive to locoregional recurrence compared to distant metastasis in NPC. Numerous findings suggest that the presence of tumour-infiltrating lymphocytes (TILs) is associated with NPC prognosis. Nevertheless, NPC is characterised by the presence of abundant TILs. This study aims to generate TIL scores in NPC from H&E-stained tissue slide images for NPC prognosis. We employed artificial intelligence and deep learning-based method for generating TIL score. Our results indicate that our methods have strong prognostic value compared to the EBV DNA copies in locoregional recurrence cases.

Abstract: Background: Locoregional recurrence of nasopharyngeal carcinoma (NPC) occurs in 10% to 50% of cases following primary treatment. However, the current main prognostic markers for NPC, both stage and plasma Epstein–Barr virus DNA, are not sensitive to locoregional recurrence. Methods: We gathered 385 whole-slide images (WSIs) from haematoxylin and eosin (H&E)-stained NPC sections (n = 367 cases), which were collected from Sun Yat-sen University Cancer Centre. We developed a deep learning algorithm to detect tumour nuclei and lymphocyte nuclei in WSIs, followed by density-based clustering to quantify the tumour-infiltrating lymphocytes (TILs) into 12 scores. The Random Survival Forest model was then trained on the TILs to generate risk score. Results: Based on Kaplan-Meier analysis, the proposed methods were able to stratify low- and high-risk NPC cases in a validation set of locoregional recurrence with a statically significant result (p < 0.001). This finding was also found in distant metastasis-free survival (p < 0.001), progression-free survival (p < 0.001), and regional recurrence-free survival (p < 0.05). Furthermore, in both univariate analysis (HR: 1.58, CI: 1.13–2.19, *p* < 0.05) and multivariate analysis (HR:1.59, CI: 1.11–2.28, *p* < 0.05), we also found that our methods demonstrated a strong prognostic value for locoregional recurrence. Conclusion: The proposed novel digital markers could potentially be utilised to assist treatment decisions in cases of NPC.

Keywords: nasopharyngeal carcinoma; tumour-infiltrating lymphocytes; locoregional recurrence; computational pathology; artificial intelligence



Citation: Wibawa, M.S.; Zhou, J.-Y.; Wang, R.; Huang, Y.-Y.; Zhan, Z.; Chen, X.; Lv, X.; Young, L.S.; Rajpoot, N. AI-Based Risk Score from Tumour-Infiltrating Lymphocyte Predicts Locoregional-Free Survival in Nasopharyngeal Carcinoma. *Cancers* **2023**, *15*, 5789. https://doi.org/ 10.3390/cancers15245789

Academic Editor: Dora Kwong

Received: 29 October 2023 Revised: 28 November 2023 Accepted: 8 December 2023 Published: 10 December 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

Nasopharyngeal carcinoma (NPC) is a rare malignancy that originates from the epithelial cells in the nasopharynx, with an epicentre in the Fossa of Rosenmuller, a pharyngeal recess within the nasal cavity. This location allows the tumour to spread through various routes [1–3]. It has a unique aetiology involving environmental factors, genetics and infection with the Epstein–Barr virus (EBV) [4–6]. The primary corrective treatment for NPC is radiation therapy (RT), which may be complemented with either concurrent, induction or adjuvant chemotherapy in advanced cases [7,8]. Clinically, advanced-stage nasopharyngeal carcinoma is typically categorized as locally advanced nasopharyngeal carcinoma (LA NPC), encompassing stage III (T3N0–2M0, T0–2N2M0) and stage IVa (T4 or N3M0) according to the UICC/AJCC Clinical Staging Criteria (8th Edition).

Distant metastases are the main cause of treatment failure. Nonetheless, locoregional recurrence occurs for 10% to 50% of the cases following primary treatment at 5 years [9,10]. This can lead to a higher risk of further recurrence or metastasis to other parts of the body. Early detection and treatment of locoregional recurrence can prevent further tumour spread and improve NPC prognosis [11].

The TNM staging system is the current primary factor for determining treatment decisions and prognosis for patients with NPC. However, despite NPC patients receiving similar treatment at the same stage, their outcomes may vary [12]. The level of circulating cell-free EBV DNA in the plasma of NPC patients has been shown to be an important prognostic marker and can aid in determining appropriate treatment [13,14]. However, EBV DNA is not as sensitive to locoregional recurrence as it is to distance metastasis [15]. It has been demonstrated that as much as 40% of local recurrence is not associated with an increase in plasma EBV DNA [16].

NPC is characterised by the presence of abundant tumour-infiltrating lymphocytes (TILs) [17]. There has been a growing interest in the potential prognostic value of TILs in various types of cancer, including NPC [18]. The presence of TILs within the tumour microenvironment has been associated with better clinical outcome in several malignancies. In NPC, the relationship between TILs and tumour progression has been the subject of numerous studies. Several investigations have reported a positive correlation between the density of TILs and favourable clinical outcomes in NPC patients. In particular, the presence of TILs has been associated with improved overall survival, disease-free survival, and distant metastasis-free survival [19–22]. These findings suggest that TILs may play a critical role in the host immune response against NPC and may serve as a prognostic biomarker for the disease. Despite this potential, it is imperative to reach a consensus on the definition and enumeration of TILs [23] to circumvent subjectivity and intra-observer variability.

Similar to other cancers, NPC diagnosis and prognosis involves a multifaceted approach. This includes symptom checking, physical exams, endoscopies, advanced imaging (MRI, CT, PET), and blood tests. To further validate the presence of tumours and refine the diagnosis, pathology tests are conducted. This involves a careful examination of tissue samples to confirm the existence of tumours and their grading. Simultaneously, radiological assessments are employed for cancer staging, aiding in determining the extent and severity of the disease. The culmination of these diagnostic efforts is a labour- and time-intensive process that extends over the course of more than one week.

Artificial intelligence (AI) and computational pathology in cancer treatment has two primary goals: automating portions of the routine clinical workflow and gaining new insights using data from that same workflow [24]. The emergence of AI and computational pathology in the diagnosis and prognosis of cancer, including NPC, is beginning to deliver promising results in several tasks [25–27]. For instance, AI algorithms have been shown to perform better than human observers in detecting nasopharyngeal mass in endoscopic images [28]. As another example in NPC diagnosis using whole-slide images (WSIs), AI algorithms have been shown to outperform junior and intermediate pathologists [29]. AI has demonstrated its potential in NPC prognosis tasks, achieving high accuracy by utilising various data modalities, such as clinicopathological data, magnetic resonance imaging

(MRI), (positron emission tomography) PET, and WSIs [30–33]. Compared to other data modalities such as MRI and PET, WSIs offer a comprehensive view of cellular and tissue structures. However, to the best of our knowledge, there are no existing studies examining AI for the detailed quantification of TILs in NPC WSIs.

The aim of this study is to categorise patients with locally advanced nasopharyngeal carcinoma into high- and low-risk groups using an AI-generated risk score derived from histology images. We examine the potential of AI and computational pathology to develop a novel digital TIL-based score for NPC (digital NPC-TILs) prognosis in terms of locoregional recurrence-free survival (LRFS). Aside from the LRFS, we also examine distant metastasis-free survival (DMFS), regional recurrence-free survival (RRFS), and progression-free survival (PFS). By adopting AI and computational pathology, our methods offer an objective assessment of TILs from widely used H&E-stained histological images. We conduct detailed analyses and show that our novel digital TIL score has prognostic value in terms of LRFS as well as the other survival endpoints. Our study offers insights into the correlation between TILs, the tumour microenvironment, and patient survival in NPC. The proposed method holds potential for aiding decision making in NPC treatment.

2. Materials and Methods

2.1. Data Collection

A total of 396 whole-slide images (WSIs) for haematoxylin and eosin (H&E)-stained NPC sections (n = 378 cases) were collected from Sun Yat-sen University Cancer Centre. All WSIs were scanned with $40 \times$ magnification at a resolution of 0.255 microns per pixel (mpp) using an Aperio scanner. Among the total of 378 cases in our cohort, 11 cases only have WSIs without any accompanying clinical information, including survival endpoints. This resulted in a final dataset of 367 cases with a total of 385 WSIs. For cases with multiple WSIs, features were aggregated as the average value. Approval was obtained from the ethics committee of the Sun Yat-sen University Cancer Centre (B2023-381-01).

The clinical data include age, sex, primary tumour (T), regional lymph node (N), stage group, and plasma EBV DNA data. M is not included in this study as our focus is on predicting the prognosis of patients with locally recurrent NPC, excluding those with metastasis to other organs. Additionally, the TNM staging status was acquired at the beginning of the study when patients exhibited early symptoms, and thus, metastasis had not yet developed. As in a previous study [34], plasma EBV DNA levels were quantified using quantitative polymerase chain reaction targeting the BAMHI-W region of the EBV genome prior to treatment. The result was reported as the concentration of EBV genome copies per millilitre of plasma. All patients in the cohort received both chemotherapy and radiotherapy.

2.2. Clinicopathological Features

The median follow-up duration in this cohort was 36 months, with a range duration of 1–79 months. The survival data were right censored by limiting the maximum observed time to 60 months (5 years). The average age of all patients was 45 with 70.84% being male. The percentage of events in LRFS, DMFS, PFS, and RRFS were 5%, 12%, 17%, and 2%, respectively. We stratified patients into high- and low-EBV DNA groups for prognosis analysis using 4000 copies/mL as a threshold (low: \leq 4000 copies/mL, high: >4000 copies/mL). This threshold was chosen to align with the latest authoritative studies [35,36]. The cohort consists of 167 high EBV DNA copies cases and 200 low EBV DNA copies. The TNM stage in the cohort was measured using the eighth edition of the International Union Against Cancer (UICC) and the American Joint Committee on Cancer (AJCC) staging system. The majority of patients were in stage 3 based on the TNM staging system (Table 1).

Variable	Sub-Variable	Count	% or Mean (SD)
Age		367	45.64 (11.30)
0	Male	260	70.84%
Sex	Female	le 107	29.16%
	T1	30	8.17%
Т	T2	49	13.35%
	T3	198	53.95%
	T4	90	24.52%
	N0	36	9.81%
N	N1	144	39.24%
N	N2	127	34.60%
	N3	60	16.35%
	Ι	8	2.18%
Stage	II	30	8.17%
	III	194	52.36%
	IV	135	36.78%
	≤ 4000	200	54.50%
EBV DNA copies	>4000	167	45.50%

Table 1. Summary of clinicopathological features in the cohort.

2.3. Tumour Nuclei Detection and Clustering

Our overall analysis pipeline, as shown in Figure 1, consists of four main steps, namely tumour nuclei detection and clustering, digital TIL quantification, risk score calculation and survival analysis. We detected, segmented, and classified nuclei in WSIs at $40 \times$ magnification using the HoverNet model [37], which was trained using PanNuke dataset [38].

In order to correct the misclassification of the nuclei, we trained a deep learning model that we term as MorphResNet that incorporates feature fusion from deep features of the Residual Neural Network (ResNet [39]) and morphology features of nuclei. We utilised features from the penultimate layer of ResNet50, which was pretrained on the ImageNet [40] dataset. We chose ResNet model due to its widespread implementation, as it not only saves computational costs and time, but also maintains predictive power without degradation [41,42]. The morphology features consist of eccentricity, convex area, contour area, equivalent diameter, major axis length, minor axis length, perimeter, solidity, orientation, radius, and bounding box area. The complete model architecture can be found in supplementary file Figure S1.

To construct nuclei dataset, we amalgamated data from PanNuke and MoNuSAC [43] to create nuclei datasets. We selectively included data labelled as "epithelial" and "inflammatory" from MoNuSAC. Each nuclei images in datasets were extracted and centred to a size of 64×64 pixels, with additional padding for those located near the image edge. Total number of nuclei images in the final dataset was 219,934. We augmented our data using the Macenko [44] method to ensure the model's generalisability in the variance of image staining. The data were split into 80-10-10 ratio for training, validation, and testing sets. We trained all layers in model using Adam optimizer for 50 epochs. The model achieved an accuracy of 0.921 and F1-score of 0.919, respectively in the test set. The model's performance was also assessed qualitatively and approved by the medical personnel involved in the study.

We clustered tumour nuclei based on their density with the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [45]. Furthermore, the tumour cluster was expanded by 60 pixels or 15 microns, resulting in a new cluster with a larger area. We set the maximum distance to 200 and the minimum number of samples to 10 in DBSCAN. All the parameters were determined empirically. These parameters allow us to avoid overlap among clusters. The initial cluster was referred to as the inner tumour



cluster (in the midst of tumour clusters), while the expanded region was designated as the outer tumour cluster.

Figure 1. Illustration of the main workflow (**A**): nuclei detection at $40 \times$ magnification, followed by tumour nuclei clustering with density-based method, which resulted in inner tumour cluster (**B**); quantification of digital TILs (**C**).

2.4. Digital TILs and Tumour Morphology

Several studies have shown that tumour microenvironment (TME) correlates with cancer survival. TME consists of stromal cells such as fibroblasts, endothelial cells, and immune cells (lymphocytes) around the tumour region. In this study, our focus was on

tumour and lymphocytic regions and their relation to survival. We formulated digitised TIL scores based on nuclei in H&E images. All lymphocyte nuclei in and around tumour clusters were defined as tumour-infiltrating lymphocytes (TILs).

We term the tumour cells and TILs located in the inner tumour cluster as intratumoural tumours cells and intratumoural TILs (Figure 2). In contrast, tumours and TIL nuclei located in the outer tumour cluster were defined as stromal tumour cells and stromal. The two types of tumour nuclei and TILs are illustrated in Figure 1. By employing measurement of the count and the area of the tumour nuclei and TILs with their respective cluster, we generated 12 digital TIL features. We called this set of features as digital NPC-TILs. The complete list of the digital NPC-TIL features and their formulae can be seen in Table S1. Apart from that, we also examined the tumour morphology in NPC survival. We calculated 9 morphology features from tumour nuclei inside the cluster area. The tumour nuclei in the expanded region of the cluster were not counted for the morphology features measurement. Next, we calculate the mean of the morphology features for each patient and utilise it for downstream analysis. The complete list of morphology features utilised in this study can be seen in Table S9.



Figure 2. Main components of digital NPC-TILs. Inner and outer tumour cluster. Intratumoural and stromal TILs.

2.5. Survival Analysis

We determined locoregional recurrence-free survival (LRFS) as the primary endpoint in our survival analysis. We also examined additional endpoints such as DMFS, RRFS, and PFS. Patients without an event at the last date of the study period were labelled as a censored event. Python 3.7 (Python Software Foundation, www.python.org, accessed on 6 February 2023) programming language was used in the whole pipeline, from nuclei detection and classification to digital TIL scoring and statistical analyses.

In this study, we sought to develop digital AI-based biomarkers that carry prognostic value in NPC. We employed four sets of features in our survival analyses, namely digital TIL scores, clinicopathological data (age, sex, TNM staging, and EBV DNA levels), tumour nuclei morphology, and a combination of digital NPC-TIL scores with tumour nuclei morphology. We trained the Random Survival Forest (RSF) model with each feature set on three-fold cross-validation. We performed a grid search to determine the optimal parameters for RSF. To assess risk score prediction from the model, we leverage Harrell's C-index (concordance index) [46] from the function concordance_index in PySurvival [47] Python package. C-index evaluates the probability of concordance between the predicted survival outcome and the actual survival outcome [48].

Subsequently, we assess the prognostic value of risk scores from the RSF model to stratify patients into low- and high-risk groups. Patients with risk scores equal to or lower than the cut-off value were considered low-risk and vice versa. The cut-off value was determined in the discovery set and calculated using Cutoff Finder [49]. We used Kaplan-Meier (KM) curves analysis with the log-rank test to evaluate the proposed biomarker in NPC prognosis in the validation set.

Furthermore, we carried out univariate and multivariate analyses to evaluate the relationship between risk factors and survival outcomes using the Cox proportional hazards (Cox PH) model. We utilised hazard ratio (HR), C-index, 95% confidence interval (CI), and *p*-value in the lifelines Python package (https://lifelines.readthedocs.io/ accessed on 6 March 2023) to appraise the Cox PH model. All statistical analyses were conducted using a two-tailed test, and a *p*-value lower than 0.05 was considered significant.

To conduct prognostic value assessment and univariate and multivariate analyses, we partitioned the data into discovery and validation sets. The validation set was set to 30% of the data, and the data were stratified based on the event of the corresponding endpoint.

3. Results

As mentioned above, we compared four sets of features in NPC prognosis: clinicopathological, digital NPC-TILs, tumour morphology, and the combination of digital NPC-TILs and tumour morphology. Three-fold cross-validation experiments were conducted to assess the generalisability of the proposed AI models. In cross-validation, data were shuffled randomly and stratified into three groups/folds. These folds were utilised for the discovery and validation of RSF models. For each validation fold, we calculated the risk score from RSF based on digital TIL features and their values as obtained from the discovery fold.

3.1. Patient Stratification with Risk Score of Digital NPC-TILs

We determined locoregional recurrence-free survival (LRFS) as the primary endpoint in our survival analysis. We also examined additional endpoints such as DMFS, RRFS, and PFS. Patients without an event at the last date of the study period were labelled as a censored event. Python 3.7 programming language was used in the whole pipeline, from nuclei detection and classification to digital TIL scoring and statistical analyses. The cut-off for stratifying patients in LRFS, DMFS, PFS, and RRFS were 4.52, 7.42, 28.20, and 3.86, respectively.

As can be observed in Figure 3, the proposed digital NPC-TIL score was able to stratify low- and high-risk NPC cases in LRFS with a statistically significant result ($p = 8.13 \times 10^{-5}$). We also obtain a similar result in other survival endpoints, such as DMFS ($p = 5.83 \times 10^{-4}$), PFS ($p = 6.24 \times 10^{-4}$), and RRFS ($p = 5.67 \times 10^{-3}$).



Figure 3. Kaplan–Meier curves for low- and high-risk groups based on the risk score of digital NPC-TILs.

3.2. Correlation Analysis of Predicted Risk Scores and Time to Event

We examined the performance of our digital NPC-TIL score in survival prediction with the RSF model and compared it with other feature sets. Our digital NPC-TIL score alone performed better compared to other feature sets in LRFS. On the discovery set, digital NPC-TIL score achieved an average C-index of 0.923 and 0.785 on the validation set. Meanwhile, other feature sets achieved lower C-index values (Table 2). Additionally, digital NPC-TIL score also delivered the highest C-index in other survival endpoints. It achieved a C-index of 0.780 in RRFS and 0.667 in PFS. In the case of DMFS, the digital NPC-TIL score achieved a C-index of 0.645, which was lower than the C-index of 0.670 associated with the clinical feature. For the complete C-index results, please refer to Table S2.

Features	Discovery Set C-Index (Mean \pm SD)	Validation Set C-Index (Mean
Clinical	0.873 ± 0.007	0.644 ± 0.124
Digital NPC-TILs	0.923 ± 0.024	0.785 ± 0.066
TM *	0.908 ± 0.017	0.624 ± 0.031
Digital NPC-TILs and Clinical	0.932 ± 0.027	0.670 ± 0.142
Digital NPC-TILs and TM *	0.936 ± 0.020	0.679 ± 0.120
Clinical and TM *	0.939 ± 0.010	0.629 ± 0.086

 0.959 ± 0.008

Table 2. Cross-validation results on LRFS prediction using RSF.

* TM: Tumour morphology.

All Features

3.3. Univariate and Multivariate Analyses

We inspected the relationship between a single risk factor of each clinicopathological data and the risk score of digital TILs with LRFS using the Cox PH regression

 \pm SD)

 0.689 ± 0.132

model (Table 3). In the univariate analysis, we found that only digital NPC-TIL score has prognostic value for LRFS (p < 0.05) with HR = 1.58 (95% CI = 1.13–2.19).

We also found that digital NPC-TIL score has a strong prognostic value in DMFS and PFS. In DMFS, the EBV DNA copies (HR: 3.88, 95% CI: 1.21–12.47, p < 0.05) and the proposed digital NPC-TIL score (HR: 1.36, 95% CI: 1.14–1.16, p < 0.001) were significantly correlated with survival outcome.

Covariate	Sub-Covariate	HR	Lower HR 95%	Upper HR 95%	<i>p</i> -Values	
Age		1.03	0.96	1.11	0.3566	
Gender	Female	references				
	Male	1.64	0.25	10.70	0.6066	
	1	references				
T	2	0.38	0.01	19.01	0.6273	
Т	3	2.89	0.29	28.75	0.3661	
	4	1.58	0.12	20.92	0.7291	
N	0		references			
	1	0.77	0.10	5.79	0.8018	
	2	0.44	0.05	3.82	0.4546	
	3	0.76	0.07	7.88	0.8145	
Stage	Ι		references			
	II	0.32	0	39.50	0.6425	
	III	1.50	0.12	18.11	0.7514	
	IV	1.36	0.11	17.03	0.8096	
EBV DNA copies	≤ 4000	references				
	>4000	1.49	0.30	7.44	0.6273	
Digital NPC-TILs		1.58	1.13	2.19	< 0.05	

Table 3. Univariate analysis of risk factors on LRFS.

In PFS, the stage N = 3 (HR:6.7, 95% CI: 1.45–30.88, p < 0.05) and digital NPC-TIL score (HR: 1.08, 95% CI: 1.03–1.14, p < 0.005) were also significantly correlated with PFS outcome. In contrast, the digital NPC-TIL score was not significantly correlated with RRFS outcome (HR: 1.65, 95% CI: 0.85–3.21) with a p-value of 0.1388. In addition, none of the other features were significantly correlated in the validation set of RRFS. The univariate analysis results on DMFS, RRFS, and PFS can be seen in Tables S3–S5 in Supplementary Materials.

The result of the multivariate analysis on LRFS is shown in Table 4. In the multivariate analysis, we found that only the risk score of digital TILs has a prognostic value in NPC (p < 0.05, HR = 1.59, 95% CI = 1.11–2.28). Digital NPC-TIL score also has prognostic value in DMFS (p < 0.005, HR = 1.35, 95% CI = 1.10–1.67) and PFS (p < 0.005, HR = 1.08, 95% CI = 1.02–1.14). Apart from digital NPC-TILs, only gender covariate has prognostic value in DMFS. The complete list of multivariate analyses can be found in Tables S6–S8.

Table 4. Multivariate analysis of risk factors on LRFS.

Covariate	Sub-Covariate	HR	Lower HR 95%	Upper HR 95%	<i>p</i> -Values
Age		1.05	0.96	1.14	0.285
Gender	Female	references			
	Male	1.20	0.15	9.62	0.866
	1		refe	rences	
Т	2	0.47	0.01	30.02	0.7200
	3	2.78	0.21	36.08	0.4352
	4	1.20	0.06	24.18	0.9073

Covariate	Sub-Covariate	HR	Lower HR 95%	Upper HR 95%	<i>p</i> -Values	
	0		references			
Ν	1	0.56	0.06	5.2	0.6107	
	2	0.38	0.04	4.13	0.4290	
	3	0.71	0.03	16.7	0.8291	
Stage	Ι	references				
	II	0.44	0	88.01	0.7593	
	III	0.96	0.05	17.89	0.9780	
	IV	1.55	0.08	30.61	0.7723	
EBV DNA copies	$\leq \! 4000$	references				
	>4000	1.04	0.17	6.34	0.963	
Digital NPC-TILs		1.59	1.11	2.28	< 0.05	

Table 4. Cont.

3.4. Digital NPC-TIL Scores in Low-Risk and High-Risk Groups

We conducted further analysis in the validation set to investigate how digital NPC-TIL score differs in low- and high-risk group. We performed either *t*-test or a Mann–Whitney U test, depending on the data distribution, to compare the mean between the two groups. As depicted in Figure 4, the low-risk cases have higher mean in 12 digital NPC-TIL features compared to the high-risk group in LRFS with statistically significant value (p < 0.05).



Figure 4. Comparison of digital NPC-TIL features on the low- and high-risk groups in LRFS. The * sign above the feature name denotes a statistically significant difference between low- and high-risk groups (*: p < 0.05; **: p < 0.001). AR = Area Ratio, DT = Density Tumour, DC = Density Cluster.

The difference in digital NPC-TILs between low- and high-risk groups was evident upon examining the WSIs; see for instance, the TILs ratio as shown in Figure 5. Red objects in the figure are cancer cell nuclei and green objects are lymphocytes nuclei. As can be seen, the quantity of lymphocytes in cases with good prognosis is significantly higher compared to those cases with poor prognosis. In this particular case, there were almost no lymphocytic infiltrates in the tumour cluster.



Figure 5. Sample high-power field images of low digital NPC-TIL score with bad prognosis (**Left**) and high digital NPC-TIL score with good prognosis (**Right**).

3.5. Tumour Nuclear Morphology Analysis

We also examined how tumour morphology differs in low- and high-risk groups of LRFS. For this purpose, we extracted nine morphological features from all tumour nuclei in the midst of the tumour clusters. We investigated the size of the tumour nuclei and the shape irregularities of tumour nuclei; higher values indicate larger tumour nuclei exhibiting increased irregularity in shape. We performed the same method as shown in Figure 4 to test for statistically significant difference between two groups. The result of this analysis can be observed in Figure 6. The features in the first two rows in Figure 6 represent tumour cell sizes, while the rest features represent the irregularity in shape. All of the nine morphology features in the low-risk group have a lower mean compared to the high-risk group. Excluding solidity and eccentricity, all the features were found to have a statistically significant difference (p < 0.05).



Figure 6. Tumour cell morphology comparison in low- and high-risk groups of LRFS. The * sign above the feature name denotes a statistically significant difference between low- and high-risk groups (p < 0.05).

3.6. Model Interpretability Analysis

Similar to other tree-based algorithms, RSF provides model interpretability via feature importance. Utilising the feature importance of RSF, we analysed the importance of each of the 12 digital TIL scores in terms of their contribution to the final digital NPC-TIL score in the validation set. As shown in Figure 7, the top 3 most important features in LRFS were stromal TIL density tumour, stromal TIL ratio, and TIL density cluster. Stromal TIL features (both tumour and ratio) refer to lymphocytes that are located at the outer edge of tumour clusters. Furthermore, the feature with the highest occurrence among the top three features across all survival endpoints (Figure S2) was the TIL density cluster, which captures information regarding the number of TILs within and around the tumour cluster.



Figure 7. Rank of feature importance in risk score of LRFS.

4. Discussion

With current limitations of the TNM staging system and clinical data [50–52], there is increasing interest in immune features of the tumour microenvironment (TME) in NPC prognosis [22,53]. Despite the promising results of immune features of the TME (such as TILs) in other types of cancers, there is a dearth of literature in exploring immune features of the TME in NPCs. In addition, most immune features are subjectively assessed and are susceptible to inter- and intra-observer variation [54].

To address this issue, this study presents a novel method to automate the scoring of TILs based on H&E WSIs in NPC using artificial intelligence and computer vision. We developed 12 digital TIL features and combined them into a single NPC-TIL score. Our analysis suggests that this score is an independent prognostic factor for NPC in terms of LRFS and other survival outcomes in our cohort.

We captured TIL information in three locations: inner tumour cluster (in the midst of tumour cluster), outer tumour cluster (in the expanded area of tumour cluster), and in both of them. Each location encompasses four sets of digital NPC TILs. TILs extracted from the inner tumour cluster were given the prefix "Intratumoural" as they describe lymphocytes inside the high-density tumour area. Furthermore, TILs extracted from the outer tumour cluster were given the prefix "Stromal" as they were found around the tumour cluster.

The first set of digital NPC TILs provides a distribution of the number of lymphocyte cells with the number of tumour cells, which were named TIL ratio, intratumoural TIL ratio, and stromal ratio. The second set provides insights into the relationship between the number of lymphocyte cells and the total area occupied by tumour cells, and these were named TIL density cluster, intratumoural TIL density cluster, and stromal TIL density cluster. The third set compares the number of lymphocyte cells with the total area of

tumour cells, named TIL density tumour, intratumoural TIL density tumour, and stromal TIL density tumour. Lastly, the fourth set delves into the spatial aspect by considering the total area covered by lymphocyte cells in relation to the total area of tumour cells, which were named the TIL area ratio, intratumoural TIL area ratio, and stromal TIL area ratio. Adopting these four feature sets across three distinct locations enhances our ability to characterize not only the quantity, but also the spatial distribution and density of TILs, providing a comprehensive understanding of their role in the intricate interplay within the tumour microenvironment.

It is known that EBV DNA is an important biomarker for distant metastasis in NPC [55,56]. However, this marker is known to be more sensitive to distant metastasis than to locoregional recurrence [57,58]. In our univariate and multivariate analyses with Cox-PH modelling, we showed that our digital NPC-TIL score is a strong prognostic factor in the LRFS of NPC. A stratification of low- and high-risk in locoregional recurrence NPC by employing this risk score yields a statistically significant result (p < 0.001).

The aim of this study is to stratify high- and low-risk groups among patients with locally advanced nasopharyngeal carcinoma through an AI-based risk score from histology images. For low-risk nasopharyngeal carcinoma patients, a "de-intensification" treatment strategy could be implemented to enhance their quality of life. Simultaneously, for high-risk locally advanced nasopharyngeal carcinoma cases, additional therapeutic measures, such as maintenance treatment with capecitabine following induction chemotherapy or incorporating full-course immunotherapy, could be considered.

We explored digital NPC-TIL scores on NPC prognosis in LRFS. As can be seen in Figure 4, digital NPC-TIL scores with good prognosis are significantly higher compared to those cases with poor prognosis. Our tumour nuclear morphology analysis suggested that digital NPC-TIL scores can capture nuclear aberration in local recurrence cases. In the poor prognosis cases, the tumour nuclei tend to have larger size and more irregular shape compared to tumour nuclei in good prognosis cases. It is known that nuclear tumour aberration is a hallmark feature of many tumours [59].

By examining model interpretability in our RSF model, we found that TIL density cluster is the most common feature in digital NPC-TILs with the highest importance across all survival endpoints. This feature captures information regarding the number of TILs within and around the tumour cluster pointing to immune response within the tumour microenvironment and confirming that the presence and quantity of TILs have implications for patient prognosis in NPC.

5. Conclusions

We have introduced a novel AI-based method that utilises routine H&E histology images and developed a set of novel digital TIL features for NPC prognosis. To the best of our knowledge, this is the first study to automate TIL scores in H&E images of NPC and to show potential prognostic markers which can help to assess patient treatment. The study provides important insights into the potential digital immune markers in NPC prognosis. The digital NPC-TIL scores could potentially be utilised to assist treatment decisions and improve patient outcomes in cases of local recurrence. Despite the promising findings about digital TILs in this study, our samples were derived solely from a single cohort. Furthermore, the absence of information on important factors such as lifestyle and comorbidities represent another constraint. Large-scale validation on multi-centric cohorts is required before the findings of this study can be implemented into clinical practice.

Supplementary Materials: The following supporting information can be downloaded at: https: //www.mdpi.com/article/10.3390/cancers15245789/s1, Figure S1. Nuclei classification model; Figure S2. Ranking of feature importance; Table S1. Digital NPC-TILs formula; Table S2. Three-fold cross-validation results of concordance index for RSF model; Tables S3–S5. Univariate analysis on DMFS, RRFS, PFS; Tables S6–S8. Multivariate analysis on DMFS, RRFS, PFS; Table S9. Tumour morphology features. **Author Contributions:** Conceptualization X.L., L.S.Y. and N.R.; methodology M.S.W.; software M.S.W.; formal analysis M.S.W.; investigation M.S.W.; data curation J.-Y.Z., Y.-Y.H., Z.Z. and X.C.; writing—original draft preparation M.S.W.; writing—review and editing M.S.W., J.-Y.Z., R.W., Y.-Y.H., Z.Z., X.C., X.L., L.S.Y. and N.R.; visualization M.S.W.; supervision: X.L., L.S.Y. and N.R.; project administration: R.W. and J.-Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: M.S.W. is funded by a PhD scholarship by the Indonesia Endowment Fund for Education (LPDP), Ministry of Finance, Republic of Indonesia, under grant number Ref: S-575/LPDP.4/2020. R.W. is supported by a PhD studentship funded by the General Charities of the City of Coventry (Grant reference: 67148) and the Computer Science Doctoral Training Centre at the University of Warwick. The Sun Yat-Sen component of the study is funded by The Natural Science Foundation of Guangdong Province, China (No. 2023B1515020044) and The National Natural Science Foundation of China (No. 82172863).

Institutional Review Board Statement: Approval was obtained from the ethics committee of the Sun Yat-sen University Cancer Centre (protocol code: B2023-381-01; date for the approval: 7 August 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study for the collection of biological samples. Patient consent was waived for the publication of data due to the anonymized records prior to analysis.

Data Availability Statement: Code for generating digital NPC-TIL score can be accessed at https: //github.com/mdsatria/npc_digital_tils (accessed on 21 November 2023) (will be available upon publication). Data can be made available upon request to the corresponding author (n.m.rajpoot@warwick.ac.uk).

Conflicts of Interest: N.R. reports a relationship with Histofy Ltd. that includes board membership. N.R. reports a relationship with GSK and AstraZeneca that includes funding grants.

References

- Young, L.S.; Dawson, C.W.; Woodman, C.B.J.; Rabkin, C.S.; Lo, K.W. Human Herpesviruses: Nasopharyngeal Carcinoma and Other Epithelial Tumors. In *Viral Infections of Humans*; Kaslow, R.A., Stanberry, L.R., Powers, A.M., Eds.; Springer: New York, NY, USA, 2022; pp. 1–40, ISBN 978-1-4939-9544-8.
- 2. Chen, Y.-P.; Chan, A.T.C.; Le, Q.-T.; Blanchard, P.; Sun, Y.; Ma, J. Nasopharyngeal Carcinoma. Lancet 2019, 394, 64–80. [CrossRef]
- Rickinson, A.B.; Lo, K.W. Nasopharyngeal Carcinoma. In Nasopharyngeal Carcinoma; Elsevier: Amsterdam, The Netherlands, 2019; pp. 1–16, ISBN 978-0-12-814936-2.
- 4. Wong, K.C.W.; Hui, E.P.; Lo, K.-W.; Lam, W.K.J.; Johnson, D.; Li, L.; Tao, Q.; Chan, K.C.A.; To, K.-F.; King, A.D.; et al. Nasopharyngeal Carcinoma: An Evolving Paradigm. *Nat. Rev. Clin. Oncol.* **2021**, *18*, 679–695. [CrossRef]
- 5. Chan, A.T.C. Nasopharyngeal Carcinoma. Ann. Oncol. 2010, 21, vii308–vii312. [CrossRef]
- Wibawa, M.S.; Lo, K.-W.; Young, L.S.; Rajpoot, N. Multi-Scale Attention-Based Multiple Instance Learning for Classification of Multi-Gigapixel Histology Images. In *Computer Vision—ECCV 2022 Workshops*; Karlinsky, L., Michaeli, T., Nishino, K., Eds.; Lecture Notes in Computer Science; Springer Nature Switzerland: Cham, Switzerland, 2023; Volume 13807, pp. 635–647, ISBN 978-3-031-25081-1.
- Poh, S.S.; Soong, Y.L.; Sommat, K.; Lim, C.M.; Fong, K.W.; Tan, T.W.; Chua, M.L.; Wang, F.Q.; Hu, J.; Wee, J.T. Retreatment in Locally Recurrent Nasopharyngeal Carcinoma: Current Status and Perspectives. *Cancer Commun.* 2021, 41, 361–370. [CrossRef] [PubMed]
- Zhang, L.; Chen, Q.Y.; Liu, H.; Tang, L.Q.; Mai, H.Q. Emerging Treatment Options for Nasopharyngeal Carcinoma. *Drug Des. Dev. Ther.* 2013, 7, 37–52. [CrossRef]
- 9. Newton, E.; Valenzuela, D.; Foley, J.; Thamboo, A.; Prisman, E. Outcomes for the Treatment of Locoregional Recurrent Nasopharyngeal Cancer: Systematic Review and Pooled Analysis. *Head Neck* **2021**, *43*, 3979–3995. [CrossRef] [PubMed]
- Au, K.H.; Ngan, R.K.C.; Ng, A.W.Y.; Poon, D.M.C.; Ng, W.T.; Yuen, K.T.; Lee, V.H.F.; Tung, S.Y.; Chan, A.T.C.; Sze, H.C.K.; et al. Treatment Outcomes of Nasopharyngeal Carcinoma in Modern Era after Intensity Modulated Radiotherapy (IMRT) in Hong Kong: A Report of 3328 Patients (HKNPCSG 1301 Study). Oral Oncol. 2018, 77, 16–21. [CrossRef]
- Chen, M.-Y.; Jiang, R.; Guo, L.; Zou, X.; Liu, Q.; Sun, R.; Qiu, F.; Xia, Z.-J.; Huang, H.-Q.; Zhang, L.; et al. Locoregional Radiotherapy in Patients with Distant Metastases of Nasopharyngeal Carcinoma at Diagnosis. *Chin. J. Cancer* 2013, *32*, 604–613. [CrossRef]
- Liu, K.; Xia, W.; Qiang, M.; Chen, X.; Liu, J.; Guo, X.; Lv, X. Deep Learning Pathological Microscopic Features in Endemic Nasopharyngeal Cancer: Prognostic Value and Protentional Role for Individual Induction Chemotherapy. *Cancer Med.* 2020, 9, 1298–1306. [CrossRef]

- Lin, J.-C.; Chen, K.Y.; Wang, W.-Y.; Jan, J.-S.; Liang, W.-M.; Tsai, C.-S.; Wei, Y.-H. Detection of Epstein-Barr Virus DNA in the Peripheral-Blood Cells of Patients With Nasopharyngeal Carcinoma: Relationship to Distant Metastasis and Survival. J. Clin. Oncol. 2016, 19, 2607–2615. [CrossRef]
- Chan, A.T.C.; Hui, E.P.; Ngan, R.K.C.; Tung, S.Y.; Cheng, A.C.K.; Ng, W.T.; Lee, V.H.F.; Ma, B.B.Y.; Cheng, H.C.; Wong, F.C.S.; et al. Analysis of Plasma Epstein-Barr Virus DNA in Nasopharyngeal Cancer after Chemoradiation to Identify High-Risk Patients for Adjuvant Chemotherapy: A Randomized Controlled Trial. JCO 2018, 36, 3091–3100. [CrossRef] [PubMed]
- 15. Lee, A.W.M.; Lung, M.L.; Ng, W.T. (Eds.) Nasopharyngeal Carcinoma: From Etiology to Clinical Practice; Academic Press, an imprint of Elsevier: London, UK; San Diego, CA, USA, 2019; ISBN 978-0-12-814936-2.
- Wei, W.I.; Yuen, A.P.W.; Ng, R.W.M.; Ho, W.K.; Kwong, D.L.W.; Sham, J.S.T. Quantitative Analysis of Plasma Cell-Free Epstein-Barr Virus DNA in Nasopharyngeal Carcinoma after Salvage Nasopharyngectomy: A Prospective Study. *Head Neck* 2004, 26, 878–883. [CrossRef]
- 17. Forder, A.; Stewart, G.L.; Telkar, N.; Lam, W.L.; Garnis, C. New Insights into the Tumour Immune Microenvironment of Nasopharyngeal Carcinoma. *Curr. Res. Immunol.* 2022, *3*, 222–227. [CrossRef]
- Almangush, A.; Ruuskanen, M.; Hagström, J.; Hirvikoski, P.; Tommola, S.; Kosma, V.-M.; Nieminen, P.; Mäkitie, A.; Leivo, I. Tumor-Infiltrating Lymphocytes Associate with Outcome in Nonendemic Nasopharyngeal Carcinoma: A Multicenter Study. *Hum. Pathol.* 2018, *81*, 211–219. [CrossRef] [PubMed]
- Lv, S.-H.; Li, W.-Z.; Liang, H.; Liu, G.-Y.; Xia, W.-X.; Xiang, Y.-Q. Prognostic and Predictive Value of Circulating Inflammation Signature in Non-Metastatic Nasopharyngeal Carcinoma: Potential Role for Individualized Induction Chemotherapy. JIR 2021, 14, 2225–2237. [CrossRef] [PubMed]
- Wang, Y.-Q.; Chen, Y.-P.; Zhang, Y.; Jiang, W.; Liu, N.; Yun, J.-P.; Sun, Y.; He, Q.-M.; Tang, X.-R.; Wen, X.; et al. Prognostic Significance of Tumor-Infiltrating Lymphocytes in Nondisseminated Nasopharyngeal Carcinoma: A Large-Scale Cohort Study: Prognostic Significance of TILs in Nondisseminated NPC. *Int. J. Cancer* 2018, 142, 2558–2566. [CrossRef] [PubMed]
- 21. Ooft, M.L.; Van Ipenburg, J.A.; Braunius, W.W.; Zuur, C.I.; Koljenović, S.; Willems, S.M. Prognostic Role of Tumor Infiltrating Lymphocytes in EBV Positive and EBV Negative Nasopharyngeal Carcinoma. *Oral Oncol.* **2017**, *71*, 16–25. [CrossRef]
- 22. Nilsson, J.S.; Sobti, A.; Swoboda, S.; Erjefält, J.S.; Forslund, O.; Lindstedt, M.; Greiff, L. Immune Phenotypes of Nasopharyngeal Cancer. *Cancers* 2020, *12*, 3428. [CrossRef]
- Liu, W.; Chen, G.; Zhang, C.; Liao, X.; Xie, J.; Liang, T.; Liao, W.; Song, L.; Zhang, X. Prognostic Significance of Tumor-Infiltrating Lymphocytes and Macrophages in Nasopharyngeal Carcinoma: A Systematic Review and Meta-Analysis. *Eur. Arch. Otorhinolaryngol.* 2022, 279, 25–35. [CrossRef] [PubMed]
- 24. Song, A.H.; Jaume, G.; Williamson, D.F.K.; Lu, M.Y.; Vaidya, A.; Miller, T.R.; Mahmood, F. Artificial Intelligence for Digital and Computational Pathology. *Nat. Rev. Bioeng.* 2023, *1*, 930–949. [CrossRef]
- Ng, W.T.; But, B.; Choi, H.C.; De Bree, R.; Lee, A.W.; Lee, V.H.; López, F.; Mäkitie, A.A.; Rodrigo, J.P.; Saba, N.F.; et al. Application of Artificial Intelligence for Nasopharyngeal Carcinoma Management—A Systematic Review. CMAR 2022, 14, 339–366. [CrossRef]
- 26. Bibault, J.-E.; Burgun, A.; Fournier, L.; Dekker, A.; Lambin, P. Artificial Intelligence in Oncology. In *Artificial Intelligence in Medicine*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 361–381, ISBN 978-0-12-821259-2.
- Abels, E.; Pantanowitz, L.; Aeffner, F.; Zarella, M.D.; Laak, J.; Bui, M.M.; Vemuri, V.N.; Parwani, A.V.; Gibbs, J.; Agosto-Arroyo, E.; et al. Computational Pathology Definitions, Best Practices, and Recommendations for Regulatory Guidance: A White Paper from the Digital Pathology Association. *J. Pathol.* 2019, 249, 286–294. [CrossRef]
- Li, C.; Jing, B.; Ke, L.; Li, B.; Xia, W.; He, C.; Qian, C.; Zhao, C.; Mai, H.; Chen, M.; et al. Development and Validation of an Endoscopic Images-Based Deep Learning Model for Detection with Nasopharyngeal Malignancies. *Cancer Commun.* 2018, 38, 59. [CrossRef] [PubMed]
- 29. 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] [PubMed]
- Zhang, F.; Zhong, L.-Z.; Zhao, X.; Dong, D.; Yao, J.-J.; Wang, S.-Y.; Liu, Y.; Zhu, D.; Wang, Y.; Wang, G.-J.; et al. A Deep-Learning-Based Prognostic Nomogram Integrating Microscopic Digital Pathology and Macroscopic Magnetic Resonance Images in Nasopharyngeal Carcinoma: A Multi-Cohort Study. *Ther. Adv. Med. Oncol.* 2020, *12*, 175883592097141. [CrossRef] [PubMed]
- Yang, Q.; Guo, Y.; Ou, X.; Wang, J.; Hu, C. Automatic T Staging Using Weakly Supervised Deep Learning for Nasopharyngeal Carcinoma on MR Images. J. Magn. Reson. Imaging 2020, 52, 1074–1082. [CrossRef]
- Peng, H.; Dong, D.; Fang, M.-J.; Li, L.; Tang, L.-L.; Chen, L.; Li, W.-F.; Mao, Y.-P.; Fan, W.; Liu, L.-Z.; et al. Prognostic Value of Deep Learning PET/CT-Based Radiomics: Potential Role for Future Individual Induction Chemotherapy in Advanced Nasopharyngeal Carcinoma. *Clin. Cancer Res.* 2019, 25, 4271–4279. [CrossRef] [PubMed]
- 33. Jing, B.; Zhang, T.; Wang, Z.; Jin, Y.; Liu, K.; Qiu, W.; Ke, L.; Sun, Y.; He, C.; Hou, D.; et al. A Deep Survival Analysis Method Based on Ranking. *Artif. Intell. Med.* **2019**, *98*, 1–9. [CrossRef] [PubMed]
- An, X.; Wang, F.-H.; Ding, P.-R.; Deng, L.; Jiang, W.-Q.; Zhang, L.; Shao, J.-Y.; Li, Y.-H. Plasma Epstein-Barr Virus DNA Level Strongly Predicts Survival in Metastatic/Recurrent Nasopharyngeal Carcinoma Treated with Palliative Chemotherapy. *Cancer* 2011, 117, 3750–3757. [CrossRef] [PubMed]
- 35. Tang, L.-L.; Guo, R.; Zhang, N.; Deng, B.; Chen, L.; Cheng, Z.-B.; Huang, J.; Hu, W.-H.; Huang, S.H.; Luo, W.-J.; et al. Effect of Radiotherapy Alone vs. Radiotherapy with Concurrent Chemoradiotherapy on Survival without Disease Relapse in Patients with Low-Risk Nasopharyngeal Carcinoma: A Randomized Clinical Trial. *JAMA* 2022, 328, 728–736. [CrossRef]

- Li, X.-Y.; Luo, D.-H.; Guo, L.; Mo, H.-Y.; Sun, R.; Guo, S.-S.; Liu, L.-T.; Yang, Z.-C.; Yang, J.-H.; Qiu, F.; et al. Deintensified Chemoradiotherapy for Pretreatment Epstein-Barr Virus DNA-Selected Low-Risk Locoregionally Advanced Nasopharyngeal Carcinoma: A Phase II Randomized Noninferiority Trial. JCO 2022, 40, 1163–1173. [CrossRef]
- 37. 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]
- 38. Gamper, J.; Alemi Koohbanani, N.; Benet, K.; Khuram, A.; Rajpoot, N. PanNuke: An Open Pan-Cancer Histology Dataset for Nuclei Instance Segmentation and Classification. In *Digital Pathology*; Reyes-Aldasoro, C.C., Janowczyk, A., Veta, M., Bankhead, P., Sirinukunwattana, K., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2019; Volume 11435, pp. 11–19, ISBN 978-3-030-23936-7.
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 770–778.
- Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 248–255.
- 41. Kim, H.E.; Cosa-Linan, A.; Santhanam, N.; Jannesari, M.; Maros, M.E.; Ganslandt, T. Transfer Learning for Medical Image Classification: A Literature Review. *BMC Med. Imaging* **2022**, *22*, 69. [CrossRef] [PubMed]
- 42. Wang, S.; Yang, D.M.; Rong, R.; Zhan, X.; Xiao, G. Pathology Image Analysis Using Segmentation Deep Learning Algorithms. *Am. J. Pathol.* **2019**, *189*, 1686–1698. [CrossRef] [PubMed]
- Verma, R.; Kumar, N.; Patil, A.; Kurian, N.C.; Rane, S.; Graham, S.; Vu, Q.D.; Zwager, M.; Raza, S.E.A.; Rajpoot, N.; et al. MoNuSAC2020: A Multi-Organ Nuclei Segmentation and Classification Challenge. *IEEE Trans. Med. Imaging* 2021, 40, 3413–3423. [CrossRef] [PubMed]
- Macenko, M.; Niethammer, M.; Marron, J.S.; Borland, D.; Woosley, J.T.; Guan, X.; Schmitt, C.; Thomas, N.E. A Method for Normalizing Histology Slides for Quantitative Analysis. In Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Boston, MA, USA, 28 June–1 July 2009; pp. 1107–1110.
- 45. Schubert, E.; Sander, J.; Ester, M.; Kriegel, H.P.; Xu, X. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Trans. Database Syst.* **2017**, *42*, 1–21. [CrossRef]
- 46. Harrell, F.E., Jr.; Lee, K.L.; Mark, D.B. Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Adequacy, and Measuring and Reducing Errors. *Stat. Med.* **1996**, *15*, 361–387. [CrossRef]
- 47. Fotso, S. Others PySurvival: Open Source Package for Survival Analysis Modeling. 2019. Available online: http://www.pysurvival.io/ (accessed on 3 April 2023).
- 48. Steck, H.; Krishnapuram, B.; Dehing-oberije, C.; Lambin, P.; Raykar, V.C. On Ranking in Survival Analysis: Bounds on the Concordance Index. *Adv. Neural Inf. Process. Syst.* **2007**, *20*, 1–8.
- 49. Budczies, J.; Klauschen, F.; Sinn, B.V.; Győrffy, 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]
- 50. Luo, W. Nasopharyngeal Carcinoma Ecology Theory: Cancer as Multidimensional Spatiotemporal "Unity of Ecology and Evolution" Pathological Ecosystem. *Theranostics* **2023**, *13*, 1607–1631. [CrossRef]
- 51. Wang, H.-Y.; Sun, B.-Y.; Zhu, Z.-H.; Chang, E.T.; To, K.-F.; Hwang, J.S.G.; Jiang, H.; Kam, M.K.-M.; Chen, G.; Cheah, S.-L.; et al. Eight-Signature Classifier for Prediction of Nasopharyngeal Carcinoma Survival. *JCO* **2011**, *29*, 4516–4525. [CrossRef]
- Zhou, J.; Deng, Y.; Huang, Y.; Wang, Z.; Zhan, Z.; Cao, X.; Cai, Z.; Deng, Y.; Zhang, L.; Huang, H.; et al. An Individualized Prognostic Model in Patients with Locoregionally Advanced Nasopharyngeal Carcinoma Based on Serum Metabolomic Profiling. *Life* 2023, 13, 1167. [CrossRef] [PubMed]
- 53. Sobti, A.; Sakellariou, C.; Nilsson, J.S.; Askmyr, D.; Greiff, L.; Lindstedt, M. Exploring Spatial Heterogeneity of Immune Cells in Nasopharyngeal Cancer. *Cancers* 2023, *15*, 2165. [CrossRef] [PubMed]
- Niazi, M.K.K.; Parwani, A.V.; Gurcan, M.N. Digital Pathology and Artificial Intelligence. *Lancet Oncol.* 2019, 20, e253–e261. [CrossRef] [PubMed]
- 55. Chua, M.L.K.; Wee, J.T.S.; Hui, E.P.; Chan, A.T.C. Nasopharyngeal Carcinoma. Lancet 2016, 387, 1012–1024. [CrossRef]
- Young, L.S.; Yap, L.F.; Murray, P.G. Epstein–Barr Virus: More than 50 Years Old and Still Providing Surprises. *Nat. Rev. Cancer* 2016, 16, 789–802. [CrossRef] [PubMed]
- 57. Lo, A.K.-F.; Dawson, C.W.; Lung, H.L.; Wong, K.-L.; Young, L.S. The Role of EBV-Encoded LMP1 in the NPC Tumor Microenvironment: From Function to Therapy. *Front. Oncol.* 2021, *11*, 640207. [CrossRef]
- 58. Hau, P.M.; Lung, H.L.; Wu, M.; Tsang, C.M.; Wong, K.-L.; Mak, N.K.; Lo, K.W. Targeting Epstein-Barr Virus in Nasopharyngeal Carcinoma. *Front. Oncol.* 2020, 10, 600. [CrossRef]
- 59. Zink, D.; Fischer, A.H.; Nickerson, J.A. Nuclear Structure in Cancer Cells. Nat. Rev. Cancer 2004, 4, 677–687. [CrossRef]

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.