Journal of Translational Medicine

RESEARCH

Open Access

Deep learning-based prediction of HER2 status and trastuzumab treatment efficacy of gastric adenocarcinoma based on morphological features

Zhida Wu[1†], Tao Wang[2†], Junlin Lan[2†], Jianchao Wang[1], Gang Chen[1], Tong Tong[2*] and Hejun Zhang[1*]

Abstract

Background First-line treatment for advanced gastric adenocarcinoma (GAC) with human epidermal growth factor receptor 2 (HER2) is trastuzumab combined with chemotherapy. In clinical practice, HER2 positivity is identified through immunohistochemistry (IHC) or fluorescence in situ hybridization (FISH), whereas deep learning (DL) can predict HER2 status based on tumor histopathological features. However, it remains uncertain whether these deep learning-derived features can predict the efficacy of anti-HER2 therapy.

Methods We analyzed a cohort of 300 consecutive surgical specimens and 101 biopsy specimens, all undergoing HER2 testing, along with 41 biopsy specimens receiving trastuzumab-based therapy for HER2-positive GAC.

Results We developed a convolutional neural network (CNN) model using surgical specimens that achieved an area under the curve (AUC) value of 0.847 in predicting HER2 amplification, and achieved an AUC of 0.903 in predicting HER2 status specifically in patients with HER2 2 + expression. The model also predicted HER2 status in gastric biopsy specimens, achieving an AUC of 0.723. Furthermore, our classifier was trained using 41 HER2-positive gastric biopsy specimens that had undergone trastuzumab treatment, our model demonstrated an AUC of 0.833 for the (CR + PR) / (SD + PD) subgroup.

Conclusion This work explores an algorithm that utilizes hematoxylin and eosin (H&E) staining to accurately predict HER2 status and assess the response to trastuzumab in GAC, potentially facilitating clinical decision-making.

Keywords HER2, Gastric adenocarcinoma, Trastuzumab, Deep learning, Efficacy

†Zhida Wu, Tao Wang and Junlin Lan are co-first authors.

*Correspondence:
Tong Tong
ttraveltong@gmail.com
Hejun Zhang
38617347@qq.com
[1]Department of Pathology, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou 350014, China
[2]College of Physics and Information Engineering, Fuzhou University, Xueyuan Road No.2, Fuzhou, Fujian 350108, China

Background

According to the GLOBOCAN 2020 global cancer statistics, gastric adenocarcinoma (GAC) is ranked fifth in incidence (5.6% of all cancer cases) and fourth (7.7%) among total cancer-related mortalities worldwide [1]. The occurrence of GAC demonstrates regional disparities, with a notably higher prevalence documented in East Asia, especially in China, where it currently ranks as the country with the highest gastric cancer incidence and mortality rates in the world [2].

Human epidermal growth factor receptor 2 (HER2) plays a critical role in tumor proliferation, differentiation, and survival, and it serves as an important therapeutic target. GAC exhibits high heterogeneity, with HER2-positive cases representing approximately 10–20% of all cases [3, 4]. Trastuzumab, a fully humanized anti-HER2 antibody, has proven effective in treating HER2-positive GAC [5]. Currently, the recommended first-line treatment for HER2-positive GAC involves combining trastuzumab with fluoropyrimidine and platinum-based chemotherapy [6], yielding overall response rates between 47.0% and 68.0% [7, 8].

Recent studies have employed convolutional neural networks (CNNs) based on histopathological data to predict the Epstein-Barr virus (EBV) and microsatellite instability (MSI) status of GAC [9, 10]. However, predictive models for HER2 status using hematoxylin and eosin (H&E) staining remain scarce. Notably, not all patients with high HER2 expression respond to trastuzumab, and predicting the response to therapy on the basis of current criteria remains a challenge [11]. Thus, there is a need for a biomarker that can predict trastuzumab response.

The aim of this study was to develop a deep learning (DL) framework for predicting the HER2 status of GAC from histopathology and evaluate the efficacy of trastuzumab therapy for HER2-positive-expression GAC. Our research demonstrates the development of a DL framework that enhances the accuracy of predicting HER2 status and assessing the response to trastuzumab treatment with a smaller sample size, focusing on GAC H&E histopathology slides.

Materials and methods

Cohort for predicting HER2 status

Data from 300 patients diagnosed with gastric adenocarcinoma (GAC) who underwent surgical resection, along with 101 biopsy samples, were reviewed. The inclusion criteria required complete surgical removal of the tumor, histological confirmation of GAC by pathologists, and the retrieval of at least 15 lymph nodes from the surgical samples. HER2 testing was performed on both the surgical and biopsy samples.

Cohort for predicting the trastuzumab treatment response

This cohort comprised 41 individuals with HER2-positive gastric biopsy specimens who received neoadjuvant targeted therapy with trastuzumab before surgery.

HER2 protein expression analysis

HER2 protein expression analysis was performed in all patients by two experienced pathologists using immunohistochemical staining on 4–5 μm thick 10% formalin-fixed paraffin-embedded tissues. The HER2 protein assay with the PATHWAY HER-2/neu rabbit monoclonal antibody (clone 4B5; Ventana Medical Systems Inc, Tucson, AZ, USA) was conducted using a BenchMark XT automated staining system (Ventana). In this study, we utilized the surgical specimen and biopsy scoring methods to evaluate the HER2 immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH) results [12]. Samples with scores of 0 and 1+were considered HER2 negative, whereas samples with an IHC score of 3+were classified as positive. Patients with a score of 2+were considered "equivocal" for HER2 overexpression. Therefore, we conducted FISH for HER2 2+cases, where "positive" was defined as an HER2/CEP17 ratio2 or an HER2/CEP17 ratio < 2 with an average HER2 copy number ≥ 6.

Treatment

The focus of this study was on patients treated with trastuzumab before surgery. The RECIST version 1.1 criteria system was employed to classify post-neoadjuvant therapy responses, including complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). CR was defined as the absence of observable tumor signs during clinical examinations lasting longer than four weeks. PR was identified as a reduction in lesion size exceeding 30%, persisting for a minimum of four weeks. SD was characterized by lesion stability with alterations between PR and PD. PD was defined as disease progression with an increase of more than 20% [13–15]. Patients with PR and CR were categorized as responders, and patients with SD and PD were categorized as non-responders.

Data preprocessing

Whole slide images (WSIs) were scanned at a magnification of 40× using a PRECICE 600 digital scanner (Unic, Beijing, China), resulting in a resolution of approximately 0.5 μm per pixel. Because of their large size, in this study, WSIs were sliced into nonoverlapping patches of 512×512 pixels at the 20× magnification scale. Patches containing irrelevant information were excluded; only those containing various tissue types such as fibers, fat, stroma, smooth muscle, and tumor tissue were retained. Stain normalization was applied to all WSI images using the method described in this study [16].

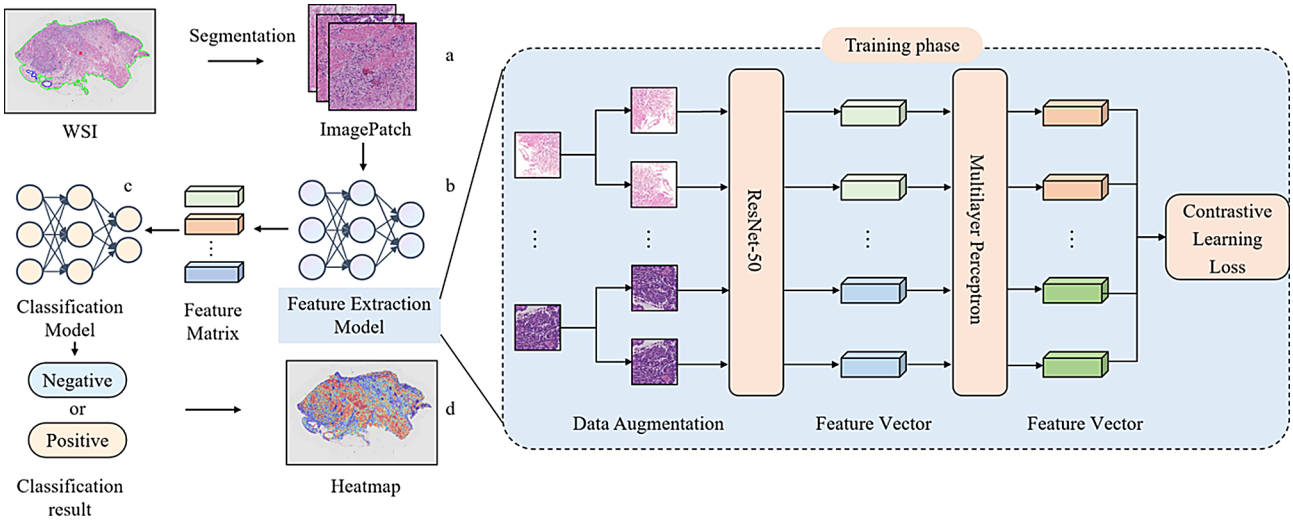

Fig. 1 Proposed annotation-free deep learning framework for HER2 status classification from WSIs: **a**, Image segmentation divides images into distinct regions; **b**, Features are extracted from tissue regions of the WSI using image patches; **c**, Pre-computed feature vectors of image patches are input into the model. An attention network consolidates patch-level information into slide-level representations, which are utilized for the final diagnostic prediction; **d**, Strongly patched (red) and weakly patched (blue) regions serve as representative samples to guide clustering layers in distinguishing between positive and negative instances of distinct classes

Table 1 Comparison with other weakly supervised algorithms on the training dataset with 10-fold cross validation

Algorithms	AUC	ACC	Sensitivity	Specificity
MIL	0.800 ± 0.133	0.786 ± 0.071	0.445 ± 0.201	0.956 ± 0.201
AttMIL	0.799 ± 0.075	0.777 ± 0.054	0.530 ± 0.090	0.900 ± 0.071
TransMIL	0.795 ± 0.094	0.812 ± 0.061	0.518 ± 0.143	0.950 ± 0.059
CLAM	0.830 ± 0.100	0.776 ± 0.082	0.546 ± 0.171	0.892 ± 0.098
CLAM_Sim(Our)	**0.847 ± 0.049**	0.792 ± 0.044	**0.613 ± 0.132**	0.880 ± 0.098

Feature extraction network

We used a pretrained ResNet-50 network in the pyramid network to extract image features, with the trained network parameters serving as network models. The baseline network was pretrained on the ImageNet dataset and was fine-tuned with the self-attention method SimCLR on a dataset of 300 patients with GAC to improve the capability feature extraction.

The main objective of the feature extractor was to automatically extract high-quality features from the patches. Deep convolutional networks exhibit an efficient feature extraction ability, enabling the formation of abstract high-dimensional features while requiring less training time through the utilization of low-dimensional features.

Additionally, to prevent overfitting and enhance model performance, four conventional data augmentation techniques were employed: horizontal flipping, color jitter, grayscale transformation, and gaussian blur. Figure 1 shows the plate layout.

Statistical analysis

The efficacy of the DL model for the classification of predicted HER2 status and the effectiveness of trastuzumab treatment were evaluated by using multiple-instance learning (MIL) and calculating the area under the curve (AUC), accuracy (ACC), sensitivity and specificity.

Results

Comparison with alternative learning algorithms

To evaluate the effectiveness of our clustering-constrained-attention multiple-instance learning (CLAM) model, we compared its performance to three weak supervision algorithms: Multiple-instance learning (MIL), attention-based MIL (AttMIL) and transformer-based MIL (TransMIL). All algorithms were evaluated using 10-fold cross-validation. The model with the highest performance was selected and evaluated on a test set. With an AUC of 0.830±0.100, our CLAM model outperformed the other state-of-the-art DL models: MIL (0.800±0.133), AttMIL (0.799±0.075), and TransMIL (0.795±0.094) (Table 1; Fig. 2).

Ablation experiments

We conducted ablation experiments to verify the superiority of the proposed framework and demonstrate significantly improved performance compared to the non-ablation algorithms in determining HER2 status. Specifically, experiments were conducted to compare the

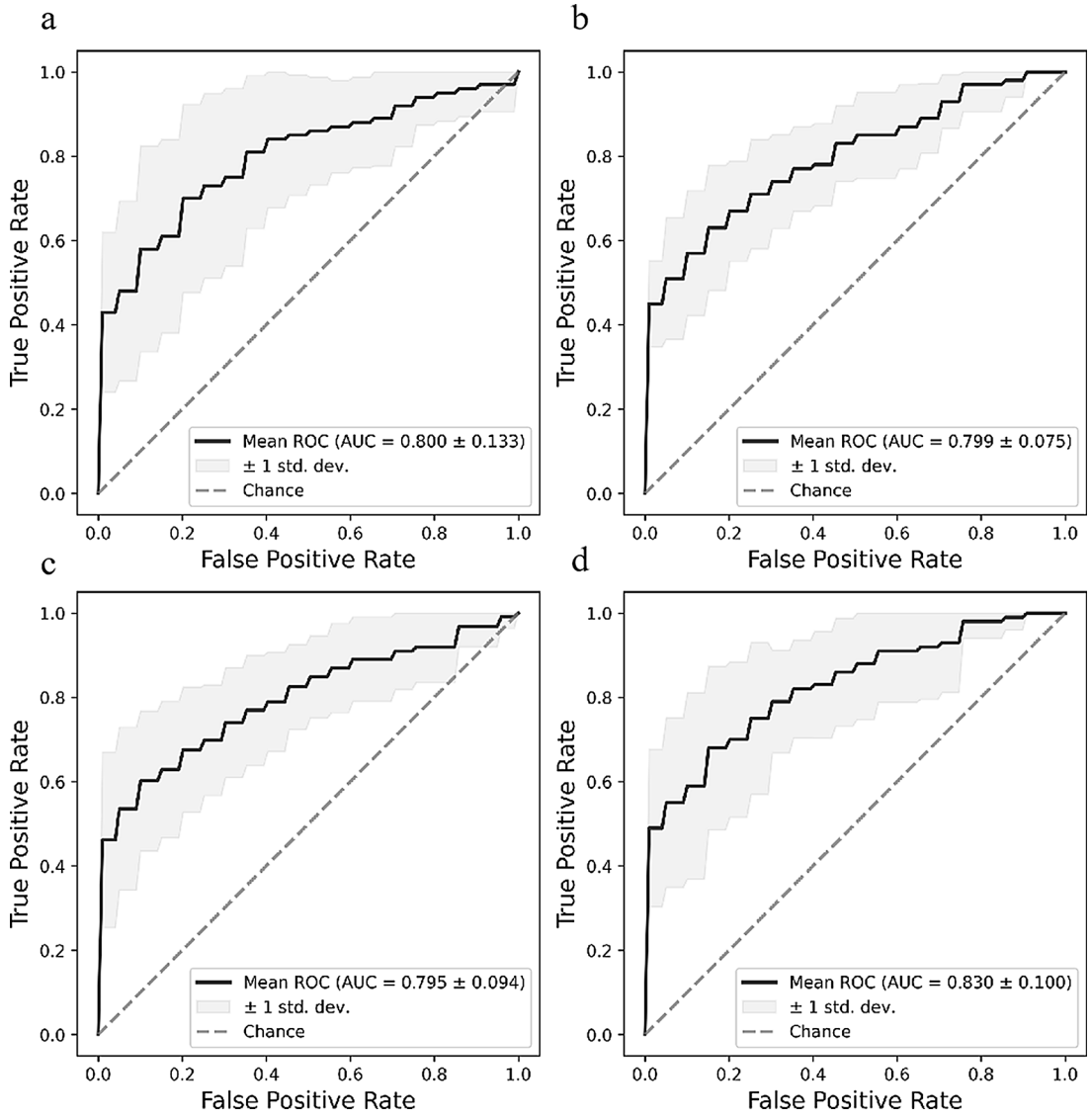

Fig. 2 AUC results for evaluated weakly supervised methods: **a**, MIL, **b**, AttMIL, **c**, TransMIL, **d**, CLAM

efficacy of feature extraction using a contrastive learning model versus the CLAM_Sim model in HER2 status prediction.

Prediction of HER2 status from tissue morphology

All WSIs from the study were reviewed by a qualified specialist for slide-level inclusion criteria and quality assurance. The dataset included both surgical resection and biopsy samples. For the surgical cohort, which comprised 97 HER2-positive cases and 203 HER2-negative cases, the proposed model achieved an AUC of 0.847, with accuracy (ACC), sensitivity, and specificity values of 0.792, 0.613, and 0.880, respectively, in the 10-fold cross-validation. Among the surgical cohort, 56 patients exhibited HER2 2+expression, including 35 cases classified as HER2-positive and 21 cases classified as HER2-negative, the predictive AUC for HER2 status in this subgroup was 0.903, with an ACC of 0.869 (Fig. 3a, b and d). Additionally, the model performed well in predicting HER2 IHC expression, with an AUC of 0.806 in distinguishing between HER2 scores of 0 and 1+/3+, and an AUC of 0.805 in distinguishing between HER2 scores of 0, 1+, and 2+/3+ (Fig. 4).

In a separate analysis of the biopsy cohort, which included 36 HER2-positive and 65 HER2-negative cases, the model achieved an AUC of 0.723 for predicting HER2 status (Fig. 3c).

Prediction of trastuzumab treatment outcomes

To further evaluate whether DL can predict the response to trastuzumab by using gastric biopsy specimens, patients were categorized into two groups according to

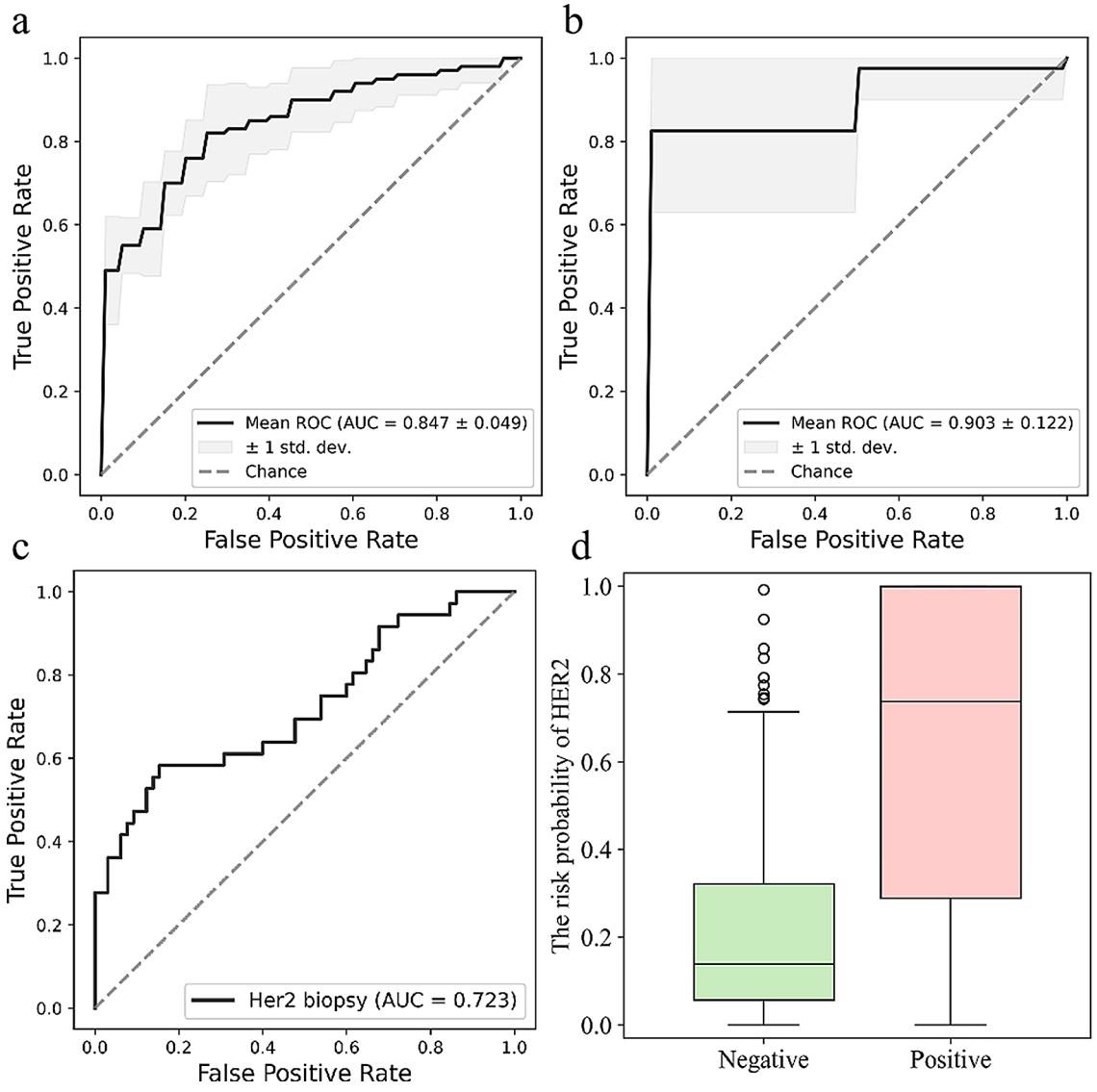

Fig. 3 Average receiver operating characteristic (ROC) curve generated through the application of stratified 10-fold cross-validation for the surgical cohort **(a)**, including the HER2 2+dataset **(b)**, with the corresponding ROC curves for the biopsy cohort shown **(c)**; **d**, The box plot was generated by consolidating the test set scores obtained from each fold for the surgical cohort

treatment response: effective group (CR+PR) and ineffective group (SD+PD). Our study also utilized the CLAM_Sim model to predict the efficacy of GAC treatment. The results show that the model achieved a predictive AUC of 0.833 for the (CR+PR)/(SD+PD) subgroup (Fig. 5).

Heat map generation

To explore the interpretability of the model, enabling pathologists to better understand how neural networks predict HER2 status on the basis of WSI images, we first obtained attention scores for each patch and converted them into percentages to visualize the importance of H&E image regions for predicting HER2 status, as a method of identifying areas that are closely associated with predicting HER2 positivity. Then, we converted attention scores into red–green–blue (RGB) color values and directly overlaid them onto these original images.

As shown in Fig. 6, we visualized the attention scores corresponding to each small patch by mapping their coordinates onto the original WSI, with high-probability regions in red and low-probability regions in blue.

The resulting heat map reveals a significant overlap between the "hot spots" from our model and the regions of the tumor. This result indicates that the model effectively identifies regions strongly correlated to HER2 status. In summary, this overlap provides valuable insights for neural network predictions of HER2 status, enhancing pathologists' confidence in the predictive accuracy of the AI model.

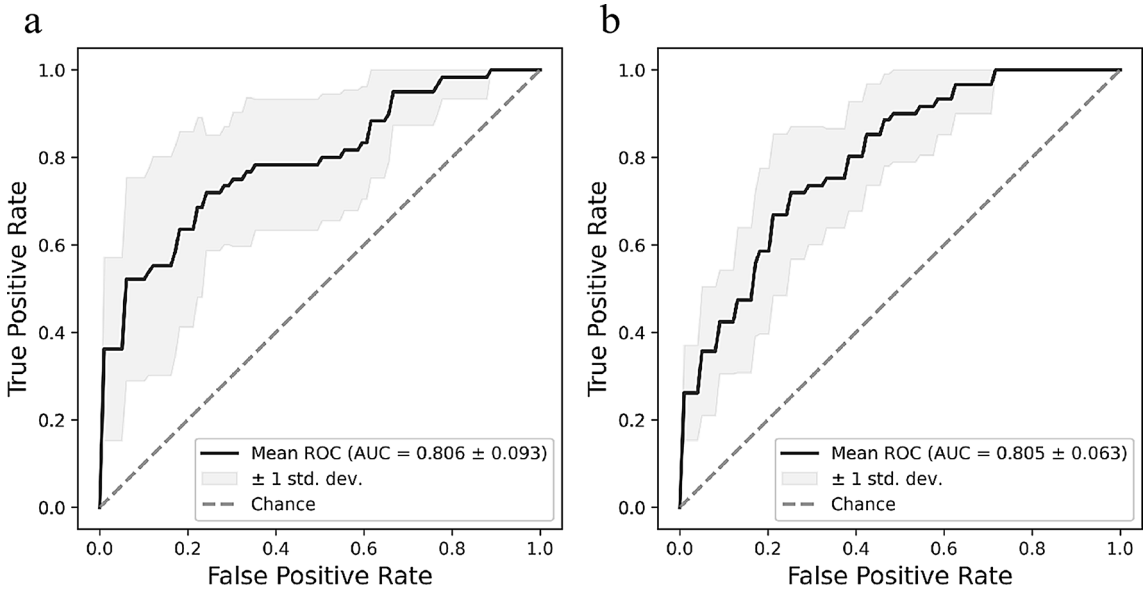

Fig. 4 HER2 IHC expression using unannotated slides: **a**, Results for the 0 and 1+/3 + groups; **b**, Results for the 0, 1+, and 2+/3 + groups

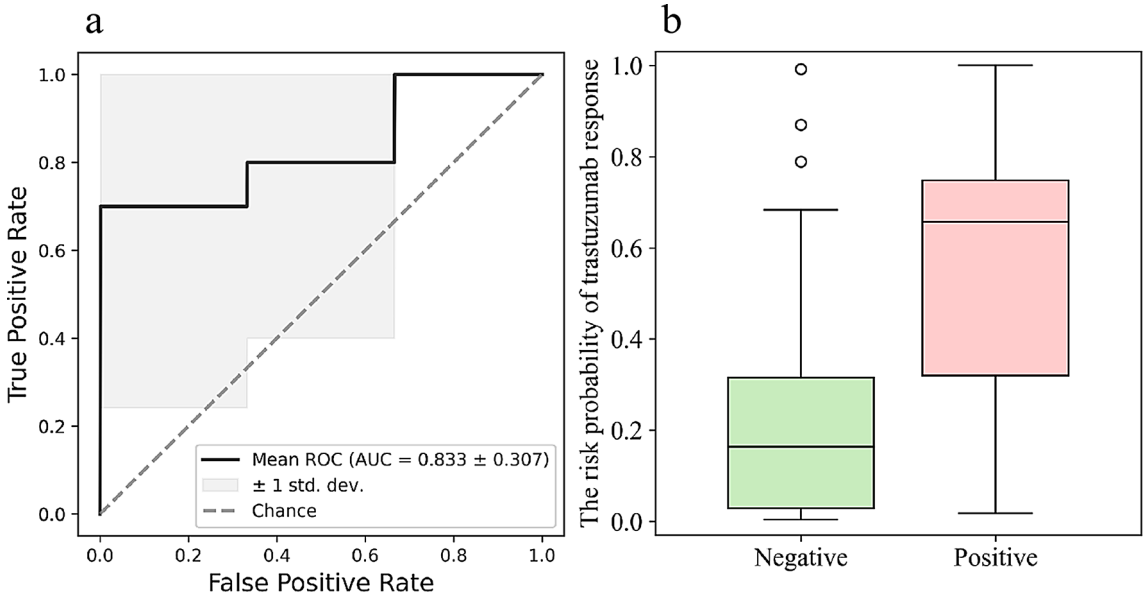

Fig. 5 **a**, ROC curve trained using 10-fold cross-validation for efficacy performance; **b**, Box plot quantified by aggregating the test set scores from each fold

Discussion

To our knowledge, this is the first study in which methods to predict HER2 status were explored for GAC patients, and to evaluate trastuzumab efficacy in biopsy specimens of HER2-positive GAC patients that had undergone initial trastuzumab treatment combined with chemotherapy. Our DL algorithm exhibited high sensitivity and specificity while using only a small sample size, thus confirming the feasibility and effectiveness of our research. In this study, we introduced a CNN trained on histological features of primary GAC tissue that can accurately predict HER2 status and assess the response to trastuzumab treatment. Additionally, we conducted ablation experiments to demonstrate the efficiency of our algorithm compared to other state-of-the-art methods.

DL, a subset of artificial intelligence (AI), can handle a large amount of patient data and recognize patterns in a diverse range of applications and industries. Recently, DL has been applied to achieve image classification in the medical imaging and computer vision fields, with particular emphasis on applications in computed tomography (CT), magnetic resonance imaging (MRI), and pathological images [17, 18].

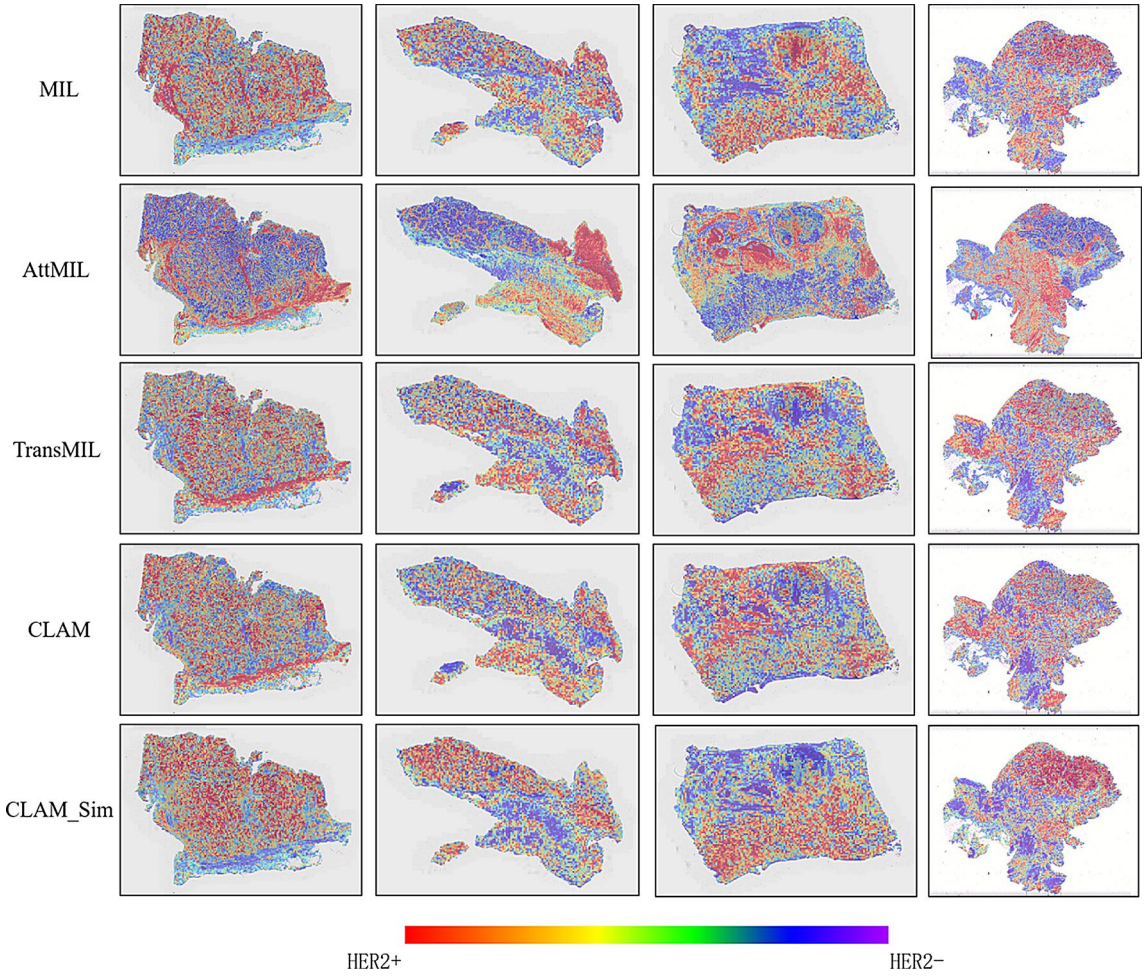

Fig. 6 Heat maps indicating the performance for the different classification algorithms

Recently, there has been a notable increase in the application of DL algorithms in the field of pathology, especially for cancer classification and estimating molecular features and cancer subtypes such as EBV and MSI status in GAC. Patient-level AUC results for EBV range from 0.8723 to 0.969 [9, 19], and for MSI, they range from 0.770 to 0.8848 [10, 20]. DL models have thus shown significant efficacy in directly identifying features, overcoming the subjective bias inherent in extracting features from H&E-stained slides.

Assessing HER2 status is Crucial in guiding clinical decisions across various cancer types. However, the widespread use of HER2 testing is limited by financial constraints associated with FISH and IHC. In HER2-positive GAC, intratumoral HER2 heterogeneity prevalence ranges from 45 to 79% by IHC and 23–54% by ISH, exceeding the rates observed in breast cancers. This variability may be associated with the diversity of GAC tissues and the biological characteristics of tumors, which may have implications for treatment efficacy and patient prognosis [21, 22]. Hence, there is an immediate requirement for a model capable of accurately predicting HER2 status in GAC.

Several studies have investigated the correlation between breast cancer HER2 status and therapeutic efficacy using DL methods but not for GAC [23, 24]. This paper presents a model to explore the potential efficacy of DL in predicting HER2 status levels in GAC, introducing an algorithm that utilizes morphological features for HER2 status prediction. In the process of model development, we evaluated the classification performance of various DL networks and found CLAM demonstrated the highest performance. Furthermore, a comparative analysis of various classification learning networks identified CLAM as the most successful classifier. Further training using a contrastive learning model led to a substantial 1.73% increase in average AUC for classification compared to a single CLAM classifier, with the accuracy improving by 1.60%.

Our findings indicate that the transformer and DL models resulted in a higher AUC value compared with that achieved by the models developed by Bychkov (0.847

vs. 0.70) [23]. Furthermore, Bychkov's research employed 712 WSI cases for training, whereas we utilized a smaller sample size of 300 WSI cases. Despite the disparity in sample size, our algorithm demonstrated superior accuracy in predicting HER2 status. Currently, cases scored as HER2 2+by IHC require additional FISH testing to confirm HER2 status. The proposed algorithm achieved a high AUC of 0.903 and an accuracy of 0.869 in predicting HER2 status for this subset of cases. Hence, we aimed to predict HER2 status specifically for cases with an IHC score of 2+. However, in the biopsy cohort, the model achieved a lower AUC of 0.723, likely due to both the smaller sample size and the limited tissue amount available, and future research with larger sample-size cohorts is needed to improve the reliability of the algorithm.

The efficacy of trastuzumab combined with chemotherapy in improving the survival rate of HER2-positive metastatic GAC patients has been demonstrated in previous studies [25]. However, the significant heterogeneity of GAC presents challenges to achieving optimal outcomes with this treatment approach. We trained a classifier using pretreatment samples obtained from patients with documented trastuzumab responses, achieving an AUC value of 0.833 by 10-fold cross-validation. To our knowledge, this study represents the first reported application of CNN-based DL algorithms to predict trastuzumab efficacy in GAC patients, showing the value of predicting anti-HER2 response efficacy.

However, there are several limitations to this study that we plan to address in future work. First, our model demonstrated lower accuracy in predicting HER2 status in gastric cancer biopsy specimens and lacked prospective samples for validation. Future efforts will involve collecting more biopsy specimens and prospective samples to further optimize the model for clinical applicability. Second, incorporating multicenter data can enable the model to learn more diverse features and enhance its generalizability. We will focus on collecting additional cases from other centers to improve the generalization capability of a model. Finally, future work will combine current imaging data with clinical data (such as tumor manifestations and circulating tumor markers) or multimodal features (such as radiomic features) to construct a multimodal DL model. Additionally, we should explore more intuitive visualization methods to elucidate the black-box properties of HER2 positivity and the efficacy of trastuzumab treatment in GAC.

Conclusion

In this study, we developed a histopathology-based model to predict HER2 status and response to trastuzumab treatment from H&E-stained histopathological images. We anticipate that this approach could be extended to other cancer types and treatments in future work, and we plan to investigate the hierarchical structure of features extracted from H&E-stained histopathological images and aim to elucidate the unknown black-box of this model for predicting HER2 status and the efficacy of trastuzumab treatment.

Abbreviations

GAC	Gastric adenocarcinoma
HER2	Human epidermal growth factor receptor 2
IHC	Immunohistochemistry
FISH	Fluorescence in situ hybridization
DL	Deep learning
CNN	Convolutional neural network
AUC	Area under the curve
H&E	Hematoxylin and eosin
EBV	Epstein-Barr virus
MSI	Microsatellite instability
CR	Complete response
PR	Partial response
SD	Stable disease
PD	Progressive disease
WSIs	Whole slide images
ACC	Accuracy
MIL	Multiple-instance learning
CLAM	Clustering-constrained-attention multiple-instance learning
AttMIL	Attention-based MIL
TransMIL	Transformer-based MIL

Acknowledgements

Thanks to the Pathology Department of Fujian Cancer Hospital for their invaluable contribution in providing both the essential research platform and profound histological expertise. Equally deserving of recognition is Fuzhou University, whose steadfast and unwavering technical support played a pivotal role in advancing our endeavors.

Author contributions

Zhida Wu: Investigation, Methodology, Data acquisition, Pathology evaluation and Writing. Tao Wang: Statistic analysis, Manuscript drafting, Deep learning-based approach designation and Implementation. Junlin Lan: Image processing, Collecting materials. Jianchao Wang: Manuscript guidance, Revision. Gang Chen: Acquisition, Analysis. Tong Tong: Resources, Supervision. Hejun Zhang: Administrative, Material support, Validation.

Funding

This research was supported by Startup Fund for Scientific Research, Fujian Medical University (Grant number 2022QH1160), Joint fund for the Science and Technology Program of Fujian Province (Grant number 2023Y9425), Scientific Research Foundation of Fujian Cancer Hospital (Grant number 2022YNG08), Fujian Provincial Health and Family Planning Research Talent Training Program (Grant number 2020GGB009), National Natural Science Foundation of China (Grant number 62171133).

Data availability

The code employed for the analysis in this paper is available online at https://github.com/ortonwang/Her2-Net.

Declarations

Ethics approval and consent to participate

The study received ethical approval from the Ethics Committee of Fujian Cancer Hospital, and the requirement to obtain written informed consent was waived due to the nature of the retrospective study.

Consent for publication

Written informed consent was obtained from all participants, and all procedures were conducted in compliance with applicable guidelines and regulations.

Competing interests
The authors declare no competing interests.

Received: 7 August 2024 / Accepted: 25 December 2024
Published online: 06 January 2025

References

1. Sung H, Ferlay J, Siegel RL, 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. Cancer J Clin. 2021;71:209–49.
2. Li H, Zhang H, Zhang H, Wang Y, Wang X, Hou H. Survival of gastric cancer in China from 2000 to 2022: a nationwide systematic review of hospital-based studies. J Global Health. 2022;12.
3. Pihlak R, Fong C, Starling N. Targeted therapies and developing Precision Medicine in Gastric Cancer. Cancers. 2023;15.
4. Pous A, Notario L, Hierro C, Layos L, Buges C. HER2-Positive gastric Cancer: the role of Immunotherapy and Novel therapeutic strategies. Int J Mol Sci. 2023;24.
5. Romond EHP, Edith A, Bryant J, Suman VJ, Geyer CE, Davidson, Nancy E, Tan-Chiu E, Martino S, Paik S. Kaufman, Peter A: Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. New Engl J Med. 2005;353:1673–84.
6. Ajani JA, D'Amico TA, Bentrem DJ, Chao J, Cooke D, Corvera C, Das P, Enzinger PC, Enzler T, Fanta P, et al. Gastric Cancer, Version 2.2022, NCCN Clinical Practice guidelines in Oncology. J Natl Compr Canc Netw. 2022;20:167–92.
7. Wakatsuki T, Yamamoto N, Sano T, Chin K, Kawachi H, Takahari D, Ogura M, Ichimura T, Nakayama I, Osumi H, et al. Clinical impact of intratumoral HER2 heterogeneity on trastuzumab efficacy in patients with HER2-positive gastric cancer. J Gastroenterol. 2018;53:1186–95.
8. Kim CG, Jung M, Kim HS, Lee CK, Jeung HC, Koo DH, Bae WK, Zang DY, Kim BJ, Kim H, et al. Trastuzumab Combined with Ramucirumab and Paclitaxel in patients with previously treated human epidermal growth factor receptor 2-Positive Advanced Gastric or Gastroesophageal Junction Cancer. J Clin Oncol. 2023;41:4394–405.
9. Zheng X, Wang R, Zhang X, Sun Y, Zhang H, Zhao Z, Zheng Y, Luo J, Zhang J, Wu H et al. A deep learning model and human-machine fusion for prediction of EBV-associated gastric cancer from histopathology. Nat Commun. 2022;13.
10. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, Marx A, Boor P, Tacke F, Neumann UP, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat Med. 2019;25:1054–6.
11. Shah MAK, Erin B, Alarcon-Rozas AE. Rajdev, Lakshmi: Immunotherapy and targeted Therapy forAdvanced Gastroesophageal Cancer:ASCO Guideline. J Clin Oncol. 2023;41:1470–91.
12. Bang Y-J, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, Lordick F, Ohtsu A, Omuro Y, Satoh T, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. Lancet. 2010;376:687–97.
13. Schwartz LH, Litiere S, de Vries E, Ford R, Gwyther S, Mandrekar S, Shankar L, Bogaerts J, Chen A, Dancey J, et al. RECIST 1.1-Update and clarification: from the RECIST committee. Eur J Cancer. 2016;62:132–7.
14. Luo Y, Chen J, Huang K, Lin Y, Chen M, Xu L, Li ZP, Feng ST. Early evaluation of sunitinib for the treatment of advanced gastroenteropancreatic neuroendocrine neoplasms via CT imaging: RECIST 1.1 or Choi Criteria? BMC Cancer. 2017;17:154.
15. Qiu ML, Guan M, Sun W-L. Yu-Ting;: Comparison of response evaluation criteria in solid tumors and tumor regression grade in evaluating the effect of preoperative systematic therapy of gastric cancer. J Clin Oncol. 2022 Jun 1;40:e16063.
16. Lan J, Cai S, Xue Y, Gao Q, Du M, Zhang H, Wu Z, Deng Y, Huang Y, Tong T, Chen G. Unpaired stain style transfer using Invertible neural networks based on Channel attention and long-range residual. IEEE Access. 2021;9:11282–95.
17. Zhang Y, Deng Y, Zou Q, Jing B, Cai P, Tian X, Yang Y, Li B, Liu F, Li Z, et al. Artificial intelligence for diagnosis and prognosis prediction of natural killer/T cell lymphoma using magnetic resonance imaging. Cell Rep Med. 2024;5:e16063–16063.
18. Schrammen PL, Ghaffari Laleh N, Echle A, Truhn D, Schulz V, Brinker TJ, Brenner H, Chang-Claude J, Alwers E, Brobeil A, et al. Weakly supervised annotation-free cancer detection and prediction of genotype in routine histopathology. J Pathol. 2021;256:50–60.
19. Vuong TTL, Song B, Kwak JT, Kim K. Prediction of Epstein-Barr Virus Status in gastric Cancer biopsy specimens using a deep learning algorithm. JAMA Netw Open. 2022;5:e2236408–18.
20. Cao R, Yang F, Ma S-C, Liu L, Zhao Y, Li Y, Wu D-H, Wang T, Lu W-J, Cai W-J, et al. Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in Colorectal Cancer. Theranostics. 2020;10:11080–91.
21. Kaito A, Kuwata T, Tokunaga M, Shitara K, Sato R, Akimoto T, Kinoshita T. HER2 heterogeneity is a poor prognosticator for HER2-positive gastric cancer. World J Clin Cases. 2019;7:1964–77.
22. Tanei T, Seno S, Sota Y, Hatano T, Kitahara Y, Abe K, Masunaga N, Tsukabe M, Yoshinami T, Miyake T et al. High HER2 Intratumoral heterogeneity is a predictive factor for poor prognosis in early-stage and locally Advanced HER2-Positive breast Cancer. Cancers (Basel). 2024;16.
23. Bychkov D, Linder N, Tiulpin A, Kücükel H, Lundin M, Nordling S, Sihto H, Isola J, Lehtimäki T, Kellokumpu-Lehtinen P-L, et al. Deep learning identifies morphological features in breast cancer predictive of cancer ERBB2 status and trastuzumab treatment efficacy. Sci Rep. 2021;11:4037–46.
24. Farahmand S, Fernandez AI, Ahmed FS, Rimm DL, Chuang JH, Reisenbichler E, Zarringhalam K. Deep learning trained on hematoxylin and eosin tumor region of interest predicts HER2 status and trastuzumab treatment response in HER2 + breast cancer. Mod Pathol. 2022;35:44–51.
25. Ha GY, Yang SH, Kang HJ, Lee HL, Kim J, Kim YJ, Yu HJ, Lee JI, Jin SH. Comparison of survival outcomes according of patients with metastatic gastric cancer receiving trastuzumab with systemic chemotherapy. Korean J Clin Oncol. 2020;16:63–70.

Publisher's note