



Article

Immunohistochemical HER2 Recognition and Analysis of Breast Cancer Based on Deep Learning

Yuxuan Che ^{1,2,3,†} , Fei Ren ^{1,†}, Xueyuan Zhang ⁴, Li Cui ¹, Huanwen Wu ^{5,*} and Ze Zhao ^{1,*} ¹ Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China² School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 101408, China³ Jinfeng Laboratory, Chongqing 401329, China⁴ Beijing Zhijian Life Technology Co., Ltd., Beijing 100036, China⁵ Department of Pathology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

* Correspondence: wuhuanwen10700@pumch.cn (H.W.); zhaoze@ict.ac.cn (Z.Z.)

† These authors contributed equally to this work.

Abstract: Breast cancer is one of the common malignant tumors in women. It seriously endangers women's life and health. The human epidermal growth factor receptor 2 (HER2) protein is responsible for the division and growth of healthy breast cells. The overexpression of the HER2 protein is generally evaluated by immunohistochemistry (IHC). The IHC evaluation criteria mainly includes three indexes: staining intensity, circumferential membrane staining pattern, and proportion of positive cells. Manually scoring HER2 IHC images is an error-prone, variable, and time-consuming work. To solve these problems, this study proposes an automated predictive method for scoring whole-slide images (WSI) of HER2 slides based on a deep learning network. A total of 95 HER2 pathological slides from September 2021 to December 2021 were included. The average patch level precision and f1 score were 95.77% and 83.09%, respectively. The overall accuracy of automated scoring for slide-level classification was 97.9%. The proposed method showed excellent specificity for all IHC 0 and 3+ slides and most 1+ and 2+ slides. The evaluation effect of the integrated method is better than the effect of using the staining result only.

Keywords: breast cancer; HER2; IHC; whole-slide image; deep learning



check for updates

Citation: Che, Y.; Ren, F.; Zhang, X.;

Cui, L.; Wu, H.; Zhao, Z.

Immunohistochemical HER2

Recognition and Analysis of Breast

Cancer Based on Deep Learning.

Diagnostics **2023**, *13*, 263.<https://doi.org/10.3390/diagnostics13020263>

diagnostics13020263

Academic Editor: Filippo Pesapane

Received: 29 November 2022

Revised: 5 January 2023

Accepted: 6 January 2023

Published: 10 January 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/](https://creativecommons.org/licenses/by/4.0/)[https://creativecommons.org/licenses/by/](https://creativecommons.org/licenses/by/4.0/)

4.0/).

1. Introduction

Breast cancer has become one of the most common cancers worldwide. According to Global Cancer Statistics 2020, there are about 2.3 million new breast cancers worldwide and about 685,000 deaths, accounting for 15.5% of female malignancies [1]. Breast cancer is also one of the important causes of female tumor-related death, which greatly affects the physical and mental health of people all over the world.

Human epidermal growth factor receptor 2 (HER2)-positive breast cancer refers to the amplification of the ERBB2/neu proto-oncogene or the overexpression of the HER2 transmembrane receptor protein. Compared with other types of breast cancer, HER2-positive breast cancer has a high degree of malignancy. It is a special breast cancer subtype with strong aggressiveness, early recurrence and metastasis, and poor prognosis [2–4].

HER2 receptor protein overexpression is generally assessed by immunohistochemistry (IHC). Normally, amplification levels of the HER2 gene were detected by fluorescence in situ hybridization (FISH) and chromogenic in situ hybridization (CISH). Guidelines of the Chinese Society of Clinical Oncology (CSCO) 2021 specify the criteria for HER2 to improve the procedures for Her2 testing and standardize the interpretation of the results [5]. According to CSCO 2021, the status of HER2 should be screened by the IHC method first for newly diagnosed breast cancer cases. If the results of the HER2 IHC staining are uncertain,

FISH detection should be performed for confirmation. As shown in Table 1, if more than 10% of the infiltrating cancer cells have strong and intact cell membranes with brown staining in IHC slides, the case displays 3+ and it is accepted as HER2-positive. If less than 10% of the infiltrating cancer cells have intact brown cell membranes or more than 10% of the infiltrating cancer cells have incomplete and/or weak to moderate membranous staining, the case is diagnosed as 2+ (HER2-equivocal) and the further ISH testing is needed to assess HER2 expression [6]. If faint/barely perceptible membrane staining is detected in more than 10% of invasive tumor cells, the case is reported as 1+ (HER2-negative). If no staining or faint/barely perceptible membrane staining is seen in less than 10% of invasive tumor cells, the case is reported as 0 (HER2-negative). The IHC evaluation criteria can be summarized into three aspects: staining intensity, circumferential membrane staining pattern, and proportion of positive cells. However, these criteria are still subjective in practice, and there is no specific and explainable numerical basis. Therefore, it is highly significant to propose an interpretable algorithm with automated IHC scoring diagnosis [7].

Table 1. Evaluation criteria for HER2 expression by IHC assay in breast cancer.

IHC Score	Staining Pattern	HER2 Expression
0	No staining or incomplete membrane staining which is faint or barely perceptible in $\leq 10\%$ of invasive tumor cells	Negative
1+	Incomplete membrane staining which is faint or barely perceptible in $>10\%$ of invasive tumor cells	Low expression
2+	(a) Weak to moderate membrane staining with uneven brownish yellow coloration in $>10\%$ of invasive tumor cells (b) $\leq 10\%$ of invasive tumor cells have circumferential membrane staining which is complete, intense, and has brownish coloration	Equivocal (low expression if the slide is ISH-negative, positive if it is ISH-positive.)
3+	$>10\%$ of invasive tumor cells have circumferential membrane staining which is complete, intense, and has brownish coloration	Positive

Computer-aided diagnosis systems have developed rapidly in the medical field [8–11]. Using computers to perform objective and quantitative analysis of medical imaging data to assist doctors in clinical diagnosis of lesions can help to improve diagnosis accuracy and efficiency [12]. The appearance of digital whole-slide images (WSIs) gives the opportunity to see and analyze more detailed information and make a great step forward in automatic metastatic breast cancer detection [13–17]. WSI is obtained by scanning and collecting traditional glass pathological sections through an automatic microscope or optical magnification system with a digital section acquisition device. It has high resolution and a large file size. In general, WSI has multiple layers, representing a pyramid structure. The different layers of the WSI correspond to different resolutions. The bottom layer of the pyramid has the highest resolution image data, while the upper layers are thumbnails of the bottom image for the pathologist to retrieve the data at low resolutions. It is worth mentioning that the length or width between layers is usually double, which makes downsampling faster and more accurate. However, since a single WSI has billions of pixels, the WSI labeling process is time-consuming for doctors. Therefore, a deep-learning-based network is used for auxiliary analysis of IHC WSIs [18]. Since the computer cannot directly process the WSI image, we need to cut the image into several patches, calculate each patch, then generate a thermal map diagnosis. We proposed an architecture to identify cancer areas in IHC images and generate corresponding probability maps.

2. Materials and Methods

As shown in Figure 1, the proposed method includes 3 stages. First, the labeled masks were extracted from the original WSIs with corresponding labels. Then, the tumor patches

and normal patches were generated randomly according to label masks. These patches were passed to a deep learning model (ResNet34) to refine the binary classification [19]. In stage two, the tissue mask was extracted from the test WSI. The patches generated from the tissue mask were passed into the model to build the probability map. Then, the binary tumor prediction was produced from the probability map with a threshold. In stage three, the test WSI was differentiated from the four subclasses: IHC 0/1+/2+/3+. This part of the work was implemented to perform an accurate and interpretable result using comprehensive judgments.

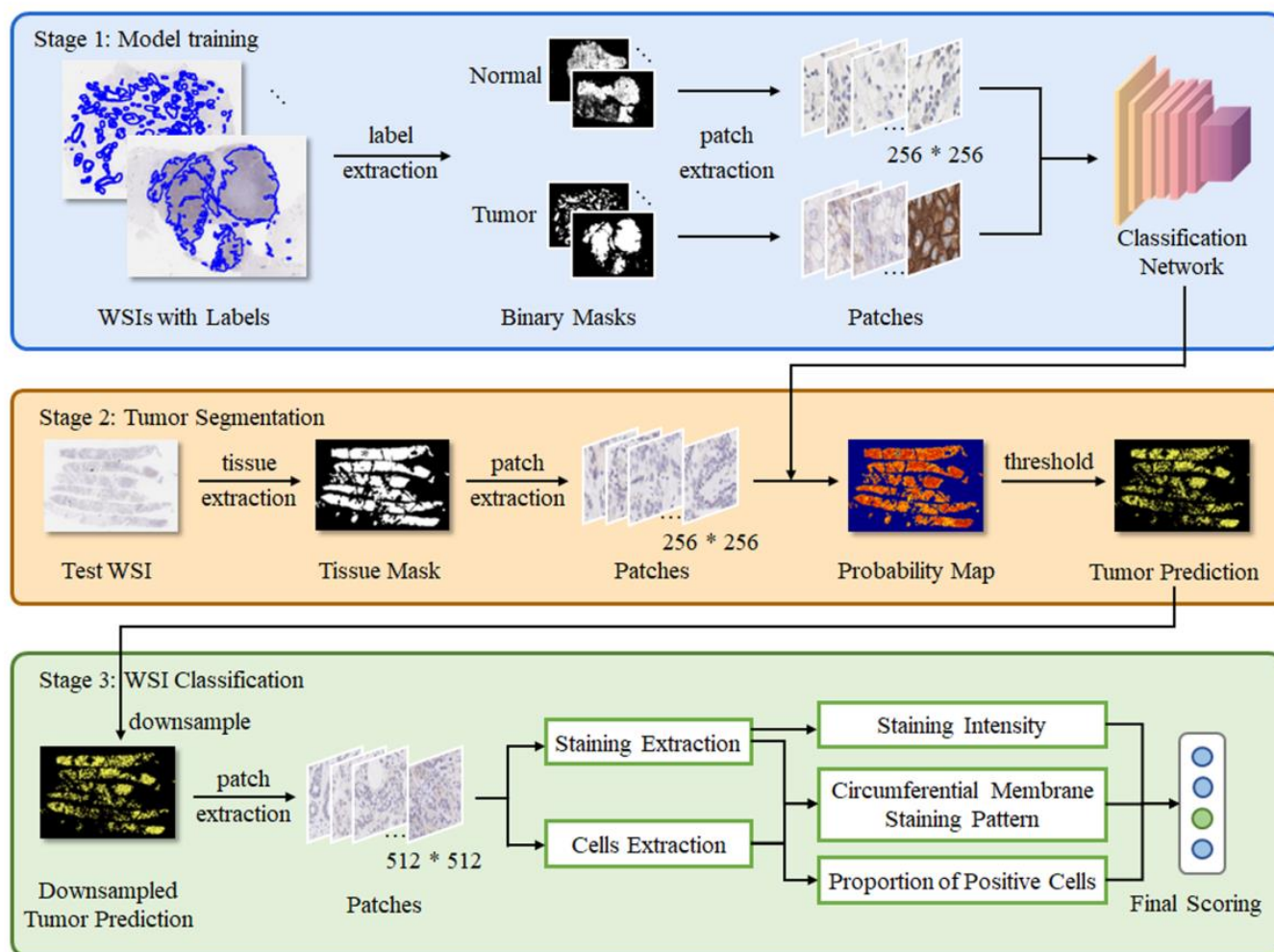


Figure 1. The proposed framework for this study.

2.1. Data Acquisition

Pathological slides from breast cancer patients in Peking Union Medical College Hospital from September to December in 2021 were retrospectively included to form the dataset of this study. These slides were scanned into WSIs using Aperio AT2 (Leica, Germany) high-throughput biopsy scanner with a 20-magnification-scale objective and bright-field illumination. The scan resolution was 0.5036 μm per pixel. The dataset consisted of a total of 95 whole slide images in Aperio format (.svs). Four categories were included, which were IHC 0, 1+, 2+, and 3+. Two pathologists selected 23 WSIs from the dataset and labeled the area with concentrated and evenly distributed tumor cells for training and testing the deep learning model. The IHC scoring of all pathological images was determined according to CSCO 2021. Figure 2 shows the samples of each category of this dataset. The data distribution performed is shown in Table 2.

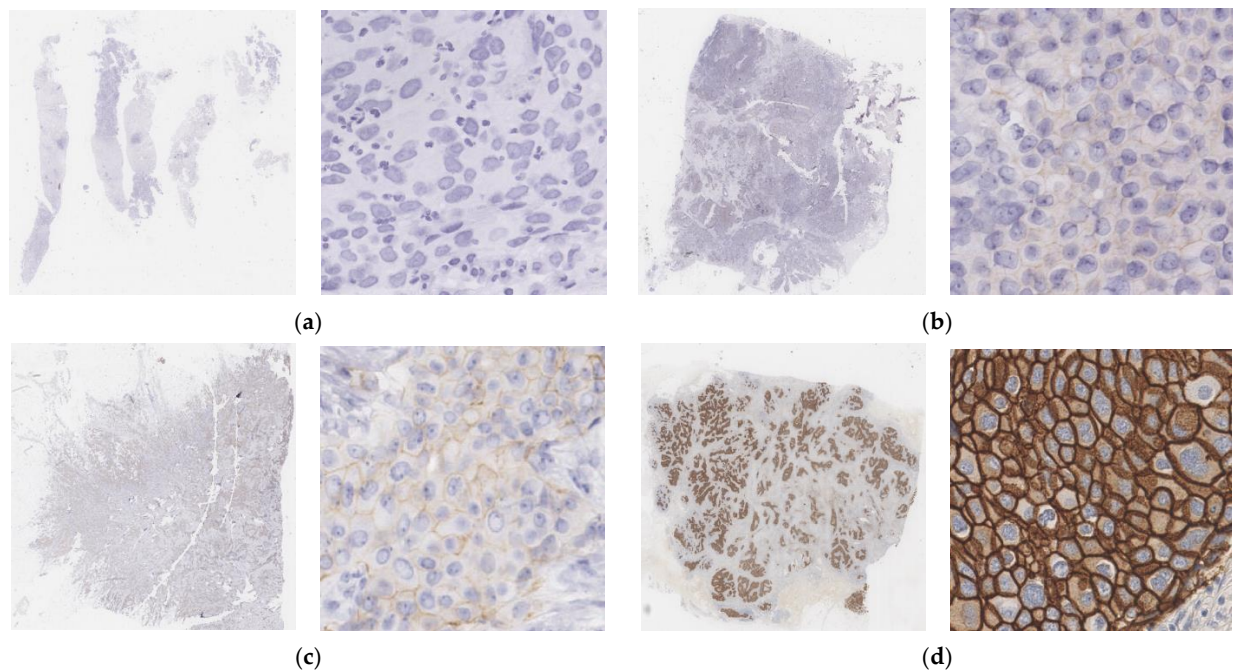


Figure 2. Sample thumbnail images (**left**) and patches (**right**) of the dataset for four categories showing typical levels of membrane staining. (a) IHC 0. (b) IHC 1+. (c) IHC 2+. (d) IHC 3+.

Table 2. Composition of the dataset.

IHC Score	No. WSIs	No. Labeled WSIs ¹	No. WSIs with 3+ control ²
0	14	2	6
1+	25	7	7
2+	36	7	24
3+	20	7	10
Total	95	23	47

¹ The labeled areas are all tumor areas rather than regions which only satisfy the IHC score. ² WSI with 3+ control means this WSI has a IHC 3+ control tissue next to the main tissue (see Figure 3).

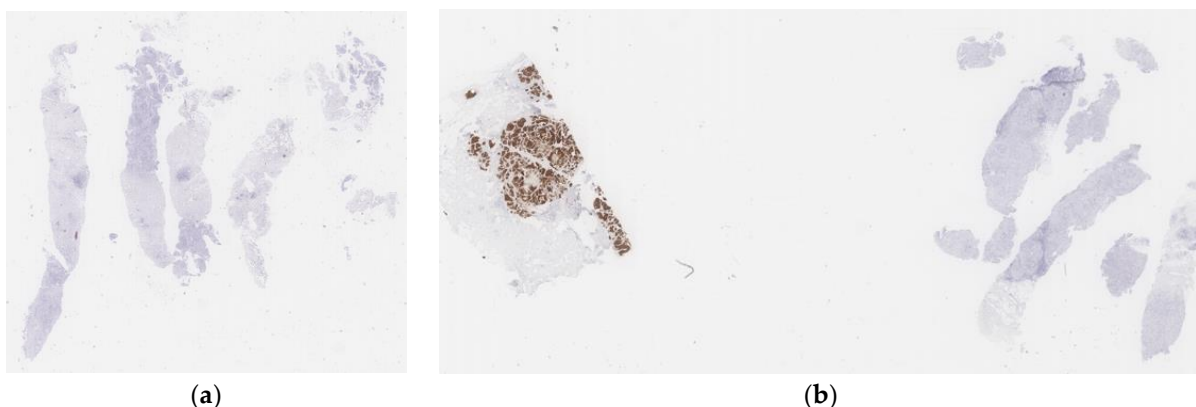


Figure 3. (a) Normal WSI. (b) WSI with 3+ control.

2.2. Image Preprocessing

Since WSIs usually have billions of pixels, either labeling or calculating WSIs at high magnification scale is an extremely tedious and time-consuming process. By observing the sections, it is found that there are tissue areas and white background areas in each WSI. Therefore, in order to reduce the computation time and complexity, it is necessary

to focus on the analysis of the tissue area and skip the white background area. In this paper, a threshold-based segmentation method is utilized to detect the background region automatically. Specifically, the original image is first transferred from the RGB color space to the HSV color space. Then, the Otsu algorithm is used to calculate the optimal threshold of each channel, and the final mask image is generated by combining the masks of the H and S channels [20]. According to the actual calculation, the Otsu method can filter about 75% of the area which belongs to the background, on average, greatly improving the efficiency of calculation.

Due to the limited number of annotated WSI, the image augmentation method can be used to amplify the existing images to increase the type and number of images. Meanwhile, it plays a certain role in inhibiting the overfitting of the model. In this study, images were rotated and flipped at random angles so as to increase the number of training patches. It is worth noting that, in the process of image extraction or image augmentation, staining standardization is not carried out. The reason is the existing dyeing standardization algorithm with good effect is not perfect in the call of GPU, and it takes too long to process a single-patch image (1–7 s/patch). In addition, staining standardization changes the color of the image to some extent. Without a color standardization process, the brown part and the blue part of IHC staining can still be distinguished by color space conversion.

Before the algorithm training was initiated, the 20-magnification-scale WSIs were cropped into $256 * 256$ pixels of patches, and 23 labeled WSIs were randomly divided into the training set and the test set. A total of 8000 patches were obtained from the training set. Lesions identified on these patches were utilized to train and test the performance of ResNet34 algorithm. Annotated patches of training and validation datasets were separated by the function embedded in the scikit-learn package (ratio 9:1), and total of 16,000 images were eventually used in stage one, among which 14,400 images were used for training and 1600 images were used for validation.

2.3. Deep Learning Structure

Deep learning has been extensively used in the diagnosis and analysis of medical images in recent years [21–27]. The convolutional neural network stands out among many deep learning networks because of its strong feature learning ability and has become a cutting-edge algorithm in the field of image classification. In stage one, ResNet34 is used as the deep learning backbone network in this paper. During the training process, $256 * 256 * 3$ patches from the tumor and non-tumor regions of WSIs were used as inputs in this stage to train the classification model to distinguish the two classes. During the training phase, the hyperparameters of the network were set as follows: the optimizer was set to SGD, the learning rate was set to 0.01, the momentum was set to 0.9, the loss function was set to nn.CrossEntropyLoss, the epoch was 50, and the patch size was 64.

2.4. Extraction of Membranes and Cells

The diagnosis of HER2 dominates the type of subsequent treatment. Therefore, this diagnosis becomes very important for breast cancer patients [28]. IHC is a special staining method for finding the HER2 protein in cancer cells based on the detection of specific antigens in tissue. The IHC staining slides are composed of a brown channel (diaminobenzidine, DAB signal) and a counterstain blue/violet channel (hematoxylin, H signal). The membrane extraction method used in this study was performed on the brown channel.

According to the CSCO 2021 guidelines, the evaluation criteria of IHC-stained slides is directly related to the membrane staining condition. Therefore, the staining intensity is presented as an evaluation indicator. Staining intensity is the most important feature for the classification of HER2 slides. The staining intensity indicates the depth of a certain color in the image. In this study, it refers to the depth of brown areas. The image is converted from RGB color space to HSV color space, and the brown area is extracted by a function in OpenCV library [29]. After that, the extracted brown areas are converted to gray level and its depth is calculated. When the staining intensity is low, it indicates that the extracted area

is lightly stained, while when the staining intensity is high, it indicates that the extracted area is deeply stained. Its value range from 0 to 1. The staining intensity of each IHC score is in a different range, and hence this evaluation index is a good candidate feature for further classification [30–33].

In this study, color deconvolution and watershed algorithms are employed to extract tumor cells. After the input RGB image is preprocessed, the RGB value is converted into optical density (OD) space, with the value range of [0, 1]. The inverse matrix of the OD matrix is the required deconvolution matrix [34]. Therefore, color deconvolution method is employed to separate and distinguish DAB and H staining [35]. Due to the overlap between cells, the watershed algorithm based on distance transformation is needed in order to process the binary image obtained by the color deconvolution method. The specific algorithm is as follows.

Firstly, the image after color deconvolution is converted to gray level and the morphology operation is performed to eliminate the interference on the boundary. Then the distance transformation of the gray image is carried out to split the adherent cells. Finally, expansion and filling methods are performed, and the connected component-based method is used to extract the cells.

After the results of staining and cell extraction are obtained, the three interpretative evaluation indicators of IHC scoring are calculated using these results. The flow diagram of the specific algorithm is shown in Figure 4. The result of staining intensity can be calculated from the mean value of the results extracted by staining. In addition, a proper threshold should be set for the results of staining extraction for morphological corrosion and expansion operation. The number of positive cells can be obtained by contour extraction. The circumferential membrane staining pattern can be obtained by dividing the number of positive cells calculated above by the total number of cell counts. For the calculation of the proportion of positive cells, traversing the pixels of the stained area to find the nearest cell of these pixels for calculation is an accurate method. However, the algorithm has high computational complexity and greatly increases the running time. In view of this, this study presents a rapid method to calculate the proportion of positive cells. The core of this method is to multiply the stain extraction mask and the cell extraction mask. If the result of one pixel is non-zero, the cell where the point is located is marked as a staining-positive cell. Otherwise, it is marked as a staining-negative cell. This algorithm has low computational complexity. In addition, the computational efficiency greatly improves.

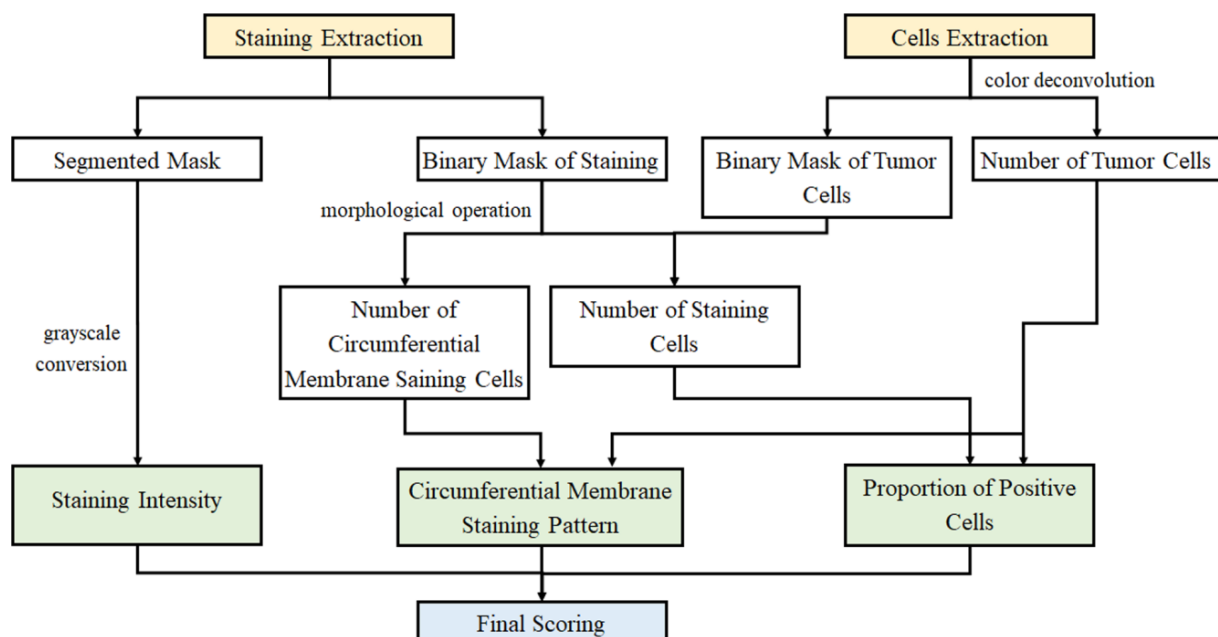


Figure 4. The detailed flow diagram of stage three.

3. Results

PyTorch was used to build and form the CNN model in this study. All experiments were conducted with a Linux server (Linux version CentOS 3.10.0 to 69.3.EL7.x86_64, CPU version Intel (R) Xeon (R) Silver 4114 @ 2.20 GHz. The name of the graphics card is NVIDIA GeForce RTX2080 Ti).

3.1. Tumor Area Classification

Figure 5 visualizes the training results of the stage one. Figure 5a is the original WSI. The black outline in Figure 5b shows the tumor areas marked by the pathologist. Due to the high resolution of WSI, patch-level classification of images is performed. The probability heat map of the tumor region output by the deep learning model is shown in Figure 5c with the original image added. Dark red color indicates areas with high probability of cancer, while light and blue areas indicate low probability of tumor. The results are basically in line with the expert annotations. A binary image of tumor areas is produced by selecting a proper threshold of the probability map. Note that small areas are filtered for visualization during the tissue area extraction phase.

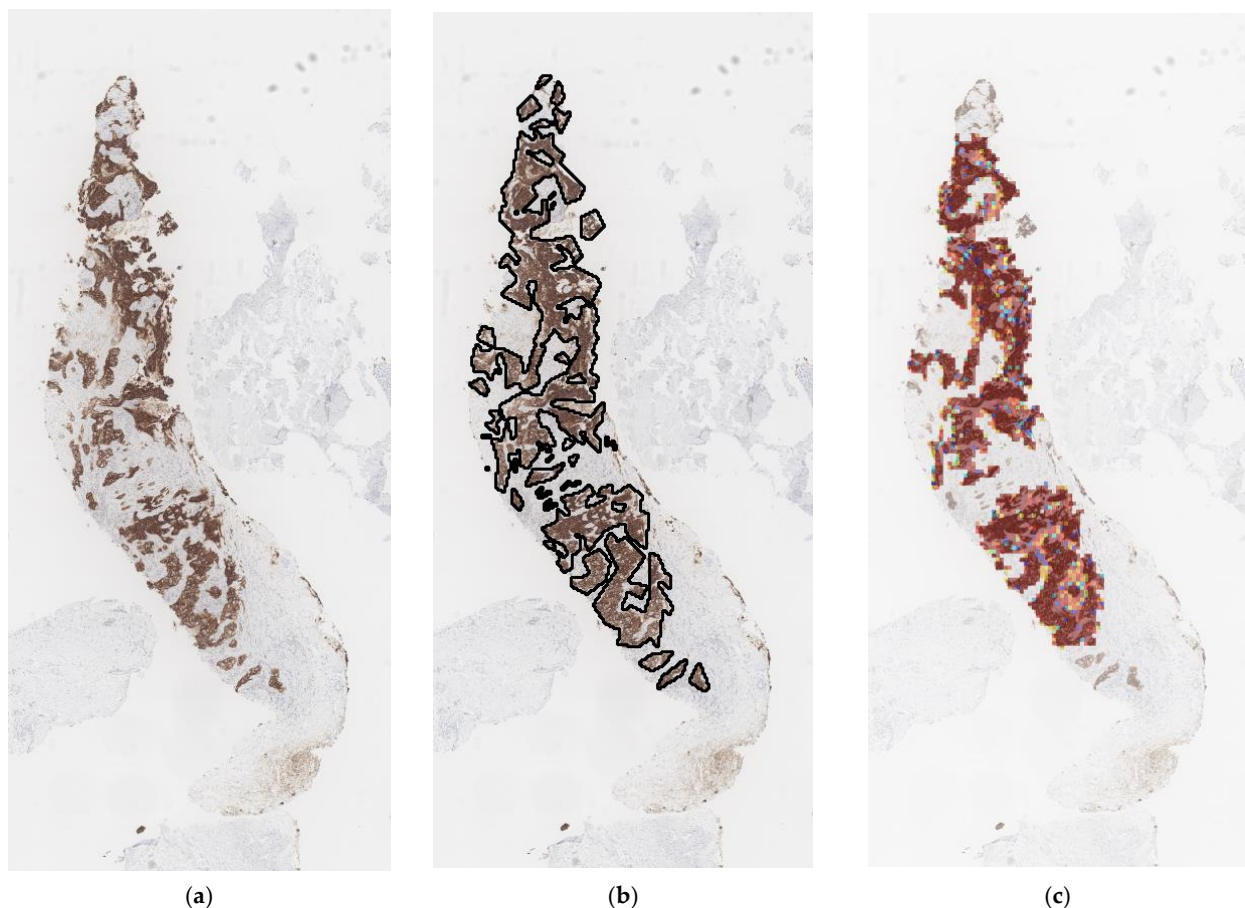


Figure 5. Visualization of model training result. (a) Original WSI. (b) WSI with black annotation. (c) The patch-level classification of tumor areas.

The dataset with annotations is randomly divided into the training set and the test set. At stage one, 16 WSIs are used for model training, and the remaining labeled WSIs are used for model inspection and evaluation. The following evaluation indexes are used to evaluate the model performance [36–41].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

where TP is true positive cases, FP stands for false positive cases, TN is true negative cases, and FN represents false negative cases. Table 3 shows the average results of patch-level data analysis. The values of false positive rate and true positive rate of the receiver operating characteristic (ROC) curve is shown in Figure 6a. The closer the ROC curve to the upper left-hand corner and the larger the area under curve (AUC) value, the better the performance of classification. Similarly, the closer the precision-recall (PR) curve to the upper right corner and the larger the average precision (AP) value, the better the performance of the model. It can be calculated that the proposed model for tumor classification provides high value of AUC (0.983) and AP (0.984).

Table 3. Patch-level classification performance on the test set.

Evaluation Indexes	Accuracy	Precision	Recall	F1 Score
	73.49%	95.77%	73.38%	83.09%

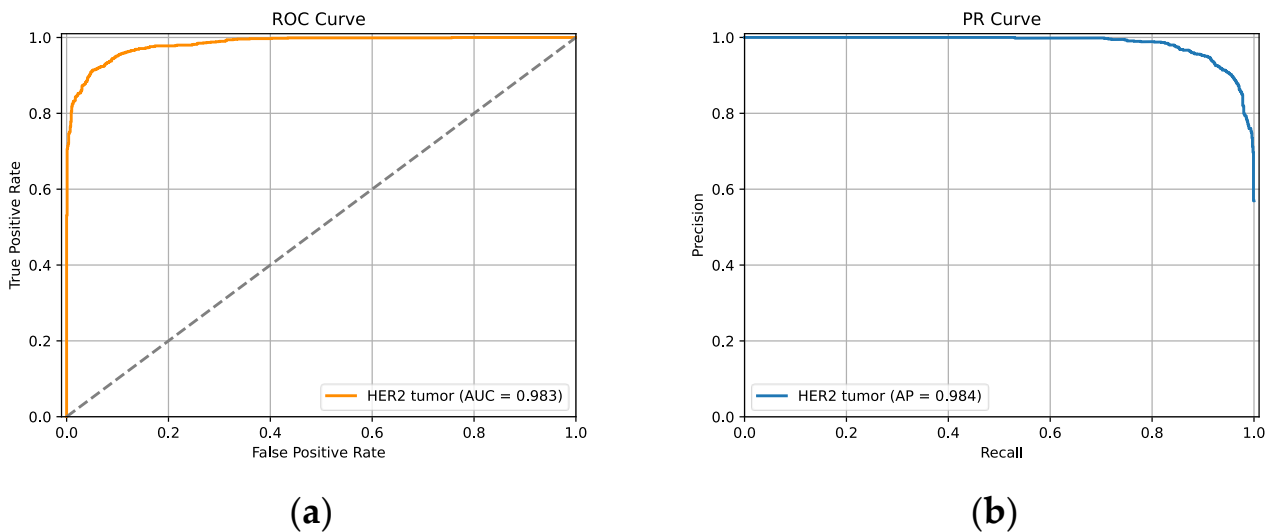


Figure 6. The ROC curve and PR curve give an estimation of the cell segmentation model's performance. (a) ROC curve. (b) PR curve.

3.2. IHC Slide Classification

In stage three, the tumor areas in each WSI need to be further processed. Firstly, downsampling is performed on tumor areas. The rate of downsampling needs to be calculated to extract non-overlapping patches of $512 * 512$ pixels. For each patch, the staining area and the number of cells is extracted. Figure 7 shows the extraction of staining regions of different IHC score. It can be observed that with the increase of IHC score, the color extracted from the staining area is darker. Therefore, the initial classification of pathological slides is constructed based on staining area by setting the proper threshold with this feature. It is also the main reference basis for IHC scoring.

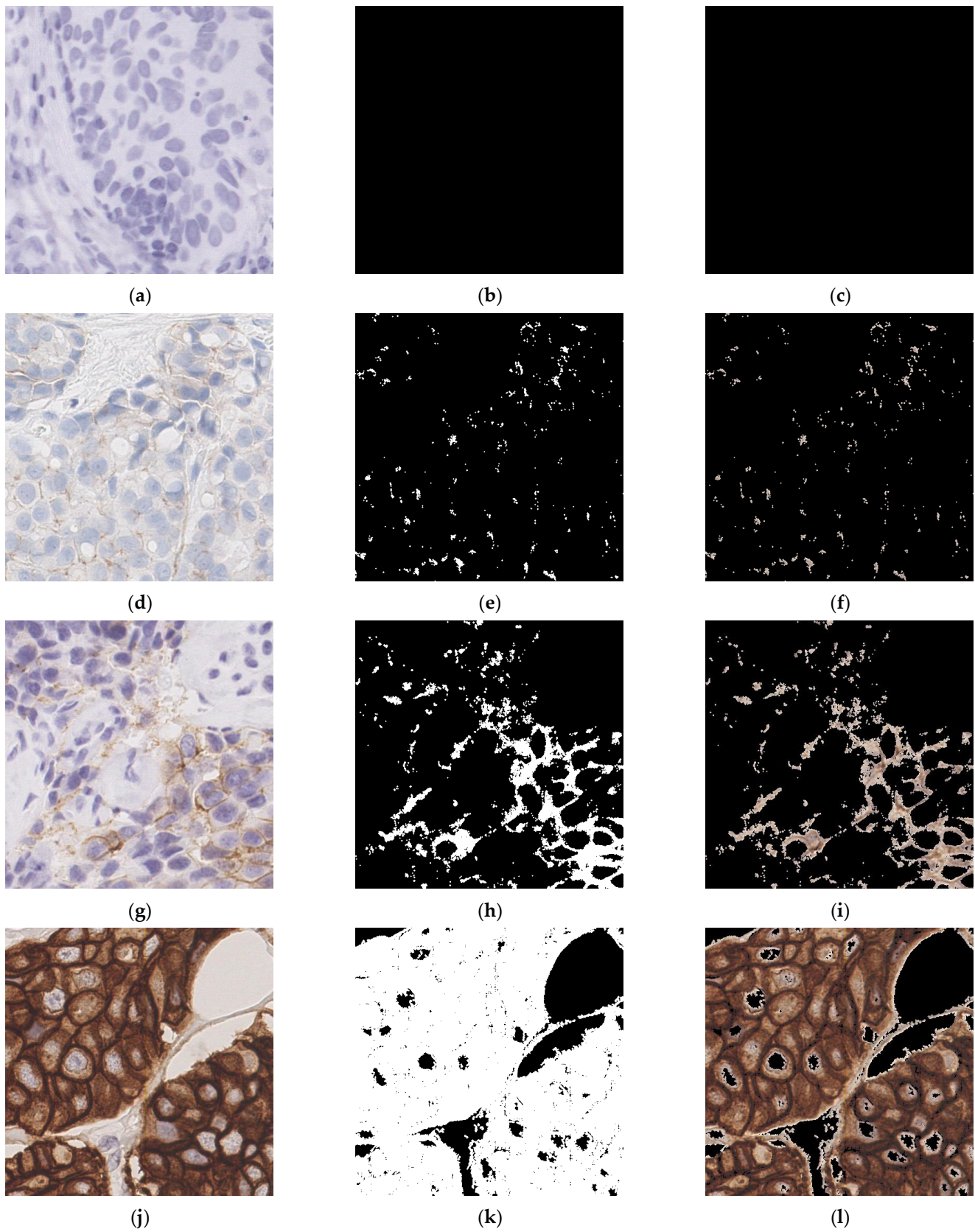


Figure 7. Visualization of staining region extraction. Each row includes the original patch (left), the binary image of extracted staining (middle), and the image of staining area with original image superimposed (right) for one IHC scoring. (a–c) IHC 0. (d–f) IHC 1+. (g–i) IHC 2+. (j–l) IHC 3+.

In this study, color deconvolution and watershed algorithm are adopted for cell extraction. For the cohesive and overlapping cells, the method based on the single-cell area is used for segmentation and statistics. As for the extraction of cells with circumferential positive membrane, a patch with IHC 3+ score is considered as an example. The schematic diagram is shown in Figure 8. The confusion matrices of slide-level IHC results which only used staining as the scoring basis and integrated the three scoring methods are illustrated in Figure 9a (accuracy: 87.4%) and Figure 9b (accuracy: 97.9%), respectively. In addition, we calculate the average time consumption of staining only method and integrated method to compute a single patch. The comparison of time cost between these two methods is shown in Figure 10. In general, the time cost of integrated method is nearly twice as much as staining only method. In addition, Table 4 shows the statistical calculation of evaluation indexes. For a single center dataset, the experimental thresholds can be set via this table.

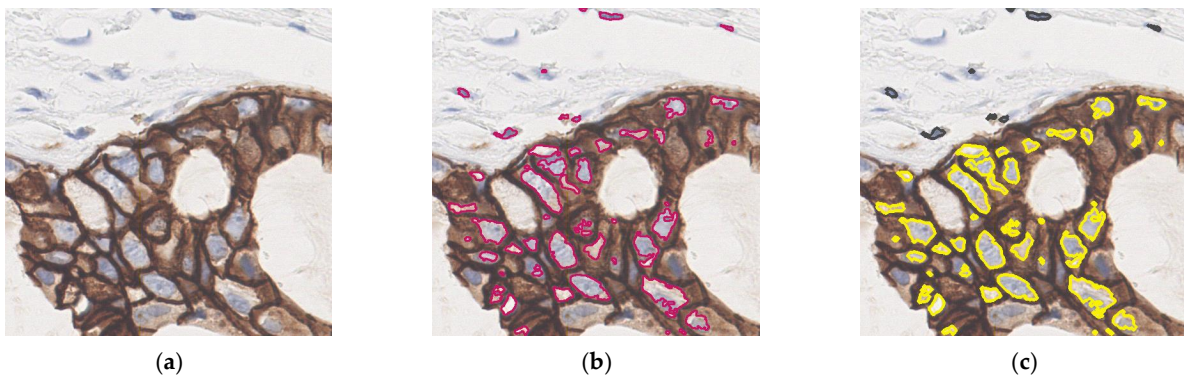


Figure 8. Visualization of cell extraction result. (a) Original patch. (b) Image with all extracted cells (red contours). (c) Divide the extracted cells to positive cells (yellow contours) and negative cells (dark blue contours).

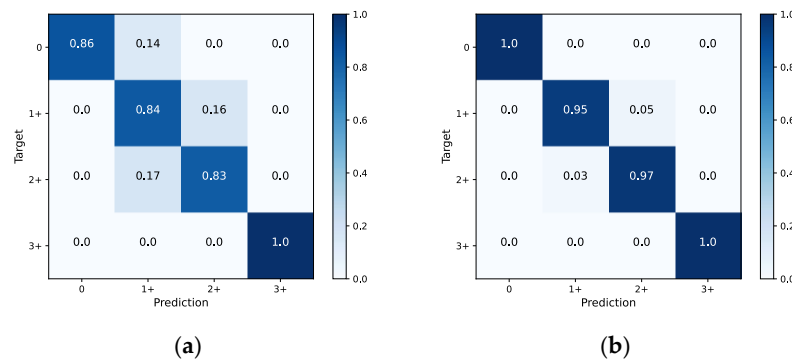


Figure 9. Confusion matrices of HER2 slide-level classification. (a) Using staining as the classification method only. (b) Using integrated methods.

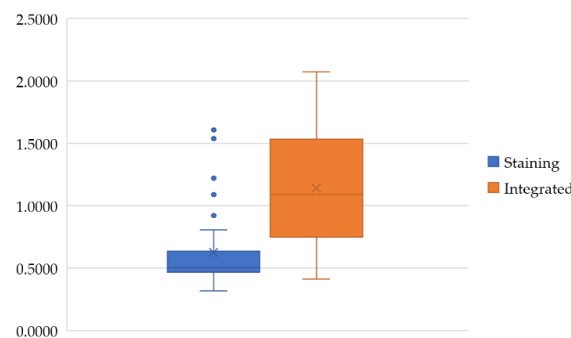


Figure 10. The time/cost comparison between staining method and integrated method for automated scoring.

Table 4. The statistical calculation of evaluation indexes.

IHC Score	0	1+	2+	3+
Staining average min	0.000	0.006	0.146	0.541
Staining average max	0.008	0.189	0.367	0.611
Positive cell ratio min	0.000	0.004	0.130	0.173
Positive cell ratio max	0.041	0.125	0.318	0.338
Circumferential membrane cell ratio min	0.000	0.011	0.454	0.461
Circumferential membrane cell ratio max	0.000	0.360	0.741	0.908

4. Discussion

This study proposed a deep-learning-based predictive framework to automatically evaluate the IHC score of HER2 WSIs in breast cancer. In total, 95 IHC section images with 23 labeled tumor areas were provided by Peking Union Medical College Hospital. The corresponding IHC scores of all HER2 slides were used in the study of this project. The predictive framework of this study is more interpretable compared with previous HER2 evaluation methods.

When a WSI is tested, the tumor area of this image is first predicted. Then, the evaluation index in tumor area is calculated and analyzed. As can be seen from the results in Table 1, the patch-level precision and f1 score of the proposed model are 95.77% and 83.09%, respectively, showing good performance in similar studies.

In stage three, three evaluation indexes of tumor patches were calculated. These calculated features are identified based on CSCO 2021 guidelines, which are referenced by pathologists when scoring IHC manually. The proposed method in stage three can extract stained areas, extract positive cells, and extract cells with complete positive cell membranes. It achieves high slide-level accuracy of HER2 score according to evaluation indexes. It is worth mentioning that the selection of the threshold is intuitive and important. Certain knowledge is needed to adjust a better threshold. Improper thresholds seriously affect the accuracy of the method. Figure 8 presents the results of cell extraction and positive cell labeling. These results are roughly consistent with the determination by manual observation, indicating the high feasibility of the prediction framework. In Table 5, we compare our method with other relevant methods. Most of these methods use manually labeled WSIs to train and test models. The proposed method can perform HER2 scoring based on WSI and obtain a high scoring accuracy.

Table 5. The comparison with related works.

	Dataset	Method	Remarks
Saha et al. [24]	752 labeled images cropped from 79 WSIs	Fully connected long short-term memory network, scoring by membrane and nuclei detection	98.33% accuracy
Vandenberghe et al. [27]	74 WSIs	Watershed segmentation, support vector machine, random forest, scoring by classifying cells	83% accordance
Kaiser et al. [26]	86 WSIs	Deep reinforcement learning, scoring by connectivity-based method	79.4% accuracy
Singh et al. [38]	1345 labeled areas from 52 WSIs	Neural network classifier, scoring by ROI-based method	91.1% accuracy
Caroline et al. [39]	2580 labeled images from 86 WSIs	K-nearest neighbor, multilayer perceptron, scoring by decision trees	90% accuracy
Khameneh et al. [41]	127 WSIs	Modified U-Net, scoring by WSI merging and membrane segmentation	94.82% segmentation and 87% classification accuracy
The proposed method	95 WSIs	ResNet, WSI segmentation, scoring by integrated calculation of staining intensity, circumferential membrane staining pattern, and proportion of positive cells	73.49% segmentation accuracy, 95.77% segmentation precision, 97.9% scoring accuracy

There are some limitations to our study. First, there may be inconsistencies in the depth of the color due to inconsistencies in the dose of dye added during the slicing process. Some deep learning methods use color standardization to better identify tumor areas. Generally, color normalization was added in the image preprocessing stage of model training and testing to make the result of classification more accurate. However, after the addition of this method, the time of execution is greatly increased, and the higher accuracy has a limited impact on the accuracy of the final IHC score. Secondly, although the processing method in stage three has a high dependence on the selection of color thresholds, this paper finds that the overall robustness of color extraction is stable after the conversion of color space. In addition, the influence of color depth on classification results can be limited by collecting HER2 slides from multiple centers.

In conclusion, this study conducted interpretable analysis and prediction of IHC scores of histological images of HER2 slides based on the deep learning method. It provides the direction for the clinical application of deep learning and promotes the development of precision therapy in the field of breast cancer.

Author Contributions: Conceptualization, Z.Z., F.R., and Y.C.; methodology, Y.C. and Z.Z.; formal analysis, Y.C.; investigation, Y.C.; resources, X.Z. and H.W.; data curation, X.Z. and H.W.; writing—original draft preparation, Y.C.; writing—review and editing, F.R. and H.W.; visualization, Y.C.; supervision, L.C. and Z.Z.; project administration, F.R. and Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by grants from the Informatization Plan of the Chinese Academy of Sciences (CAS-WX2021SF-0101), the National Key Research and Development Program of China (2021YFF1201005), and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA16021400).

Institutional Review Board Statement: The study was approved by the Ethics Review Committee of Peking Union Medical College Hospital (protocol code: S-K311).

Informed Consent Statement: Patient consent was waived due to the retrospective design of this study.

Data Availability Statement: The data are not available for public access because of patient privacy concerns but are available from the corresponding author on reasonable request.

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

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
2. The Cancer Genome Atlas Network. Comprehensive Molecular Portraits of Human Breast Tumours. *Nature* **2012**, *490*, 61–70. [[CrossRef](#)] [[PubMed](#)]
3. Caldarella, A.; Crocetti, E.; Bianchi, S.; Vezzosi, V.; Urso, C.; Biancalani, M.; Zappa, M. Female Breast Cancer Status According to ER, PR and HER2 Expression: A Population Based Analysis. *Pathol. Oncol. Res. POR* **2011**, *17*, 753–758. [[CrossRef](#)] [[PubMed](#)]
4. Wolff, A.C.; Hammond, M.E.H.; Hicks, D.G.; Dowsett, M.; Mcshane, L.M.; Allison, K.H.; Allred, D.C.; Bartlett, J.M.S.; Bilous, M.; Fitzgibbons, P. Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Update. *Arch. Pathol. Lab. Med.* **2013**, *31*, 3997. [[CrossRef](#)] [[PubMed](#)]
5. Hao, J.; Li, J. *Guidelines of Chinese Society of Clinical Oncology (CSCO) Breast Cancer 2021*, 1st ed.; People's Medical Publishing House: Beijing, China, 2021.
6. Razavi, S.; Hatipoglu, G.; Yalcin, H. Automatically Diagnosing HER2 Amplification Status for Breast Cancer Patients Using Large FISH Images. In Proceedings of the Signal Processing & Communications Applications Conference, Antalya, Turkey, 15–18 May 2017.
7. Qaiser, T.; Mukherjee, A.; Pb, C.R.; Munugoti, S.D.; Tallam, V.; Pitkaho, T.; Lehtimki, T.; Naughton, T.; Berseth, M.; Pedraza, A. Her2 Challenge Contest: A Detailed Assessment of Automated Her2 Scoring Algorithms in Whole Slide Images of Breast Cancer Tissues. *Histopathology* **2018**, *72*, 227–238. [[CrossRef](#)] [[PubMed](#)]
8. Lecun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436. [[CrossRef](#)]

9. Havaei, M.; Davy, A.; Warde-Farley, D.; Biard, A.; Courville, A.; Bengio, Y.; Pal, C.; Jodoin, P.M.; Larochelle, H. Brain Tumor Segmentation with Deep Neural Networks. *Med. Image Anal.* **2017**, *35*, 18–31. [[CrossRef](#)]
10. Teresa, A.; Guilherme, A.; Eduardo, C.; José, R.; Paulo, A.; Catarina, E.; António, P.; Aurélio, C.; Anna, S. Classification of Breast Cancer Histology Images Using Convolutional Neural Networks. *PLoS ONE* **2017**, *12*, e0177544.
11. Rakha, E.A.; Pinder, S.E.; Bartlett, J.; Ibrahim, M.; Ellis, I.O. Updated UK Recommendations for HER2 Assessment in Breast Cancer. *J. Clin. Pathol.* **2014**, *68*, 93–99. [[CrossRef](#)]
12. Akbar, S.; Jordan, L.B.; Purdie, C.A.; Thompson, A.M.; McKenna, S.J. Comparing Computer-Generated and Pathologist-Generated Tumour Segmentations for Immunohistochemical Scoring of Breast Tissue Microarrays. *Br. J. Cancer* **2015**, *113*, 1075–1080. [[CrossRef](#)]
13. Xu, Y.; Li, Y.; Liu, M.; Wang, Y.; Fan, Y.; Lai, M.; Chang, I.C. Gland Instance Segmentation by Deep Multichannel Neural Networks. *arXiv* **2016**, arXiv:1607.04889.
14. Sharma, H.; Zerbe, N.; Klempert, I.; Hellwich, O.; Hufnagl, P. Deep Convolutional Neural Networks for Automatic Classification of Gastric Carcinoma Using Whole Slide Images in Digital Histopathology. *Comput. Med. Imaging Graph.* **2017**, *61*, 2–13. [[CrossRef](#)]
15. Bardou, D.; Zhang, K.; Ahmad, S.M. Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks. *IEEE Access* **2018**, *6*, 24680–24693. [[CrossRef](#)]
16. Volynskaya, Z.; Evans, A.J.; Asa, S.L. Clinical Applications of Whole-Slide Imaging in Anatomic Pathology. *Adv. Anat. Pathol.* **2017**, *24*, 215–221. [[CrossRef](#)] [[PubMed](#)]
17. Pantanowitz, L.; Farahani, N.; Parwani, A. Whole Slide Imaging in Pathology: Advantages, Limitations, and Emerging Perspectives. *Pathol. Lab. Med. Int.* **2015**, *7*, 23–33. [[CrossRef](#)]
18. Wang, D.; Khosla, A.; Gargeya, R.; Irshad, H.; Beck, A.H. Deep Learning for Identifying Metastatic Breast Cancer. *arXiv* **2016**, arXiv:1606.05718.
19. 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.
20. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
21. Bejnordi, B.E.; Zuidhof, G.; Balkenhol, M.; Hermsen, M.; Bult, P.; Van Ginneken, B.; Karssemeijer, N.; Litjens, G.; van der Laak, J. Context-Aware Stacked Convolutional Neural Networks for Classification of Breast Carcinomas in Whole-Slide Histopathology Images. *J. Med. Imaging* **2017**, *4*, 044504. [[CrossRef](#)] [[PubMed](#)]
22. Litjens, G.; Sánchez, C.I.; Timofeeva, N.; Hermsen, M.; Nagtegaal, I.; Kovacs, I.; Christina, H.; Bult, P.; Ginneken, B.V.; Jeroen, V. Deep Learning as a Tool for Increased Accuracy and Efficiency of Histopathological Diagnosis. *Sci. Rep.* **2016**, *6*, 26286. [[CrossRef](#)]
23. Wu, Z.; Wang, L.; Li, C.; Cai, Y.; Liu, Y. DeepLRHE: A Deep Convolutional Neural Network Framework to Evaluate the Risk of Lung Cancer Recurrence and Metastasis from Histopathology Images. *Front. Genet.* **2020**, *11*, 768. [[CrossRef](#)]
24. Saha, M.; Chakraborty, C. Her2Net: A Deep Framework for Semantic Segmentation and Classification of Cell Membranes and Nuclei in Breast Cancer Evaluation. *IEEE Trans. Image Process.* **2018**, *27*, 2189–2200. [[CrossRef](#)] [[PubMed](#)]
25. Yan, X.; Jia, Z.; Wang, L.B.; Ai, Y.; Fang, Z.; Lai, M.; Chang, I.C. Large Scale Tissue Histopathology Image Classification, Segmentation, and Visualization via Deep Convolutional Activation Features. *BMC Bioinform.* **2017**, *18*, 281.
26. Qaiser, T.; Rajpoot, N.M. Learning Where to See: A Novel Attention Model for Automated Immunohistochemical Scoring. *IEEE Trans. Med. Imaging* **2019**, *38*, 2620–2631. [[CrossRef](#)] [[PubMed](#)]
27. Vandenbergh, M.E.; Scott, M.L.J.; Scorer, P.W.; Sderberg, M.; Barker, C. Relevance of Deep Learning to Facilitate the Diagnosis of HER2 Status in Breast Cancer OPEN. *Sci. Rep.* **2017**, *7*, 45938. [[CrossRef](#)]
28. Brüggemann, A.; Eld, M.; Lelkaitis, G.; Nielsen, S.; Grunkin, M.; Hansen, J.D.; Foged, N.T.; Vyberg, M. Digital Image Analysis of Membrane Connectivity Is a Robust Measure of HER2 Immunostains. *Breast Cancer Res. Treat.* **2012**, *132*, 41–49. [[CrossRef](#)] [[PubMed](#)]
29. Hou, L.; Samaras, D.; Kurc, T.M.; Gao, Y.; Davis, J.E.; Saltz, J.H. Patch-Based Convolutional Neural Network for Whole Slide Tissue Image Classification. In Proceedings of the Computer Vision & Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
30. Mukundan, R. A Robust Algorithm for Automated HER2 Scoring in Breast Cancer Histology Slides Using Characteristic Curves. In *Annual Conference on Medical Image Understanding and Analysis*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 386–397.
31. Masmoudi, H.; Hewitt, S.M.; Petrick, N.; Myers, K.J.; Gavrielides, M.A. Automated Quantitative Assessment of HER-2/Neu Immunohistochemical Expression in Breast Cancer. *IEEE Trans. Med. Imaging* **2009**, *28*, 916–925. [[CrossRef](#)]
32. Muhimmah, I.; Heksaputra, D.; Indrayanti, Ma’Mun, S.; Tamura, H.; Purnomo, M.R.A. Color Feature Extraction of HER2 Score 2+ Overexpression on Breast Cancer Using Image Processing. In *MATEC Web of Conferences*; EDP Sciences: Les Ulis, France, 2018; p. 03016.
33. Mukundan, R. Image Features Based on Characteristic Curves and Local Binary Patterns for Automated HER2 Scoring. *J. Imaging* **2018**, *4*, 35. [[CrossRef](#)]
34. Zheng, Y.; Jiang, Z.; Zhang, H.; Xie, F.; Shi, J.; Xue, C. Adaptive Color Deconvolution for Histological WSI Normalization. *Comput. Methods Programs Biomed.* **2019**, *170*, 107–120. [[CrossRef](#)]
35. Geijs, D.; Intezar, M.; Litjens, G.; Laak, J. Automatic Color Unmixing of IHC Stained Whole Slide Images. In Proceedings of the Digital Pathology, Houston, TX, USA, 11–12 February 2018.

36. Choi, S.; Chu, J.; Kim, B.; Sang, Y.H.; Kim, S.T.; Lee, J.; Kang, W.K.; Han, H.; Sohn, I.; Kim, K.M. Tumor Heterogeneity Index to Detect Human Epidermal Growth Factor Receptor 2 Amplification by Next-Generation Sequencing. *J. Mol. Diagn.* **2019**, *21*, 612–622. [[CrossRef](#)]
37. Yang, J.; Ju, J.; Guo, L.; Ji, B.; Shi, S.; Yang, Z.; Gao, S.; Yuan, X.; Tian, G.; Liang, Y.; et al. Prediction of HER2-Positive Breast Cancer Recurrence and Metastasis Risk from Histopathological Images and Clinical Information via Multimodal Deep Learning. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 333–342. [[CrossRef](#)]
38. Singh, P.; Mukundan, R. A Robust HER2 Neural Network Classification Algorithm Using Biomarker-Specific Feature Descriptors. In Proceedings of the 2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP), Vancouver, BC, Canada, 29–31 August 2018; IEEE: Vancouver, BC, Canada, 2018; pp. 1–5.
39. Cordeiro, C.Q.; Ioshii, S.O.; Alves, J.H.; Oliveira, L.F.D. An Automatic Patch-Based Approach for HER-2 Scoring in Immunohistochemical Breast Cancer Images Using Color Features. *arXiv* **2018**, arXiv:1805.05392.
40. Choudhury, K.R.; Yagle, K.J.; Swanson, P.E.; Krohn, K.A.; Rajendran, J.G. A Robust Automated Measure of Average Antibody Staining in Immunohistochemistry Images. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* **2010**, *58*, 95–107. [[CrossRef](#)] [[PubMed](#)]
41. Khameneh, F.D.; Razavi, S.; Kamasak, M. Automated Segmentation of Cell Membranes to Evaluate HER2 Status in Whole Slide Images Using a Modified Deep Learning Network. *Comput. Biol. Med.* **2019**, *110*, 164–174. [[CrossRef](#)] [[PubMed](#)]

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.