IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 31 May 2023, accepted 11 June 2023, date of publication 14 June 2023, date of current version 21 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3286309

RESEARCH ARTICLE

Prediction of Microsatellite Instability From Gastric Histological Images Based on Residual Attention Networks With Non-Local Modules

SUNG-NIEN YU[1,2], (Member, IEEE), SHIH-CHIANG HUANG[3,4], WEI-CHEN WANG[1], YU PING CHANG[1], KUANG-HUA CHEN[3], AND TSE-CHING CHEN[3]

[1]Department of Electrical Engineering, National Chung Cheng University, Chiayi, Minxiong 621301, Taiwan
[2]Center for Innovative Research on Aging Society (CIRAS), National Chung Cheng University, Chiayi, Minxiong 621301, Taiwan
[3]Department of Anatomic Pathology, Linkou Chang Gung Memorial Hospital, College of Medicine, Chang Gung University, Taoyuan City 333423, Taiwan
[4]Graduate Institute of Clinical Medical Sciences, College of Medicine, Chang Gung University, Taoyuan City 333423, Taiwan

Corresponding author: Sung-Nien Yu (ieesny@ccu.edu.tw)

This work was supported in part by the Center for Innovative Research on Aging Society (CIRAS) from the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE), Taiwan; in part by the Ministry of Science and Technology (MOST), Taiwan, under Contract MOST 109–2320-B-182A-010-MY3 and Contract MOST 111-2221-E-194-057; in part by the Ministry of Health and Welfare (MOHW), Taiwan, under Contract MOHW 112-TDU-B-222-124011; and in part by the Chang Gung Memorial Hospital under Contract CMRPG5L0142.

ABSTRACT Gastric cancer can be classified into different subtypes according to their genetic expression. Microsatellite instability (MSI) is one of these subtypes and an important clinical marker for prognosis and consideration for immunotherapy. Since genetic testing is relatively expensive and laborious, this study tackles the challenge of using deep neural networks (DNNs) to identify MSI based on analyzing histomorphologic features of gastric whole-slide images (WSIs). A two-stage patch-wise framework was proposed, which first differentiates the tumor regions from normal, then predicts MSI status from the tumorous patches. The proposed deep learning architecture enhances the residual attention network with non-local modules and visual context fusion modules, thereby allowing both local fine-grained details and coarse long-range dependencies to be captured. Image post-processing procedures were also proposed to better align the region segmentation with pathologist annotations. The model was applied to a three-way classification task, namely normal tissue, microsatellite stable (MSS), and MSI, using a private dataset gathered by Chang Gung Memorial Hospital and achieved 91.95% slide-wise accuracy. We also studied the feasibility of transfer learning by fine tuning on the TCGA-STAD public dataset, where we attained a high accuracy of 96.53% and an AUC of 0.99, outperforming previous literature.

INDEX TERMS Image post-processing, microsatellite instability, non-local neural networks, residual attention network, whole slide image.

I. INTRODUCTION

Gastric cancer is the fifth most common type of cancer and the third leading cause of cancer-related deaths worldwide according to Global Cancer Observatory (GLOBOCAN) 2020 estimates [1]. There were 1.09 million new cases (5.6% of all cancers) and 0.77 million new cancer-related deaths (7.7% of all cancers) due to gastric cancer in 2020. Gastric cancer exhibits significant heterogeneity in histological appearance and molecular signatures [2], [3]. The Cancer Genome Atlas (TCGA) classifies gastric cancer into four molecular subtypes: Epstein-Barr virus (EBV)-positive, microsatellite instability (MSI), genomically stable (GS), and chromosomal instability (CIN) [4]. The Asian Cancer Research Group (ACRG) uses expression profiling to divide gastric cancers into four subtypes: MSI, MSS tumor protein p53-inactive (MSS/TP53-), MSS tumor protein p53-active (MSS/TP53+), and MSS epithelial-mesenchymal

The associate editor coordinating the review of this manuscript and approving it for publication was Carmelo Militello.

transition (MSS/EMT) [5]. Microsatellites, also known as short tandem repeats (STRs) or short sequence repeats (SSRs), are strings of DNA motifs (1-10 base pairs) that are often repeated tens or hundreds of times. Microsatellite instability (MSI) is a condition in which these repeating sequences in the DNA of cells have lengths different from normal and are known to be caused by impaired mismatch repair (MMR) proteins, namely, MLH1, MSH2, MSH6, and PMS2. During the DNA replication and repair process, DNA polymerase slippage occurs such that DNA motifs are repeatedly inserted or deleted, leading to varying lengths in the microsatellite [6]. These genomic mutations, known as replication errors (RERs), reflect genetic instability and are an important pathway of carcinogenesis [7]. Nowadays, MSI has become an important clinical biomarker for the prognosis of cancer patients and their response to immunotherapy. Colorectal and gastric cancer patients with positive MSI condition have been shown in literature to be suitable for immune checkpoint inhibitors [8], [9].

However, the detection of MSI requires expensive and complex genetic testing or alternative methods such as immunohistochemistry of MMR proteins. Interestingly, MSI-associated colorectal or gastric cancer may exhibit certain distinctive medullary-like or solid histomorphologic features, including poor differentiation, expansive growth patterns, and abundant inflammatory stroma [10], [11]. These unique histological appearances of MSI-associated cancers imply the potential applicability of diagnosing MSI directly from histological slides by pathologists or artificial intelligence.

In 1966, Prewitt et al. [12] proposed methods for measuring cells from histological images and methods for object detection, feature extraction, and classification functions, which are still used in modern times. In the early days, cytology and pathological histology could only be observed in narrow fields of view using conventional microscopes, and collection of these image slices was a time-consuming and laborious task [13]. With the advent of digital pathology scanners, tissue slices can be digitized, paving the way for deep learning research using digital pathology slices [14], [15].

However, the analysis of these digital whole-slide images (WSIs) presents enormous challenges. As WSIs are pyramid images with extremely high resolution with tens of billions of pixels, they cannot be supplied directly to models for training because of hardware limitations. The most common solution is to cut WSIs into many small patches to meet hardware limitations. However, when WSIs are broken down into patches, the spatial relations between the features within different patches of the same WSI are lost.

To address this problem, researchers such as Li et al. [16] employed dilated convolutions of different receptive fields to retain as much spatial information as possible. Kosaraju et al. [17] combined features extracted from two different WSI resolutions to obtain both local and global information. Zhu et al. [18] proposed feature balanced module which combined channel attention and spatial attention modules to extract information from the different dimensions. In this study, we propose using overlapping training patches and non-local network module to retain features between patches.

Specific to the task of identifying MSI status in gastric tissue slices, various model architectures have been explored in the literature. For example, Kather et al. [19] and Su et al. [20] employed ResNet18, Muti et al. [21] employed ShuffleNet, and Hinata et al. [22] compared VGG16, VGG19, ResNet50, and EfficientNetB0 and found that VGG16 gave the best result. Saldanha et al. [23] employed clustering-guided contrastive learning (RetCCL) [24] as the model architecture and further explored distributed learning in the form of swarm learning to merge models trained on decentralized datasets. These researches had employed different datasets, such as TCGA-STAD [4], BERN [25], LEEDS [26], and TUM [27], with TCGA-STAD being the most common. The results were mostly presented in terms of area under the curve of the receiver operating characteristic curve (AUC), achieving values ranging from 0.80 to 0.87.

To further improve the classification result for MSI detection in gastric tissue slices, we proposed a two-stage classification framework based on residual attention network enhanced with non-local network and visual context fusion modules to capture both fine-grained details and coarse spatial information within and between patches. The first stage differentiates tumorous patches from normal patches. Then, the tumorous patches are fed to stage two network to identify their MSS or MSI status. Image post-processing procedures are also proposed to better align predictions with pathologist annotations. Various experiments using a private dataset gathered by Chang Gung Memorial Hospital (CGMH) and TCGA-STAD public dataset were designed to test the efficacy of the proposed framework.

The contributions of this study can be summarized as follows:

1. We proposed a two-stage MSI classification framework employing residual attention network, non-local network, and visual context fusion modules to capture information across patches. The framework achieved 96.53% overall accuracy and an AUC of 0.99 for MSI detection on gastric tissue slices using the public TCGA-STAD dataset, outperforming current literature.
2. We demonstrated that transfer learning by pretraining using our private CGMH dataset and then fine-tuning on the public TCGA-STAD dataset improved overall accuracy. The proposed post-processing improved tumor localization IoU.
3. By visualizing the model's attentive regions using Score-CAM, we showed that the regions aligned with clinical expectations of how MSI status could be identified.

The rest of this paper is organized as follows: In Section II, we explain the methodology of the proposed framework. Section III details the experimental procedure and discusses

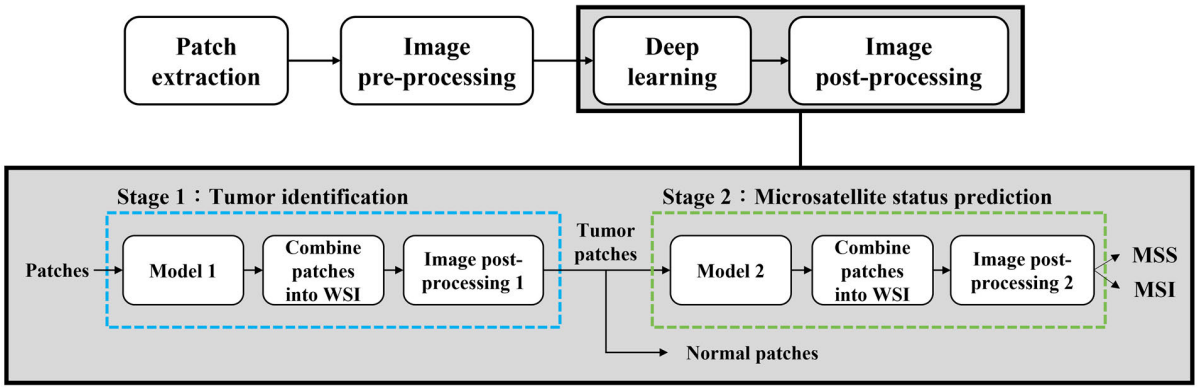


FIGURE 1. MSI detection framework.

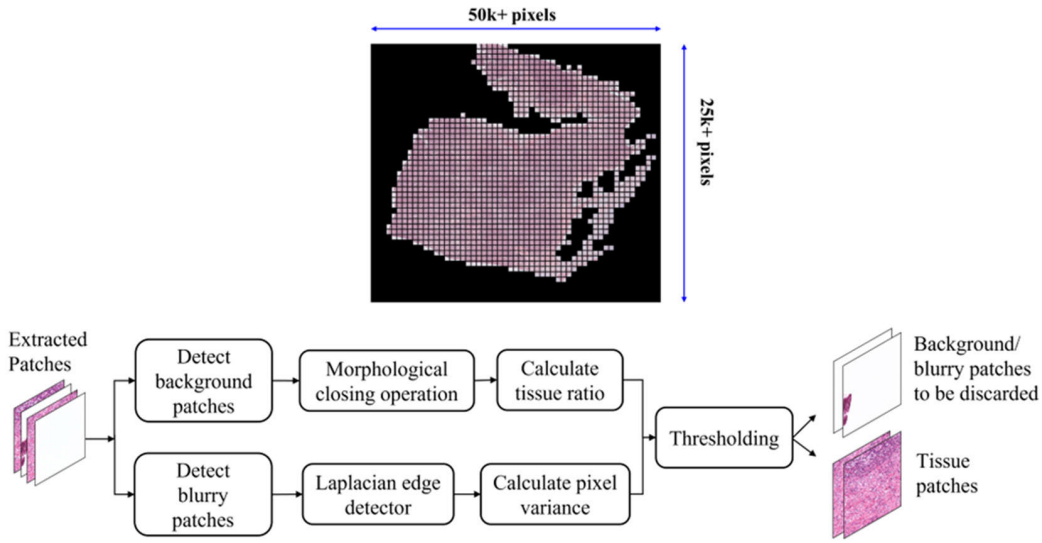


FIGURE 2. Patch extraction procedure.

the results. Comparisons with other relevant studies are presented in Section IV, and Section V concludes the paper.

II. METHODOLOGY

The proposed MSI detection framework includes four phases: patch extraction, image pre-processing, deep learning, and image post-processing. The deep learning and post-processing phases were further divided into two stages: tumor identification and microsatellite status prediction, as shown in Fig. 1.

In stage 1 of the deep learning phase, model 1 was used to classify the patches as normal or tumorous tissues. These classified patches were then combined into the WSIs for post-processing. A threshold was set to determine whether the WSI should be considered healthy or tumorous. The tumor patches from tumorous WSIs were then fed into stage 2 model for binary classification into MSS or MSI. The classified patches were again combined into WSIs for post-processing and finally output as MSI detection results.

We describe in detail each step of the framework in the following subsections.

A. PATCH EXTRACTION

The WSIs used in this study had resolutions of approximately 200k $\times$ 100k pixels, with sizes ranging from 2 to 5 GB. Limited by hardware capacity, it was not possible to directly feed these original-resolution WSIs into the deep learning network for training. Therefore, we downsampled the $\times 10$ magnification WSIs by a factor of four and then cut the resulting image into patches of 512 $\times$ 512 pixels in a sliding window fashion. For the training set, these patches were extracted with a 50% overlap. This not only increases the number of samples available for training but also retains more spatial information between different patches [28]. As for the test set, non-overlapping patches were extracted so that the patches can be pieced back into the original whole slides to evaluate slide-wise metrics.

Next, we employ traditional morphological closing operation and Laplacian edge detector [29] to eliminate background and blurry patches, as outlined in Fig. 2. The closing operation involves expansion then erosion to fill the smaller voids and connect narrow gaps often seen in adipose tissue and mucus. These tissues appear sparse on the whole slide

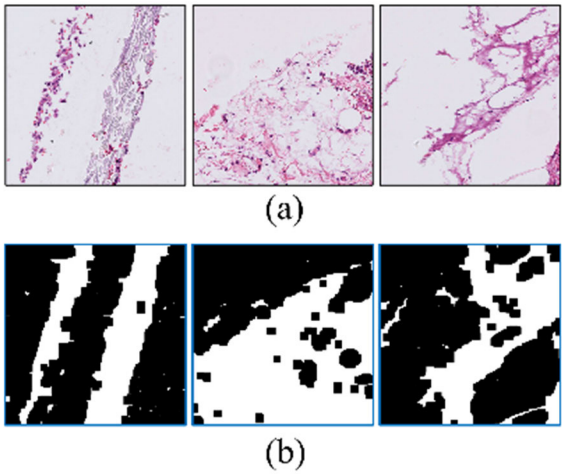


FIGURE 3. (a) Sparse tissue, (b) resulting mask after closing operation.

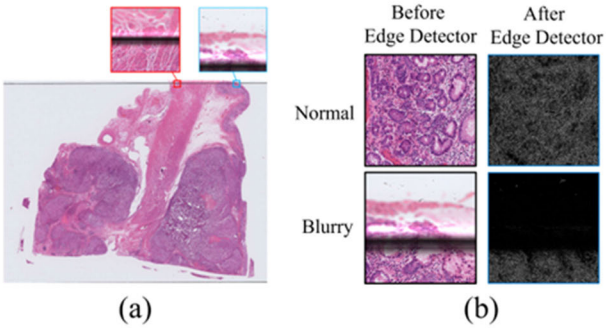


FIGURE 4. (a) Blurry patches found on the edges of glass slides. (b) The patch on the top is normal, whereas the patch below is affected by blurry artifacts. The same patches after the Laplacian edge detection are shown on the right. The affected patches exhibited a lower pixel variance.

and may be mistaken for a background, as shown in Fig. 3(a). The closing operation enhanced their presence, resulting in the masking images shown in Fig. 3(b). Patches with ratio of tissue to patch area lower than 0.2 were treated as background and discarded.

Some patches may also be affected by the edges of the glass slides, resulting in a blurry line that laid across the images, as seen in Fig. 4(a). To prevent these flawed patches from affecting the model, we employed the Laplacian edge detector to the images, as shown in Fig. 4(c). The blurry regions of the patches have much fewer edges detected and thus have a lower variance in the pixel value overall. We set the variance threshold to 500 and patches with variances lower than the threshold were considered blurry and discarded.

B. IMAGE PREPROCESSING

In traditional pathological examinations, tissues removed from the human body are chemically stained to identify the different parts of the tissues. In this study, samples were stained with hematoxylin and eosin (H&E) [30]. Hematoxylin stains basophilic structures (e.g., nuclei) blue, and eosin stains eosinophilic structures (e.g., intracellular and extracellular proteins, and connective tissue) to different shades of pink [31]. However, owing to differences in scanning instruments and dye manufacturers, the color distribution of stained tissues can vary. Color normalization was performed to reduce the differences, as shown in Fig. 5.

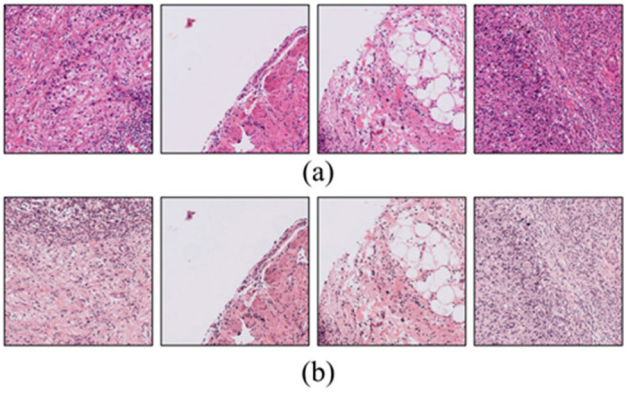


FIGURE 5. (a) The original images, (b) color-normalized images.

We employed the color normalization method proposed by Macenko et al. [32], which converts RGB values in tissue samples into optical density (OD) space. Each pixel is assumed to be a linear combination of two staining vectors corresponding to the two dye colors. The staining vectors for each sample image can then be determined by finding the edges of the pixel distribution in the OD space.

C. DEEP LEARNING MODEL

One difference between medical images and natural images is that lesions in medical images often have both fine-grained details and coarse complex structures. To extract both fine and coarse features from the WSI patches, we proposed a double-branched architecture which includes a residual attention network branch that extracts fine local features and a non-local network branch to obtain coarse structural information across different parts of the image.

1) RESIDUAL ATTENTION NETWORK

The architecture proposed in this study is based on the residual attention network proposed by Wang et al. [33], which has achieved state-of-the-art results in many object recognition benchmarks. The residual attention network reformulated the attention mechanism in a residual and feedforward fashion so that a very deep residual attention network could be easily trained end-to-end. This allows for very fine features to be learned.

The residual attention network module has two branches: (1) a soft mask branch and (2) a trunk branch. The trunk branch receives the input $\boldsymbol{x}$ and extracts feature output $T(\boldsymbol{x})$ via convolution. The soft mask branch uses a concept similar to the encoder-decoder used in image segmentation tasks to form top-down and bottom-up structures. In the encoder section, multiple convolution and pooling layers reduced the image resolution and increased the receptive

field size, thereby extracting high-level features. These high-level features represent the features in the sample images that the model should pay attention to. Subsequently, the decoder section upsamples these high-level feature maps through deconvolution layers with bilinear interpolation to the original image size. This is followed by two 1×1 convolution layers, a sigmoid activation function, and finally outputted as $M(\boldsymbol{x})$.

This soft mask output $M(\boldsymbol{x})$ is combined with the output of trunk branch $T(\boldsymbol{x})$ in a residual fashion, where an identity mapping is added so that feature selection by the soft mask branch either improves the result, or is not worse than without attention, as shown in (1).

$$H_{i,c}(\boldsymbol{x}) = (1 + M_{i,c}(\boldsymbol{x})) \cdot T_{i,c}(\boldsymbol{x}) \tag{1}$$

2) NON-LOCAL NEURAL NETWORKS

To further enhance the residual attention network in identifying long-range dependencies, we incorporated the non-local network proposed by Wang et al. [34]. The non-local network originates from the non-local mean filter [35], which performs denoising by first calculating the similarity among pixel points of the image. Pixels with a higher similarity were assigned higher weights. Then, a weighted average of all pixels was calculated, leading to a denoised image that retained important sharp features.

In a conventional CNN model, convolutional layers are used to extract important features from the input image. However, the size of the receptive field is restricted by the convolution-kernel size. By employing the concept of non-local mean filter, the non-local network takes the full input image as the receptive field and is weighted by its similarity with the current pixel. The basic formulation is shown in (2):

$$\boldsymbol{y}_i = \frac{1}{C(\boldsymbol{x})} \sum_{\forall j} f\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) g\left(\boldsymbol{x}_j\right) \tag{2}$$

where i is the current output pixel index, j is the index of all pixels, $\boldsymbol{x}$ is input and $\boldsymbol{y}$ the output. The dimensions of $\boldsymbol{x}$ and $\boldsymbol{y}$ are identical. $f(\cdot)$ calculates similarity between pixels i and j. $g(\cdot)$ calculates the feature value at pixel location j. Finally, the output is normalized by $C(\boldsymbol{x})$.

In [34], $g(\cdot)$ was defined as a linear embedding as formulated in (3).

$$g\left(\boldsymbol{x}_j\right) = \boldsymbol{W}_g \boldsymbol{x}_j \tag{3}$$

where $\boldsymbol{w}_g$ was the weight matrix to be learned through 1×1 convolution. The similarity function $f(\cdot)$ had four different formulations: Gaussian, embedded Gaussian, dot product, and concatenation. It was shown in the paper that the four formulations had similar results. For our architecture, we selected the embedded Gaussian formulation, as shown in (4):

$$f\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = e^{\theta(\boldsymbol{x}_i)^T \phi(\boldsymbol{x}_j)} \tag{4}$$

where $\theta(\boldsymbol{x}_i)$, $\phi\left(\boldsymbol{x}_j\right)$, and $C(\boldsymbol{x})$ are formulated respectively as follows:

$$\theta(\boldsymbol{x}_i) = \boldsymbol{W}_\theta \boldsymbol{x}_i \tag{5}$$

$$\phi\left(\boldsymbol{x}_j\right) = \boldsymbol{W}_\phi \boldsymbol{x}_j \tag{6}$$

$$C(\boldsymbol{x}) = \sum_{\forall j} f\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) \tag{7}$$

Substituting (4) and (7) into (2), we obtain (8), which is equivalent to calculating the softmax with respect to j, as shown in (9).

$$\boldsymbol{y}_i = \sum_{\forall j} \frac{e^{\theta(\boldsymbol{x}_i)^T \phi(\boldsymbol{x}_j)}}{\sum_{\forall j} e^{\theta(\boldsymbol{x}_i)^T \phi(\boldsymbol{x}_j)}} g(\boldsymbol{x}_j) \tag{8}$$

$$\boldsymbol{y} = Softmax(\boldsymbol{x}^T \boldsymbol{W}_\theta^T \boldsymbol{W}_\phi \boldsymbol{x}) g(\boldsymbol{x}) \tag{9}$$

To allow (2) to be applied to any model, the author added the concept of shortcut connection and redefined the equation as a non-local block, as shown in (10).

$$z_i = \boldsymbol{W}_z \boldsymbol{y}_i + \boldsymbol{x}_i \tag{10}$$

By setting the initial value of $\boldsymbol{W}_z$ to zero and let $\boldsymbol{x}_i$ be the identity mapping, this non-local block can be inserted into any pretrained model.

Inspired by Wen et al. [36], we attached the non-local visual context fusion module after the non-local neural network. This module applies three different operations to the input feature map, namely the average pooling, maximum pooling, and bilinear interpolation. These extract different perspectives of the input while down-sampling the feature map to the same size as the output from the residual attention network branch so that outputs from the two branches can be merged.

3) FULL ARCHITECTURE

The full deep learning architecture proposed in this study is shown in Fig. 6.

D. IMAGE POST-PROCESSING

As shown previously in Fig. 1. The MSI detection framework comprises two stages. The first stage differentiates between normal and tumor WSIs. The second stage differentiates tumor patches in tumorous WSIs into MSI or MSS categories.

After each of the two deep learning stages, we propose the application of image post-processing. In the first stage where we locate tumor regions we observed that, while the WSI images were annotated by professional laboratory physicians, to reduce the labeling time, physicians may choose to label the overall lesion area rather than selecting one by one the smaller individual regions at a higher magnification. As a result, the annotated tumor region may contain some normal tissue, as shown in Fig. 7.

Therefore, when the outputs of stage one deep learning model were pieced back into the WSIs, the predicted tumorous regions could often be scattered with bits of predicted normal patches. To better align the output result with the physician's annotation, we applied the connected-component

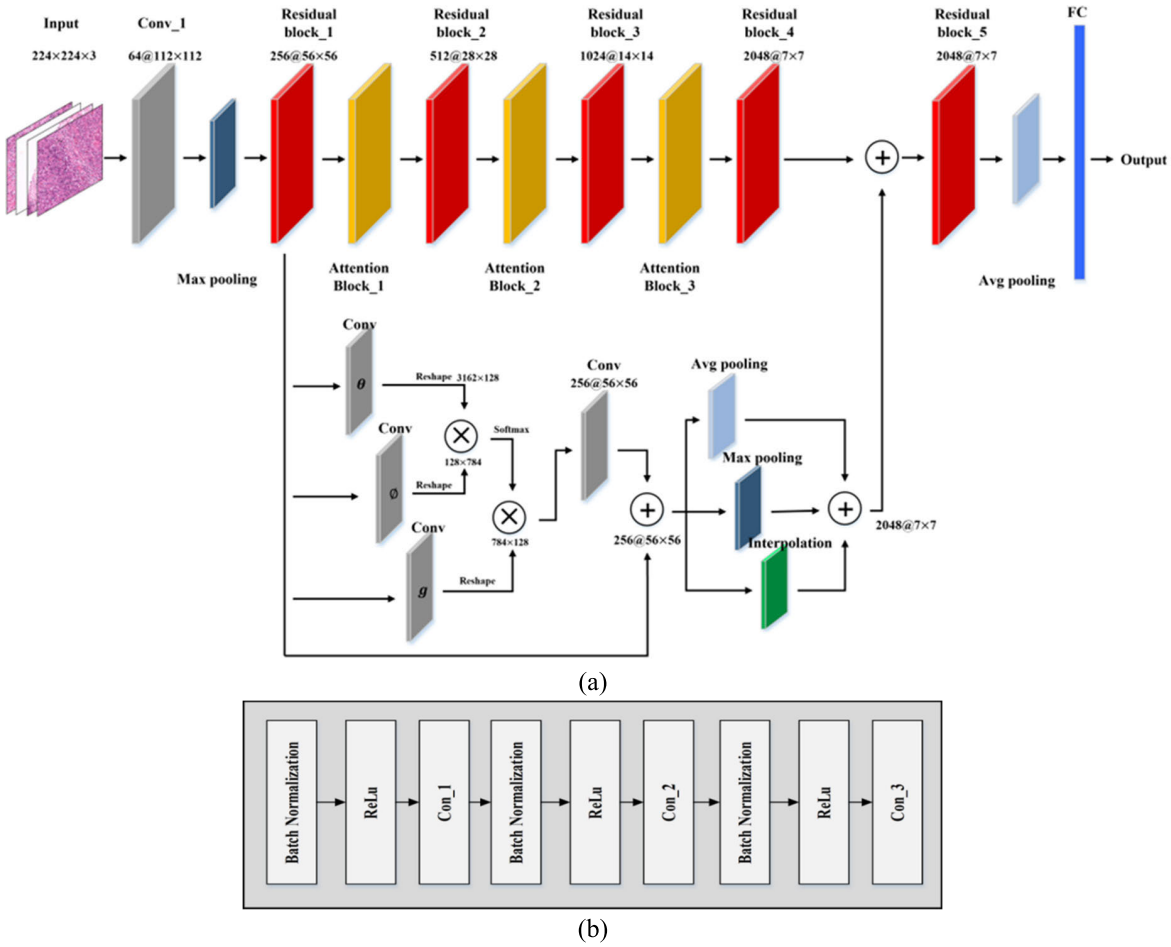


FIGURE 6. (a) Full deep learning architecture proposed in this study. (b) The residual block.

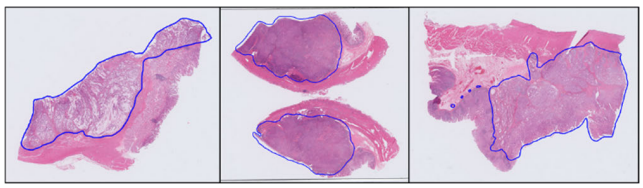


FIGURE 7. WSI annotation tends to be broad strokes.

labeling algorithm (CCL) [37] to eliminate the sporadic noises. Tumor regions formed by two or fewer patches were considered noise and removed. The holes surrounded by tumor tissues were then filled using flood-fill algorithm [38], [39], as shown in Fig. 8.

In the second stage of deep learning where we differentiate between MSI and MSS, we leverage the fact that MSI and MSS categories are mutually exclusive. An output WSI image may contain both MSI and MSS features. Depending on which features are more dominant, the whole WSI image would be considered either an MSI or MSS sample, and all tumorous patches within that WSI image would carry the solely MSI or MSS annotation.

Therefore, image post-processing was applied after the second deep learning stage. After the predicted outputs were pieced back into the WSIs, the numbers of MSI and MSS patches within the same WSI were calculated. As shown in Fig. 9, the pink, orange, and brown regions were predicted normal, MSI, and MSS patches, respectively. The number of orange MSI patches was much higher than that of brown MSS patches. Therefore, the whole WSI sample would be considered an MSI sample, and the MSI or MSS labels of all tumorous patches within that sample would be changed to MSI (orange), and vice versa.

III. EXPERIMENT AND RESULTS

The experiment was split into the tumor identification stage and microsatellite status recognition stage. In both stages, the results are presented in terms of patch-wise and whole-slide metrics. In addition, ablation comparison among four models, including baseline models ResNet18 and residual attention network (RA), and models proposed in this study, namely the residual attention with non-local module (RA-NL) and the addition of post-processing (RA-NL-PP), will be discussed.

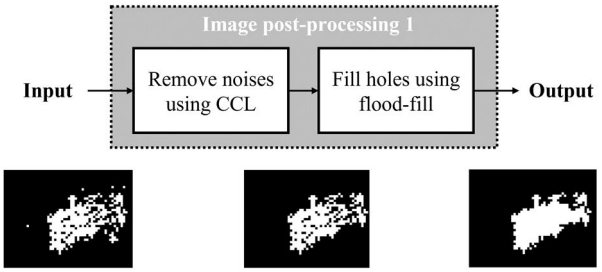


FIGURE 8. Stage 1 post-processing procedures.

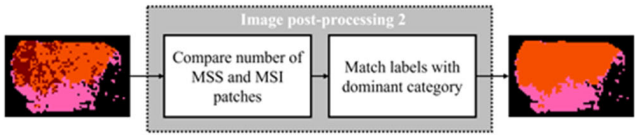


FIGURE 9. Stage 2 post-processing procedures.

TABLE 1. Number of CGMH WSI samples.

	Total	Train	Test
Normal	193	155	38
MSS	152	123	29
MSI	103	83	20
Total	448	361	87

A. DATASETS

Two datasets were employed in this study. The first was a private dataset gathered by Chang Gung Memorial Hospital (CGMH) and was referred to as the CGMH dataset. The CGMH dataset was paraffin-embedded gastric biopsy provided by CGMH and annotated by Dr. SC Huang and KH Chen from the Department of Anatomic Pathology at CGMH. All slice images were stained using the standard Hematoxylin-Eosin staining technique. A total of 448 WSIs with 193 normal, 152 MSS, and 103 MSI tissue sections were included. An 8:2 training to validation ratio was used, as listed in Table 1.

The second is the public TCGA-STAD dataset which contains 78,189 MSI and 140,391 MSS histological image patches extracted from the gastric WSIs of the Cancer Genome Atlas (TCGA) [4]. All images were 224 × 224 pixels in size and were processed using the same color normalization method proposed by Macenko et al. [32].

The efficacy of the proposed framework was tested primarily on the CGMH dataset. The TCGA-STAD public dataset was used to test the efficacy of the proposed method across datasets and to compare the performance with other relevant studies on a fairer basis.

B. SOFTWARE ENVIRONMENT

The code was written in Python version 3.8. The deep learning model was built using PyTorch version 1.9.1, running under CUDA version 11.4. In terms of other libraries employed, as the WSIs were not stored in a standard format as those in traditional images, but in an image format known as NDPI developed by Japanese instrument developer Hamamatsu (Shizuoka Prefecture, Japan), they were loaded using the Openslide version 1.1.2 library. The other hyperparameters of the model are listed in Table 2.

TABLE 2. Model hyperparameters.

Parameters	Values
Epoch	40
Learning rate	0.0001
LR_decay	0.01
Optimizer	Ranger
Batch size	128

TABLE 3. Tumor recognition in terms of patches.

	Accuracy	Precision	Recall	Specificity
ResNet18	91.24%	83.43%	88.16%	92.54%
RA	92.88%	88.12%	88.01%	94.95%
RA-NL (Ours)	92.91%	89.23%	86.74%	95.54%
RA-NL-PP (Ours)	**94.33%**	**90.93%**	**90.04%**	**96.16%**

C. STAGE 1: TUMOR IDENTIFICATION

In this stage, we first distinguished between normal and tumorous samples. The training WSI samples were regrouped into 193 normal samples and 255 gastric cancer samples (including both MSS and MSI samples). This is now a two-way classification problem. Patches were extracted from these WSI samples following the procedures described in Section II, resulting in a total of 1,925,337 patches, with 1,387,697 normal and 537,640 tumorous patches. Eighty percent of these patches were used to train the proposed models, and twenty percent tested. Table 3 presents the results.

For this two-way classification, the performance measures were defined as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{11}$$

$$Precision = \frac{TP}{TP + FP} \tag{12}$$

$$Recall = \frac{TP}{TP + FN} \tag{13}$$

$$Specificity = \frac{TN}{FP + TN} \tag{14}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{15}$$

where TP, TN, FP, FN were true positive, true negative, false positive, and false negative respectively. The tumor class was considered positive.

TABLE 4. Tumor recognition in terms of slides.

	Accuracy	Precision	Recall	Specificity	IoU
ResNet18	90.80%	85.96%	100.00%	78.95%	0.7310
RA	89.66%	84.48%	100.00%	76.32%	0.7549
RA-NL (Ours)	**94.25%**	**90.74%**	**100.00%**	**86.84%**	**0.7617**
RA-NL-PP (Ours)	**94.25%**	**90.74%**	**100.00%**	**86.84%**	**0.8005**

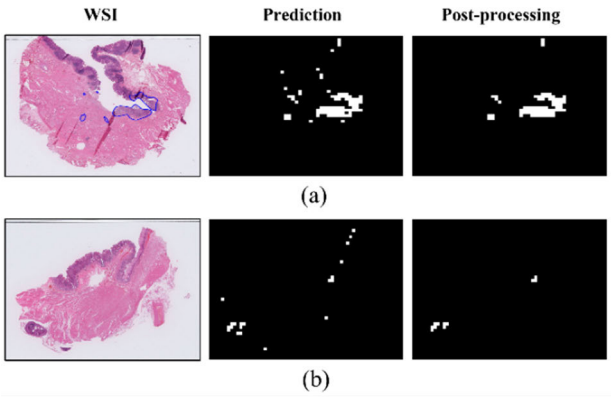


FIGURE 10. Original WSI, predicted tumor region, and after post-processing for (a) actual tumorous WSI where blue circled regions are ground truth tumor region, and (b) normal healthy WSI. Although post-processing can improve the alignment of prediction to physician annotation, it does not alter the slide-wise predicted label.

TABLE 5. Miscrosatellite status recognition in terms of patches.

	Micro Accuracy	Macro Precision	Macro Recall	Macro Specificity
ResNet18	87.13%	76.66%	78.64%	93.12%
RA	88.58%	80.27%	77.22%	93.47%
RA-NL (Ours)	89.19%	80.78%	80.22%	93.57%
RA-NL-PP (Ours)	**92.90%**	**89.96%**	**88.79%**	**95.02%**

Without image post-processing, the residual attention model with a non-local network module (RA-NL) outperformed the baseline in most patch-wise classification metrics, indicating the efficacy of the proposed model. Although the recall rate appeared inferior to that of the baseline, the proposed post-processing was able to correct the errors and improve all metrics by a large margin.

The strength of the proposed architecture stands out when evaluated using whole slides instead of patches. As shown in Table 4, the slide-wise metrics for both RA-NL and RA-NL-PP outperformed the baseline by a large margin.

The predicted tumor region, when compared to the ground truth, also achieved a much higher IoU than the baseline. This indicates that the proposed addition of the non-local module was able to successfully capture long-range dependencies between patches, achieving better regional prediction. The

TABLE 6. Miscrosatellite status recognition in terms of slides.

	Micro Accuracy	Macro Precision	Macro Recall	Macro Specificity
ResNet18	85.06%	86.89%	85.17%	92.68%
RA	83.91%	87.60%	83.77%	92.03%
RA-NL (Ours)	**91.95%**	**92.28%**	**92.53%**	**95.93%**
RA-NL-PP (Ours)	**91.95%**	**92.28%**	**92.53%**	**95.93%**

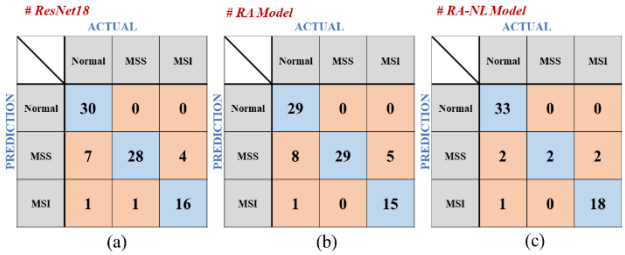


FIGURE 11. Confusion matrix of slide-wise prediction for baseline (a) ResNet18, (b) RA model, and (c) proposed RA-NL model.

metrics before and after post-processing were similar because when deciding whether a slide should be considered tumorous, the rule was set that if a contiguous region of five or more patches were tumorous, then the whole slide would be considered tumorous. Because the post-processing procedure removes noisy predictions that contain two or fewer tumorous patches, it does not affect the predicted label of the slide. As shown in Fig. 10, while post-processing can remove sporadic noises in prediction, contiguous predicted tumorous patches of more than five were still present, so normal tissue images may still be classified as tumors. Nevertheless, post-processing improved the IoU metric.

D. STAGE 2: MICROSATELLITE STATUS RECOGNITION

The second stage extends from the first stage, performing another 2-way classification on the predicted tumorous patches from stage 1 to distinguish whether the tumorous patches were MSS or MSI tissues. The combined result is a 3-way classification system consisting of two independent stages. To train this stage 2 model, the same patches used in the previous stage were employed, including 324,817 MSS and 212,823 MSI patches, totaling 537,640 tumorous patches. The results are presented in Tables 5 and 6.

The evaluation metrics were calculated in terms of the 3-way classification, including micro accuracy, macro precision, macro recall, and macro specificity. Each class was in turn taken as the positive class, while all other classes were negative. Here, micro-accuracy was calculated as the number of samples correctly classified over the total number of samples. The macro-average measures were defined as

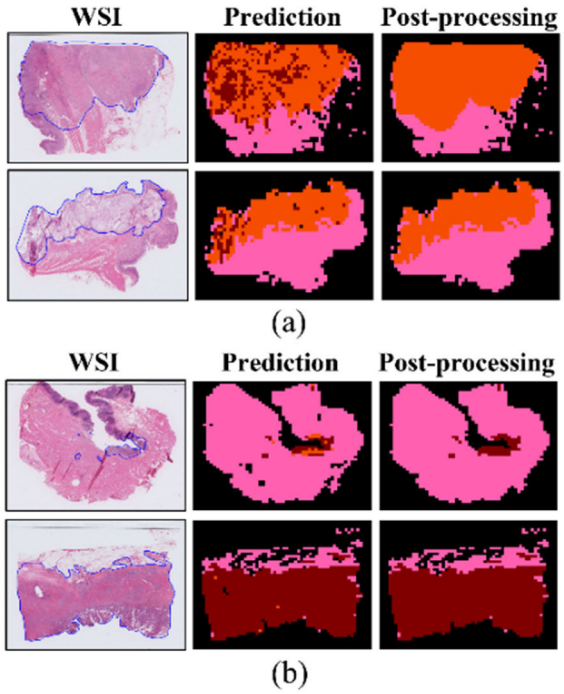


FIGURE 12. Ground truth and predicted tumor regions before and after post-processing for (a) MSI and (b) MSS samples.

follows, taking precision as an example:

$$MacroPrecision = \frac{\sum_c Precision_c}{n} \tag{16}$$

where c was the index of classes and n the number of classes.

The proposed RA-NL-PP model significantly outperformed the other models in both patch-wise and slide-wise 3-way classification. Based on the confusion matrices shown in Fig. 11, RA-NL successfully recalled more MSI samples than the baseline. In Fig. 12, we present some of the prediction outputs for MSI and MSS samples.

Furthermore, we applied Score-CAM [40] to the MSI and MSS patches for the proposed RA-NL model to highlight the areas that contributed to the model's prediction. The red and blue regions in the heatmap in Fig. 13 indicate high and low attention, respectively.

The histological patches of high attention in MSI-associated gastric cancer shown in Fig. 13(a) align with regions of poorly differentiated cancer cells, geographic necrosis, and prominent lymphoplasmacytic infiltration, consistent with the findings of previous studies [10], [11]. In contrast, the highlighted patches in the MSS cases shown in Fig. 13(b) represent glandular formation, little peritumoral inflammatory reaction, infiltrating borders, and poorly cohesive carcinoma.

IV. COMPARISON WITH RELEVANT STUDIES

To compare with previous studies, we performed four additional experiments. The experimental design and comparative results are summarized as follows, please also refer to Table 7.

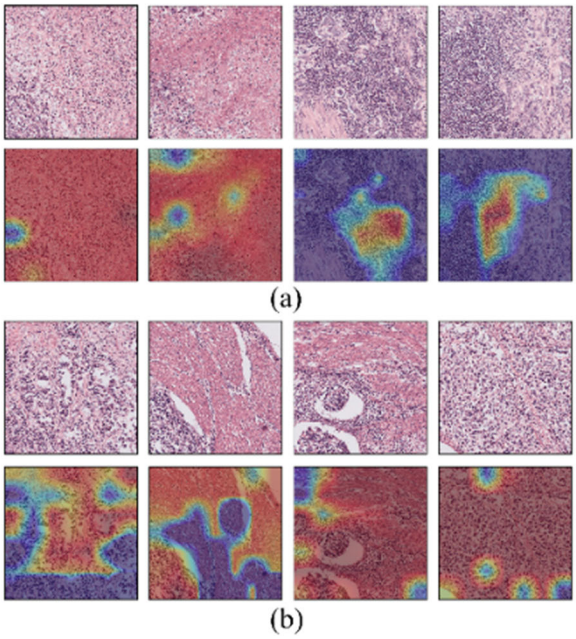


FIGURE 13. (a) Score-CAM heatmap for MSI patches, (b) Score-CAM heatmap for MSS patches.

Experiment 1: We tested the proposed RA-NL model using our private CGMH dataset, as detailed in Section III. Our model achieved an overall accuracy of 81.13%, MSS accuracy of 86.15%, and MSI accuracy of 72.94%. AUC was 0.86.

Experiment 2: We changed the validation procedure to the same as in [19], that is, 70% training, 15% validation, and 15% testing. The private CGMH dataset was again employed, and the model achieved an overall accuracy of 79.32%, where the MSS and MSI accuracies were 83.75% and 72.07 %, respectively. AUC was 0.86. Compared with the 8:2 validation method in Experiment 1, while the AUC score was the same, the overall accuracy decreased slightly.

Experiment 3: The TCGA-STAD dataset was employed, similar to that used in [19]. Again, 70% training, 15% validation, and 15% testing were used. Our RA-NL model achieved an overall accuracy of 95.86%, with an MSS accuracy of 96.38% and an MSI accuracy of 94.92%. AUC was 0.99. Compared with previous literature, ours was much higher.

Experiment 4: Transfer learning was applied using the same dataset and procedure as in Experiment 3. The model was first pretrained using our private dataset and then fine-tuned on the TCGA-STAD dataset. The overall accuracy achieved was 96.53%, where the MSS and MSI accuracies were 97.09% and 95.52 %, respectively. AUC was 0.99. Compared with Experiment 3, the application of transfer learning improved the overall performance of the model.

The results of the four experiments and other relevant studies are summarized in Table 7 for comparison. The comparative results show that the proposed model applied to gastric histopathology images for distinguishing MSI

TABLE 7. Comparison with relevant studies.

Experiments	Method	Dataset	AUC	Overall accuracy	MSS accuracy	MSI accuracy	Validation Method
Kather *et al.* [19]	ResNet18	TCGA-STAD	0.81	-	-	-	Train: 70% Val：15% Test：15%
Su *et al.* [20]	ResNet18	Beijing Cancer Hospital	-	-	-	86.36%	Train: 80% Test：20%
Saldanha *et al.* [23]	Swarm learning and RetCCL	Train: BERN[25], LEEDS[26], TUM[27] Test: TCGA-STAD	0.8092	-	-	-	Cross dataset
Muti *et al.* [21]	ShuffleNet	TCGA	0.836	-	-	-	3-fold
Hinata *et al.* [22]	VGG16	Train: UTokyo, TCGA Test: TCGA	0.87	-	-	-	Cross dataset
Experiment 1 (Ours)	RA-NL	CGMH	0.86	81.13%	86.15%	72.94%	Train: 80% Test：20%
Experiment 2 (Ours)	RA-NL	CGMH	0.86	79.32%	83.75%	72.07%	Train: 70% Val：15% Test：15%
Experiment 3 (Ours)	RA-NL	TCGA-STAD	**0.99**	95.86%	96.38%	94.92%	Train: 70% Val：15% Test：15%
Experiment 4 (Ours)	RA-NL (Transfer learning)	TCGA-STAD	**0.99**	**96.53%**	**97.09%**	**95.52%**	Train: 70% Val：15% Test：15%

and MSS statuses is effective. Transfer learning can further improve performance.

V. CONCLUSION AND FUTURE WORK

Digital pathological slices could not be fed directly into an ordinary neural network for training because of their high resolution; therefore, patches were extracted from the WSIs. However, this also destroys the spatial relationship between different parts of the same WSI, and the CNN model can only identify the local features present in each patch. In this study, we address this issue by extracting overlapping training patches and combining local residual attention model with non-local modules to obtain better spatial information. Post-processing was also proposed to eliminate noisy predictions that could have been caused by bits of normal tissue present inside the annotated tumor region.

A two-stage classification framework was proposed. Employing the CGMH private dataset, the first stage classified WSI samples as normal or tumorous, achieving a high slide-wise overall accuracy of 94.25%, precision of 90.74%, recall of 100%, and specificity of 86.84%. The proposed post-processing improved tumor localization IoU from 0.7617 to 0.8005. The model performed well for both tumor identification and segmentation. The second stage further classified the tumorous patches into MSI or MSS, achieving a 91.95% slide-wise overall accuracy. Using transfer learning, we fine-tuned the model further on the public TCGA-STAD dataset, achieving 96.53% overall accuracy and an AUC of 0.99 for MSI detection on gastric tissue slices, outperforming current literature.

This study is only the beginning. Owing to hardware limitations, we only employed $\times 10$ resolution slices, while $\times 40$ resolution slices were available. If patches could be extracted from these much higher resolution slices, or information from both high- and low-resolution slices could be considered simultaneously, we believe that the model could be further improved.

In addition, in this study, the two stages employed the same architecture and same WSI samples for comparison with other relevant works, but this need not be the case. Different models and sample resolutions can be employed for these two stages. For example, in the first stage, where normal and tumorous samples are classified, a simpler model using lower-resolution samples may suffice to speed up the training and inference time.

ACKNOWLEDGMENT

The results are in part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

(Sung-Nien Yu and Shih-Chiang Huang contributed equally to this work.)

REFERENCES

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, A Cancer J. Clinicians*, vol. 71, no. 3, pp. 209–249, May 2021.

[2] B. Hu, N. El Hajj, S. Sittler, N. Lammert, R. Barnes, and A. Meloni-Ehrig, "Gastric cancer: Classification, histology and application of molecular pathology," *J. Gastrointestinal Oncol.*, vol. 3, no. 3, p. 251, 2012.

[3] S. Huang, K. Ng, T. Yeh, C. Cheng, J. Lin, Y. Liu, H. Chuang, and T. Chen, "Subtraction of Epstein–Barr virus and microsatellite instability genotypes from the Lauren histotypes: Combined molecular and histologic subtyping with clinicopathological and prognostic significance validated in a cohort of 1,248 cases," *Int. J. Cancer*, vol. 145, no. 12, pp. 3218–3230, Dec. 2019.

[4] A. J. Bass et al., "Comprehensive molecular characterization of gastric adenocarcinoma," *Nature*, vol. 513, no. 7517, p. 202, 2014.

[5] R. Cristescu et al., "Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes," *Nature Med.*, vol. 21, no. 5, pp. 449–456, 2015.

[6] J. L. Weber and P. E. May, "Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction," *Amer. J. Hum. Genet.*, vol. 44, no. 3, p. 388, 1989.

[7] S. Semba, H. Yokozaki, S. Yamamoto, W. Yasui, and E. Tahara, "Microsatellite instability in precancerous lesions and adenocarcinomas of the stomach," *Cancer*, vol. 77, no. S8, pp. 1620–1627, Apr. 1996.

[8] R. Bonneville, M. A. Krook, E. A. Kautto, J. Miya, M. R. Wing, H.-Z. Chen, J. W. Reeser, L. Yu, and S. Roychowdhury, "Landscape of microsatellite instability across 39 cancer types," *JCO Precis. Oncol.*, no. 1, pp. 1–15, Nov. 2017.

[9] D. T. Le et al., "PD-1 blockade in tumors with mismatch-repair deficiency," *New England J. Med.*, vol. 372, no. 26, pp. 2509–2520, 2015.

[10] J. Rüschoff, W. Dietmaier, J. Lüttges, G. Seitz, T. Bocker, H. Zirngibl, J. Schlegel, H. K. Schackert, K. W. Jauch, and F. Hofstaedter, "Poorly differentiated colonic adenocarcinoma, medullary type: Clinical, phenotypic, and molecular characteristics," *Amer. J. Pathol.*, vol. 150, no. 5, pp. 1815–1825, May 1997.

[11] B.-J. Lu, M. Lai, L. Cheng, J.-Y. Xu, and Q. Huang, "Gastric medullary carcinoma, a distinct entity associated with microsatellite instability-H, prominent intraepithelial lymphocytes and improved prognosis," *Histopathology*, vol. 45, no. 5, pp. 485–492, Nov. 2004, doi: 10.1111/j.1365-2559.2004.01998.x.

[12] J. M. S. Prewitt and M. L. Mendelsohn, "The analysis of cell images," *Ann. New York Acad. Sci.*, vol. 128, no. 3, pp. 1035–1053, 1966.

[13] S. Mukhopadhyay et al., "Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: A multicenter blinded randomized noninferiority study of 1992 cases (pivotal study)," *Amer. J. Surgical Pathol.*, vol. 42, no. 1, pp. 39–52, 2018.

[14] C. L. Srinidhi, O. Ciga, and A. L. Martel, "Deep neural network models for computational histopathology: A survey," *Med. Image Anal.*, vol. 67, Jan. 2021, Art. no. 101813.

[15] S. Ai, C. Li, X. Li, T. Jiang, M. Grzegorzek, C. Sun, M. M. Rahaman, J. Zhang, Y. Yao, and H. Li, "A state-of-the-art review for gastric histopathology image analysis approaches and future development," *BioMed Res. Int.*, vol. 2021, pp. 1–19, Jun. 2021.

[16] Y. Li, X. Li, X. Xie, and L. Shen, "Deep learning based gastric cancer identification," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 182–185.

[17] S. C. Kosaraju, J. Hao, H. M. Koh, and M. Kang, "Deep-Hipo: Multi-scale receptive field deep learning for histopathological image analysis," *Methods*, vol. 179, pp. 3–13, Jul. 2020.

[18] Z. Zhu, X. Ding, D. Zhang, and L. Wang, "Weakly-supervised balanced attention network for gastric pathology image localization and classification," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1–4.

[19] J. N. Kather, A. T. Pearson, N. Halama, D. Jäger, J. Krause, S. H. Loosen, A. Marx, P. Boor, F. Tacke, U. P. Neumann, H. I. Grabsch, T. Yoshikawa, H. Brenner, J. Chang-Claude, M. Hoffmeister, C. Trautwein, and T. Luedde, "Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer," *Nature Med.*, vol. 25, no. 7, pp. 1054–1056, Jul. 2019.

[20] F. Su, J. Li, X. Zhao, B. Wang, Y. Hu, Y. Sun, and J. Ji, "Interpretable tumor differentiation grade and microsatellite instability recognition in gastric cancer using deep learning," *Lab. Invest.*, vol. 102, no. 6, pp. 641–649, Jun. 2022, doi: 10.1038/s41374-022-00742-6.

[21] H. S. Muti et al., "Development and validation of deep learning classifiers to detect Epstein–Barr virus and microsatellite instability status in gastric cancer: A retrospective multicentre cohort study," *Lancet Digit. Health*, vol. 3, no. 10, pp. e654–e664, Oct. 2021, doi: 10.1016/S2589-7500(21)00133-3.

[22] M. Hinata and T. Ushiku, "Detecting immunotherapy-sensitive subtype in gastric cancer using histologic image-based deep learning," *Sci. Rep.*, vol. 11, no. 1, p. 22636, Nov. 2021, doi: 10.1038/s41598-021-02168-4.

[23] O. L. Saldanha, H. S. Muti, H. I. Grabsch, R. Langer, B. Dislich, M. Kohlruss, G. Keller, M. van Treeck, K. J. Hewitt, F. R. Kolbinger, G. P. Veldhuizen, P. Boor, S. Foersch, D. Truhn, and J. N. Kather, "Direct prediction of genetic aberrations from pathology images in gastric cancer with swarm learning," *Gastric Cancer*, vol. 26, no. 2, pp. 264–274, Mar. 2023, doi: 10.1007/s10120-022-01347-0.

[24] X. Wang, Y. Du, S. Yang, J. Zhang, M. Wang, J. Zhang, W. Yang, J. Huang, and X. Han, "RetCCL: Clustering-guided contrastive learning for whole-slide image retrieval," *Med. Image Anal.*, vol. 83, Jan. 2023, Art. no. 102645, doi: 10.1016/j.media.2022.102645.

[25] B. Dislich, N. Blaser, M. D. Berger, B. Gloor, and R. Langer, "Preservation of Epstein–Barr virus status and mismatch repair protein status along the metastatic course of gastric cancer," *Histopathology*, vol. 76, no. 5, pp. 740–747, Apr. 2020, doi: 10.1111/his.14059.

[26] T. Hayashi, T. Yoshikawa, K. Bonam, H. M. Sue-Ling, M. Taguri, S. Morita, A. Tsuburaya, J. D. Hayden, and H. I. Grabsch, "The superiority of the seventh edition of the TNM classification depends on the overall survival of the patient cohort: Comparative analysis of the sixth and seventh TNM editions in patients with gastric cancer from Japan and the united kingdom," *Cancer*, vol. 119, no. 7, pp. 1330–1337, Apr. 2013, doi: 10.1002/cncr.27928.

[27] M. Kohlruss, B. Grosser, M. Krenauer, J. Slotta-Huspenina, M. Jesinghaus, S. Blank, A. Novotny, M. Reiche, T. Schmidt, L. Ismani, A. Hapfelmeier, D. Mathias, P. Meyer, M. M. Gaida, L. Bauer, K. Ott, W. Weichert, and G. Keller, "Prognostic implication of molecular subtypes and response to neoadjuvant chemotherapy in 760 gastric carcinomas: Role of Epstein–Barr virus infection and high- and low-microsatellite instability," *J. Pathol., Clin. Res.*, vol. 5, no. 4, pp. 227–239, Oct. 2019, doi: 10.1002/cjp2.137.

[28] S. F. Qadri, L. Shen, M. Ahmad, S. Qadri, S. S. Zareen, and S. Khan, "OP-convNet: A patch classification-based framework for CT vertebrae segmentation," *IEEE Access*, vol. 9, pp. 158227–158240, 2021, doi: 10.1109/ACCESS.2021.3131216.

[29] R. Gonzalez and R. Woods, *Digital Image Processing*, 4th ed. New York, NY, USA: Pearson, 2018.

[30] J. D. Bancroft and C. Layton, "The hematoxylins and eosin," in *Bancroft's Theory and Practice of Histological Techniques*, 7th ed. Amsterdam, The Netherlands: Elsevier, 2012, pp. 173–186.

[31] P. Kleczek, J. Jaworek-Korjakowska, and M. Gorgon, "A novel method for tissue segmentation in high-resolution H&E-stained histopathological whole-slide images," *Computerized Med. Imag. Graph.*, vol. 79, Jan. 2020, Art. no. 101686.

[32] M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, X. Guan, C. Schmitt, and N. E. Thomas, "A method for normalizing histology slides for quantitative analysis," in *Proc. IEEE Int. Symp. Biomed. Imag., Nano Macro*, Jun. 2009, pp. 1107–1110.

[33] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6450–6458.

[34] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7794–7803.

[35] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 60–65.

[36] Y. Wen, L. Chen, H. Chen, X. Tang, Y. Deng, Y. Chen, and C. Zhou, "Non-local attention learning for medical image classification," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2021, pp. 1–6, doi: 10.1109/ICME51207.2021.9428267.
[37] L. He, Y. Chao, K. Suzuki, and K. Wu, "Fast connected-component labeling," *Pattern Recognit.*, vol. 42, no. 9, pp. 1977–1987, Sep. 2009.
[38] R. Khudeev, "A new flood-fill algorithm for closed contour," in *Proc. Siberian Conf. Control Commun.* Tomsk, Russia: IEEE, 2005, pp. 170–174, doi: 10.1109/SIBCON.2005.1611214.
[39] E.-M. Nosal, "Flood-fill algorithms used for passive acoustic detection and tracking," in *Proc. New Trends Environ. Monitor. Using Passive Syst.*, Oct. 2008, pp. 1–5, doi: 10.1109/PASSIVE.2008.4786975.
[40] H. Wang, Z. Wang, M. Du, F. Yang, Z. Zhang, S. Ding, P. Mardziel, and X. Hu, "Score-CAM: Score-weighted visual explanations for convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 111–119.

SUNG-NIEN YU (Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1987 and 1991, respectively, and the Ph.D. degree in biomedical engineering from Case Western Reserve University, Cleveland, OH, USA, in 1996.

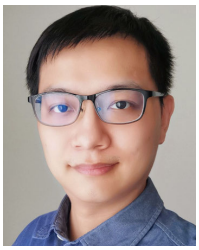
He joined the Faculty of the Department of Electrical Engineering, National Chung Cheng University (NCCU), Chiayi, Taiwan, as an Assistant Professor, in 1999. He was promoted to Associate Professor, in August 2006. Since August 2010, he has been a Professor with the Department of Electrical Engineering, NCCU, where he has been the Department Chair, since August 2021. His research interests include biomedical signal processing, biomedical image processing, and the application of artificial intelligence to biomedical problems. He is a Life Member of the Taiwanese Society of Biomedical Engineering.

SHIH-CHIANG HUANG received the M.D. degree from National Taiwan University, Taipei, Taiwan, in 2004. He is currently pursuing the Ph.D. degree with the Graduate Institute of Clinical Medical Sciences, College of Medicine, Chang Gung University, Taoyuan, Taiwan.

From 2014 to 2016, he was a Research Fellow with the Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. He finished the resident training with the Department of Anatomical Pathology, Linkou Chang Gung Memorial Hospital, Taoyuan, where he is currently a Certificated Practicing Pathologist. His research interests include bone and soft tissue pathology, gastrointestinal and pancreaticobiliary pathology, molecular pathology, and the application of artificial intelligence in diagnostic pathology.

Dr. Huang is a member of the Taiwan Society of Pathology. He received awards and honors, including the Best Paper Award of Residents (Chang Gung Memorial Hospital, Linkou), the Best Abstract Award from the International Association of Chinese Pathologists, the Ta-You Wu Memorial Award, and the National Innovation Award of the Taiwan Government.

WEI-CHEN WANG received the B.S. degree in automatic control engineering from Feng Chia University, Taichung, Taiwan, in 2017, and the M.S. degree in electrical engineering from National Chung Cheng University, Chiayi, Taiwan, in 2021. From 2018 to 2019, he was an Automation Engineer for a year. During work, he has been involved in image recognition and machine learning. His research interests include biomedical image processing and the development of image recognition systems based on artificial intelligence.



YU PING CHANG received the B.S. degree in physics from the National University of Singapore, Singapore, in 2015, and the M.A. degree in religion and humanity from Tzu Chi University, Hualien, Taiwan, in 2018. He is currently pursuing the M.S. degree in electrical engineering with National Chung Cheng University (NCCU), Chiayi, Taiwan.

From 2021 to 2022, he was a Research Assistant with NCCU. His research interest includes the development of an explainable neuro-symbolic AI system with application in medical diagnosis.

KUANG-HUA CHEN received the M.D. degree from Kaohsiung Medical University, in 2010.

From 2010 to 2014, she was a Resident with the Department of Anatomic Pathology, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan. After obtaining the Certificate of Anatomic Pathology, she has been an Attending Physician with the Linkou Chang Gung Memorial Hospital, since December 2014. She was a Visiting Scholar with the Department of Pathology, Division of Dermatopathology, University of California at San Francisco, San Francisco, CA, USA, from 2017 to 2018. She is currently a Surgical Pathologist with a subspecialty in dermatopathology.

Dr. Chen is a member of the Taiwan Society of Pathology.

TSE-CHING CHEN received the M.D. degree from the Chung Shan Medical College, Taichung, Taiwan, in 1988, and the D.Phil. degree in immunology from the Sir William Dunn School of Pathology, University of Oxford, U.K., in 2004.

After two years of military service, he began resident training with the Department of Anatomic Pathology, Chang Gung Memorial Hospital, Taoyuan, Taiwan, in 1990. He was an Attending Physician, in 1994, and got the promotion to the Director of the Department of Anatomical Pathology and a Professor, in 2007. He was a Professor with the School of Medicine, Chang Gung University, Taoyuan, in 2012. He has published more than 280 SCI papers, covering pathology, immunology, molecular biology, oncology, and artificial intelligence in medicine.

Prof. Chen is a member of the Taiwan Society of Pathology and the U.S. and Canadian Society of Pathology. His department received the Gold Award in Health Smart Taiwan, Joint Commission of Taiwan, in 2021.