AI-Assisted Detection of Gastric Intestinal Metaplasia: Design and Validation of the "IntelliMeta" Algorithm

Abstract

Background: Gastric intestinal metaplasia (GIM) is a potential precancerous lesion that significantly increases the risk of gastric cancer. Its accurate detection requires expertise in digestive pathology and remains challenging due to histological complexity and interobserver variability. Artificial intelligence (AI) represents a promising tool to support early and precise diagnosis.

Methods: We developed *IntelliMeta*, an AI-based algorithm designed to automatically detect GIM on digitized gastric biopsy slides. A dataset of 229 histological slides (173 normal, 56 with GIM) collected at the Hassan II Regional Hospital of Agadir was digitized using an APERIO LV1 scanner. After expert annotation, a total of 902 histological images were processed. The algorithm, based on a Visual Geometry Group **(**VGG) transfer learning model, was trained and validated using data preprocessing, augmentation, and cross-validation. Key functionalities include automatic segmentation, multi-region quantification, and binary classification (focal vs diffuse GIM).

Results: The transfer learning V1 model achieved the most balanced performance, with an overall accuracy of 56.5% and a processing speed of 547 ms/step, outperforming a custom CNN and a slower transfer learning V2 model. Despite the limited dataset size, IntelliMeta successfully detected GIM regions, provided confidence scores, and quantified lesion extent. The system also integrated a user-friendly interface for visualization and interpretation.

Conclusion:*IntelliMeta* represents the first national and continental contribution to AI-assisted detection of GIM. Although limited by dataset size, the algorithm demonstrates promising efficiency for supporting pathologists in the early diagnosis of gastric precancerous lesions. Further improvements, including dataset expansion and threshold optimization, could enhance clinical applicability.

Key words: Gastric Intestinal Metaplasia; Artificial Intelligence; Deep Learning; Digital Pathology; Transfer Learning; Convolutional Neural Networks; Computer-Aided Diagnosis

Introduction

Gastric cancer remains one of the leading causes of cancer-related mortality worldwide, particularly in regions with a high prevalence of *Helicobacter pylori* infection [1]. Gastric intestinal metaplasia (GIM) represents a persistent and irreversible precancerous lesion characterized by the replacement of the gastric epithelium with an intestinal-type epithelium [2]. Its early and accurate detection is crucial, as GIM significantly increases the risk of progression to gastric adenocarcinoma [3].

Histopathological examination of gastric biopsies is considered the gold standard for diagnosing GIM [4]. However, the process requires substantial expertise to distinguish between goblet cells and pseudo-goblet cells, and may sometimes necessitate ancillary techniques such as special stains or immunohistochemistry. These additional methods are costly, time-consuming, and not always available in routine practice [5]. Furthermore, interobserver variability and sampling limitations may hinder the reproducibility and accuracy of GIM detection [6].

Recent advances in artificial intelligence (AI), particularly deep learning approaches, have shown great promise in medical image analysis, including radiology, dermatology, and pathology [7,8]. In digital pathology, AI-based algorithms have demonstrated their capacity to detect subtle morphological changes, quantify histological features, and assist in diagnostic standardization [9]. Convolutional neural networks (CNNs), in particular, have become powerful tools for the classification and segmentation of histopathological images, reducing subjectivity and improving diagnostic efficiency [10].

In this study, we developed *IntelliMeta*, an AI-based algorithm designed to automatically detect GIM in digitized gastric biopsies. To our knowledge, this represents the first initiative at both the national (Morocco) and continental (Africa) levels to address this diagnostic challenge through AI. By integrating automatic segmentation, multi-region quantification, and binary classification (focal vs diffuse GIM), IntelliMeta provides a novel approach to support pathologists in the early identification of precancerous gastric lesions.

Materials and Methods

Data Collection and Preparation :

Gastric biopsy slides were retrospectively collected from the Department of Pathology at Hassan II Regional Hospital in Agadir between 2023 and the first semester of 2025. A total of 229 histological slides were included, comprising 173 normal gastric mucosa cases and 56 cases with confirmed gastric intestinal metaplasia (GIM). All samples were fixed in formalin, embedded in paraffin, sectioned, and stained with hematoxylin–eosin (H&E) according to standard pathology protocols.

Slide Digitization and Annotation :

All slides were digitized using the APERIO LV1 scanner at ×40 magnification, generating whole-slide images (WSIs) in SVS format. The WSIs were subsequently validated for technical quality. Pathologists manually annotated representative areas of GIM and normal mucosa using QuPath software, producing a dataset of 902 histological image patches (319 GIM and 583 normal). Annotations were reviewed and validated by an expert gastrointestinal pathologist to ensure diagnostic accuracy.

Preprocessing and Data Augmentation

Image preprocessing included resizing to 224×224 pixels, conversion to RGB, and normalization within the [0,1] range. To address data imbalance and improve generalization, data augmentation was applied using random rotations, horizontal and vertical shifts, zooming, and flipping.

Algorithm Architecture :

The *IntelliMeta* algorithm was based on a transfer learning approach using a pre-trained Visual Geometry Group (VGG) convolutional neural network, adapted for binary classification (normal vs GIM). The final architecture included modified fully connected layers to output prediction probabilities with a decision threshold set at 0.5.

Segmentation and Quantification :

Following classification, a segmentation module was implemented to localize GIM regions within WSIs. The pipeline combined color-space analyses (HSV, LAB), adaptive thresholding, and morphological operations to enhance region detection. Quantitative metrics such as the percentage of GIM surface area, mean lesion size, and number of regions were computed.

Performance Evaluation

Three models were trained and compared:

- Transfer Learning V1 (adopted model)
- Transfer Learning V2
- Custom CNN

Performance was assessed using accuracy, sensitivity, specificity, positive predictive value, and negative predictive value, derived from confusion matrices. Processing speed was also evaluated.

User Interface Development :

A graphical user interface (GUI) was developed to facilitate interaction with the algorithm. The GUI included functions for slide uploading, automated detection, lesion localization, and real-time visualization of confidence scores and quantitative metrics.

Results

Dataset and Annotations:

A total of 229 gastric biopsy slides were collected from the Pathology Department of Hassan II Regional Hospital, Agadir (173 normal gastric mucosa, 56 with confirmed GIM). All slides were digitized, and after expert annotation, 902 histological image patches were generated (583 normal, 319 GIM). Annotation focused on morphologic hallmarks of GIM, particularly the presence of goblet cells, pseudo-goblet cells, and architectural changes.

Model Training and Classification Performance :

Three models were evaluated: Transfer Learning V1, Transfer Learning V2, and a custom CNN. Their comparative performances are summarized in **Table 1**.

- **Transfer Learning V1** achieved the most balanced performance, with an overall accuracy of 56.5%. The confusion matrix revealed moderate sensitivity and specificity, with a recall of 35% for positive cases (GIM) and 66% for negatives. The processing speed (547 ms/step) makes it suitable for near real-time applications.
- **Custom CNN** showed severe classification bias, systematically labeling all cases as GIM, which resulted in poor discrimination between normal and pathological slides (accuracy 34.5%).
- **Transfer Learning V2** reached a slightly higher accuracy (60.0%), but with very low recall for GIM-positive cases (8%). Despite acceptable performance for normal slides, its processing time (2 s/step) was considerably slower, limiting practical usability.

Table 1. Comparative performance of the three tested models.

Model	Accuracy (%)	Processing Speed	Positive Case Recall	Negative Case Recall	Comments
Transfer Learning V1	56.5	547 ms/step	35%	66%	Most balanced performance; adopted model
Custom CNN	34.5	558 ms/step	100% (all cases as GIM)	0%	Severebias; poor discrimination
Transfer Learning V2	60.0	2 s/step	8%	~90%	Slightly higher accuracy but weak sensitivity

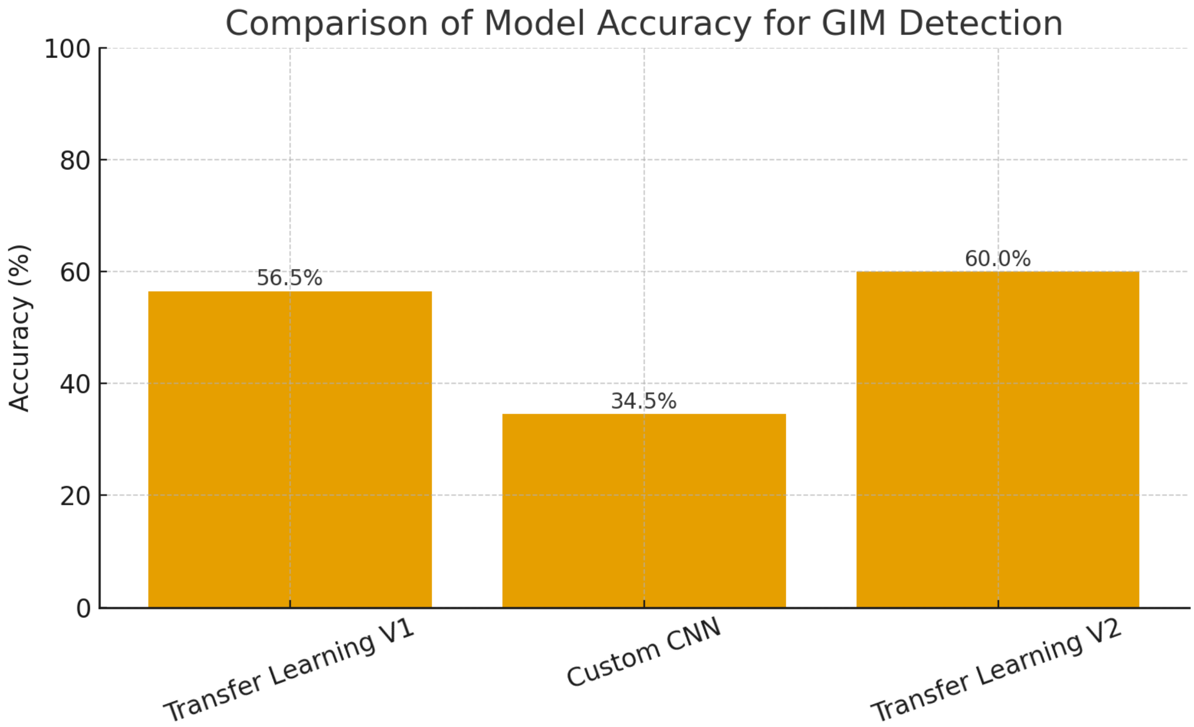

Figure 1. Accuracy comparison of the three tested models.(Bar chart showing accuracy for Transfer Learning V1, Custom CNN, and Transfer Learning V2, with Transfer Learning V1 chosen as the adopted model.)

Segmentation and Quantification of GIM :

The segmentation module was able to accurately highlight GIM regions within whole-slide images. Post-processing with color-space analysis and morphological filters enhanced region detection, providing a clear distinction between affected and unaffected mucosa.

Quantitative metricsgeneratedincluded:

- **Percentage of GIM surface area**: 6.9% in focal lesions versus up to 82.1% in diffuse lesions.
- **Number of detected regions**: from as few as 3 (focal GIM) to as many as 23 (diffuse GIM).
- **Mean lesion area per region**, which helped differentiate small focal foci from extensive diffuse involvement.

User Interface Performance

The IntelliMeta graphical interface allowed pathologists to interact with the algorithm in real time:

- **Slide uploading and validation** ensured compatibility of SVS files.
- **Classification results** were displayed with a confidence score. For example:
  - A focal GIM biopsy: predicted as “GIM” with 63.1% confidence, 6.9% surface involvement, and 3 regions detected.
  - A diffuse GIM biopsy: predicted as “GIM” with 96.8% confidence, 82.1% surface involvement, and 23 regions detected.
- **Segmentation overlay** allowed visualization of GIM regions circled in red on the WSI, facilitating rapid review and validation by the pathologist.

Discussion

Gastric intestinal metaplasia (GIM) is a well-recognized precancerous lesion that increases the risk of gastric cancer, particularly in populations with a high prevalence of *Helicobacter pylori* [1–3]. Accurate detection of GIM is clinically essential, yet remains challenging in practice. Histopathological assessment is the diagnostic gold standard [4], but it requires substantial expertise to reliably distinguish goblet from pseudo-goblet cells [5], and interobserver variability remains a major limitation [6]. This is particularly relevant in low- and middle-income countries, where access to ancillary tests such as AB/PAS staining or immunohistochemistry may be limited [5,11].

In this study, we present *IntelliMeta*, the first AI-based system developed at a national (Morocco) and continental (Africa) level to assist in the detection of GIM on digitized gastric biopsies. By combining transfer learning with automatic segmentation and quantification, IntelliMeta introduces a comprehensive approach that goes beyond binary classification, offering lesion extent analysis (focal vs diffuse) and confidence scores. Such quantitative support is of high clinical value since the extent of GIM has been linked to higher malignant potential [2,12].

Our results demonstrate that the Transfer Learning V1 model achieved the most balanced performance (accuracy 56.5%, moderate sensitivity and specificity) compared to a biased CNN and a slower transfer learning variant. While the accuracy remains modest, the robustness and processing speed of Transfer Learning V1 highlight its potential for integration in pathology workflows. These findings are consistent with previous AI applications in gastrointestinal pathology, where deep learning models achieved variable performance depending on dataset size and annotation quality [6,13,14].

The segmentation and quantification module represents a major added value of IntelliMeta. Beyond classification, it provides reproducible metrics such as surface percentage of GIM and number of affected regions. This approach parallels the recent shift in pathology toward “computational quantification” of lesions, which has been shown to reduce interobserver variability and standardize reporting [9,15]. Furthermore, the GUI interface enhances interpretability, which is critical for pathologists’ acceptance of AI tools [10,16].

Nevertheless, several limitations must be acknowledged. The relatively small and imbalanced dataset (229 slides, 902 patches) restricted the algorithm’s learning capacity. Previous work has shown that AI models in pathology require large, diverse datasets—often in the thousands of slides—to achieve clinically acceptable performance [7,14,17]. The inability to further expand the dataset due to technical issues (scanner malfunction) also limited training. Additionally, the moderate recall for GIM-positive cases reflects a need for improved threshold optimization and augmentation strategies. Future studies should therefore focus on dataset expansion, multicentric validation, and incorporation of molecular/immunohistochemical data as multimodal inputs [12,18].

From a broader perspective, IntelliMeta reflects the growing role of AI in digital pathology. Convolutional neural networks and transfer learning approaches have already been successfully applied in prostate, breast, and colorectal pathology [9,13,14], and their extension to gastric precancerous lesions is timely. In resource-limited settings such as Morocco, where gastric cancer remains a public health concern [1,19], AI-assisted diagnostic systems could help mitigate workforce shortages and improve early detection strategies.

In conclusion, while IntelliMeta currently shows moderate accuracy, it provides a proof-of-concept for AI-assisted detection of GIM, with segmentation and quantification capabilities that enhance its clinical relevance. With dataset expansion and multicentric validation, this approach has the potential to significantly improve diagnostic accuracy, reduce variability, and support precision prevention of gastric cancer.

References:

[1]. Huang RJ, Choi AY, Truong CD, Yeh MM, Hwang JH. Diagnosis and management of gastric intestinal metaplasia: Current status and future directions. *Gut Liver*. 2019;13(6):596-603.

[2]. Sugano K, Moss SF, Kuipers EJ. Gastric intestinal metaplasia: Real culprit or innocent bystander as a precancerous condition for gastric cancer? *Gastroenterology*. 2023;165(6):1352-61.

[3]. Tjandra D, Busuttil RA, Boussioutas A. Gastric intestinal metaplasia: Challenges and the opportunity for precision prevention. *Cancers (Basel)*. 2023;15(15):3913.

[4]. Shah SC, Piazuelo MB, Kuipers EJ, Li D. AGA Clinical Practice Update on the diagnosis and management of atrophic gastritis. *Gastroenterology*. 2021;161(4):1325-32.

[5]. Kiliçarslan A, Varli G. Should special staining (AB/PAS) be routinely performed in gastric biopsies for the detection of intestinal metaplasia? *Int J SurgPathol*. 2024;32(2):217-22.

[6]. Iwaya M, Hayashi Y, Sakai Y, et al. Artificial intelligence for evaluating the risk of gastric cancer: reliable detection and scoring of intestinal metaplasia with deep learning algorithms. *GastrointestEndosc*. 2023;98(6):925-33.

[7]. Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med*. 2019;25(6):954-61.

[8]. Rahmatizadeh S, Dabbagh A, Shabani F. Foundations of artificial intelligence: Transforming health care now and in the future. *Anesthesiol Clin*. 2025;43(2):201-14.

[9]. Zhu L, Pan J, Mou W, et al. Harnessing artificial intelligence for prostate cancer management. *Cell Rep Med*. 2024;5(4):101506.

[10]. Sun R, Deutsch E, Fournier L. Artificial intelligence and medical imaging. *Bull Cancer*. 2022;109(1):83-8.

[11]. Zullo A, Hassan C, Romiti A, et al. Follow-up of intestinal metaplasia in the stomach: When, how and why. *World J Gastrointest Oncol*. 2012;4(3):30-6.

[12]. Shah SC, Gawron AJ, Mustafa R, Piazuelo MB. Histologic subtyping of gastric intestinal metaplasia: Overview and clinical considerations. *Gastroenterology*. 2020;158(3):745-50.

[13]. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24(10):1559-67.

[14]. Echle A, Grabsch HI, Quirke P, et al. Clinical-grade detection of microsatellite instability in colorectal cancer using deep learning. *Nat Med*. 2020;26(6):918-24.

[15]. Komura D, Ishikawa S. Machine learning methods for histopathological image analysis. *Comput Struct Biotechnol J*. 2018;16:34-42.

[16]. Steiner DF, MacDonald R, Liu Y, et al. Evaluation of the use of combined artificial intelligence and pathologist assessment to review and grade prostate biopsies. *JAMA Netw Open*. 2020;3(11):e2023267.

[17]. Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole-slide images. *Nat Med*. 2019;25(8):1301-9.

[18]. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60-88.

[19]. World Health Organization. Global Cancer Observatory: Cancer Today. Lyon: International Agency for Research on Cancer; 2023.

Conflict of Interest Statement

The authors declare that they have no conflicts of interest related to this work.

Acknowledgments

The authors would like to thank the Faculty of Medicine and Pharmacy of Agadir for its support in the development and implementation of the escape room project for pathology laboratory education.

Authors' Contributions

A. Miry: Conceptualization, Literature Review, Scenario Design, Draft Writing, Supervision
F. Bazhar: Data Collection, Literature Review, Draft Writing
M. Tbouda: Literature Review, Draft Writing, Scenario Design
K. Oqbani: Scenario Design, Data Collection, Draft Review
H. Elagouri: Data Collection, Draft Review, Final Editing
S. Abbaoui: Literature Review, Final Editing, Supervision
All authors read and approved the final version of the manuscript.

Funding statement: This work received no specific funding.