
Overall Survival Prediction of Diffuse large B-cell Lymphoma Using Deep Learning and Hematoxylin and Eosin Histological Images

[Joaquim Carreras](#)*

Posted Date: 2 January 2026

doi: 10.20944/preprints202601.0065.v1

Keywords: DLBCL; prognosis; overall survival; deep learning; CNN; XAI; computer vision; AI



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Communication

Overall Survival Prediction of Diffuse large B-cell Lymphoma Using Deep Learning and Hematoxylin and Eosin Histological Images

Joaquim Carreras

Department of Pathology, School of Medicine, Tokai University. 143 Shimokasuya, Isehara, Kanagawa, 259-1193, Japan; joaquim.carreras@tokai.ac.jp

Abstract

Diffuse large B-cell lymphoma (DLBCL) is one of the most frequent subtypes of non-Hodgkin lymphoma (NHL). In approximately 40% of the patients, the prognosis and clinical evolution is unfavorable. This study is a proof-of-concept computer vision exercise to support the feasibility of predicting the prognosis of DLBCL using only hematoxylin and eosin (H&E) histological images and deep learning. A conventional series of DLBCL of 114 cases was split into two prognostic groups according to the overall survival (curve fitting and slope analysis): patients who died before the first 2 years ("Dead 2-years", $b1 = -0.054$), and the others ($b1 = -0.003$). Twenty different convolutional neural networks (CNN) were used, and explainable artificial intelligence (XAI) was used to identify the areas of the images that the network used for classification. The final model based on DarkNet-19 predicted prognosis groups with high performance (test set accuracy = 96.26%). The other performance parameters were precision (94.46%), recall (95.02%), false positive rate (3.07%), specificity (96.93%), and F1 score (94.74%). XAI, including grad-CAM, occlusion sensitivity, and image-LIME confirmed that the CNN was focusing on the correct areas. Correlation with the clinicopathological characteristics found that the Dead < 2-years group was correlated with stage III-IV, International Prognostic Index (IPI) High + High/intermediate, progressive disease, non-GCB cell-of-origin, CD10-, BCL2+, and EBER+. Analysis of the immune microenvironment, cell cycle, and germinal center markers showed that Dead < 2-years had higher IL10, PD-L1, and CD163, and lower E2F1 protein expressions. In conclusion, the overall survival of DLBCL can be predicted using H&E histological images and deep learning. The trained CNN could be used as pre-trained CNN model for transfer learning in the future.

Keywords: DLBCL; prognosis; overall survival; deep learning; CNN; XAI; computer vision; AI

1. Introduction

Diffuse large B-cell lymphoma (DLBCL) is one of the most frequent histological subtypes of non-Hodgkin lymphoma (NHL), accounting for approximately 25% of adult NHL cases [1]. Most patients present with a rapid enlarging symptomatic mass with nodal enlargement. In approximately 60% of the cases, DLBCL present with advanced-stage (III or IV) and high serum lactate dehydrogenase (LDH), and in 30% with fever, weight loss, night sweats, and bone marrow infiltration [2–7]. DLBCL is cured in 60% of the cases with current therapy, particularly in patients who achieve complete response with first-line treatment. However, in 40% of the cases, the clinical evolution is unfavorable [5,8,9].

DLBCL is a heterogeneous clinicopathological entity. It is derived from germinal center B-cells (centroblasts) or post-germinal activated B-cells (immunoblasts) [10,11]. The tumor cells of DLBCL are large (e.g. nuclei twice the size of a small lymphocyte and larger than the nucleus of a macrophage) [10–17]. Centroblasts are large, noncleaved cells with round or oval nuclei, vesicular chromatin, often with multiple peripheral nucleoli, and a narrow rim of basophilic cytoplasm [10–

17]. Immunoblasts are usually larger cells with prominent nucleoli and more abundant cytoplasm, often with plasmacytoid characteristics. Some cases are mixed, and other morphological variants exist [10,11,13,14]. The pathogenesis of DLBCL is complex and shows a heterogeneous landscape [1].

DLBCL, not otherwise specified (NOS) is defined as having a mature B-cell phenotype and large cell morphology, but having none of the criteria that define specific large B-cell lymphoma subtypes. Using either immunohistochemistry or gene expression studies of DLBCL, the cell-of-origin (COO) status can be determined, including germinal center B-cell (GCB) versus activated B-cell (ABC) subtypes [2,5]. Other large B-cell lymphoma subtypes are primary mediastinal, T-cell/histiocyte-rich, plasmablastic, primary cutaneous leg-type, immune privileged sites, intravascular, associated with chronic inflammation, IRF4 rearranged, ALK-positive, with 11q aberration, primary effusion lymphoma, and Epstein-Barr virus (EBV)-positive DLBCL [2,5].

Nodal DLBCL can spread to other organs, such as the liver, kidneys, lung, bone marrow, and central nervous system. Extranodal/extramedullary involvement is frequently present in early-stage disease (stage I/II). The most frequent primary extranodal presentation is the gastrointestinal tract, but virtually any tissue could be affected, including the testis, bone, thyroid, salivary glands, tonsil, and skin [1,18–20].

Several prognostic models are used in DLBCL. The International Prognostic Index (IPI) and its variants are the main prognostic variables that are routinely used. The original IPI included the following factors and age >60 years, high serum LDH, Eastern Cooperative Oncology Group (ECOG) performance status ≥ 2 , clinical stage III-IV, and number of extranodal sites >1 associated with poor prognosis [21]. The COO status determined by immunohistochemistry using CD10, BCL6, and MUM1, or by gene expression profiling, correlated with non-GCB/ABC-like with poorer prognosis of the patients [21–27]. *MYC* rearrangement is seen in 5-15% of the cases and can be associated with *BCL2* and *BCL6* rearrangements [28]. Double-hit *MYC* and *BCL2* cases have a worse prognosis [2,5,29–35]. Other techniques such as deep sequencing have confirmed the DLBCL heterogeneity and identified driver mutations with different clinical outcomes [2,5,6,17,36–38].

The histological characteristics of the tissue reflect the proteomic, transcriptomic and genomic pathological background. This study aimed to predict the prognosis of patients with DLBCL based only on the histological evaluation of hematoxylin and eosin staining. The 2-years timepoint shows a change in the slope of the survival curve. Therefore, the DLBCL cases were classified according to the overall survival, using as an event the presence of death within the first 2 years (“Dead 2-years”), which is similar to progression of disease 24 (POD24). Using several models, the prognostic was predicted with high performance, and several explainable artificial intelligence (XAI) techniques were used to visually highlight the regions that are most important for the deep neural network’s classification decision.

2. Materials and Methods

2.1. Patients and samples

A series of 114 patients with a histological diagnosis of conventional diffuse large B-cell lymphoma (DLBCL) from 2007 to 2011 were selected from the Department of Pathology, Tokai University, School of Medicine. The cases were diagnosed according to the current classifications of hematolymphoid neoplasia [2,5,10], which included the evaluation of hematoxylin and eosin, immunophenotype, and molecular techniques when required [2,5,6]. Based on the overall survival plot, two groups were defined: Dead < 2-years and the Others. The rationale is that at the 2-years time point the slope of the overall survival curve changes, passing from aggressive ($b1 = -0.054$) to less aggressive clinical behavior ($b1 = -0.003$).

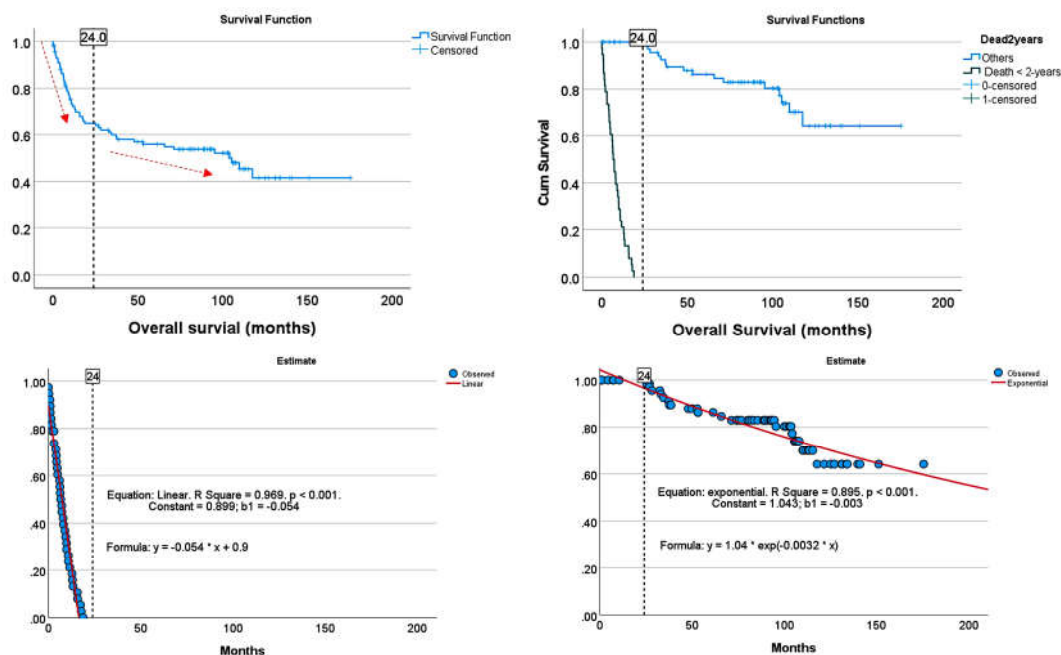


Figure 1. Overall survival groups. The cases were grouped according to their survival based on the slope of the survival curve and the 2-years timeline: cases who died within the first 24 months (Dead < 2-years), and the Other cases. A more detailed analysis was performed using curve estimation.

This study was conducted following the guidelines of the Declaration of Helsinki of the World Medical Association (WMA) and ethical principles for medical research involving human participants. The Institutional Review Board of Tokai University approved the study (IRB20-156).

Immune microenvironment data were retrieved from our previous publications [39–41], and the immunohistochemistry reanalyzed. For CD163, IL10 and PD-L1, new immunohistochemistry (IHC) was performed as the number of cases in this study increased from the previous publications. The IHC methodology, including primary antibody details, is shown in Appendix B.

2.2. Deep learning image classification

Whole-tissue sections were stained with H&E and digitized using a slide scanner (NanoZoomer S360, C13220-01, Hamamatsu Photonics K.K.). The whole neoplastic areas were identified, digitally extracted at 20× magnification and 150 dpi, and split into image patches of 224×224×3 resolution. After splitting, image patches of size different from 224×224 and with less than 80% of tissue were discarded.

First, the series was slit into a training set (70%) and a validation set (10%) to help prevent overfitting during training and a test set (20%). The training set patches were shuffled before training. No augmentation options were used during the training.

The training options were as follows: sgdm solver, 0.001 initial learning rate, constant learning rate schedule, 128 minibatch size, 5 max epochs, and 50 validation frequency. In the NasNet-Large CNN, due to hardware limitations to run the analysis, the minibatch size was set at 16.

Image patches were classified using transfer learning and 20 different types of CNN, including AlexNet, DarkNet-19, DarkNet-53, DenseNet-201, EfficientNet-b0, GoogleLeNet, GoogleLeNet_places365, Inception-ResNet-v2, Inception-v3, MobileNet-v2, NasNet-Mobile, NasNet-Large, ResNet-18, ResNet-50, ResNet-101, Shufflenet, SqueezeNet, VGG-16, VGG-19, and Xception.

The following hardware and software were used: NanoZoomer S360, #C13220-01 (Hamamatsu Photonics K.K.); 12-Core AMD Ryzen 9 5900X CPU, 4900 MHz; 49075 MB DDR4-3200 SDRAM; NVIDIA GeForce RTX 4080 SUPER GPU; MATLAB R2023b Update 10 (23.2.0.2859533), 64-bit (win64); PhotoScape v3.7; IBM SPSS Statistics version 27 (Release 27.0.1.0, 64-bit edition); NDP.view 2.9.29 (RUO) 2022/01/14 (Hamamatsu Photonics K.K.).

Figures 2–7 depict the methodological design.

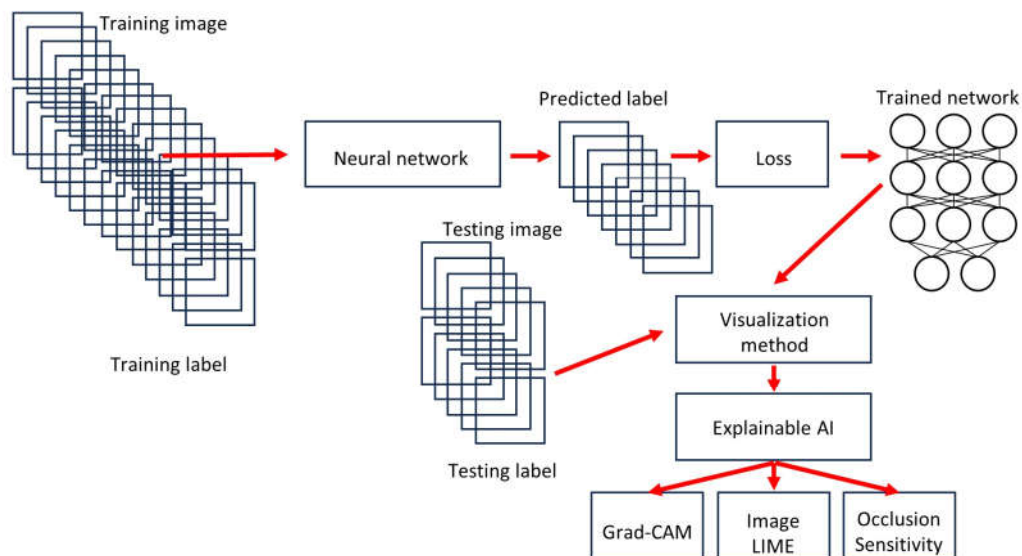


Figure 2. Methodology. The series was randomly slit into a training (70%) and validation (10%) sets for training, and test (20%) set using new data. Explainable artificial intelligence (XAI) was performed using Grad-CAM, image LIME, and occlusion sensitivity.



Figure 3. Architectures of the main different types of convolutional neural networks used in this study.

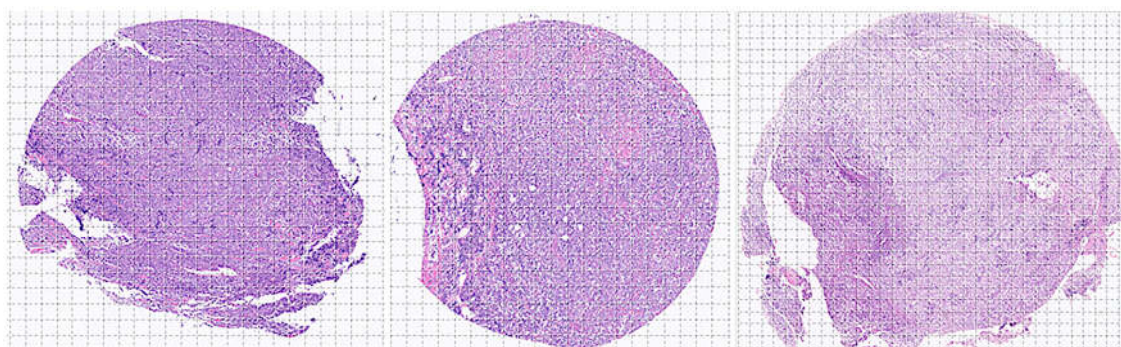


Figure 4. Image splitting. Histological images stained with hematoxylin and eosin (H&E) were first digitalized at 200×magnification and 150 dpi. Later, the images were split at 224×224×3 resolution, which is suitable for convolutional neural network processing. This image shows the splitting of 3 cases of DLBCL obtained from the H&E of a tissue microarray.

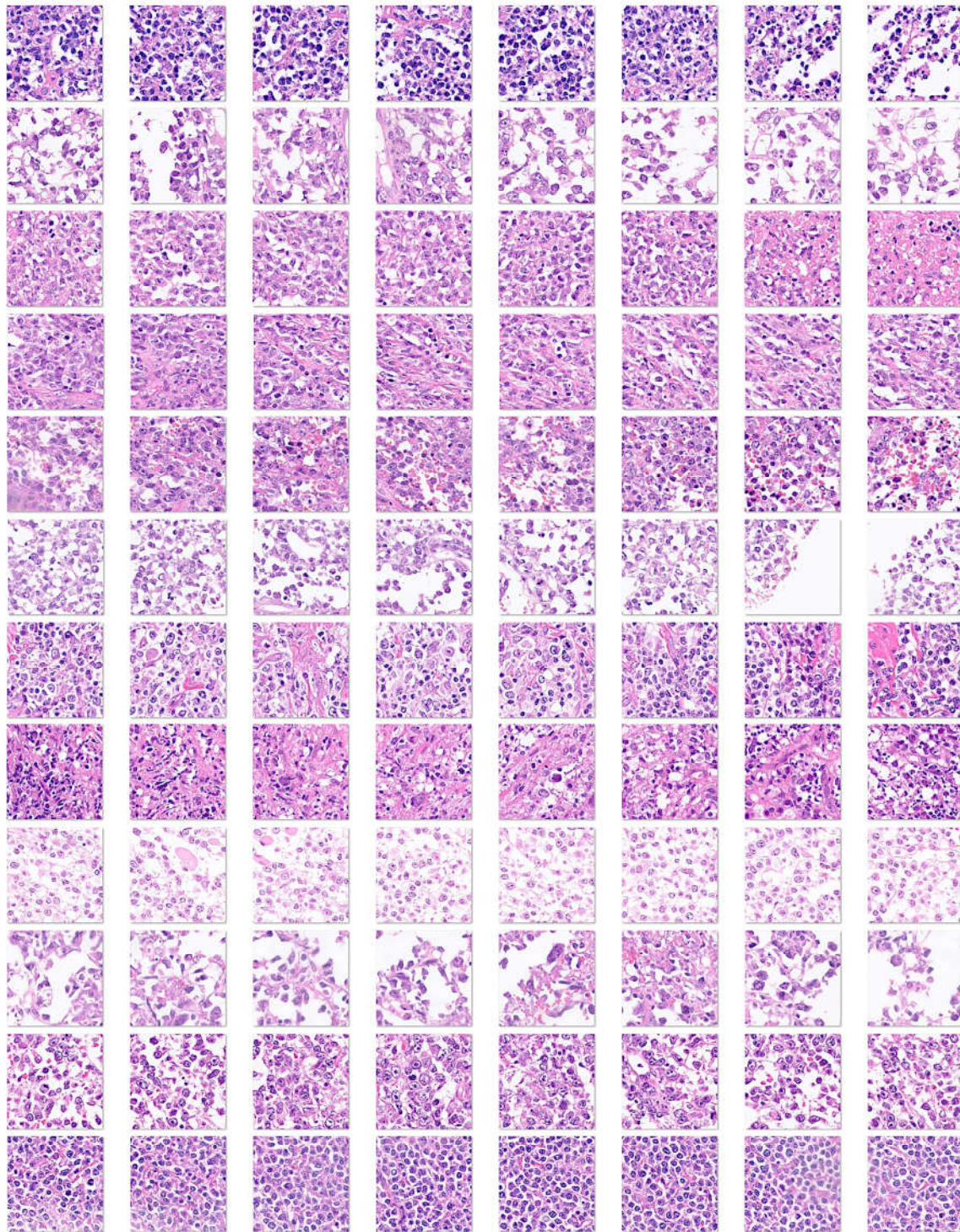


Figure 5. Image splitting of the Dead < 2-years group. These DLBCL cases were characterized by death events before the first 2 years of overall survival. Histological images stained with hematoxylin and eosin (H&E) were first digitalized at 200×magnification and 150 dpi. Later, the images were split at 224×224×3 resolution, which is suitable for convolutional neural network processing. This image shows the splitting of 12 cases of DLBCL, which is a heterogeneous clinicopathologic entity. DLBCL is derived from germinal center B-cells (centroblasts) or post-germinal activated B-cells (immunoblasts).

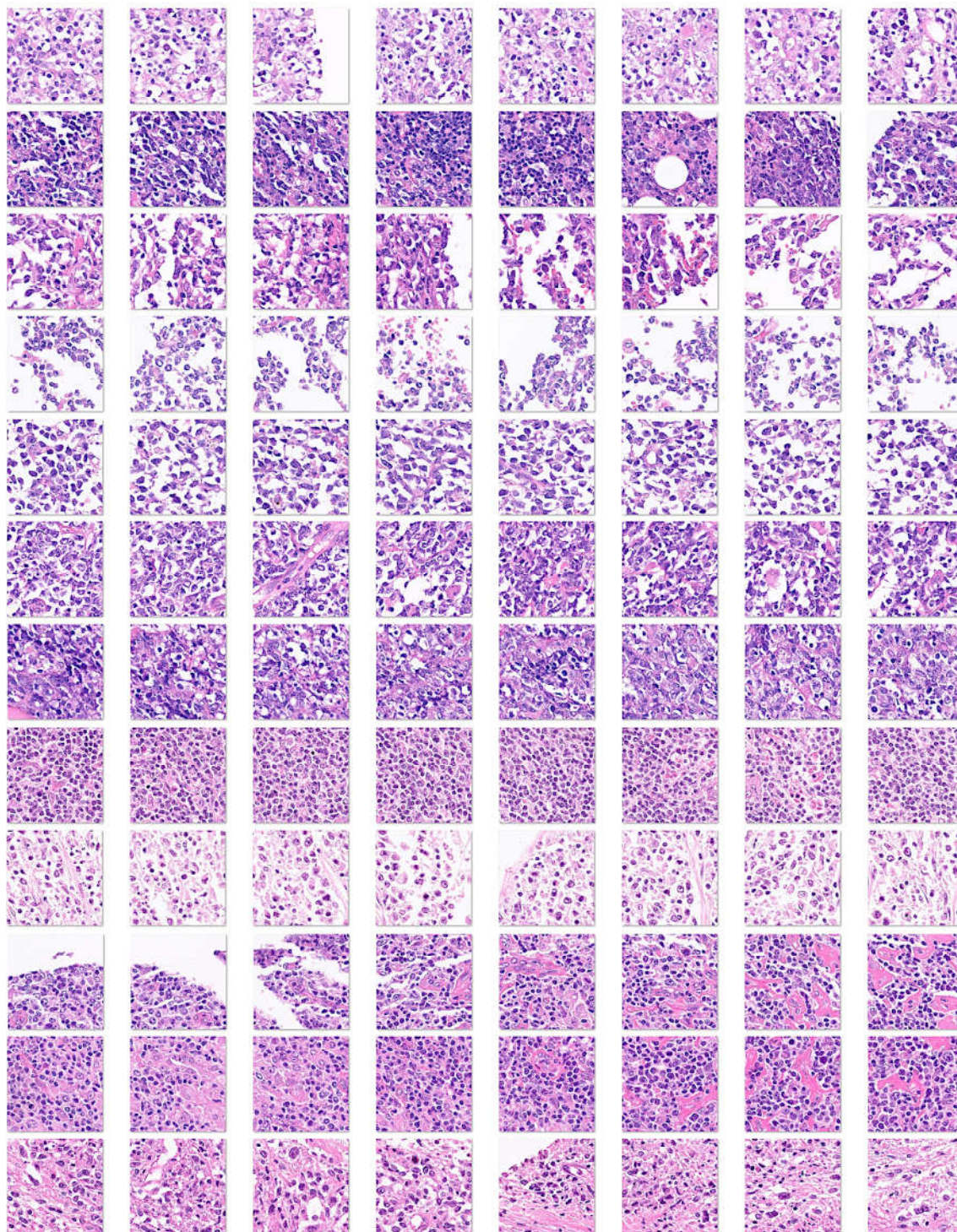


Figure 6. Image splitting of the Others group). This image shows examples of DLBCL cases that did not have a death event before the 2 years of overall survival follow-up (i.e. less aggressive cases, “others”). Histological images stained with hematoxylin and eosin (H&E) were first digitalized at 200×magnification and 150 dpi. Later, the images were split at 224×224×3 resolution, which is suitable for convolutional neural network processing. This image shows the splitting of 12 cases of DLBCL. DLBCL is a heterogeneous clinicopathologic entity. DLBCL is derived from germinal center B-cells (centroblasts) or postgerminal activated B-cells (immunoblasts).

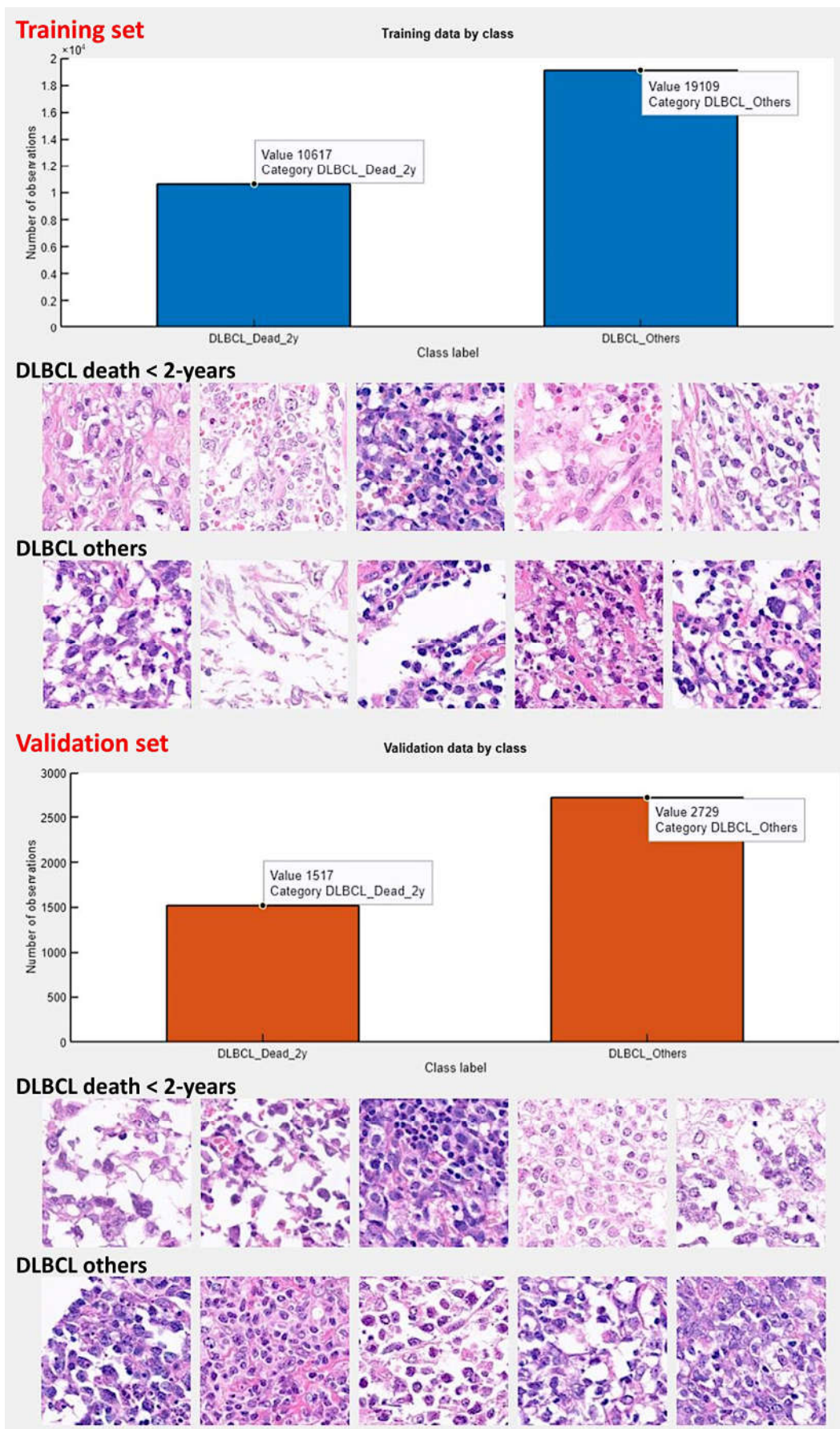


Figure 7. Training and validation datasets. The series was slit into a training set (70%), validation set (10%) to help prevent over-fitting, and test set (20%). No augmentation options were used during the training. This figure shows the number of image patches and random example images of the training and validation sets.

3. Results

3.1. Image patch-based CNN classification of the 2-years cut-off OS groups

In the training/validation set, DarkNet-19 had the best validation accuracy (95.71%), and the accuracy/training time ratio was moderate (0.114). The most efficient architecture was ResNet-18, which had a validation accuracy of 92.11% and a ratio of 0.426.

Table 1. Training/validation set CNN performance.

CNN model	Learn ables	Laye rs	Connec tions	Image Input	Training time	Validatio n accuracy (%)	Efficien cy	Relativ e time
NasNet-Large	84.9M	1243	1462	331×331×3	1429 min	96.42	0.001	564.2
DarkNet-19	20.8M	64	63	256×256×3	14 min 1	95.71	0.114	5.5
DarkNet-53	41.6M	184	206	256×256×3	120 min 35	95.6	0.013	47.6
DenseNet-201	20M	708	805	224×224×3	255 min 9	93.9	0.006	100.7
ResNet-101	44.6M	347	379	224×224×3	114 min 21	93.31	0.014	45.1
Inception-v3	23.8M	315	349	299×299×3	53 min 15	92.25	0.029	21.0
ResNet-50	25.5M	177	192	224×224×3	14 min 3	92.25	0.109	5.5
ResNet-18	11.6M	71	78	224×224×3	3 min 36	92.11	0.426	1.4
VGG-16	138.3	41	40	224×224×3	155 min 10	92.11	0.009	61.3
MobileNet-v2	3.5M	154	163	224×224×3	12 min 55	90.86	0.117	5.1
Inception-ResNet- VGG-19	55.8M	824	921	299×299×3	509 min 8	89.83	0.003	201.0
EfficientNet-b0	5.3M	290	363	224×224×3	54 min 0	89.07	0.075	21.3
GoogleLeNet-	5.9M	144	170	224×224×3	5 min 30	88.37	0.268	2.2
GoogleLeNet	6.9M	144	170	224×224×3	5 min 21	88.34	0.275	2.1
Shufflenet	1.4M	172	187	224×224×3	5 min 42	88.22	0.258	2.3
NasNet-Mobile	5.3M	913	1072	224×224×3	30 min 19	87.73	0.048	12.0
Xception	22.9M	170	181	299×299×3	522 min 52	87.28	0.003	206.4
AlexNet	60.9M	25	24	227×227×3	2 min 32	83.94	0.552	1.0
SqueezeNet	1.2M	68	75	227×227×3	2 min 44	79.35	0.484	1.1

Efficiency: Validation accuracy/time (% / sec). Relative time: faster CNN (AlexNet) as reference.

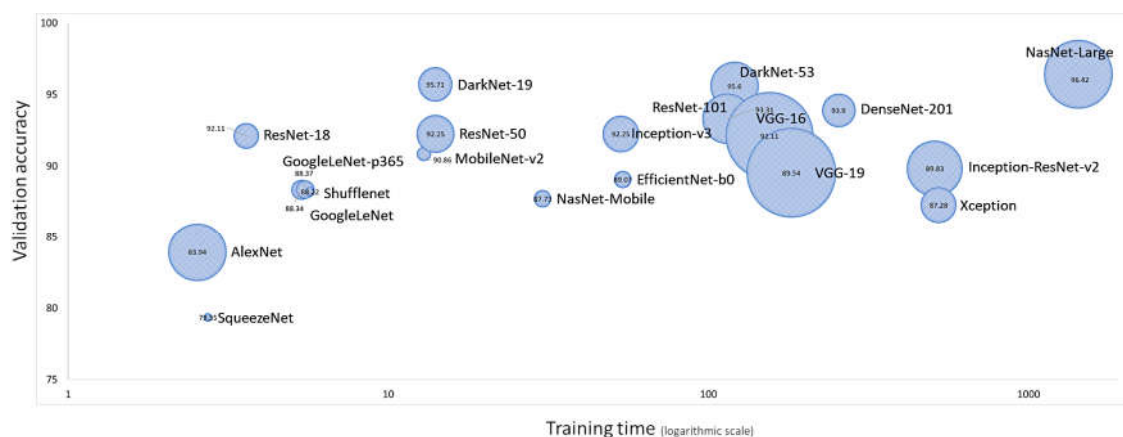


Figure 8. Neural network training performances. The most important characteristics are the neural network accuracy (x-axis), speed (y-axis), and size (circle). Choosing a neural network is a tradeoff between these characteristics. NasNet-Large had the best validation accuracy (96.52%), followed by DarkNet-19 (95.71%). However, in relation with AlexNet that was the fastest, NasNet-Large took 564.2 more time to compute (i.e. relative time).

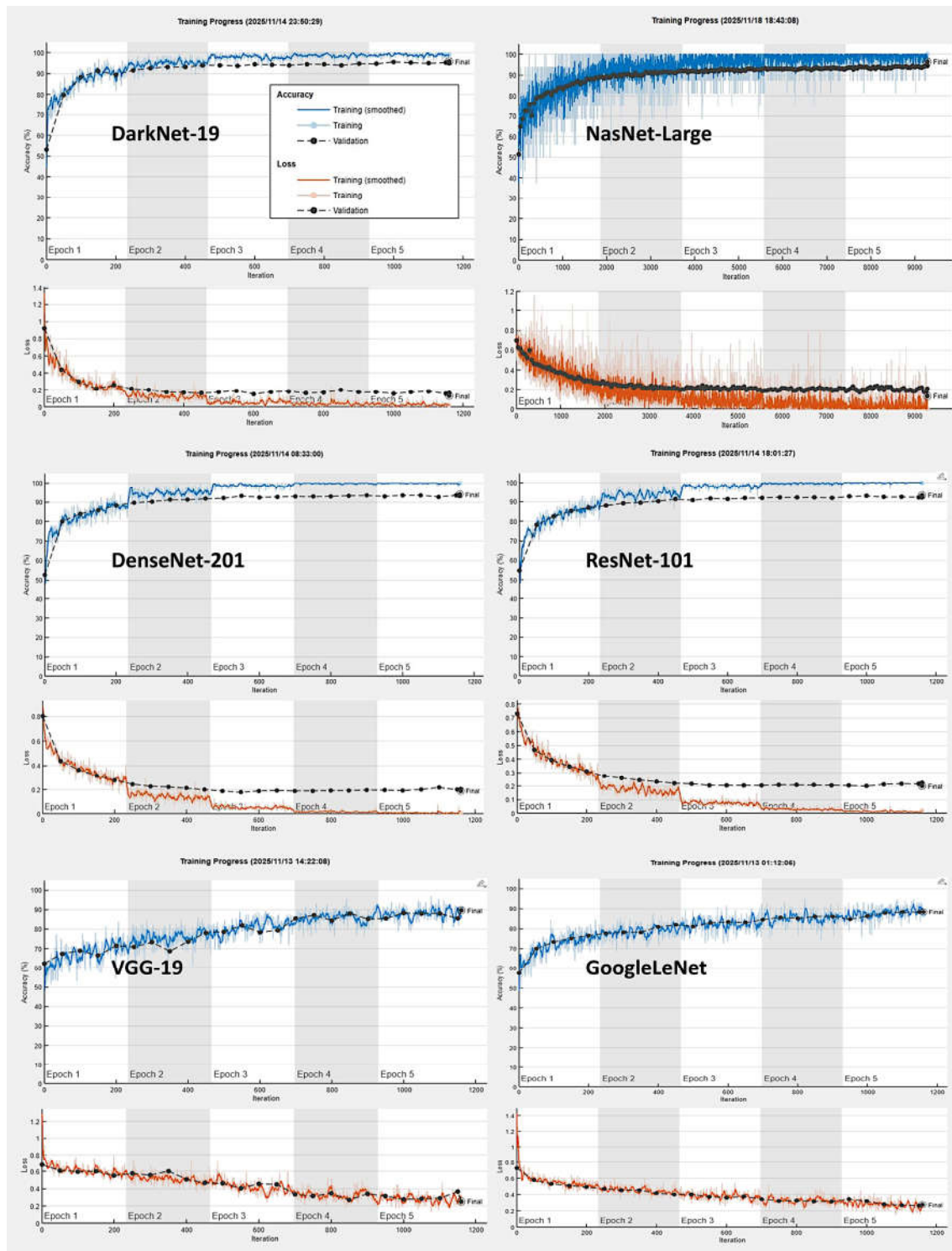


Figure 9. CNN training plots.

In the validation set, DarkNet-19 had the best accuracy (96.26%), followed by NasNet-Large (96.21%), DarkNet-53 (95.47%), and DenseNet-201 (93.67%). The confusion charts of the models with

higher performance are shown in Figures 10 and 11. All the performance parameters of the models in the test set are shown in Table 2 and Figure 12.

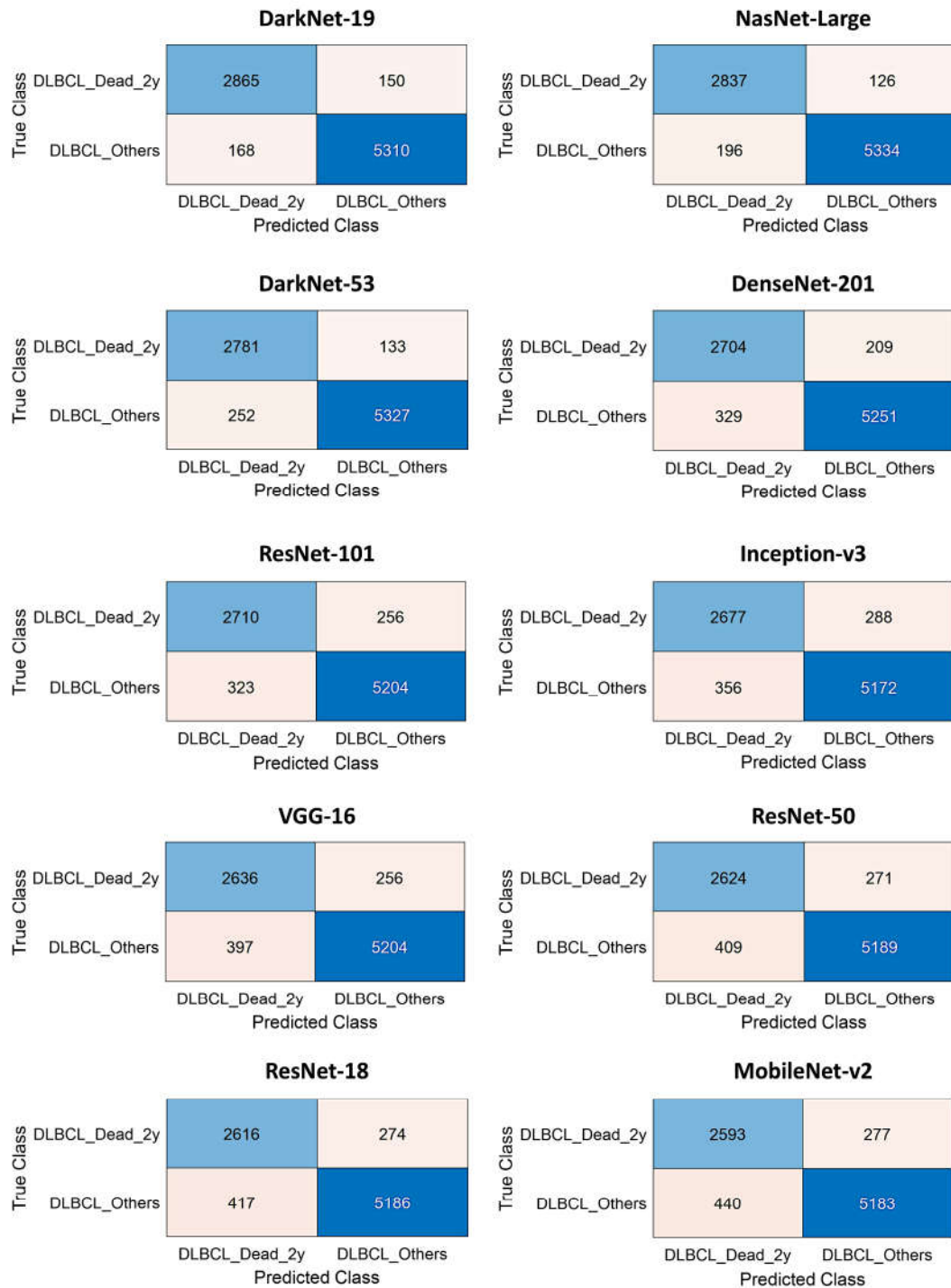


Figure 10. Test set confusion charts.

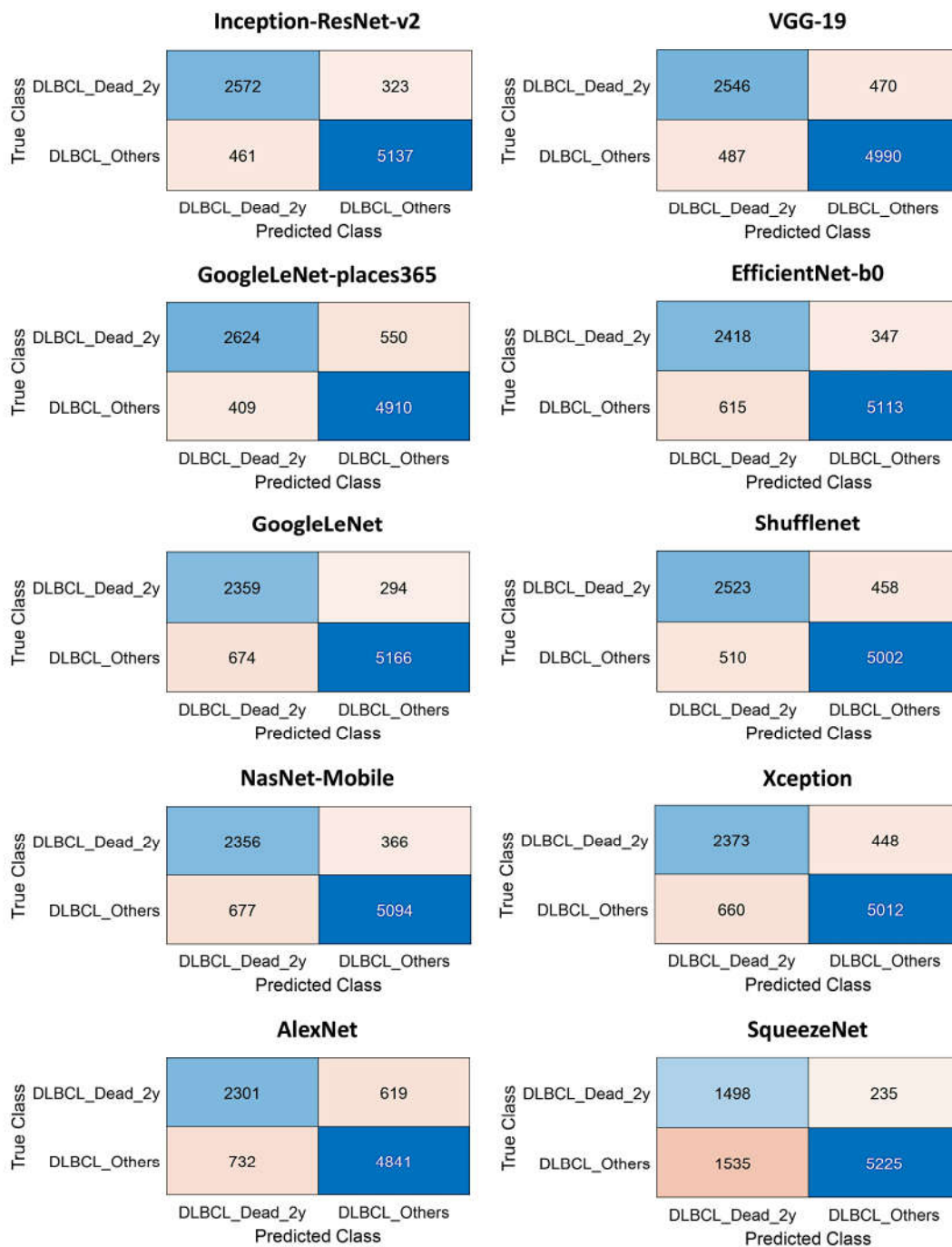


Figure 11. The test set confusion charts, continuation.

Table 2. The test set CNN performance parameters.

CNN	Accuracy	Precision	Recall	False Positive Rate	Specificity	F1 score
DarkNet-19	96.26	94.46	95.02	3.07	96.93	94.74
NasNet-Large	96.21	93.54	95.75	3.54	96.46	94.63
DarkNet-53	95.47	91.69	95.44	4.52	95.48	93.53
DenseNet-201	93.67	89.15	92.83	5.9	94.1	90.95
ResNet-101	93.18	89.35	91.37	5.84	94.16	90.35
Inception-v3	92.42	88.26	90.29	6.44	93.56	89.26
VGG-16	92.31	86.91	91.15	7.09	92.91	88.98
ResNet-50	91.99	86.52	90.64	7.31	92.69	88.53
ResNet-18	91.86	86.25	90.52	7.44	92.56	88.33

MobileNet-v2	91.56	85.49	90.35	7.83	92.17	87.85
Inception-ResNet-	90.77	84.8	88.84	8.24	91.76	86.77
VGG-19	88.73	83.94	84.42	8.89	91.11	84.18
GoogleLeNet-	88.71	86.52	82.67	7.69	92.31	84.55
EfficientNet-b0	88.67	79.72	87.45	10.74	89.26	83.41
GoogleLeNet	88.6	77.78	88.92	11.54	88.46	82.98
Shufflenet	88.6	83.18	84.64	9.25	90.75	83.9
NasNet-Mobile	87.72	77.68	86.55	11.73	88.27	81.88
Xception	86.95	78.23	84.11	11.64	88.36	81.07
AlexNet	84.09	75.87	78.8	13.13	86.87	77.31
SqueezeNet	79.16	49.39	86.44	22.71	77.29	62.86

Recall equals to Sensitivity and True Positive Rate. False Positive Rate equals to 1-Specificity.

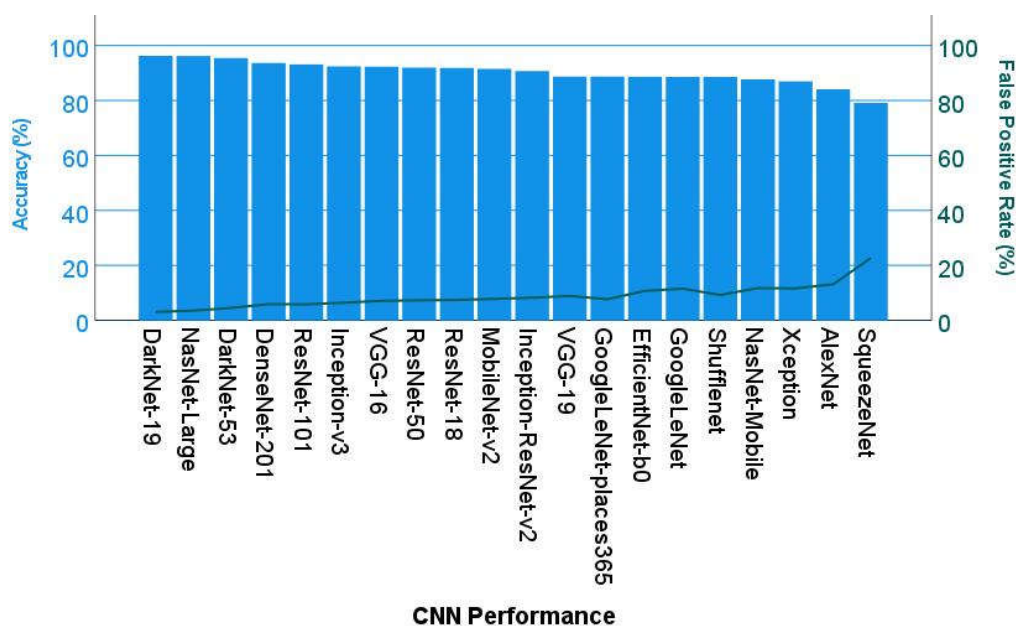


Figure 12. Test set CNN performance. This figure shows the performance parameters of accuracy and false positive rate of the different CNN.

Explainable artificial intelligence (XAI) was used to identify the areas of the images that the network DarkNet-19 used for classification. Figures 13 and 14 show the Grad-CAM, occlusion sensitivity, and image Lime analysis.

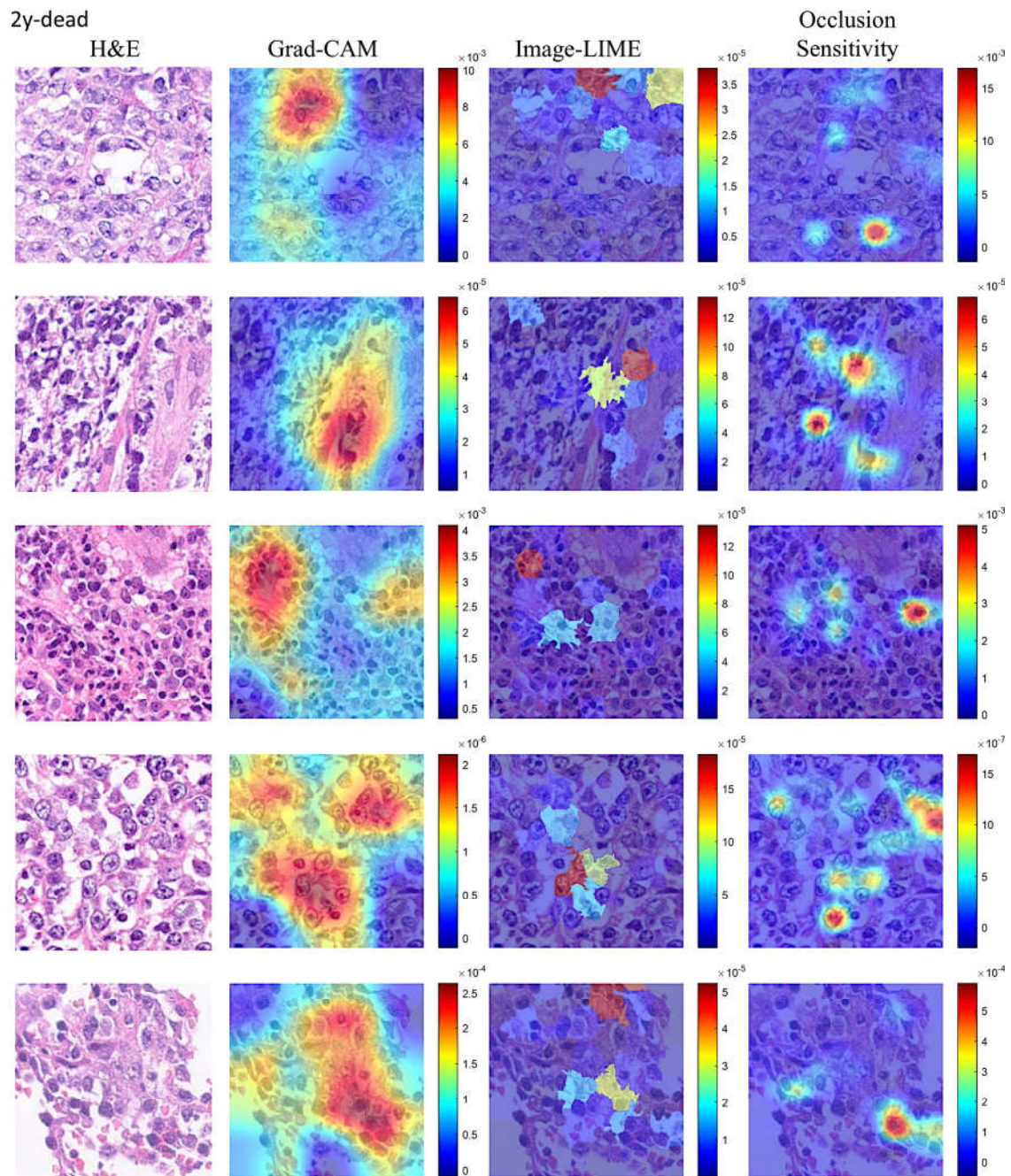


Figure 13. Explainable artificial intelligence (XAI) of the 2-years dead group. This figure shows 4 different image patches analyzed with XAI techniques, including Grad-CAM, ImageLIME, and OcclusionSensitivity.

XAI techniques were used to identify the areas of an image that the network DarkNet-19 used for classification. XAI showed that the convolutional neural networks focused on the epithelial component of the neoplasia.

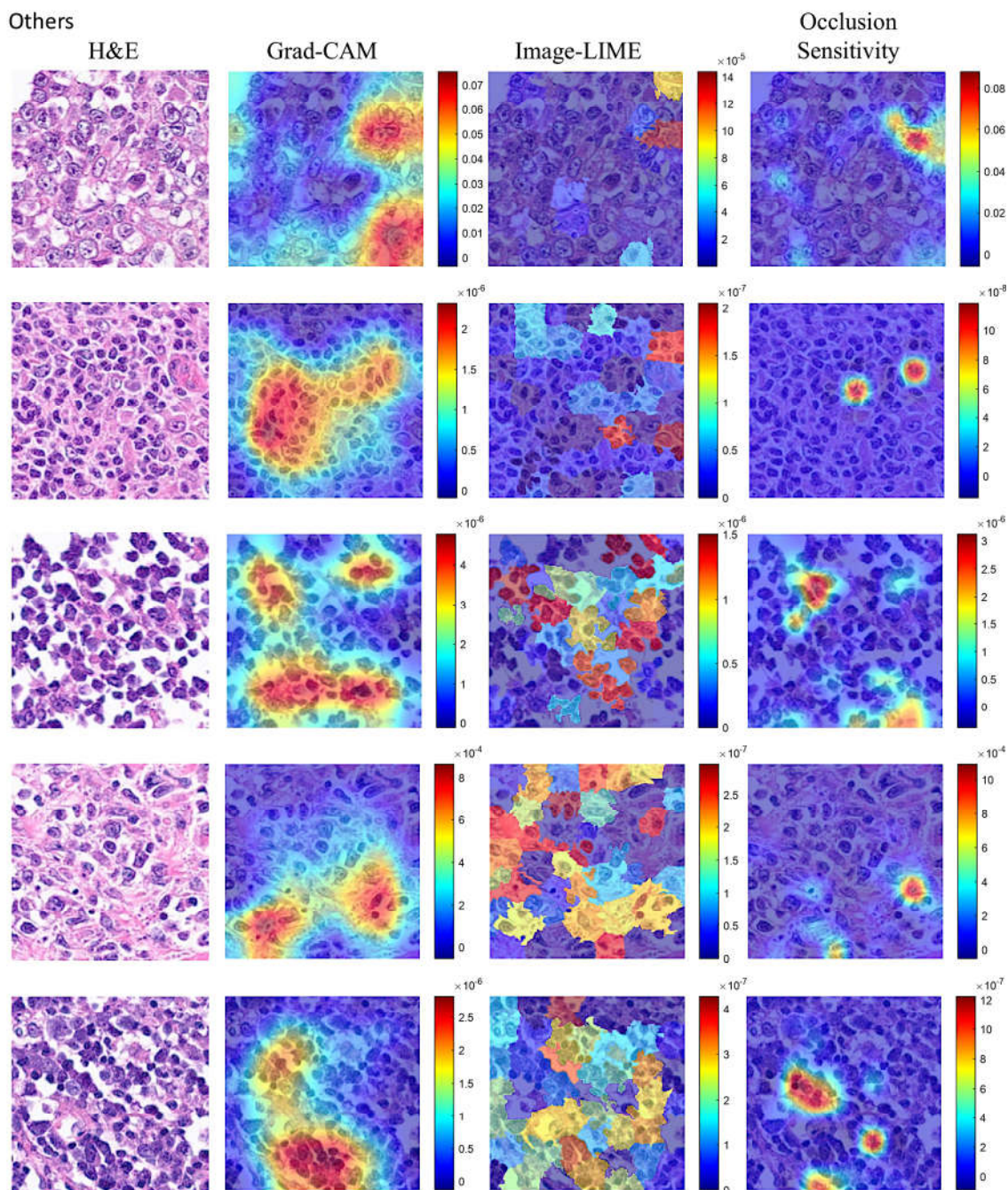


Figure 14. Explainable artificial intelligence (XAI) of the Others group. This figure shows 4 different image patches analyzed with XAI techniques, including Grad-CAM, ImageLIME, and OcclusionSensitivity. XAI techniques were used to identify the areas of an image that the network Darknet-19 used for classification. XAI showed that the convolutional neural networks focused on the epithelial component of the neoplasia.

3.3. Clinicopathological Characteristics

The Dead < 2-years group was correlated with several clinicopathological characteristics of the patients. The Dead < 2-years patients correlated with stage III-IV, International Prognostic Index (IPI) High + High/intermediate, absence of clinical response to treatment, non-GCB cell-of-origin, CD10-, BCL2+, and EBER+. Analysis of the immune microenvironment, cell cycle, and germinal center markers showed that Dead < 2-years had higher CD163, IL10, PD-L1, and lower E2F1 protein expressions (all p values < 0.05) (Table 3, Figure 15).

Table 3. Correlation with the clinicopathological characteristics.

Variable	All cases	Dead 2-years	Others	P value
Frequency	114	38/114 (33.3%)	76/114 (66.7%)	-
Clinical characteristics				
Age > 60 years	81/114 (71.1%)	30/38 (78.9%)	51/76 (67.1)	0.273
Male	60/114 (52.6%)	19/38 (50%)	41/76 (53.9%)	0.697
Location				
Nodal (+Spleen)	58/114 (50.9%)	16/38 (42.1%)	42/76 (55.3%)	0.430
Waldeyer's ring	11/114 (9.6%)	3/38 (7.9%)	8/76 (10.5%)	
Gastrointestinal	13/114 (11.4%)	5/38 (13.2%)	8/76 (10.5%)	
Other extranodal	32/114 (28.1%)	14/38 (36.8%)	18/76 (23.7%)	
Stage III-IV	46/97 (47.4%)	18/28 (64.3%)	28/69 (40.6%)	0.044
IPI High+High/Intermediate	31/91 (34.1%)	14/27 (51.9%)	17/64 (26.6%)	0.029
RCHOP/RCHOP-like treatment	93/98 (94.9%)	26/28 (92.9%)	67/70 (95.7%)	0.513
Clinical response	68/92 (73.9%)	5/24 (20.8%)	63/68 (92.5%)	< 0.001
Hight sIL2R	79/99 (79.8%)	27/29 (93.1%)	52/70 (74.3%)	0.052
Pathological characteristics				
CD3+	0/114 (0%)	0/38 (0%)	0/76 (0%)	1.0
CD20+	114/114 (100%)	38/38 (100%)	76/76 (100%)	1.0
CD5+	13/113 (11.5%)	4/38 (10.5%)	9/75 (12.0%)	1.0
CD10+	33/113 (29.2%)	2/38 (5.3%)	31/75 (41.3%)	< 0.001
BCL6+	76/113 (67.3%)	26/38 (68.4%)	50/75 (66.7%)	1.0
MUM1+	93/113 (82.3%)	33/38 (86.8%)	60/75 (80%)	0.442
Non-GCB	77/114 (67.5%)	35/38 (92.1%)	42/76 (55.3%)	< 0.001
BCL2+	89/113 (78.8%)	36/38 (94.7%)	53/75 (70.7%)	0.003
MYC rearrangement	9/98 (9.2%)	2/29 (6.9%)	7/69 (10.1%)	1.0
EBER+	28/114 (25%)	15/37 (40.5%)	13/75 (17.3%)	0.011
Ki67	16.1% +/- 14.2	15.3% +/- 12.2	16.5% +/- 14.9	0.959
Immune microenvironment				
IL10	12.2% +/- 15.8 (n =	18.6% +/- 19.6	9.2% +/- 12.8	0.006
PD-L1 (CD274)	12.2% +/- 15.8% (n =	18.5% +/- 19.6	9.1% +/- 12.8	0.026
CSF1R	33.5% +/- 27.5 (n = 94)	28.7% +/- 25.4	35.8% +/- 28.3	0.247
CD163	39.2% +/- 25.9 (n =	48.2% +/- 24.5	34.6% +/- 25.6	0.008
CASP8	6.7% +/- 8.4 (n = 94)	6.0% +/- 9.4	7.1% +/- 8.0	0.268
TNFAIP8	41.3% +/- 25.6 (n=93)	46.2 +/- 24.0	39.3% +/- 26.1	0.223
Cell cycle / GC-related				
LMO2	2.6% +/- 3.5 (n = 92)	2.4% +/- 3.9	2.7% +/- 3.4	0.051
MYC	5.4% +/- 5.7 (n = 93)	6.5% +/- 6.4	4.9% +/- 5.5	0.318
MDM2	10.8% +/- 8.1 (n = 93)	9.7% +/- 6.1	11.3% +/- 8.8	0.594
CDK6	5.1% +/- 7.4 (n = 93)	3.6% +/- 5.3	5.7% +/- 8.1	0.056
E2F1	1.8% +/- 1.8 (n = 93)	1.2% +/- 0.9	2.0% +/- 1.9	0.020
BCL2	6.8% +/- 9.7 (n = 93)	3.4% +/- 4.5	8.1% +/- 10.9	0.087
TP53	5.2% +/- 8.1 (n = 94)	6.6% +/- 10.3	4.6% +/- 7.0	0.128

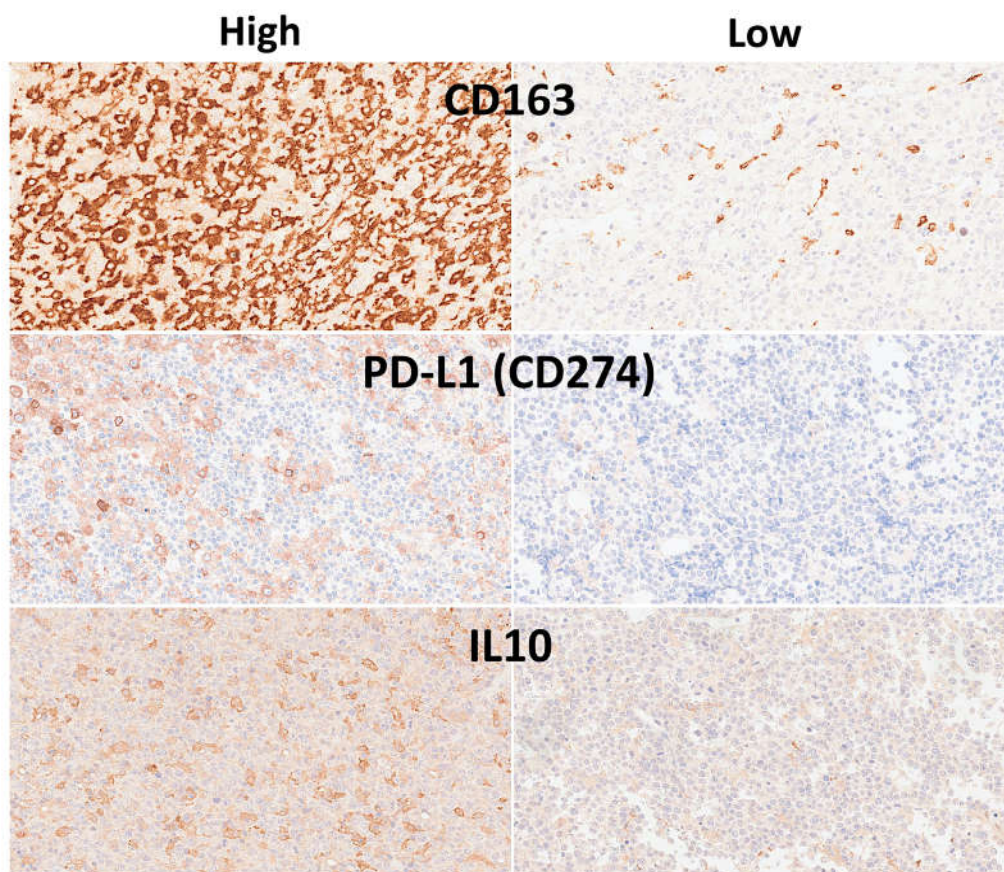


Figure 15. Most relevant immune microenvironment immunohistochemical markers. Analysis of the immune microenvironment, cell cycle, and germinal center markers showed that Dead < 2-years had higher CD163, IL10, and PD-L1 (all p values < 0.05).

4. Discussion

DLBCL is one of the most common diagnostic categories of non-Hodgkin lymphoma (NHL), accounting for approximately 25% of NHL cases in the developed world [2,5]. Histologically, it is characterized by large transformed B cells that depict a diffuse growth pattern and efface the normal architecture of the underlying histological structure [2,5]. The diagnostic category of DLBCL, NOS, refers to conventional DLBCL with a mature B-cell phenotype and large cell morphology, but lacking none of the criteria that define specific large B-cell lymphoma subtypes (i.e., other large B-cell variants) [2,5,10].

This study used a series of 114 cases of conventional large B-cell lymphoma. Overall, it included not only NOS cases but also 28 cases of EBER-positive DLBCL. Most of the cases were nodal (58/114, 50.9%), followed by other extranodal (28.1%) and gastrointestinal (11.4%). The aim of the project was to identify cases with poor prognosis using only H&E staining and CNN. After the evaluation of the overall survival curve, two groups were defined: patients who died before the 2 years and the others. Correlation with the clinicopathological characteristics found that the Dead < 2-years group was correlated with stage III-IV, International Prognostic Index (IPI) High + High/intermediate, progressive disease, non-GCB cell-of-origin, CD10-, BCL2+, and EBER+. Analysis of the immune microenvironment, cell cycle, and germinal center markers showed that Dead < 2-years had higher IL10, CD163, PD-L1, and lower E2F1 expressions. Therefore, the CNN managed to identify the histological features that correlated with the prognosis of the patients. Features that were not very obvious under conventional histological examination under the optical microscope.

The deep learning workflow includes preprocessing data, importing and building the network, training the network, tuning the network, and visualizing the results. In this study, we used transfer learning to take advantage of the knowledge provided by a pre-trained network to learn new patterns

in new data. Fine-tuning a pretrained network with transfer learning is typically much faster and easier than training from the beginning. The reuse of the pretrained network includes the following steps: load the pretrained network, replace the final layers, train the network, predict and assess the network accuracy, and deploy the results. This study used 20 pre-trained networks, including AlexNet, DarkNet-19, DarkNet-53, DenseNet-201, EfficientNet-b0, GoogleLeNet, GoogleLeNet-places365, Inception-ResNet-v2, Inception-v3, MobileNet-v2, NasNet-Large, NasNet-Mobile, ResNet-101, ResNet-18, ResNet-50, Shufflenet, SqueezeNet, VGG-16, VGG-19, and Xception. The CNN architectures were different and the performances ranged from 79.16% of SqueezeNet to 96.26% of DarkNet-19. The training times differed as well, being 1429 min for NasNet-Large and 2 min for AlexNet. The final analysis was performed using DarkNet-19. The final model based on DarkNet-19 predicted prognosis groups with high performance (test set accuracy = 96.26%). The other performance parameters were precision (94.46%), recall (95.02%), false positive rate (3.07%), specificity (96.93%), and F1 score (94.74%).

A machine learning model is often referred to as a "black box" model because understanding how the model makes predictions can be difficult. Interpretability tools help you overcome this aspect of machine learning algorithms and reveal how predictors contribute (or do not contribute) to predictions. Moreover, you can validate whether the model uses the correct evidence for its predictions and find model biases that are not immediately apparent. In this study, explainable artificial intelligence (XAI) was performed using Grad-CAM, occlusion sensitivity and image Lime on the network DarkNet-19. Overall, XAI showed that the convolutional neural networks (CNNs) were focusing on the neoplastic B-lymphocytes but some components of the microenvironment may also have played an influence. Of note, current XAI techniques do not allow a more detailed analysis of this type of histological tissue

In conclusion, narrow artificial intelligence (i.e., trained to perform a specific or a set of closely related tasks) can predict the prognosis of DLBCL based on the computer vision CNN histological analysis of H&E images, but it is process-specific and operates within limited constraints.

Author Contributions: All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded to J.C. by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the Japan Society for the Promotion of Science (JSPS), grant numbers KAKEN 15K19061, 18K15100, and 23K06454.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of TOKAI UNIVERSITY, SCHOOL OF MEDICINE (protocol codes IRB21R-096, IRB14R-080, and IRB20-156).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: All data is available upon request to Joaquim Carreras (joaquim.carreras@tokai.ac.jp). Data is also uploaded to CERN OpenAire ZENODO repository:

Acknowledgments: N/A.

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

Abbreviations

The following abbreviations are used in this manuscript:

CNIO	Spanish National Cancer Research Center
CNN	Convolutional neural network
DLBCL	Diffuse large B-cell lymphoma
H&E	Hematoxylin and eosin
NHL	Non-Hodgkin lymphoma
XAI	Explainable artificial intelligence

Appendix A

Useful MATLAB R2023b Update 10 (23.2.0.2859533) code:

```
%L1 is image 1

%% Resize to 299x299x3 because trainedNetwork is Inceptionv3
B = imresize (L1, [299 NaN])

%% Check size of image
sizesz = size(B)

%% Grad-CAM
imds = B
label = classify(trainedNetwork_1,imds)
scoreMap = gradCAM(trainedNetwork_1,imds,label);
figure
imshow(imds)
hold on
imagesc(scoreMap,'AlphaData',0.5)
colormap jet
colorbar

%% imageLIME
label = classify(trainedNetwork_1,B)
scoreMap = imageLIME(trainedNetwork_1,B,label);
figure
imshow(B)
hold on
imagesc(scoreMap,'AlphaData',0.5)
colormap jet
colorbar

%% OcclusionSensitivity
label = classify(trainedNetwork_1,B)
scoreMap = occlusionSensitivity(trainedNetwork_1,B,label);
figure
imshow(B)
hold on
imagesc(scoreMap,'AlphaData',0.5)
colormap jet
colorbar

%% To classify and show image
net = trainedNetwork_1
classes = net.Layers(end).Classes;
[YPred, scores] = classify(net, B);
[score, classIdx] = max(scores);

predClass = classes (classIdx);

imshow(B);
title(sprintf("%s (%.2f)",string(predClass),score));

%% To display an image
imshow(B)

%%
imageLoc = "1104802"
imds = imageDatastore(imageLoc)
imds2 = augmentedImageDatastore([299 299 3], imds);
YPred = classify(trainedNetwork_1,imds2)
scores = predict(trainedNetwork_1,imds2)
```

Appendix B

Immunohistochemistry was performed using Leica Bond-Max system and reagents following the manufacturer's instruction. Slides were digitalized using a NanoZoomer S360 digital slide scanner (Hamamatsu Photonics). Digital image quantification was performed using Fiji software (ImageJ). The primary antibodies and the staining conditions for the target markers were the following:

Table A1. Immunohistochemical markers and primary antibodies

Marker	Target	Clone	Company
CD3	T lymphocytes	LN10	Novocastra (Leica)
CD20	B lymphocytes	L26	Novocastra (Leica)
CD5	T lymphocytes	4C7	Novocastra (Leica)
CD10	Germinal center	56C6	Novocastra (Leica)
BCL6	Germinal center	LN22	Novocastra (Leica)
MUM1 /	Plasma cell differentiation	EAU32	Novocastra (Leica)
BCL2	Apoptosis	Bcl2/10/D5	Novocastra (Leica)
EBER	EBV-encoded mRNA	BP0589/AR0833	Novocastra (Leica)
Ki67	Cell proliferation	MM1	Novocastra (Leica)
IL10	Immuno-oncology	LS-B7432	Lifespan Bioscience
PD-L1	Immuno-oncology	E1J2	Cell Signaling
CSF1R	Immuno-oncology	FER216	CNIO
CD163	Tumor-associated	10D6	Novocastra (Leica)
CASP8	Active subunit p18	11B6	Novocastra
TNFAIP8	Apoptosis	14559-MM0	Sino Biological
LMO2	Hematopoietic	299B	CNIO
MYC	Proto-oncogene	Y69	Abcam
MDM2	Proto-oncogene	IF2	Invitrogen
CDK6	Cell cycle	98D	CNIO
E2F1	Cell cycle	Agro368V	CNIO
TP53	Cell regulation	DO-7	Novocastra (Leica)

CNIO, Spanish National Cancer Research Center (CNIO), Madrid, Spain.

References

1. Brown JR, Aster JC, Lister A, Rosmarin A. Pathobiology of diffuse large B cell lymphoma and primary mediastinal large B cell lymphoma. In UpToDate. Website: www.uptodate.com (last accessed on 19th November 2025).
2. Alaggio, R.; Amador, C.; Anagnostopoulos, I.; Attygalle, A.D.; Araujo, I.B.O.; Berti, E.; Bhagat, G.; Borges, A.M.; Boyer, D.; Calaminici, M.; et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia* **2022**, *36*, 1720-1748, doi:10.1038/s41375-022-01620-2.
3. Arber, D.A.; Campo, E.; Jaffe, E.S. Advances in the Classification of Myeloid and Lymphoid Neoplasms. *Virchows Arch* **2023**, *482*, 1-9, doi:10.1007/s00428-022-03487-1.
4. Campo, E. The 2022 classifications of lymphoid neoplasms : Keynote. *Pathologie (Heidelb)* **2023**, *44*, 121-127, doi:10.1007/s00292-023-01247-9.
5. Campo, E.; Jaffe, E.S.; Cook, J.R.; Quintanilla-Martinez, L.; Swerdlow, S.H.; Anderson, K.C.; Brousset, P.; Cerroni, L.; de Leval, L.; Dirnhofer, S.; et al. The International Consensus Classification of Mature Lymphoid Neoplasms: a report from the Clinical Advisory Committee. *Blood* **2022**, *140*, 1229-1253, doi:10.1182/blood.2022015851.
6. de Leval, L.; Alizadeh, A.A.; Bergsagel, P.L.; Campo, E.; Davies, A.; Dogan, A.; Fitzgibbon, J.; Horwitz, S.M.; Melnick, A.M.; Morice, W.G.; et al. Genomic profiling for clinical decision making in lymphoid neoplasms. *Blood* **2022**, *140*, 2193-2227, doi:10.1182/blood.2022015854.

7. Song, J.Y.; Dirnhofer, S.; Piris, M.A.; Quintanilla-Martinez, L.; Pileri, S.; Campo, E. Diffuse large B-cell lymphomas, not otherwise specified, and emerging entities. *Virchows Arch* **2023**, *482*, 179-192, doi:10.1007/s00428-022-03466-6.
8. Li, S.; Young, K.H.; Medeiros, L.J. Diffuse large B-cell lymphoma. *Pathology* **2018**, *50*, 74-87, doi:10.1016/j.pathol.2017.09.006.
9. Crump, M. Management of Relapsed Diffuse Large B-cell Lymphoma. *Hematol Oncol Clin North Am* **2016**, *30*, 1195-1213, doi:10.1016/j.hoc.2016.07.004.
10. Swerdlow, S.H.; Campo, E.; Pileri, S.A.; Harris, N.L.; Stein, H.; Siebert, R.; Advani, R.; Ghielmini, M.; Salles, G.A.; Zelenetz, A.D.; et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood* **2016**, *127*, 2375-2390, doi:10.1182/blood-2016-01-643569.
11. Xie, Y.; Pittaluga, S.; Jaffe, E.S. The histological classification of diffuse large B-cell lymphomas. *Semin Hematol* **2015**, *52*, 57-66, doi:10.1053/j.seminhematol.2015.01.006.
12. de Leval, L.; Jaffe, E.S. Lymphoma Classification. *Cancer J* **2020**, *26*, 176-185, doi:10.1097/PPO.0000000000000451.
13. Willemze, R.; Cerroni, L.; Kempf, W.; Berti, E.; Facchetti, F.; Swerdlow, S.H.; Jaffe, E.S. The 2018 update of the WHO-EORTC classification for primary cutaneous lymphomas. *Blood* **2019**, *133*, 1703-1714, doi:10.1182/blood-2018-11-881268.
14. Jaffe, E.S. Diagnosis and classification of lymphoma: Impact of technical advances. *Semin Hematol* **2019**, *56*, 30-36, doi:10.1053/j.seminhematol.2018.05.007.
15. Quintanilla-Martinez, L.; Swerdlow, S.H.; Tousseyn, T.; Barrionuevo, C.; Nakamura, S.; Jaffe, E.S. New concepts in EBV-associated B, T, and NK cell lymphoproliferative disorders. *Virchows Arch* **2023**, *482*, 227-244, doi:10.1007/s00428-022-03414-4.
16. Hans, C.P.; Weisenburger, D.D.; Greiner, T.C.; Gascoyne, R.D.; Delabie, J.; Ott, G.; Muller-Hermelink, H.K.; Campo, E.; Braziel, R.M.; Jaffe, E.S.; et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* **2004**, *103*, 275-282, doi:10.1182/blood-2003-05-1545.
17. Collinge, B.; Hilton, L.K.; Wong, J.; Alduaij, W.; Ben-Neriah, S.; Slack, G.W.; Farinha, P.; Boyle, M.; Meissner, B.; Cook, J.R.; et al. High-grade B-cell lymphoma, not otherwise specified: an LLMP study. *Blood Adv* **2025**, *9*, 5409-5422, doi:10.1182/bloodadvances.2025016651.
18. Herlevic, V.; Reynolds, S.B.; Morris, J.D. Gastric Lymphoma. In *StatPearls*; Treasure Island (FL), 2025.
19. Grainger, B.T.; Cheah, C.Y. Primary testicular lymphoma. *Cancer Treat Rev* **2025**, *136*, 102927, doi:10.1016/j.ctrv.2025.102927.
20. Chen, S.Y.; Xu, P.P.; Feng, R.; Cui, G.H.; Wang, L.; Cheng, S.; Mu, R.J.; Zhang, H.L.; Wei, X.L.; Song, Y.P.; et al. Extranodal diffuse large B-cell lymphoma: Clinical and molecular insights with survival outcomes from the multicenter EXPECT study. *Cancer Commun (Lond)* **2025**, *45*, 919-935, doi:10.1002/cac2.70033.
21. International Non-Hodgkin's Lymphoma Prognostic Factors, P. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med* **1993**, *329*, 987-994, doi:10.1056/NEJM199309303291402.
22. Scott, D.W. Cell-of-Origin in Diffuse Large B-Cell Lymphoma: Are the Assays Ready for the Clinic? *Am Soc Clin Oncol Educ Book* **2015**, e458-466, doi:10.14694/EdBook_AM.2015.35.e458.
23. Choi, W.W.; Weisenburger, D.D.; Greiner, T.C.; Piris, M.A.; Banham, A.H.; Delabie, J.; Braziel, R.M.; Geng, H.; Iqbal, J.; Lenz, G.; et al. A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. *Clin Cancer Res* **2009**, *15*, 5494-5502, doi:10.1158/1078-0432.CCR-09-0113.
24. Alizadeh, A.A.; Eisen, M.B.; Davis, R.E.; Ma, C.; Lossos, I.S.; Rosenwald, A.; Boldrick, J.C.; Sabet, H.; Tran, T.; Yu, X.; et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **2000**, *403*, 503-511, doi:10.1038/35000501.
25. Rosenwald, A.; Wright, G.; Chan, W.C.; Connors, J.M.; Campo, E.; Fisher, R.I.; Gascoyne, R.D.; Muller-Hermelink, H.K.; Smeland, E.B.; Giltnane, J.M.; et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* **2002**, *346*, 1937-1947, doi:10.1056/NEJMoa012914.

26. Wright, G.; Tan, B.; Rosenwald, A.; Hurt, E.H.; Wiestner, A.; Staudt, L.M. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A* **2003**, *100*, 9991-9996, doi:10.1073/pnas.1732008100.
27. Berglund, M.; Thunberg, U.; Amini, R.M.; Book, M.; Roos, G.; Erlanson, M.; Linderroth, J.; Dictor, M.; Jerkeman, M.; Cavallin-Stahl, E.; et al. Evaluation of immunophenotype in diffuse large B-cell lymphoma and its impact on prognosis. *Mod Pathol* **2005**, *18*, 1113-1120, doi:10.1038/modpathol.3800396.
28. Aster JC, Abramson JS, Lister A, Rosmarin AG. Prognosis of diffuse large B cell lymphoma. In UpToDate (last accessed on 20th November, 2025).
29. Horn, H.; Ziepert, M.; Becher, C.; Barth, T.F.; Bernd, H.W.; Feller, A.C.; Klapper, W.; Hummel, M.; Stein, H.; Hansmann, M.L.; et al. MYC status in concert with BCL2 and BCL6 expression predicts outcome in diffuse large B-cell lymphoma. *Blood* **2013**, *121*, 2253-2263, doi:10.1182/blood-2012-06-435842.
30. Johnson, N.A.; Slack, G.W.; Savage, K.J.; Connors, J.M.; Ben-Neriah, S.; Rogic, S.; Scott, D.W.; Tan, K.L.; Steidl, C.; Sehn, L.H.; et al. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J Clin Oncol* **2012**, *30*, 3452-3459, doi:10.1200/JCO.2011.41.0985.
31. Green, T.M.; Young, K.H.; Visco, C.; Xu-Monette, Z.Y.; Orazi, A.; Go, R.S.; Nielsen, O.; Gadeberg, O.V.; Mourits-Andersen, T.; Frederiksen, M.; et al. Immunohistochemical double-hit score is a strong predictor of outcome in patients with diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J Clin Oncol* **2012**, *30*, 3460-3467, doi:10.1200/JCO.2011.41.4342.
32. Hu, S.; Xu-Monette, Z.Y.; Tzankov, A.; Green, T.; Wu, L.; Balasubramanyam, A.; Liu, W.M.; Visco, C.; Li, Y.; Miranda, R.N.; et al. MYC/BCL2 protein coexpression contributes to the inferior survival of activated B-cell subtype of diffuse large B-cell lymphoma and demonstrates high-risk gene expression signatures: a report from The International DLBCL Rituximab-CHOP Consortium Program. *Blood* **2013**, *121*, 4021-4031; quiz 4250, doi:10.1182/blood-2012-10-460063.
33. Somasundaram, E.; Abramson, J.S. Double hit lymphoma: contemporary understanding and practices. *Leuk Lymphoma* **2025**, *66*, 26-33, doi:10.1080/10428194.2024.2412295.
34. Karmali, R.; Shouse, G.; Torka, P.; Moyo, T.K.; Romancik, J.; Barta, S.K.; Bhansali, R.; Cohen, J.B.; Shah, N.N.; Zurko, J.; et al. Double hit & double expressor lymphomas: a multicenter analysis of survival outcomes with CD19-directed CAR T-cell therapy. *Blood Cancer J* **2025**, *15*, 43, doi:10.1038/s41408-025-01250-8.
35. Qiu, L.; Medeiros, L.J.; Li, S. High-grade B-cell lymphomas: Double hit and non-double hit. *Hum Pathol* **2025**, *156*, 105700, doi:10.1016/j.humpath.2024.105700.
36. Reddy, A.; Zhang, J.; Davis, N.S.; Moffitt, A.B.; Love, C.L.; Waldrop, A.; Leppa, S.; Pasanen, A.; Meriranta, L.; Karjalainen-Lindsberg, M.L.; et al. Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. *Cell* **2017**, *171*, 481-494 e415, doi:10.1016/j.cell.2017.09.027.
37. Schmitz, R.; Wright, G.W.; Huang, D.W.; Johnson, C.A.; Phelan, J.D.; Wang, J.Q.; Roulland, S.; Kasbekar, M.; Young, R.M.; Shaffer, A.L.; et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N Engl J Med* **2018**, *378*, 1396-1407, doi:10.1056/NEJMoa1801445.
38. Chapuy, B.; Stewart, C.; Dunford, A.J.; Kim, J.; Kamburov, A.; Redd, R.A.; Lawrence, M.S.; Roemer, M.G.M.; Li, A.J.; Ziepert, M.; et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med* **2018**, *24*, 679-690, doi:10.1038/s41591-018-0016-8.
39. Carreras, J.; Kikuti, Y.Y.; Roncador, G.; Miyaoka, M.; Hiraiwa, S.; Tomita, S.; Ikoma, H.; Kondo, Y.; Ito, A.; Shiraiwa, S.; et al. High Expression of Caspase-8 Associated with Improved Survival in Diffuse Large B-Cell Lymphoma: Machine Learning and Artificial Neural Networks Analyses. *BioMedInformatics* **2021**, *1*, 18-46. <https://doi.org/10.3390/biomedinformatics1010003>.
40. Carreras, J.; Kikuti, Y.Y.; Miyaoka, M.; Hiraiwa, S.; Tomita, S.; Ikoma, H.; Kondo, Y.; Ito, A.; Shiraiwa, S.; Hamoudi, R.; et al. A Single Gene Expression Set Derived from Artificial Intelligence Predicted the Prognosis of Several Lymphoma Subtypes; and High Immunohistochemical Expression of TNFAIP8 Associated with Poor Prognosis in Diffuse Large B-Cell Lymphoma. *AI* **2020**, *1*, 342-360. <https://doi.org/10.3390/ai1030023>

41. Carreras, J.; Kikuti, Y.Y.; Miyaoka, M.; Roncador, G.; Garcia, J.F.; Hiraiwa, S.; Tomita, S.; Ikoma, H.; Kondo, Y.; Ito, A.; et al. Integrative Statistics, Machine Learning and Artificial Intelligence Neural Network Analysis Correlated CSF1R with the Prognosis of Diffuse Large B-Cell Lymphoma. *Hemato* 2021, 2, 182-206. <https://doi.org/10.3390/hemato2020011>.

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.