International Journal of Medical Informatics 215 (2026) 106455

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

Original Research Article

Artificial intelligence-based reclassification of gastric adenocarcinoma enables prognostic stratification via diffuse-type patch proportion

Cătălin Andraș [a,b], Corina-Elena Minciună [a,b,*], Dragos Stan [c], Stefan Tudor [a,b], Teodora Manuc [b,d], Mihnea P. Dragomir [e,f,g], Ovidiu Bitere [a,b], Alexandru Micu [a,b], Florina Almarii [b,h], Gabriela Droc [b,i], George Calin [j,k], Vlad Herlea [b,h,*], Catalin Vasilescu [a,b]

[a] General Surgery Department, Fundeni Clinical Institute, Bucharest, Romania
[b] "Carol Davila" University of Medicine and Pharmacy, Bucharest, Romania
[c] Businessdata Systems SRL, Bucharest, Romania
[d] Department of Gastroenterology, Fundeni Clinical Institute, Bucharest, Romania
[e] Institute of Pathology, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin 10117 Berlin, Germany
[f] German Cancer Consortium (DKTK), Partner Site Berlin, and German Cancer Research Center (DKFZ), Heidelberg, Germany
[g] Berlin Institute of Health (BIH), Berlin, Germany
[h] Department of Pathology, Fundeni Clinical Institute, Bucharest, Romania
[i] Department of Anesthesiology and Intensive Care I, Fundeni Clinical Institute, 022328 Bucharest, Romania
[j] Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[k] Center for RNA Interference and Non-Coding RNAs, The University of Texas MD Anderson Cancer Center, Houston, TX 77054, USA

ARTICLE INFO

Keywords:
Artificial Intelligence
Gastric Cancer
Laurén Classification
Survival Predictor
Risk Assessment

ABSTRACT

Background: Gastric adenocarcinoma (GAC) remains a major global health burden with marked heterogeneity, complicating diagnosis and prognostic assessment. The Laurén classification, though widely used, suffers from interobserver variability, particularly in defining the mixed subtype. Artificial intelligence (AI)-driven image analysis may improve standardization and prognostic assessment in GAC.

Methods: We retrospectively analyzed 404 patients with resected GAC (2015–2022) from Fundeni Clinical Institute. Whole-slide images (WSIs) were annotated by pathologists with expertise and processed into patches for training a two-stage deep learning pipeline based on YOLO26m-cls. The first model (GAC-I) distinguished malignant from non-malignant tissue, while the second (GAC-ST) classified malignant patches as intestinal or diffuse. We developed the diffuse prognostic score (DPS), defined as the proportion of diffuse patches relative to total patches, and correlated it with overall survival (OS).

Results: GAC-I and GAC-ST achieved high diagnostic performances, with accuracies of 0.9437 ± 0.0317 (F1 score: 0.9456 ± 0.0243) and 0.8080 ± 0.0833 (F1 score: 0.7528 ± 0.1094). DPS ≥ 0.5 was significantly associated with lower median OS (16.1 months) compared to DPS < 0.5 (42.067 months), association confirmed by multivariate Cox-regression analysis (HR 3.88, $p < 0.001$) and matched case-control analysis. Groups were balanced across all variables except tumor differentiation, which was more frequently high-grade in DPS ≥ 0.5. After adjustment, DPS ≥ 0.5 remained an independent predictor of mortality (HR 2.684, $p = 0.027$).

Conclusion: We developed and validated a robust AI-based framework for automated GAC classification and prognostic stratification using H&E WSIs. DPS is an independent, reproducible marker of OS, supporting its potential integration into clinical pathology workflows to guide personalized treatment.

Abbreviations: AI, artificial intelligence; CI, confidence interval; CLAM, Clustering-constrained Attention Multiple Instance Learning; CNN, Convolutional Neural Network; DL, deep learning; DPS, Diffuse prognostic score; DSMIL, Dual-Stream Multiple Instance Learning; GAC, gastric adenocarcinoma; H&E, Hematoxylin-eosin staining; MIL, Multiple Instance Learning; OS, overall survival; SD, standard deviation; TCGA, The Cancer Genome Atlas Program; WSI, Whole Slide Image.

* Corresponding authorsat: "Carol Davila" University of Medicine and Pharmacy, Bucharest, Romania.
E-mail addresses: corina.minciuna@umfcd.ro (C.-E. Minciună), vlad.herlea@umfcd.ro (V. Herlea).

https://doi.org/10.1016/j.ijmedinf.2026.106455
Received 29 October 2025; Received in revised form 20 April 2026; Accepted 21 April 2026
Available online 24 April 2026

1. Introduction

Gastric cancer remains a significant burden on the healthcare systems worldwide despite recent advances in diagnosis, as well as neoadjuvant and adjuvant therapies, including the introduction of immunotherapy [1]. It remains a major global health concern, ranking fifth in incidence and cancer-related mortality [2]. With most of the gastric cancer being gastric adenocarcinomas (GAC) [3] and considering the varied responses to immunotherapy, such differences were attributed to the extensive tumor heterogeneity [4]. Since then, several classifications emerged for GAC: the first one used, instated in 1965, was Laurén classification [5] soon followed by the WHO [6], Japanese [7], South-Korean [8] classifications and also complemented by molecular ones: the intrinsic subtypes, Lei subtypes, The Cancer Genome Atlas (TGCA), Asian Cancer Research Group (ACRG) and others [9–11].

Although the Laurén classification, which categorizes GAC into intestinal and diffuse types, has been widely used for decades, it underwent a major revision in 1995 when F. Carneiro introduced the mixed subtype [12]. This addition was based on the observation that many GACs exhibit both intestinal and diffuse morphological features. Carneiro defined the mixed type by the existence of more than one morphologic component (intestinal and diffuse), provided that the smaller component occupies at least 5% of the reviewed slides [12].

Despite its clinical relevance, there is still no universally accepted definition of the mixed subtype. Ongoing debate persists regarding the precise proportion of intestinal versus diffuse components necessary for a tumor to be classified as mixed, highlighting the complexity and heterogeneity of GAC [12–15]. The absence of standardized criteria for mixed GAC makes prognostic evaluation and survival analysis challenging, reinforcing the need for improved, quantifiable frameworks to better understand and stratify Laurén subtypes.

The histopathological assessment of GAC is highly dependent on the pathologist's expertise and inherently time-consuming, even in expert hands. In practice, this involves reviewing clinical information, performing a gross examination of the gastric tissue, selecting representative areas for sectioning, all this before evaluating the basic hematoxylin and eosin (H&E) slide. For the Laurén classification, particular attention must be paid to architectural and cytological features to distinguish between intestinal, diffuse, and mixed types, a task that can be nuanced and subjective, further emphasizing the need for standardization of the Laurén subtypes. Each added layer of classification increases the diagnostic burden, contributing to an already demanding workload that includes high case volumes, tight turnaround times, and the expectation of a precise diagnostic. Even though the recent molecular classifications emerged to aid the stratification of GAC, the associated costs remain prohibitively high.

The introduction of digital pathology aided the implementation of artificial intelligence (AI) and machine learning, especially deep learning methods like Convolutional Neural Networks (CNNs) in pathology. Deep-learning (DL) classification systems have the potential to solve digital pathology issues. The current models are typically based on datasets from publicly available images, most widely used being TCGA [3,16,17]. Until recently, these systems depended on well annotated and large sets of data, but new methods have been developed that can perform a fully automated demarcation of any tumor type [1]. The development of digital pathology with the aid of AI may be the key for a proper and cost-efficient evaluation, possibly with some key findings that may be already found on the histopathological slide that are missing to the human eye.

Despite these advancements, it is well established that CNNs are subject to multiple limitations arising from both technical and human-related factors, including data quality, annotation variability, and image acquisition inconsistencies [18]. In our study, similar challenges are anticipated; therefore, efforts will be made to mitigate their impact through the application of rigorous exclusion criteria and the involvement of experienced image engineering. Although the CNNs can be trained to recognize specific features of cancerous tissue, their ability to predict clinical outcomes, such as survival rates, response to chemotherapy or immunohistopathological markers are yet to be evaluated [19].

Another notable review ascertains that although digital pathology is a rapidly expanding field with a growing number of studies published each year, the results are rarely translated into routine diagnostic practice [20], highlighting the need for more clinically oriented validation studies.

Summing up, this study aims to deepen the understanding of GAC through the application of advanced AI-driven image analysis methods. The primary objective of this study was to develop a CNN-based model capable of accurately identifying GAC on WSIs and subsequently classify it into either the intestinal or diffuse subtype according to the Laurén classification [5]. By focusing on these two well-defined categories, the model bypasses the ongoing controversy and variability in the literature surrounding the definition and interpretation of the mixed subtype. The secondary objective was to develop a prognostic score for patients with GAC, aiming to provide a reliable and efficient tool for clinical practice with the goal of improving patient care.

2. Methods

2.1. Patient cohort

This retrospective observational study included all patients diagnosed with GAC that underwent surgical resection between 2015–2022 in the General Surgery Department of Fundeni Clinical Institute. The relevant clinical data were recorded (e.g. age, sex, time of operation, biological tumoral markers, histopathologic characteristics). From the initial 450 patients, those with poor slide quality and multiple tumor types on the same slide (as reviewed by the pathology team) were excluded. Another exclusion criteria used was the postoperative demise as defined by the current literature, less than one month from surgery [21]. The final study cohort comprised 404 patients.

2.2. Image acquisition

For each patient, the formalin-fixed, paraffin-embedded (FFPE) tissue block exhibiting the deepest tumor infiltration was selected by a pathologist. From each block, 5-μm-thick sections were prepared and stained with H&E. High-resolution WSIs of gastric tissue were obtained using a digital slide scanner, Grundium Ocus40. These slides are digitized at diagnostic magnification levels of 40x, capturing rich morphological details necessary for analysis.

2.3. Datasets and preprocessing

Following acquisition, preprocessing steps have been performed to prepare the images for computational analysis. These steps include color normalization to mitigate staining variability across samples, removal of scanning artifacts or out-of-focus regions, and division of WSIs into smaller, manageable image patches, which serve as the basic input units for model training.

The original images, in *.svs format, were converted to *.png or *.tif for downstream processing. To reduce computational load, images were resized to two levels: Level 0 (106,496 × 98,304 pixels) and Level 1 (26,624 × 24,576 pixels). Preliminary verification was performed on Level 1 images to confirm the consistency of annotations.

To enable model training, each image was split into smaller patches. Following a comparative analysis of patch sizes (224 × 224, 640 × 640, and 1280 × 1280), two pathologists concluded independently that 640 × 640 pixels was optimal for preserving sufficient diagnostic detail. Patches that were entirely grey or white were discarded, as they lacked tissue content and hence diagnostic value. Patches with less than 2% of their area annotated were labeled as "unknown." In cases where

multiple tissue classes were present in a single patch, the patch was classified according to the majority class, with one exception: if a patch contained both Normal and Malignant (Intestinal or Diffuse) tissue, it was automatically labeled as Malignant.

2.4. Expert annotation

A team composed of 2 pathologists with expertise performed manual annotation of the WSIs. Specific regions of interest are labeled to differentiate between cancerous and non-cancerous tissue areas based on histopathological criteria. These expert annotations are used to generate ground truth labels for supervised learning, ensuring that the CNN models are trained on accurate datasets. This step is essential to guide the networks toward learning diagnostically meaningful patterns within the histological patterns.

A total of 58 WSIs were independently annotated by two pathologists, blinded to clinical outcomes (F.A. and V.H.), according to four primary tissue classes: Normal, Normal Non-mucosa, Intestinal Subtype (Malignant), and Diffuse Subtype (Malignant). In cases of disagreement, a joint consensus review was performed and final labels were assigned accordingly. Standardized Laurén classification criteria and predefined annotation rules were applied, and annotation quality was verified after patch extraction to ensure consistency and reduce variability. Annotation geometries included both polygons and multi-polygons and were initially saved in *.geofson format, later converted to *.txt files to ensure compatibility with the training pipeline.

Two distinct datasets were created. Dataset A, used for training GAC identification model (GAC-I), aimed to differentiate between Positive (adenocarcinoma, including both subtypes) and Negative (Normal and Normal non-mucosa) patches. Dataset B, used for training GAC subclassifying model (GAC-ST), focused on classifying the Positive patches with high confidence into either the Intestinal or Diffuse subtype. Each dataset contained 58 WSIs for GAC-I with a total of 291.000 patches and 55 WSIs for GAC-ST with a total of 111.775 patches.

2.5. Development of CNN models

Model Overview To achieve the aims of this study, a YOLO26m-cls classification model, developed by Ultralytics, was utilized. YOLO (You Only Look Once) is a state-of-the-art real-time object detection algorithm originally introduced by J. Redmon et al [22]. YOLO26m-cls offers significant advancements in both speed and accuracy over its predecessors and is particularly well-suited for tasks involving large histological images.

At the outset, a single CNN model, referred to as GAC-O, was tested to perform multi-class categorization, distinguishing between Normal, Intestinal, and Diffuse tissue subtypes. Despite increasing the number of training images and applying extensive data augmentation, GAC-O did not achieve satisfactory performance. Consequently, the architecture was restructured into a two-step cascade model: GAC-I designed to distinguish between Normal and Malignant tissue, and GAC-ST intended to further classify the malignant patches identified by GAC-I into Intestinal and Diffuse subtypes. This cascade approach allowed for improved control of classification accuracy and addressed the limitations observed in the initial model.

Architecture The architecture of YOLO26m-cls was selected for its ability to optimize spatial information processing while maintaining high-speed inference, a key requirement for whole-slide image analysis. The backbone of YOLO26m-cls includes convolutional blocks, bottleneck structures, and a Cross Stage Partial Focus (CSP-Focus) mechanism designed to preserve feature maps efficiently. The C3K2 block handles deep feature extraction effectively.

The neck of the architecture incorporates Spatial Pyramid Pooling Fast (SPFF) and up sampling layers to aggregate context across multiple scales. Attention mechanisms, particularly the C2PSA block, are employed to enhance the model's sensitivity to important image regions. The C2PSA introduces position-sensitive attention, allowing the model to focus on diagnostically relevant areas while maintaining a balance between computational cost and detection accuracy. The head of the model supports multi-scale predictions, enabling the detection of tissue anomalies at various levels of resolution.

Training and Augmentation To ensure methodological robustness and prevent data leakage, a five-fold cross-validation scheme was implemented at the slide level using YOLO26m-cls configuration. To improve model robustness and generalization, several data augmentation techniques were applied. These included color manipulations through HSV adjustments (hue = 0.015, saturation = 0.7, brightness = 0.4), random image rotations up to 45 degrees, and scaling by a factor of 0.1. Additional transformations included vertical and horizontal flipping (each with a 50% probability), mosaic augmentation (merging four training images into one), and full central cropping. These augmentations simulated variability in tissue presentation, lighting, and orientation, thereby enhancing the model's capacity to generalize to unseen samples.

Data splitting was performed at the whole-slide image (WSI) level to prevent data leakage between training and validation datasets. Specifically, all image patches extracted from a given WSI were assigned exclusively to a single fold during the five-fold cross-validation procedure. This approach ensures that no patches originating from the same slide are present in both training and validation sets, thereby maintaining full data independence and enabling a realistic evaluation of model performance on previously unseen patient slides.

Given that the goal of the study is clinical evaluation of a pipeline rather than architectural innovation, extensive ablation of the underlying model was not performed. Instead, robustness is assessed through slide-level cross-validation and multi-fold consistency.

Extensive information regarding the model architecture and performance metrics can be found in the Supplementary AI Data.

2.6. Statistical analysis

Categorical variables were reported as frequencies and percentages and were compared by Chi-Square test. For categorical variables with less than 5 cases per subcategory Fisher's Exact Test was used. Continuous variables were presented as mean ± standard deviation (SD) or median and interquartile range, after checking for normality.

Overall survival (OS) was defined as the time interval from surgery to either the patient's death or the date of the last follow-up if the patient was still alive at that time. The OS rates were estimated with Kaplan-Meyer method, and they were compared between different groups by Log-rank test. Univariate analysis and multivariate analysis have been performed to assess the prognostic factors associated with OS using Cox regression analysis. All variables associated with a p value ≤ 0.05 in univariate analysis were included in multivariate analysis, that was realized with a stepwise Cox regression using the Backward: Wald method taking into consideration the size of the database.

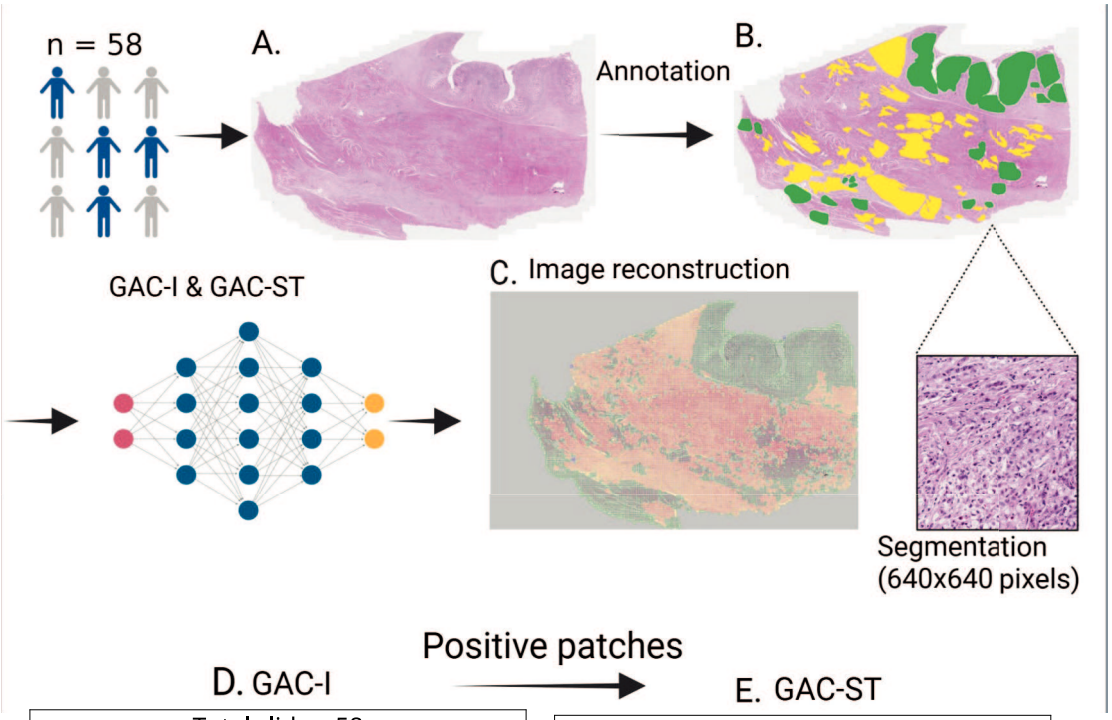
The differences with a p-value < 0.05 were considered statistically significant. All statistical analyses were performed using IBM SPSS Statistics for Windows (version 26.0, IBM, Chicago, IL).

3. Results

3.1. Classifier development

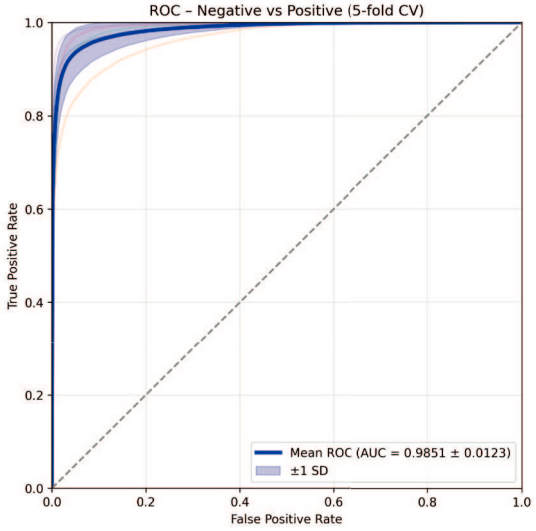
To train the CNN models, WSIs from 58 patients were initially reviewed and annotated by two pathologists with expertise (Fig. 1**A** and **B**). The annotated regions were then subdivided into non-overlapping patches of 640 × 640 pixels (Fig. 1**C**). From the total of 58 annotated cases, we ended up with 291.000 patches. Five mutually exclusive folds were constructed, with approximately 46–47 slides used for training and 11–12 slides for validation in each iteration, and each slide serving exactly once as validation.

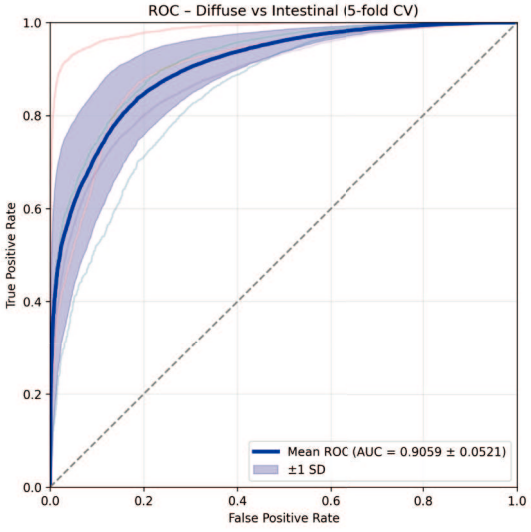
The first model, termed GAC-I, was developed to perform binary



Total slides: 58	
Total patches: ≈ 291.000	
Metrics (mean ± SD across folds)	
Accuracy	94.37% ± 3.17%
Precision	95.90% ± 1.07%
Sensitivity	93.28% ± 3.76%
Specificity	95.55% ± 2.16%
F1-score	94.56% ± 2.43%
AUC	0.9851 ± 0.0123
AP	0.9876 ± 0.0067

Total slides: 55	
Total patches: 111.775	
Metrics (mean ± SD across folds)	
Accuracy	80.80% ± 8.33%
Precision	76.52% ± 23.59%
Sensitivity	79.36% ± 11.48%
Specificity	86.13% ± 4.18%
F1-score	75.28% ± 10.94%
AUC	0.9059 ± 0.0521
AP	0.8987 ± 0.1187





(caption on next page)

Fig. 1. Overview of the study pipeline and performance of deep learning models for gastric adenocarcinoma classification. GAC-I and GAC-ST form a two-stage deep learning pipeline enabling automated and high-confidence subclassification of gastric adenocarcinoma. Created in https://BioRender.com. **A.** Original whole slide image (WSI) stained with hematoxylin and eosin (H&E) demonstrating a representative gastric adenocarcinoma sample prior to annotation or model processing. **B.** Annotated image used for training the models. Regions are labeled by pathologists with expertise as normal tissue (green) and malignant intestinal-type tissue (yellow). These annotations provided the ground truth for training the GAC-I and GAC-ST models. **C.** Reconstructed image generated by applying the trained convolutional neural networks (CNNs) to unseen WSIs. Only image patches with prediction confidence ≥ 90% were used to ensure classification robustness. These reconstructions served to visually confirm the model's ability to replicate histological architecture. **D: GAC-I**: Initial binary classification CNN designed to differentiate malignant (positive) from non-malignant (negative) gastric tissue. Achieved an accuracy of 0.9437 ± 0.0317 and an F1-score of 0.9456 ± 0.0243. **E: GAC-ST**: Subsequent subtype-classification CNN that analyzed only cancer-positive patches from GAC-I. It classified each patch into intestinal or diffuse Laurén subtypes. Achieved an accuracy of 0.8080 ± 0.0833 and F1-score of 0.7528 ± 0.1094. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

classification, distinguishing between cancerous and non-cancerous gastric tissue. This model served as the primary diagnostic filter, identifying relevant tumor regions for subsequent subtype classification. GAC-I was trained using the annotated patches and validated on an independent subset to evaluate achieving an accuracy across five validation folds of 0.9437 ± 0.0317 and an F1-score of 0.9456 ± 0.0243, AUC of 0.9851 ± 0.0123 and AP of 0.9876 ± 0.0067 indicating high reliability in GAC identification (Fig. 1**D**).

To improve model performance and diagnostic reliability, a stepwise, two-stage analysis pipeline was adopted. Image patches classified as cancerous by the initial binary model (GAC-I) were subsequently analyzed by a second CNN, termed GAC-ST. This model was specifically designed to distinguish between histological subtypes of GAC—namely, intestinal and diffuse types. Like GAC-I, GAC-ST was trained and validated using expert-annotated image patches, but it operated exclusively on cancer-positive regions. This focused approach enabled the model to learn subtype-specific histological features with higher fidelity. GAC-ST achieved an overall accuracy of across five validation folds 0.8080 ± 0.0833, F1-score of 0.7528 ± 0.1094, AUC of 0.9059 ± 0.0521 and an AP of 0.8987 ± 0.1187 (Fig. 1**E**).

3.2. Testing the classifier

After independent training and validation, the two-stage pipeline was applied to the full cohort of 404 GAC patients. For each case, the WSI was processed sequentially through GAC-I and GAC-ST, allowing for automated localization of tumor regions followed by histological subtyping.

To further validate model reliability, reconstructed WSIs were generated using only those patches that achieved classification confidence above 90% (Fig. 1**C**). These reconstructions were then visually compared with the original slides (Fig. 2), providing qualitative confirmation of the models' precision. Together, the GAC-I and GAC-ST models form a highly accurate, confidence-aware pipeline for automated GAC identification and subtyping based on WSIs.

3.3. AI classifier score predicts OS

Among the variables analyzed in the Cox univariate analysis, the following were identified as statistically significant predictors of overall survival: gender, residual tumor status (R), TNM classification, tumor differentiation grade, Dindo-Clavien classification, positive surgical margin, total number of retrieved lymph nodes, number of positive lymph nodes, and tumor diameter (Fig. 3**A-E**, Supplementary Table 3).

After extensive testing and analysis of the values generated by the CNN models, we evaluated various prognostic markers to identify the factor that best correlates with OS, the corresponding data can be found in Supplementary Table 3 and also in Supplementary Data 2.

Because no significance was achieved for the intestinal and diffuse annotations alone of the CNN classifier, we looked for a novel factor to integrate the potential predictive power of both models, termed the Diffuse Prognostic Score (DPS). This score was selected based on its higher hazard ratio (HR).

$$DPS = \frac{Number\ of\ diffuse\ patches}{Number\ of\ positive\ patches + \ Number\ of\ negative\ patches}$$

The DPS score demonstrated a significant association with OS (p-value < 0.0001), highlighting its potential as a strong predictor of survival outcomes. These findings emphasize the prognostic importance of specific factors, such as the presence of positive and diffuse patches, in predicting OS, whereas other factors, including total intestinal patches and total negative patches, do not appear to have a significant impact on survival (Fig. 3**D** and **E**).

Independent predictive factors for improved long-term outcomes were identified by incorporating variables with a p-value ≤ 0.05 from the univariate analysis into the multivariate analysis. To minimize collinearity, variables representing overlapping or dependent clinical information (e.g., TNM components and related composite variables) were excluded from simultaneous inclusion in the multivariate Cox regression model, ensuring model stability and interpretability. Notably, gender, tumor stage, Dindo-Clavien classification, and the DPS score were found to be significant predictors of OS regardless of any other factors (p < 0.005) (Fig. 3**F** and Supplementary Table 4).

3.4. Characterization of the DPS high cases, and DPS prognostic power

To further investigate the clinical relevance of the DPS, we used a threshold value of 0.5 to stratify cases and evaluate its impact on OS using Kaplan–Meier analysis. Additionally, we aimed to explore potential histological patterns associated with high and low DPS. A threshold of 0.5 was selected a priori without optimization to avoid potential bias; future studies will explore data-driven approaches for threshold determination and validation.

Pathologists' reanalysis of representative H&E-stained WSIs from cases with DPS < 0.5 and DPS > 0.5 showed that tumors with DPS > 0.5 exhibited more infiltrative growth patterns, discohesive cells, and poorly defined glandular structures—features typical of diffuse-type adenocarcinoma (Fig. 4**A**). In contrast, cases with DPS < 0.5 showed better gland formation and more cohesive tumor architecture, consistent with the intestinal subtype (Fig. 4**A**). Kaplan–Meier survival analysis revealed a significant difference in OS between patients with DPS > 0.5 and those with DPS < 0.5, further supporting the prognostic value of this image-derived marker (Fig. 4**B**). These observations support the biological plausibility of DPS as a meaningful quantitative marker, bridging morphological subtype with survival outcomes.

To assess its impact on survival, a matched case-control analysis was performed. Matching was performed on 33 cases using a combination of exact and fuzzy criteria. Exact matching variables included sex, tumor stage, and residual tumor status, while fuzzy matching allowed a tolerance of ± 1 for Dindo-Clavien classification and ± 10 years for age at diagnosis. Additional information regarding the matching can be found in Supplementary Data 1.

Tumor differentiation grade differed significantly between groups. A chi-square analysis revealed a strong association between tumor differentiation grade and group status ($\chi^2 = 16.216$, df = 2, p = 0.000). In the control group (Group 0- DPS < 0.5), most tumors were moderately differentiated (grade 2, 34.5%) or poorly differentiated (grade 3,

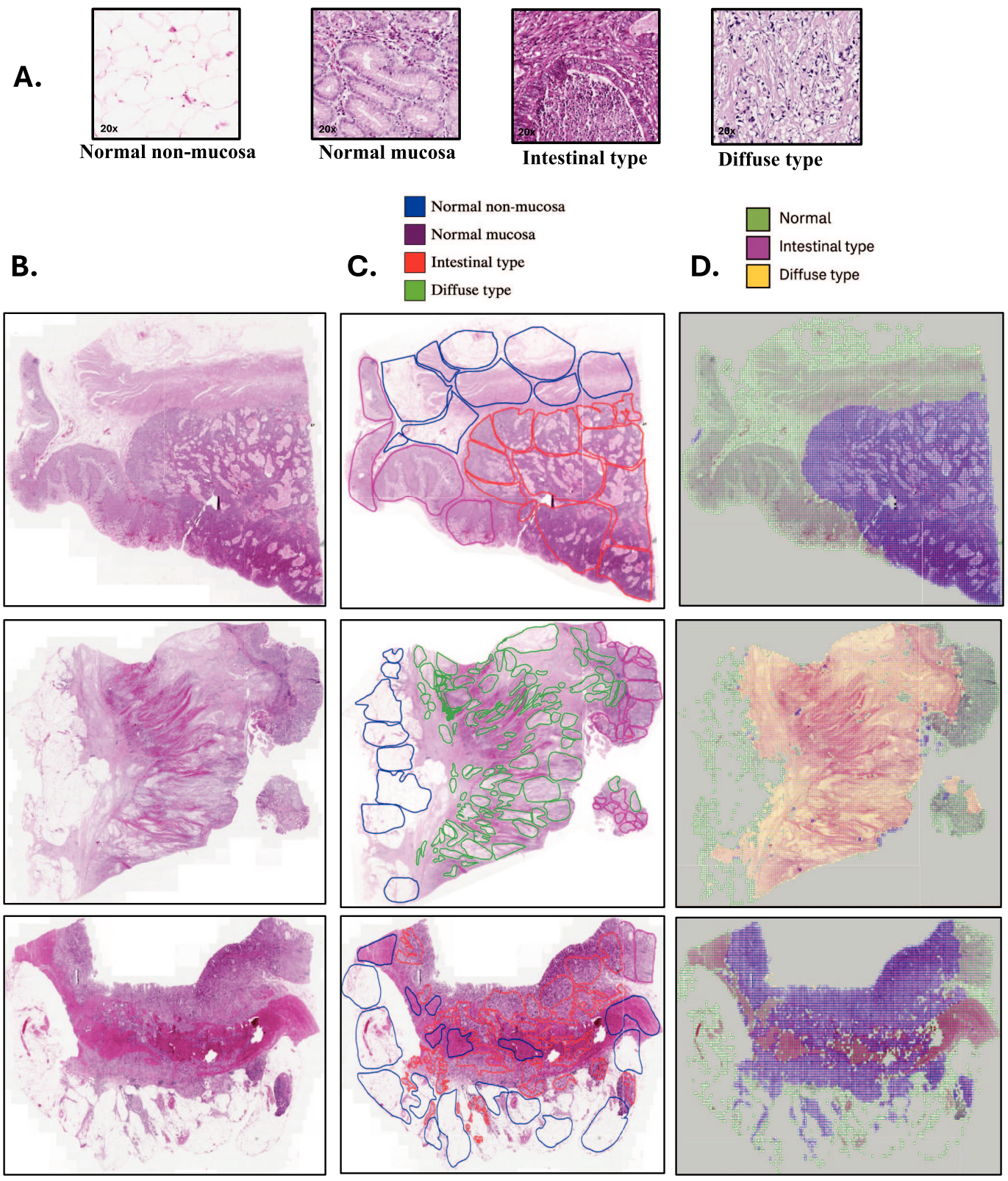


Fig. 2. Histopathological classification and annotation of gastric tissue types. **A.** Representative high-magnification (20 ×) H&E-stained image patches illustrating four gastric tissue types: Normal non-mucosa, Normal mucosa, Intestinal-type gastric cancer, and Diffuse-type gastric cancer. **B.** Whole-slide images (WSIs) of H&E-stained gastric tissue sections. **C.** Manual annotations outlining different tissue types: Normal non-mucosa (blue), Normal mucosa (magenta), Intestinal-type (red), and Diffuse-type (green). Automated model predictions overlaid on WSIs, showing classification into three major categories: Normal (green), Intestinal-type (magenta), and Diffuse-type (yellow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

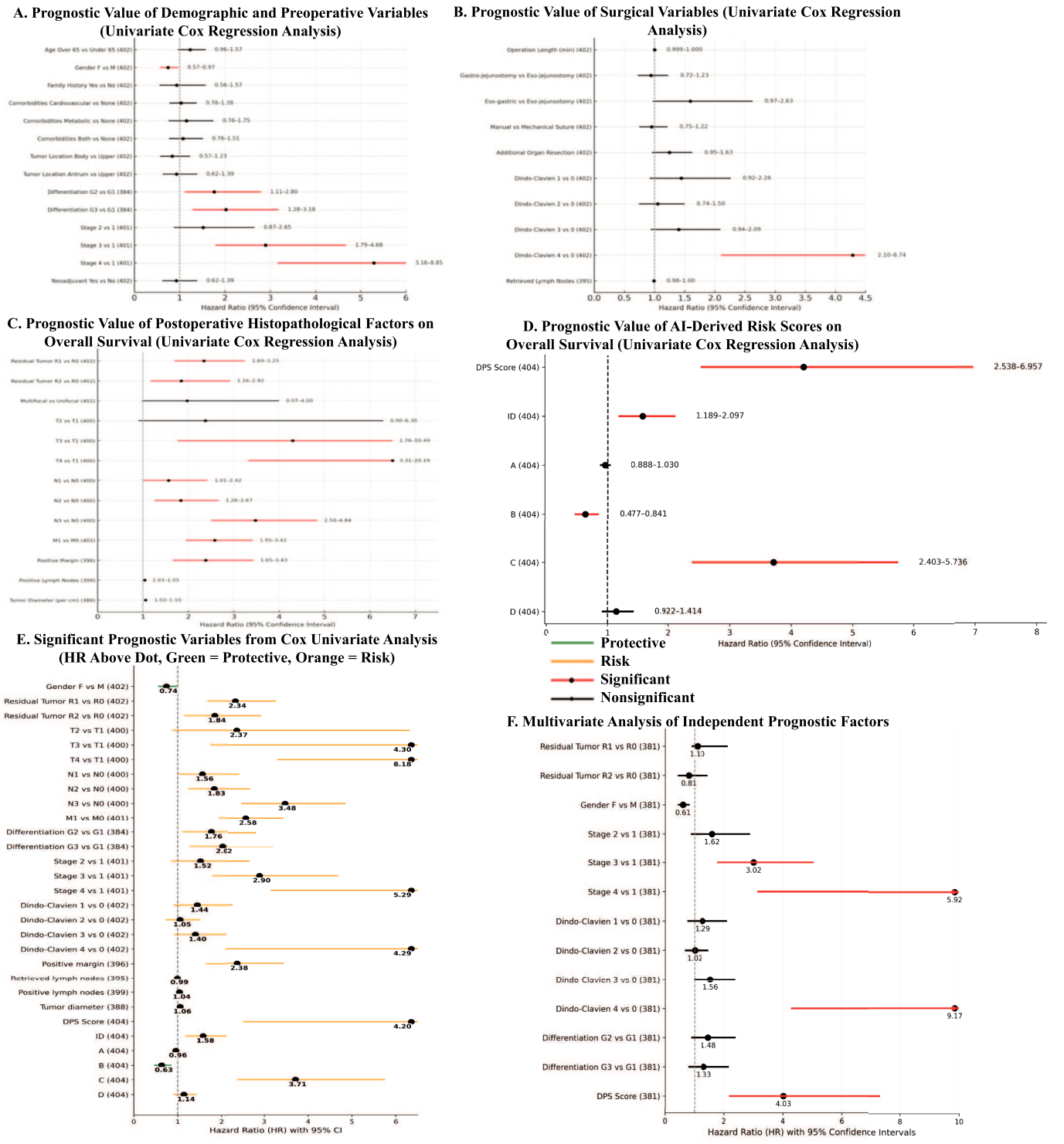
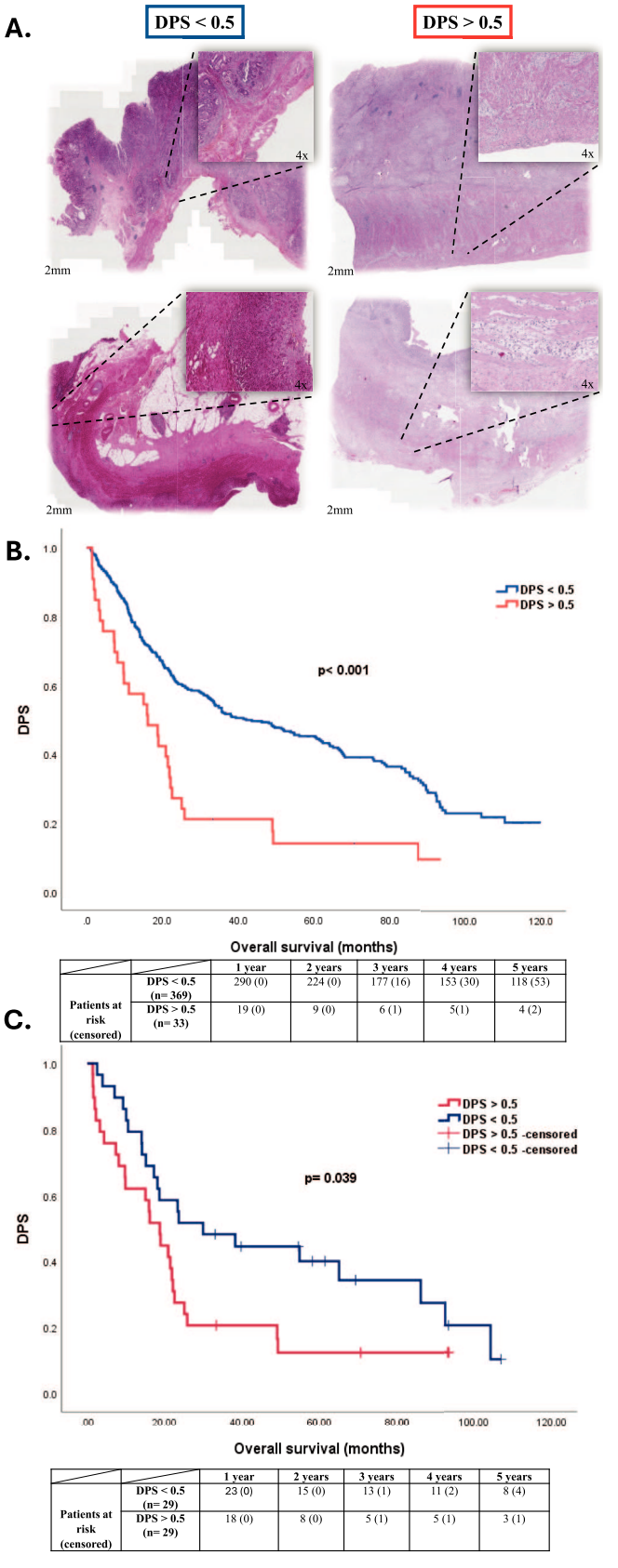


Fig. 3. Prognostic Factors Influencing Overall Survival in Gastric Adenocarcinoma Patients. This figure summarizes the prognostic value of clinical, pathological, and AI-derived variables for overall survival (OS) in patients with gastric adenocarcinoma using Cox regression analyses. **A.** Univariate Cox analysis of demographic and preoperative variables shows that tumor stage III–IV (HR: 2.895 in stage III, HR: 5.290 in stage IV, $p < 0.001$) and poor differentiation (G3) were significantly associated with reduced survival. **B.** Univariate analysis of surgical variables highlights that Dindo-Clavien grade 4 complications (HR: 4.286) were significantly associated with decreased survival. **C.** Postoperative histopathological factors, such as residual tumor (R1), advanced T-stage (T4), nodal involvement (N3), and the number of positive lymph nodes, were significant predictors of poor OS. **D.** AI-derived risk scores (DPS and sub-models A–D) demonstrated strong prognostic value, with higher risk scores significantly predicting worse outcomes (e.g., DPS Score HR: 4.202, CI: 2.538–6.957). **E.** Summary of significant prognostic variables from univariate analyses. Hazard ratios (HRs) above 1 (orange bars) indicate increased risk, while those below 1 (green bars) indicate protective effects. **F.** Multivariate Cox regression analysis identified residual tumor status, sex, advanced tumor stage, Dindo-Clavien complications, differentiation grade, and the DPS score as independent prognostic factors for overall survival. Hazard ratios are presented with 95% confidence intervals. Red markers indicate statistically significant risk factors, while black denotes nonsignificant results. AI-derived scores were computed based on deep learning risk stratification models trained on clinicopathologic features. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



		1 year	2 years	3 years	4 years	5 years
Patients at risk (censored)	DPS < 0.5 (n= 369)	290 (0)	224 (0)	177 (16)	153 (30)	118 (53)
	DPS > 0.5 (n= 33)	19 (0)	9 (0)	6 (1)	5(1)	4 (2)

		1 year	2 years	3 years	4 years	5 years
Patients at risk (censored)	DPS < 0.5 (n= 29)	23 (0)	15 (0)	13 (1)	11 (2)	8 (4)
	DPS > 0.5 (n= 29)	18 (0)	8 (0)	5 (1)	5 (1)	3 (1)

(caption on next column)

Fig. 4. Histopathological Features and Prognostic Significance of the Diffuse Prognostic Score (DPS). **A.** Representative H&E–stained WSIs of GAC cases stratified by DPS. Cases with DPS < 0.5 exhibit preserved glandular architecture, moderate cellular atypia, and organized stromal components. In contrast, cases with DPS > 0.5 display diffuse growth patterns, poorly differentiated tumor cells, and prominent desmoplastic stroma. Insets at 4 × magnification highlight key morphological differences captured by the AI-based DPS algorithm. **B.** Kaplan–Meier survival curve comparing overall survival between patients with DPS < 0.5 (blue line, n = 369) and DPS > 0.5 (red line, n = 33). Patients with high DPS scores had significantly poorer survival (p < 0.001). The accompanying risk table indicates the number of patients at risk and censored at yearly intervals up to 5 years post-diagnosis. **C.** Kaplan–Meier survival analysis of a matched subset of patients (n = 58) stratified by DPS < 0.5 (blue line) vs DPS > 0.5 (red line). DPS < 0.5 remained associated with significantly better overall survival (p = 0.039, log-rank test). Numbers at risk and censored are shown in the table below. DPS was computed using a deep learning pipeline trained on histological data. Survival differences suggest that DPS captures prognostically relevant morphologic features not evident in traditional grading systems. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

48.3%), with only 5 well-differentiated cases (grade 1, 17.2%). In contrast, the case group (Group 1- DPS ≥ 0.5) exhibited a marked predominance of poorly differentiated tumors (grade 3, 96.3%), with a small proportion of well-differentiated differentiated tumors (grade 1, 3.7%, one case) and no moderately differentiated tumors. This distribution indicates that poorly differentiated tumors were significantly more frequent in the case group. While the assumption of expected cell frequencies was partially violated (50% of cells had expected counts less than 5), the association remained statistically significant, as confirmed by the Pearson chi-square (p = 0.000), likelihood ratio test (p = 0.000), and a significant linear-by-linear association (p = 0.001). To account for residual confounding due to tumor differentiation grade, which remained imbalanced post-matching, we included it as a covariate in the multivariable Cox regression model. This approach allowed for adjustment of its potential effect on survival outcomes.

A Cox proportional hazards model was employed to evaluate the effect of DPS group status on OS, with adjustment for tumor differentiation grade. The model demonstrated statistical significance (Omnibus $\chi^2(3) = 7.996$, p = 0.046), indicating that the included variables contributed meaningfully to survival prediction. Group DPS high emerged as a significant independent predictor of survival (p = 0.027), with a hazard ratio (HR) of 2.68 (95% CI: 1.12–6.43). Tumor differentiation grade did not significantly influence survival in the model (p = 0.203); however, interpretation of its individual HR was limited due to extreme coefficient estimates and large standard errors, likely attributable to sparse data and quasi-complete separation in certain subgroups.

Survival outcomes were further evaluated using Kaplan–Meier estimates stratified by DPS. The median survival time for the case group (DPS ≥ 0.5) was 18.7 months (95% CI: 13.426–23.974), significantly shorter than the control group (DPS < 0.5), which had a median survival of 29.9 months (95% CI: 4.359–55.441) (Fig. 4**C**). The mean survival times also reflected this difference, with estimates of 25.8 months (95% CI: 15.391–36.287) for cases and 49.1 months (95% CI: 34.401–63.876) for controls. The difference in survival distributions between groups was statistically significant, as confirmed by the log-rank test ($\chi^2 = 4.272$, df = 1, p = 0.039), indicating that patients in the case group experienced significantly poorer survival compared to controls.

4. Discussion

With the advent of whole slide imaging and the rise of digital pathology, AI quickly demonstrated significant potential in assisting pathologists. The past decade has witnessed an exponential increase in published research articles on AI and digital pathology, with the first study addressing GC emerging in 2017 [23]. Since then, recent studies have highlighted the potential of deep learning in GC diagnosis and

prognosis assessments. For instance, Huang *et al.* [16] introduced two models, GastroMIL for diagnosing GC and MIL-GC for prognosis prediction, which not only outperformed junior pathologists in diagnostic accuracy but also demonstrated significant predictive power for OS. The key distinction is that the MIL-GC model does not follow to any established GC classifications, compared with our article; instead, it identifies prognostically relevant features directly from the images and uses survival data for its training and automatically calculating a prognostic score. Similarly, Jang et al. [24] developed CNN models to automatically subclassify GC, achieving high precision in distinguishing between differentiated vs. undifferentiated tumors, as well as mucinous vs. non-mucinous subtypes. While this paper provides a quantitative tool for subclassification according to the South-Korean classification, which may indirectly relate to prognosis prediction, it does not clearly correlate the model's output to OS. Additionally, Veldhuizen et al. [3] which, to the best of our knowledge, represents the first DL classifier of GC based on the Laurén classification, proved that a deep learning-based approach to Laurén classification could stratify patients more effectively in terms of OS compared to traditional, pathologist-driven methods, highlighting the potential of AI to reduce interobserver variability. There are several key differences compared to our approach. First, their training and validation were performed using GC slides from the TCGA database, providing a larger dataset for training and internal validation (n = 166), whereas our models have been trained and validated on an internal database annotated by pathologists with expertise. Second, although the other study initially aimed to classify mixed-type GC, only two cases fell into this category, leading to its exclusion. In contrast, our models do not require the whole slide to be classified as either intestinal or diffuse; instead, they classify individual patches as either intestinal or diffuse without committing to an overall final class. A recent study by Ren et al. [25] presents a highly accurate, interpretable, and generalizable artificial intelligence framework for gastric cancer diagnosis, achieving approximately 99.2% accuracy, a macro F1-score of 0.991, and an AUC of 0.996. The differences observed in performance metrics between their study and the present work can be attributed to key methodological distinctions. Notably, Ren et al. [25] employed a patch-based approach using curated, publicly available datasets, whereas our model is trained on whole slide images (WSIs) derived from real-world clinical settings. This introduces greater variability, including heterogeneous staining, artifacts, and diagnostically ambiguous regions, which more closely reflect routine clinical practice.

The first developed model, GAC-I, demonstrated high accuracy in distinguishing between normal and malignant tissue. GAC-ST was based on the original Laurén classification, as described by Pekka Laurén in 1965, given that it is the oldest and most widely used system worldwide [5] and has a well-established correlation with OS [12,26–28].

By addressing a key gap in the literature—namely, the ongoing debate regarding the specific percentage thresholds of intestinal versus diffuse components required for a tumor to be classified as mixed—we recognized that the mixed type is more of a conventional entity rather than a distinct pathological structure. With the exception of F. Carneiro and the South-Korean classification, the scientific community has not established a definitive threshold for defining the mixed type. The difference between F. Carneiro who stated that as little as 5% of the minor component, be it intestinal or diffuse, is sufficient for a tumor to be classified as mixed type, in contrast to the South-Korean guidelines that requires roughly equal proportions of both intestinal and diffuse elements. This marked discrepancy adds to the confusion surrounding the diagnosis of an already challenging disease in terms of prognosis. Therefore, we chose to eliminate the mixed class entirely and instead used the number of intestinal and diffuse patches per slide for our data analysis. Several studies have trained CNN models to distinguish between intestinal and diffuse GC subtypes with high accuracy [3,29], and some have even successfully classified the mixed type [30]. We advanced this approach by developing the DPS score and statistically validating its significance through its strong association with OS, thereby identifying a subgroup of GC patients who may require additional treatment or closer monitoring. Our findings are consistent with previous studies, including multiple reports that have clearly demonstrated and confirmed that diffuse GC exhibits a more aggressive clinical behavior, as reflected by OS outcomes [12,26–28]. By developing a quantitative score derived from the distribution of histological patches on standard H&E slides, our method provides an objective metric that may surpass traditional classifications in predicting patient outcomes. Moreover, it has the potential to identify a subpopulation of GC patients with extremely poor prognosis, thereby potentially facilitating more personalized therapeutic strategies. Our findings contribute to the growing evidence that deep learning can significantly improve tumor classification and patient stratification in GC.

Despite these encouraging results, our study has several limitations. The most significant is its retrospective design, which may introduce selection bias. To fully determine the clinical utility of the CNN-based prognostic score, DPS, prospective clinical trials will be required. Annotation was performed using a consensus-based approach with standardized criteria, interobserver variability remains an inherent limitation of histopathological assessment, particularly in Laurén classification, and may introduce a degree of residual subjectivity. Although the model was developed and internally validated on a large dataset, external validation using independent cohorts was not feasible due to the limited availability of gastric cancer whole-slide images with comprehensive clinical annotations, particularly survival data, and standardized Laurén subtype labeling. Furthermore, the current analysis is based solely on histopathological images and clinical data.

In contrast to Multiple Instance Learning (MIL) approaches such as CLAM (Clustering-constrained Attention Multiple Instance Learning) and DSMIL (Dual-Stream Multiple Instance Learning), and transformer-based architectures that leverage global slide-level context, the proposed pipeline adopts a patch-level classification strategy that enables dense spatial prediction and quantitative analysis of tumor heterogeneity. While MIL and transformer models may capture broader contextual relationships, they typically provide slide-level outputs and require more complex training pipelines. Our approach prioritizes scalability, interpretability, and the ability to derive biologically meaningful metrics, such as the DPS, directly from spatially resolved predictions.

Nevertheless, the study offers several important strengths. Only 9.9% of patients received neoadjuvant therapy, thereby minimizing treatment-related variability in tumor morphology, despite the lack of formal regression scoring. Notably, the model was developed using a large, proprietary, and previously unpublished dataset, in contrast to most existing studies that rely on publicly available databases. In addition, the cohort is relatively homogeneous, composed predominantly of white Eastern European patients, which reduces inter-population variability and supports internal consistency in the findings.

By capturing the continuous distribution of histological subtypes at the patch level, the DPS may provide a more refined and biologically relevant assessment of tumor heterogeneity than conventional Laurén classification, which relies on categorical, slide-level assignment. Future studies should formally compare these approaches to determine whether DPS offers superior prognostic stratification.

Future research should also focus on external validation of the proposed CNN-based prognostic score through multicenter prospective studies involving diverse and larger patient cohorts. Incorporating additional data sources—such as immunohistochemical markers [31], molecular profiles, and regression scores following neoadjuvant therapy—could significantly enhance the model's predictive accuracy and clinical applicability. Moreover, evaluating the utility of AI in predicting therapeutic responses, including sensitivity to chemotherapy, targeted agents, and immunotherapy, may further support its role in personalized treatment planning. Integrating these elements will be essential to refine the prognostic model and fully transpose it into clinical practice.

5. Conclusion

This study presents a deep learning–based pipeline for the classification and prognostic assessment of gastric adenocarcinoma using routine histopathological slides. By leveraging the Laurén classification, the proposed models achieved high diagnostic performance and enabled the derivation of a novel AI-based metric, the DPS, which was associated with overall survival in this single-center cohort.

The use of a large, previously unpublished and relatively homogeneous dataset supports the internal consistency of the findings. However, external validation in independent, multicenter cohorts is required to confirm the generalizability and clinical utility of the DPS.

Overall, these results highlight the potential of AI-driven image analysis in digital pathology and suggest that such approaches may contribute to improved risk stratification and clinical decision-making in the future, pending further validation.

Ethical approval

The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Fundeni Clinical Institute, Bucharest, Romania (protocol number 21194/16.04.2021).

Author contribution

CEM, VH and CV conceptualized the study. CA, CEM and DS performed the data curation. CEM developed the methodology of the study. VH and FA performed validation (histopathology expertise). DS provided software and computer vision expertise. CA and CEM performed the formal data analysis. All the authors contributed to interpretation of the results. CA and CEM wrote the first draft of the manuscript and all the authors critically revised the manuscript. CV, GD and VH supervised the project. All the authors approved the final version of the manuscript. All the authors decided to submit this study and agreed to be accountable for all aspects of the work as recommended by the International Committee of Medical Journals Editors (ICMJE) authorship criteria.

CRediT authorship contribution statement

Cătălin Andraș: Writing – original draft, Methodology, Investigation, Formal analysis. **Corina-Elena Minciună:** Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Dragos Stan:** Software, Formal analysis. **Stefan Tudor:** Writing – review & editing, Resources. **Teodora Manuc:** Investigation, Data curation. **Mihnea P. Dragomir:** Writing – review & editing, Visualization. **Ovidiu Bitere:** Investigation, Data curation. **Alexandru Micu:** Investigation, Data curation. **Florina Almarii:** Methodology, Investigation, Data curation. **Gabriela Droc:** Writing – review & editing. **George Calin:** Writing – review & editing, Supervision, Conceptualization. **Vlad Herlea:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Catalin Vasilescu:** Writing – review & editing, Supervision, Resources, Formal analysis, Conceptualization.

Funding

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS – UEFISCDI, project number PN-III-P4-PCE-2021-1068, within PNCDI III (to C.E.M., H.V, S.T., T.M., C.V.). Dr. Dragomir is supported by the Berlin Institute of Health Clinician-Scientist (CS Program) and by Deutsches Konsortium für Translationale Krebsforschung (DKTK) Berlin (Young Investigator Grant 2022).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2026.106455.

References

[1] C.E. Minciuna, M. Tanase, T.E. Manuc, S. Tudor, V. Herlea, M.P. Dragomir, G. A. Calin, C. Vasilescu, The seen and the unseen: molecular classification and image based-analysis of gastrointestinal cancers, Comput. Struct. Biotechnol. J. 12 (20) (2022) 5065–5075, https://doi.org/10.1016/j.csbj.2022.09.010. PMID: 36187924; PMCID: PMC9489806.

[2] J. Ferlay, M. Ervik, F. Lam, M. Laversanne, M. Colombet, L. Mery, M. Piñeros, A. Znaor, I. Soerjomataram, F. Bray, Global cancer observatory: cancer today, International Agency for Research on Cancer, Lyon, France, 2024.

[3] Veldhuizen GP, Röcken C, Behrens HM, Cifci D, Muti HS, Yoshikawa T, Arai T, Oshima T, Tan P, Ebert MP, Pearson AT, Calderaro J, Grabsch HI, Kather JN. Deep learning-based subtyping of gastric cancer histology predicts clinical outcome: a multi-institutional retrospective study. Gastric Cancer. 2023 Sep;26(5):708-720. Doi: 10.1007/s10120-023-01398-x. Epub 2023 Jun 3. PMID: 37269416; PMCID: PMC10361890.

[4] Mariette C, Carneiro F, Grabsch HI, van der Post RS, Allum W, de Manzoni G; European Chapter of International Gastric Cancer Association. Consensus on the pathological definition and classification of poorly cohesive gastric carcinoma. Gastric Cancer. 2019 Jan;22(1):1-9. Doi: 10.1007/s10120-018-0868-0. Epub 2018 Aug 25. Erratum in: Gastric Cancer. 2019 Mar;22(2):421. Doi: 10.1007/s10120-019-00925-z. PMID: 30167905.

[5] P. Laurén, The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. an attempt at a histo-clinical classification, Acta Pathol. Microbiol. Scand. 64 (1965) 31–49, https://doi.org/10.1111/apm.1965.64.1.31. PMID: 14320675.

[6] Nagtegaal ID, Odze RD, Klimstra D, Paradis V, Rugge M, Schirmacher P, Washington KM, Carneiro F, Cree IA; WHO Classification of Tumours Editorial Board. The 2019 WHO classification of tumours of the digestive system. Histopathology. 2020 Jan;76(2):182-188. Doi; 10.1111/his.13975. Epub 2019 Nov 13. PMID: 31433515; PMCID: PMC7003895.

[7] Japanese Gastric Cancer Association. Japanese classification of gastric carcinoma: 3rd English edition. Gastric Cancer. 2011 Jun;14(2):101-12. Doi: 10.1007/s10120-011-0041-5. PMID: 21573743.

[8] Kim TH, Kim IH, Kang SJ, Choi M, Kim BH, Eom BW, Kim BJ, Min BH, Choi CI, Shin CM, Tae CH, Gong CS, Kim DJ, Cho AE, Gong EJ, Song GJ, Im HS, Ahn HS, Lim H, Kim HD, Kim JJ, Yu JI, Lee JW, Park JY, Kim JH, Song KD, Jung M, Jung MR, Son SY, Park SH, Kim SJ, Lee SH, Kim TY, Bae WK, Koom WS, Jee Y, Kim YM, Kwak Y, Park YS, Han HS, Nam SY, Kong SH; Development Working Groups for the Korean Practice Guidelines for Gastric Cancer 2022 Task Force Team. Korean Practice Guidelines for Gastric Cancer 2022: An Evidence-based, Multidisciplinary Approach. J Gastric Cancer. 2023 Jan;23(1):3-106. Doi: 10.5230/jgc.2023.23.e11. Erratum in: J Gastric Cancer. 2023 Apr;23(2):365-373. Doi: 10.5230/jgc.2023.23.e20. PMID: 36750993; PMCID: PMC9911619.

[9] I. Songun, C.J. van de Velde, J.W. Arends, P. Blok, A.J. Grond, G.J. Offerhaus, J. Hermans, J.H. van Krieken, Classification of gastric carcinoma using the Goseki system provides prognostic information additional to TNM staging, Cancer 85 (10) (1999) 2114–2118, https://doi.org/10.1002/(sici)1097-0142(19990515)85:10<2114::aid-cncr3>3.0.co;2-u. PMID: 10326687.

[10] S.C. Ming, Gastric carcinoma. A pathobiological classification, Cancer 39 (6) (1977) 2475–2485, https://doi.org/10.1002/1097-0142(197706)39:6<2475::aid-cncr2820390626>3.0.co;2-l. PMID: 872047.

[11] Wang Q, Liu G, Hu C. Molecular classification of gastric adenocarcinoma. Gastroenterology Res. 2019 Dec;12(6):275-282. 10.14740/gr1187. Epub 2019 Nov 21. PMID: 31803306; PMCID: PMC6879029.

[12] F. Carneiro, M. Seixas, M. Sobrinho-Simões, New elements for an updated classification of the carcinomas of the stomach, Pathol. Res. Pract. 191 (6) (1995) 571–584, https://doi.org/10.1016/S0344-0338(11)80878-2. PMID: 7479380.

[13] J.H. Pyo, H. Lee, B.H. Min, J.H. Lee, M.G. Choi, J.H. Lee, T.S. Sohn, J.M. Bae, K. M. Kim, S. Yeon, S.H. Jung, J.J. Kim, S. Kim, Early gastric cancer with a mixed-type Laurén classification is more aggressive and exhibits greater lymph node metastasis, J. Gastroenterol. 52 (5) (2017 May) 594–601, https://doi.org/10.1007/s00535-016-1254-5. Epub 2016 Sep 2 PMID: 27590416.

[14] Y.C. Chen, W.L. Fang, R.F. Wang, C.A. Liu, M.H. Yang, S.S. Lo, C.W. Wu, A.F. Li, Y. M. Shyr, K.H. Huang, Clinicopathological variation of laurén classification in gastric cancer, Pathol. Oncol. Res. 22 (1) (2016) 197–202, https://doi.org/10.1007/s12253-015-9996-6. Epub 2015 Oct 27 PMID: 26502923.

[15] X. Zhao, Y. Li, Z. Yang, H. Zhang, H. Wang, J. Lin, J. Liu, Q. Zhao, Adenocarcinoma with mixed subtypes in the early and advanced gastric cancer, Can. J. Gastroenterol. Hepatol. 29 (2021) (2021) 8497305, https://doi.org/10.1155/2021/8497305. PMID: 34746042; PMCID: PMC8570884.

[16] Huang B, Tian S, Zhan N, Ma J, Huang Z, Zhang C, Zhang H, Ming F, Liao F, Ji M, Zhang J, Liu Y, He P, Deng B, Hu J, Dong W. Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: a retrospective multicentre study. EBioMedicine. 2021 Nov;73:103631. Doi: 10.1016/j.ebiom.2021.103631. Epub 2021 Oct 19. PMID: 34678610; PMCID: PMC8529077.

[17] Y. Wang, C. Hu, T. Kwok, C.A. Bain, X. Xue, R.B. Gasser, G.I. Webb, A. Boussioutas, X. Shen, R.J. Daly, J. Song, DEMoS: a deep learning-based ensemble approach for predicting the molecular subtypes of gastric adenocarcinomas from histopathological images, Bioinformatics 38 (17) (2022) 4206–4213, https://doi.org/10.1093/bioinformatics/btac456. PMID: 35801909.

[18] P.C. Rizzo, I. Girolami, S. Marletta, L. Pantanowitz, P. Antonini, M. Brunelli, N. Santonicco, P. Vacca, N. Tumino, L. Moretta, A. Parwani, S. Satturwar, A. Eccher, E. Munari, Technical and diagnostic issues in whole slide imaging published validation studies, Front. Oncol. 12 (2022) 918580, https://doi.org/10.3389/fonc.2022.918580. PMID: 35785212; PMCID: PMC9246412.

[19] P. Jin, X. Ji, W. Kang, Y. Li, H. Liu, F. Ma, S. Ma, H. Hu, W. Li, Y. Tian, Artificial intelligence in gastric cancer: a systematic review, J. Cancer Res. Clin. Oncol. 146 (9) (2020) 2339–2350, https://doi.org/10.1007/s00432-020-03304-9. Epub 2020 Jul 1 PMID: 32613386.

[20] P.C. Rizzo, A. Caputo, E. Maddalena, N. Caldonazzi, I. Girolami, A.P. Dei Tos, A. Scarpa, M. Sbaraglia, M. Brunelli, S. Gobbo, S. Marletta, L. Pantanowitz, V. Della Mea, A. Eccher, Digital pathology world tour, Digit Health 29 (9) (2023) 20552076231194551, https://doi.org/10.1177/20552076231194551. PMID: 37654717; PMCID: PMC10467307.

[21] D. Nepogodiev, J. Martin, B. Biccard, A. Makupe, A. Bhangu, National institute for health research global health research unit on global surgery, Global Burden of Postoperative Death. Lancet. 393 (10170) (2019) 401, https://doi.org/10.1016/S0140-6736(18)33139-8. PMID: 30722955.

[22] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.

[23] H. Sharma, N. Zerbe, I. Klempert, O. Hellwich, P. Hufnagl, Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology, Comput. Med. Imaging Graph. 61 (2017) 2–13, https://doi.org/10.1016/j.compmedimag.2017.06.001. Epub 2017 Jun 16 PMID: 28676295.

[24] H.J. Jang, I.H. Song, S.H. Lee, Deep learning for automatic subclassification of gastric carcinoma using whole-slide histopathology images, Cancers (Basel) 13 (15) (2021) 3811, https://doi.org/10.3390/cancers13153811. PMID: 34359712; PMCID: PMC8345042.

[25] T. Ren, V. Govindarajan, S. Bourouis, X. Wang, S. Ke, An interpretable hybrid deep learning framework for gastric cancer diagnosis using histopathological imaging, Sci. Rep. 15 (1) (2025) 34204, https://doi.org/10.1038/s41598-025-15702-5. PMID: 41034364; PMCID: PMC12489133.

[26] Cuschieri A, Talbot IC, Weeden S; MRC Upper GI Cancer Working Party. Influence of pathological tumour variables on long-term survival in resectable gastric cancer. Br J Cancer. 2002 Mar 4;86(5):674-9. 10.1038/sj.bjc.6600161. PMID: 11875724; PMCID: PMC2375301.

[27] Luu C, Thapa R, Woo K, Coppola D, Almhanna K, Pimiento JM, Chen DT, Marquez DD, Hodul PJ. Does histology really influence gastric cancer prognosis? J Gastrointest Oncol. 2017 Dec;8(6):1026-1036. 10.21037/jgo.2017.09.08. PMID: 29299363; PMCID: PMC5750182.

[28] C.T. Tang, L. Zeng, J. Yang, C. Zeng, Y. Chen, Analysis of the incidence and survival of gastric cancer based on the Laurén classification: a large population-based study using SEER, Front. Oncol. 10 (2020) 1212, https://doi.org/10.3389/fonc.2020.01212. PMID: 32850357; PMCID: PMC7416646.

[29] F. Kanavati, M. Tsuneki, A deep learning model for gastric diffuse-type adenocarcinoma classification in whole slide images, Sci. Rep. 11 (1) (2021) 20486, https://doi.org/10.1038/s41598-021-99940-3. PMID: 34650155; PMCID: PMC8516929.

[30] X. Ning, R. Liu, N. Wang, X. Xiao, S. Wu, Y. Wang, C. Yi, Y. He, D. Li, H. Chen, Development of a deep learning-based model to diagnose mixed-type gastric cancer accurately, Int. J. Biochem. Cell Biol. 162 (2023) 106452, https://doi.org/10.1016/j.biocel.2023.106452. Epub 2023 Jul 21 PMID: 37482265.

[31] J. Wang, S. Zhang, J. Li, M. Deng, Z. Zeng, Z. Dong, F. Chen, W. Liu, L. Wu, H. Yu, Development and clinical validation of deep learning-based immunohistochemistry prediction models for subtyping and staging of gastrointestinal cancers, BMC Gastroenterol. 25 (1) (2025) 494, https://doi.org/10.1186/s12876-025-04045-0. PMID: 40597706; PMCID: PMC12211442.