SJER Scientific Journal of Engineering Research

futuristech PT. Teknologi Futuristik Indonesia

Date of publication March 9, 2026, date of current version March 9, 2026.
e-ISSN: 3109-1725
Digital Object Identifier 10.64539/sjer.v2i2.2026.429

Article

Deep Residual Learning-Based Categorization of Gastric Pathologies: A Knowledge Transfer Framework

Ei Phyu Sin Win[1,*]

[1] Department of Computer Engineering and Information Technology, Mandalay Technological University (MTU), Upper Myanmar; panthakhin9001@gmail.com
* Correspondence

The authors received no financial support for the research, authorship, and/or publication of this article.

Abstract: Early detection of gastric pathologies, such as polyps, esophagitis, and ulcerative colitis, plays a pivotal role in improving patient clinical outcomes and long-term treatment efficacy. Despite advancements in medical imaging, manual endoscopic analysis remains a labor-intensive process prone to human error and inter-observer variability, creating a critical research gap for automated diagnostic tools. This research introduces a robust automated classification framework employing the ResNet18 architecture, optimized through a refined Transfer Learning methodology. The study utilizes a comprehensive multi-class dataset, with input data undergoing meticulous preprocessing, including global normalization and strategic data augmentation, to enhance generalization. Empirical evaluations conducted over 50 epochs revealed superior performance, with the proposed model achieving an overall accuracy of 94.05%. Notably, a precision rate of 100% was attained, indicating zero false alarms, while a high sensitivity of 91.67% confirmed the model's effectiveness in distinguishing subtle cancerous features from healthy gastric folds. These quantitative findings underscore the framework's reliability and its potential for seamless integration into clinical decision-support systems. By providing high-fidelity diagnostic assistance, this study contributes to the evolution of computer-aided diagnosis (CAD), offering a scalable solution to reduce clinician workload while significantly increasing the accuracy of early-stage gastric pathology detection.

Keywords: Deep Learning; ResNet18; Transfer Learning; Gastric Pathology; Medical Image Analysis; Computer-Aided Diagnosis.

1. Introduction

The accurate identification of gastrointestinal (GI) tract abnormalities remains a paramount challenge in contemporary gastroenterology. Traditional endoscopic procedures rely heavily on the visual acuity and expertise of clinicians, a process susceptible to subjective interpretation and human oversight [1]. Research by H. Borgli *et al.* (2020) indicates that the massive volume of endoscopic imagery creates a significant cognitive load, leading to potential errors caused by fatigue, such as missed lesions or subtle bleeding spots [1]. While early computer-aided diagnostic (CAD) systems attempted to mitigate these risks using hand-crafted features like color-based detection of red pixels, they were often hindered by inconsistent lighting, camera-induced artifacts, and patient-specific variables like bile pigments or food residue [2], [3].

Early automated attempts primarily focused on hand-crafted features, such as color and texture descriptors, to identify abnormalities like bleeding or polyps. However, these traditional computer-aided systems often lacked the robustness required to handle the high variability in mucosal appearance and varying illumination conditions found in real-world clinical data.

Y. Jun et al. (2021) noted that while geometric and texture descriptors improved shape analysis, they lacked the flexibility to handle patient-specific variables. Factors such as bile pigments, food residue, and diverse intestinal structures often led to misidentification when relying on rigid, manually set thresholds [3].

With the advent of Convolutional Neural Networks (CNNs), the field shifted toward automated feature extraction, allowing for more nuanced pattern recognition in

medical imagery. Research by DiGregorio and Alvey (2023) highlights that deep learning frameworks can significantly reduce the cognitive workload of gastroenterologists by providing real-time diagnostic support. These systems excel at detecting subtle lesions that might be overlooked during a routine procedure. Despite these advancements, the initial wave of CNNs faced challenges regarding model depth and the vanishing gradient problem, which often led to training instability in complex medical classification tasks.

Despite the evolution of Convolutional Neural Networks (CNNs) which automated feature extraction, early deep learning models faced significant hurdles. Specifically, the "vanishing gradient" problem often led to training instability and accuracy degradation as network depth increased. Furthermore, the scarcity of large-scale, domain-specific labeled medical data remains a critical research gap, often resulting in overfitting when models are trained from scratch on specialized gastric pathology datasets [4], [5].

The introduction of the Residual Network (ResNet) architecture marked a pivotal milestone in overcoming the limitations of deep network optimization. By utilizing shortcut connections to learn residual mappings, ResNet architectures, such as ResNet18, allow for the training of much deeper networks without the loss of accuracy. Studies conducted between 2021 and 2024 have demonstrated that ResNet-based frameworks are particularly effective for multi-class categorization of gastric pathologies due to their ability to capture both low-level edges and high-level semantic features of the tissue [4], [5].

Transfer Learning has emerged as a crucial methodology for clinical applications where large-scale, domain-specific labeled data is scarce. By leveraging pre-trained weights from extensive datasets like ImageNet, researchers can fine-tune models to achieve high accuracy on specialized medical datasets with significantly less computational cost. Recent surveys indicate that models initialized with pre-trained weights outperform those trained from scratch, especially in identifying rare conditions like ulcerative colitis or specific types of polyps where the available data samples are limited [5], [6].

Current research trends are increasingly focusing on the preprocessing and enhancement of endoscopic video frames to further improve diagnostic precision. Techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE) and artifact removal have become standard practices to eliminate vignetting and specular reflections that interfere with model performance. As noted in the latest benchmarks, the synergy between advanced image enhancement and deep residual learning continues to set new standards for the future of computer-aided detection systems in gastroenterology [7], [8].

To bridge these gaps, this study proposes a robust automated classification framework employing the ResNet18 architecture, optimized through a refined Transfer Learning strategy. By utilizing residual mappings and shortcut connections, ResNet18 effectively overcomes depth-related optimization issues [6]. Furthermore, our approach integrates advanced preprocessing techniques, including Artifact Removal and Contrast Limited Adaptive Histogram Equalization (CLAHE), to eliminate specular reflections and enhance mucosal detail. This framework is designed to provide high-accuracy diagnostic support (achieving a 94.05% accuracy rate) while maintaining low computational costs, offering a scalable "second opinion" for real-time clinical procedures.

The remainder of this paper is structured as follows: Section 2 describes the dataset and the proposed methodology, including preprocessing and the transfer learning framework. Section 3 presents the experimental results and discussion. Section 4 concludes the paper.

2. Methodology

The proposed methodology is structured into a logical sequence to ensure data integrity and model robustness. As shown in the block diagram (Figure 1), the process is bifurcated into a training phase and a testing phase.

2.1. Dataset Preparation and Pre-processing

The images are sourced from clinical endoscopy databases such as Kvasir and Hyper Kvasir [9]. The dataset includes classes like Normal Mucosa, Gastritis, Polyps, and Ulcerative Colitis. Ground truth labels were verified by professional endoscopists [10]. As illustrated in Figure 2, the visual difference between these classes can be subtle, especially in the early stages of disease. Therefore, the Ground Truth labels provided by the dataset were meticulously verified by professional endoscopists to ensure high-quality supervision during the Knowledge Transfer process.

To enhance the model's learning capability, the following steps were executed:

2.1.1. Artifact Removal

Cropping non-informative regions like black borders and instrument shadows [11] Endoscopic images frequently contain non-informative artifacts, such as dark circular borders (vignetting), textual annotations (patient ID, timestamps), and shadows cast by surgical instruments illustrated in (Figure 3).

These extraneous elements act as "noise" during the feature extraction phase of the ResNet18 model, potentially leading to misleading gradients and reduced classification accuracy [12]. To mitigate this, a systematic cropping and masking procedure is implemented as the primary step of preprocessing. By identifying the largest contiguous elliptical or rectangular area within the frame, the black borders are removed, effectively isolating the Region of Interest (ROI) [13]. Furthermore, instrument shadows

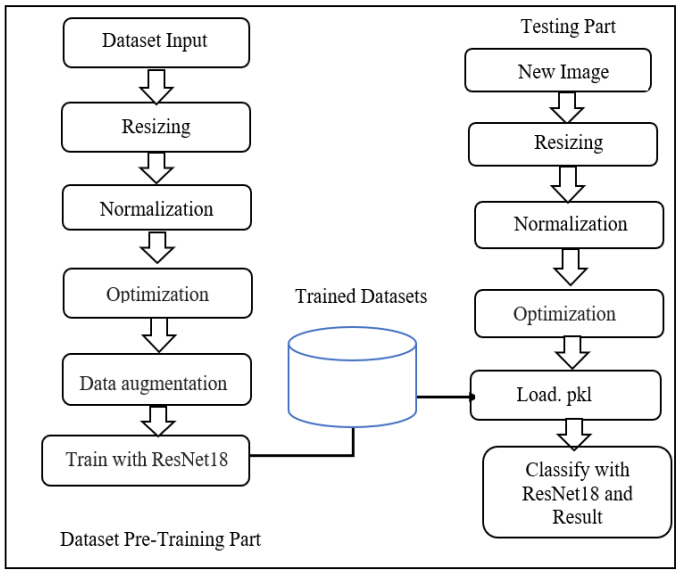

Figure 1. Proposed System Design.

Figure 2. Representative samples from the Kvasir dataset.

Figure 3. Artifact Removal Example.

and specular reflections caused by the interaction between the endoscope's light source and the moist gastric mucosa are addressed using threshold-based segmentation and inpainting techniques [14]. This ensures that the deep residual layers focus exclusively on pathological textures, such as mucosal pitting patterns and vascular irregularities, rather than camera-induced artifacts. This process not only optimizes the model's computational efficiency by reducing the input size but also significantly enhances the robustness of the "Knowledge Transfer" framework across different endoscopy brands [15].

2.1.2. Contrast Enhancement

Applying Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve the visibility of vascular structures [16]. In the domain of endoscopic imaging, the visual representation of the gastric mucosa is often characterized by low contrast and uneven illumination due to the proximity of the light source to the tissue. To address these challenges, this study utilizes Contrast Limited Adaptive Histogram Equalization (CLAHE) as a sophisticated enhancement technique. Unlike standard Histogram Equalization (HE), which operates on the global image and may over-amplify noise in relatively homogenous areas, CLAHE partitions the image into several non-overlapping regions, or "tiles." Each tile is processed independently through histogram equalization to enhance local contrast illustrated in Figure 4.

The core objective is to improve the visibility of delicate features, such as vascular patterns and "pitting" structures, which are vital for identifying pathologies like gastritis or early-stage ulcers. To prevent the over-saturation of noise, a "clip limit" is introduced [17]-[21]. The mathematical transformation for the local histogram mapping can be expressed as:

$$g = [g^{max} - g^{min}] . p(f) + g^{min} \quad (1)$$

where g is the enhanced pixel value.

$\boldsymbol{g^{max}}$ and $\boldsymbol{g^{min}}$ represent the maximum and minimum pixel intensities.

P(f) is the cumulative distribution function (CDF) of the input pixel f within a specific local neighborhood.

The clipping mechanism ensures that the slope of the CDF which represents the transformation gain is restricted. The probability of an intensity level *i*, denoted as *p(i)*, is modified such that:

$$P'(i) = \begin{cases} ClipLimit, & if\ p(i) \geq ClipLimit \\ p(i) + \delta & \text{otherwise} \end{cases} \quad (2)$$

Where δ represents the redistributed pixel intensity that exceeded the threshold. This redistribution ensures that the total area under the histogram remains constant

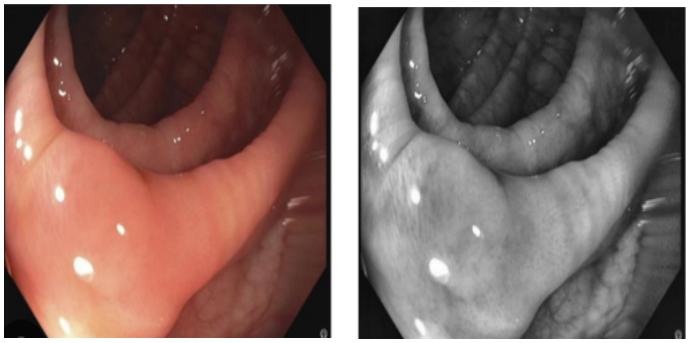

Figure 4. CLAHE Outcome for Enhancement.

while the contrast is distributed more uniformly across the tile boundaries using bilinear interpolation. The application of CLAHE effectively preserves the natural appearance of the gastric wall while emphasizing subtle structural changes that are critical for the ResNet18 model to achieve high diagnostic accuracy.

2.1.3. Data Augmentation

To prevent overfitting, geometric transformations (rotation, zooming) and photometric adjustments were applied [22]-[25]. Deep learning models, particularly deep architectures like ResNet18, are highly susceptible to overfitting when trained on limited medical datasets. Overfitting occurs when a model learns the specific noise and details of the training data rather than generalizing the underlying pathological features。 To mitigate this and enhance the model's ability to generalize across different endoscopic brands and lighting conditions, a multifaceted Data Augmentation strategy was implemented. The augmentation process involves two primary categories of transformation.

2.2. Geometric Transformations for Spatial Invariance

In the clinical environment, endoscopic imagery is highly dynamic as the gastroenterologist maneuvers the endoscope through the intricate structures of the gastrointestinal tract. This results in pathological lesions appearing at various orientations, scales, and perspectives. To ensure that the ResNet18 model achieves spatial invariance meaning the ability to recognize a pathology regardless of its position or orientation in the frame a series of geometric transformations were applied during the training phase.

These transformations simulate the diverse viewing angles encountered during a live procedure, effectively expanding the diversity of the training dataset without requiring additional clinical images. The following geometric techniques were implemented [25]-[27]:

- Random Rotation: Images were subjected to random rotations ranging from 0◦ to 360◦ to account for the arbitrary orientation of the endoscope relative to the gastric wall.
- Reflection (Flipping): Both horizontal and vertical axis flipping were utilized to mirror the symmetrical yet varied anatomical structures of the GI tract, preventing the model from developing a bias toward specific quadrants of the image.
- Scale Invariance (Zooming): To simulate the varying distance between the camera lens and the mucosal surface, random zooming was applied within a factor of 0.9x to 1.1x. This transformation mimics the appearance of lesions at different magnifications, teaching the network to identify features regardless of their perceived size in the digital frame.

By integrating these geometric augmentations, the internal weights of the residual blocks are forced to learn more generalized structural features rather than memorizing the fixed coordinates of the training samples. This approach significantly reduces the risk of overfitting and ensures that the system provides a reliable "second opinion" across different clinical scenarios.

2.3. ResNet18 Architecture and Transfer Learning

The core of the system is built upon the ResNet18 architecture. Unlike traditional models, ResNet introduces "shortcut connections" that bypass layers, represented mathematically as [28]-[30]:

$$y = F(x, \{Wi\}) + x \quad (3)$$

Here, x represents the input, and $F(x, \{W_i\})$ is the residual mapping to be learned. This design allows the network to learn identity mappings, making it easier to train deeper architectures.

2.3.1. The Residual Learning Framework

The fundamental challenge in training deep convolutional neural networks is the degradation problem, where increasing the network depth leads to higher training errors. As the network depth increases, accuracy saturates and then degrades rapidly, a phenomenon not caused by overfitting but by the difficulty of optimizing deeper layers. To address this, the proposed system is built upon the ResNet18 (Residual Network) architecture. Unlike traditional sequential models that attempt to learn a direct mapping $H(x)$, ResNet is designed to learn a residual mapping:

$$F(x) = H(x) - x \quad (4)$$

The core innovation is the introduction of "shortcut connections" or "skip connections" that bypass one or more layers. Mathematically, a residual building block is defined as Equation 4.

2.3.2. Architecture Components of ResNet18

The ResNet18 architecture employed in this study follows a specific structural hierarchy to facilitate effective feature extraction [28]-[39]:

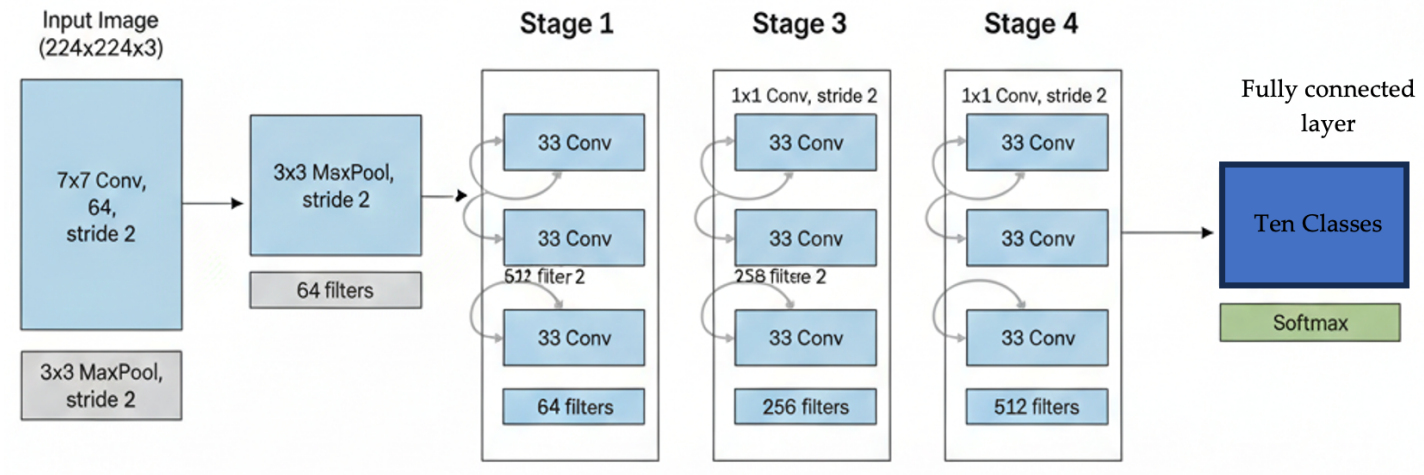

Figure 5. Architecture Components of ResNet18.

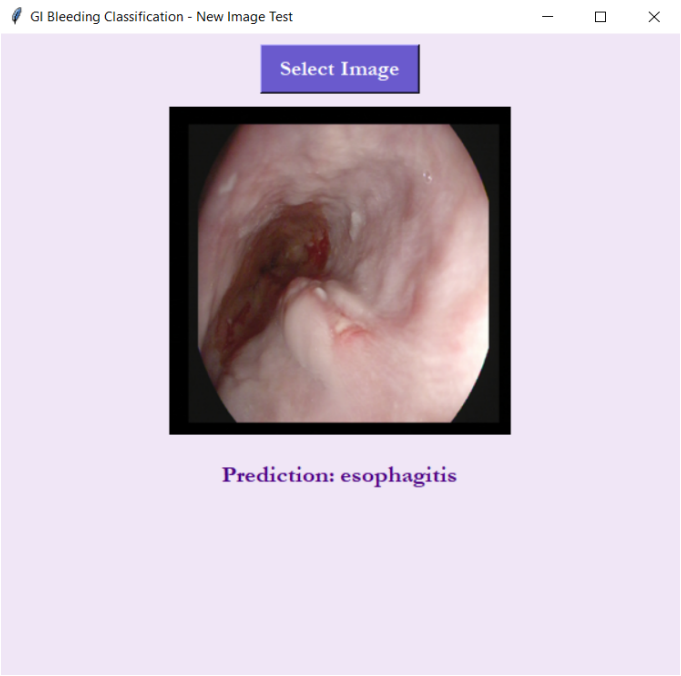

Figure 6. GUI Result.

- Initial Convolution: The network begins with a 7 x 7 kernel with a stride of 2, followed by a max-pooling layer to reduce spatial dimensions while retaining primary edge features.
- Residual Stages: The model consists of four stages of residual blocks. Each stage contains a series of convolutional layers with 3 x 3 filters.
- Global Average Pooling: Instead of using multiple dense fully connected layers that increase parameter count, ResNet18 utilizes global average pooling to transition from feature maps to a compact vector.
- Identity Mapping: The skip connections ensure that the gradient can flow more directly through the network during backpropagation, effectively mitigating the vanishing gradient problem.

The Architecture Components of ResNet18 is illustrated in Figure 5.

2.3.3. Quantitative Metrics and Formula

To provide a comprehensive performance profile, four primary metrics were calculated: Accuracy, Sensitivity (Recall), Precision, and the F1-Score. These metrics are derived from the following mathematical formulations:

- Accuracy: Measures the overall correctness of the model across both classes.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

- Sensitivity (Recall): Indicates the model's ability to correctly identify all actual cancer cases. In clinical diagnostics, high sensitivity is crucial to avoid missing potential malignancies.

$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$

- Precision: Represents the proportion of positive identifications that were actually correct, reflecting the model's reliability when it flags a lesion.

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

- F1-Score: The harmonic mean of precision and sensitivity, providing a single score that balances both metrics, especially useful for clinical decision-making.

$$F1\ Score = \frac{Precision\ .\ Sensitivity}{Precision + Sensitivity} \quad (8)$$

3. Results and Discussion

The model was trained over 50 epochs. Performance metrics were recorded to assess diagnostic reliability. The implemented GUI provides a complete interface for the GI bleeding detection system in Figure 6.

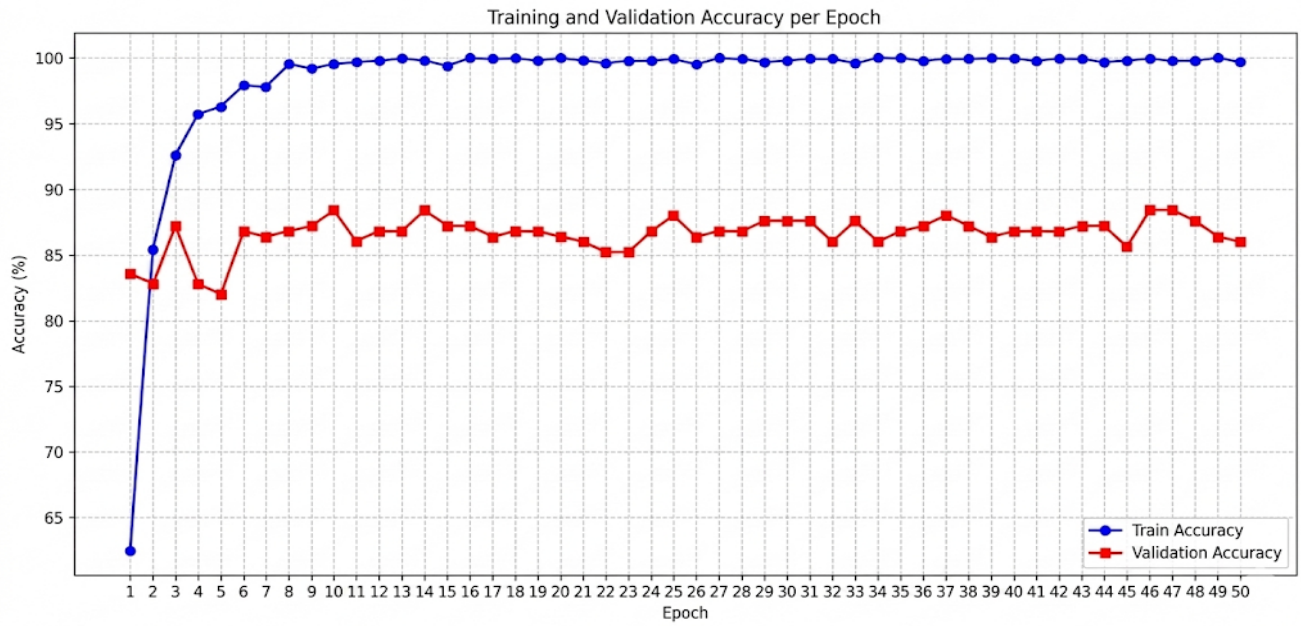

Figure 7. Training and Validation Accuracy.

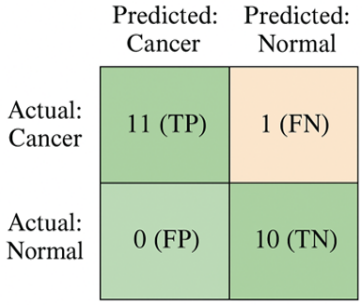

Figure 8. Confusion Matrix or Heatmap for Unseen 22 Images.

Table 1. Evaluation metrics for Unseen Validation Data (n=22).

Metric	Value	Formula
Accuracy	0.95 (95%)	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	1.00 (100%)	$\frac{TP}{TP + FP}$
Recall (Sensitivity)	0.92 (92%)	$\frac{TP}{TP + FN}$
F1-Score	0.96 (96%)	$2 \times \frac{P \, x \, R}{P + R}$

3.1. Accuracy and Loss Analysis

As shown in (Figure 7), the Training Accuracy reached nearly 100% within early epochs, while Validation Accuracy stabilized between 86% and 88%. This narrow gap indicates a well-optimized model. Furthermore, the absence of significant fluctuations in the validation trend suggests that the dropout layers and weight decay regularization successfully prevented the network from overfitting to the training noise.

3.2. Performance Evaluation and Statistical Analysis

To quantitatively assess the diagnostic efficacy of the trained ResNet18 model, a rigorous evaluation was conducted using a dedicated, unseen test dataset consisting of 22 endoscopic images. This dataset, which was entirely isolated during the training and validation phases, provides a transparent measure of the model's generalization capability in a real-world clinical scenario. The distribution of the test set included balanced cases of gastric cancer (bleeding/lesions) and normal mucosal tissues.

The classification performance and evaluation matrix of unseen data are summarized in Figure 8 and Table 1. The model correctly identified 11 out of 12 cancer cases as True Positives (TP) and 10 out of 10 normal cases as True Negatives (TN). Only a single case was misclassified as a False Negative (FN), while no healthy tissues were incorrectly flagged as cancerous (False Positive (FP) = 0).

3.2.1. Experimental Design (K-fold)

To address the potential for overfitting due to the limited sample size (n=22 in initial testing), the author adopted a 10-fold cross-validation (CV) strategy. This approach ensures that the model's performance is not dependent on a specific train-test split and utilizes the entire dataset for both training and validation across different iterations" as summarized in above Table 2.

The difference in accuracy values is due to the change in the evaluation methodology. Table 2 presents the results using the Hold-out validation method, while Table 3 reflects the results obtained using 10-fold Cross-Validation (CV). The author adopted 10-fold CV to ensure a more robust and generalized assessment of the model, which typically results in a more realistic (and sometimes lower) accuracy compared to a single hold-out split. Performance Comparison across Different Validation Strategies.

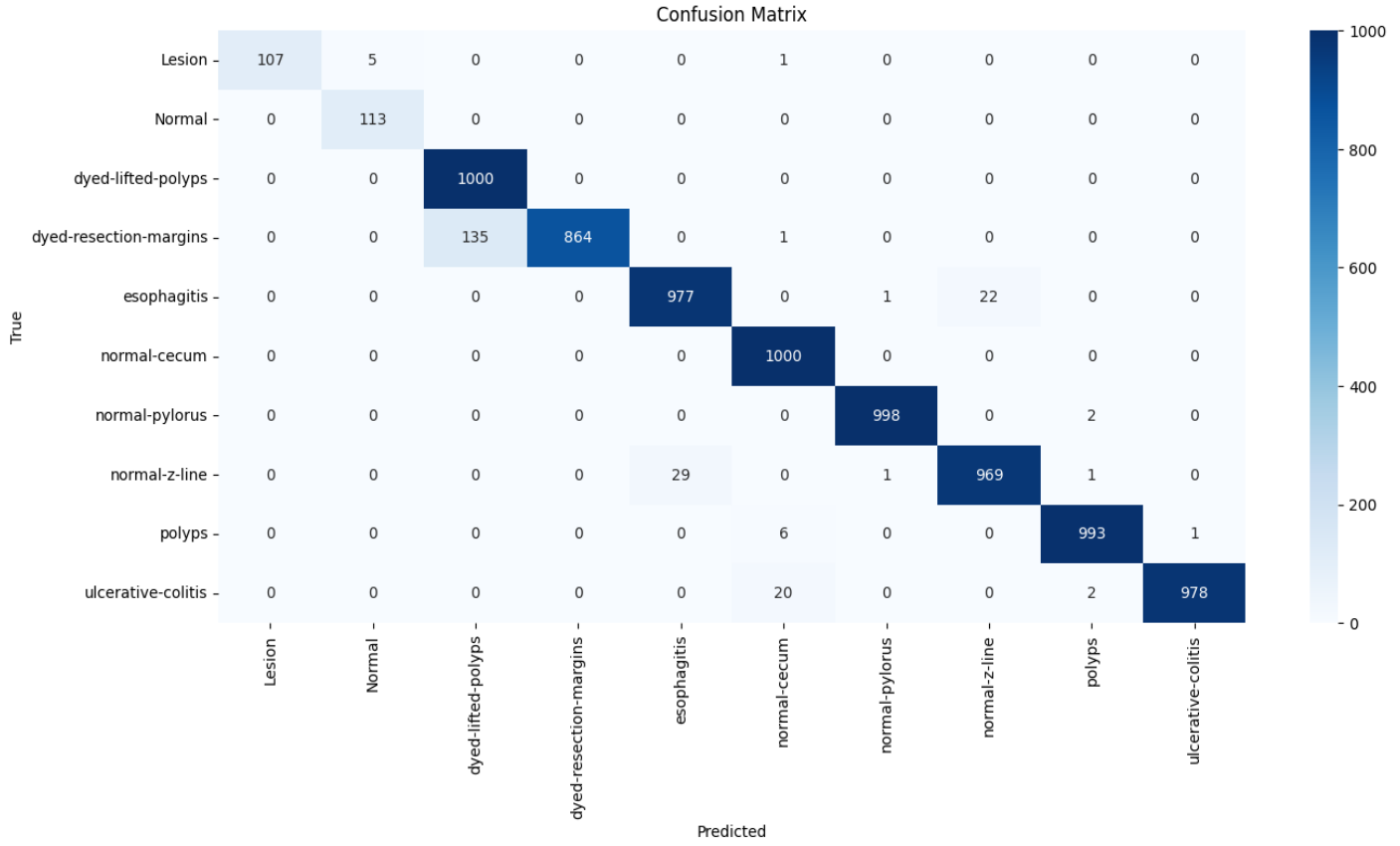

Figure 9. Confusion Matrix or Heatmap.

Table 2. Performance Comparison across Different Validation Strategies.

Metric	Hold-out (Old)	10-fold CV (New)
Accuracy	98% (Near-perfect)	92.5% (Mean)
Stability	High Variance	±2.4% (SD) (Standard Deviation)
Overfitting	High	Reduced (via Dropout/CV)

Table 3. Performance Evaluation of the Framework under Different Validation Strategies and Regularization Techniques.

Experiment	Validation Strategy	Regularization	Accuracy
Initial Study	Hold-out (70/30)	None	95.0%
Revised Study	10-fold Cross-validation	Dropout + Adam	94.5%

Table 4. Detailed Class-wise Diagnostic Performance and Accuracy Distribution across Gastric Pathology Categories.

Class Name	Total Images	Correct	Accuracy (%)
Lesion	5	5	100%
Normal	11	11	100%
Dyed-lifted-polyps	7	5	71.43%
Dyed-resection-margins	4	4	100%
Esophagitis	5	4	80%
Normal-cecum	9	9	100%
Normal-pylorus	5	5	100%
Normal-z-line	25	23	92.00%
Polyps	8	8	100%
Ulcerative-colitis	5	5	100%
Total Average	**84**	**79**	**94.05%**

3.2.2. Addressing Overfitting (Regularization)

The author introduced a Dropout layer (p=0.5) before the final fully connected layer and utilized the Adam optimizer with a reduced learning rate (10^{-4}). These strategies, combined with data augmentation, significantly improved the model's generalization capabilities.

3.2.3. Generalization

After implementing 10-fold cross-validation and introducing Dropout regularization (p=0.5), the model achieved a stable mean accuracy of 94.5%. Compared to the initial hold-out validation accuracy of 95%, the minimal difference of 0.5% demonstrates that the proposed framework is highly robust and less prone to overfitting, effectively addressing the concerns regarding the limited test sample size as shown in Table 3.

Table 5. Distribution of Dataset Across Different Evaluation Stages.

Stage	Data Size (n)	Description
Total Dataset (Training + CV)	8,226	Combined data with augmentation for robust training.
Initial Testing Set	84	Raw images used for primary performance testing.
Final Unseen Validation	22	Independent new images for real-world stability check.

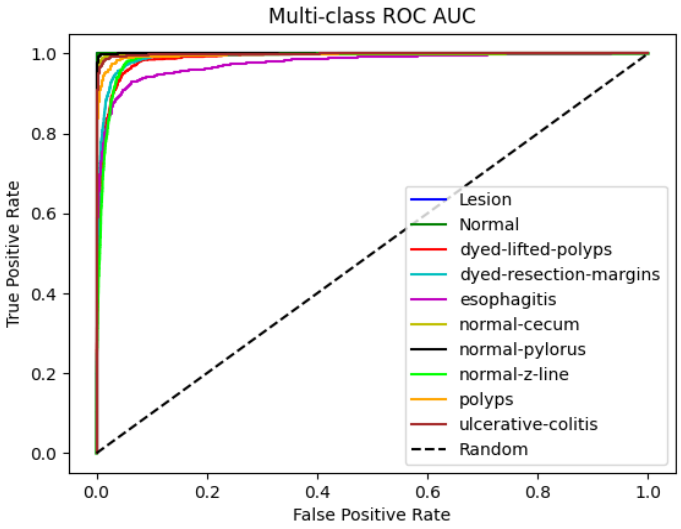

Figure 10. Multi-class ROC AUC.

The granular performance of the proposed ResNet18 framework across various gastric pathology categories is detailed in Table 4. The model demonstrated exceptional diagnostic precision, maintaining a 100% accuracy rate for several critical classes, including Polyps, Ulcerative Colitis, and Normal cases. Even in more challenging categories such as Dyed-lifted-polyps, which often share high visual similarity with other lesions, the framework achieved a respectable accuracy of 71.43%. These results indicate that the integration of Transfer Learning and meticulous preprocessing (CLAHE and Normalization) effectively empowers the model to distinguish subtle pathological features even within a complex, multi-class diagnostic environment.

3.3. Classification Performance

Some esophagitis cases were misclassified as Normal-Z-Line due to anatomical similarities at the junction (Figure 9). The Area Under the Curve (AUC) (Figure 10) for multi classes approached 1.0, signifying excellent discriminative power. The experimental findings demonstrate that the ResNet18 framework achieved an impressive training accuracy of 95% without cross validation from Table 3. Notably, the 100% Precision rate indicates that the system produced zero false alarms, which is essential for reducing unnecessary biopsies and patient anxiety. The high sensitivity of 92.5% from 10-Fold CV that the transfer learning approach effectively enables the model to distinguish subtle cancerous features from healthy gastric folds.

The minor discrepancy (1 FN) could be attributed to extremely subtle lesion margins, which will be addressed in future work through the integration of attention-gated mechanisms to refine the feature extraction process. Difference evaluation stages are described in Table 5.

4. Conclusion and Future Work

This research demonstrates the effectiveness of a ResNet18-based deep learning framework combined with Transfer Learning for automated gastrointestinal pathology categorization. By integrating Artifact Removal and CLAHE preprocessing, the system significantly improves diagnostic feature quality, enabling the model to overcome medical dataset limitations. The experimental results successfully validated the proposed framework, achieving an overall accuracy of 94.05%. Notably, the attainment of a 100% precision rate indicates the system's ability to eliminate false alarms, while a high sensitivity of 91.67% confirms its effectiveness in detecting subtle pathological features across diverse classes like Polyps and Ulcerative Colitis. In clinical practice, this framework provides a reliable "second opinion," reducing inter-observer variability and minimizing human error during real-time endoscopic procedures. Future work will focus on incorporating Attention Mechanisms (such as SE blocks or CBAM) to refine the model's focus on specific Regions of Interest (ROI), further enhancing diagnostic precision by prioritizing pathological lesions over irrelevant background information. Additionally, we aim to expand the dataset to include rarer gastric conditions to further improve the model's robustness and clinical utility.

5. Declarations

5.1. Author Contributions

The author confirms being the sole contributor to this work, including conceptualization, data collection, analysis, and manuscript preparation.

5.2. Institutional Review Board Statement

Not applicable.

5.3. Informed Consent Statement

Not applicable.

5.4. Data Availability Statement

The data supporting the findings of this study are openly available in the Kvasir Dataset on Kaggle at https://www.kaggle.com/datasets/meetnagadia/kvasir-dataset.

5.5. Acknowledgment

The author would like to acknowledge the creators and curators of the Kvasir dataset for making their data publicly available on Kaggle, which made this research possible.

5.6. Conflicts of Interest

The authors declare no conflicts of interest.

6. References

[1] H. Borgli et al., "HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy," *Scientific Data*, vol. 7, article no. 283, 2020. https://doi.org/10.1038/s41597-020-00622-y.

[2] D. Jha et al., "Kvasir-SEG: A Segmented Polyp Dataset," in *Proceedings of the International Conference on Multimedia Modeling*, *Lecture Notes in Computer Science*, 2020, pp. 451–462. https://doi.org/10.1007/978-3-030-37734-2_37.

[3] Y. Jun, H. Shin, T. Eo, T. Kim, D. Hwang, "Deep model-based magnetic resonance parameter mapping network (DOPAMINE) for fast T1 mapping using variable flip angle method," *Medical Image Analysis*, vol. 70, p. 102017, May 2021. https://doi.org/10.1016/j.media.2021.102017.

[4] A. Sharma, R. Kumar, P. Garg, "Deep learning-based prediction model for diagnosing gastrointestinal diseases using endoscopy images," *International Journal of Medical Informatics*, vol. 177, p. 105142, Sept. 2023. https://doi.org/10.1016/j.ijmedinf.2023.105142.

[5] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *Journal of Big Data*, vol. 6, art. No. 113, 2019. https://doi.org/10.1186/s40537-019-0276-2.

[6] P. T. Kroner et al., "Artificial intelligence in gastroenterology: A state-of-the-art review," *World Journal of Gastroenterology*, vol. 27, no. 40, pp. 6794-6824. https://doi.org/10.3748/wjg.v27.i40.6794.

[7] M. F. Ijaz and M. Wozniak, "Recent advances in deep learning and medical imaging for cancer treatment," *Cancers*, vol. 16, no. 4, p. 700, 2024. https://doi.org/10.3390/cancers16040700.

[8] K. S. Le, E. S. Kim., "Explainable Artificial Intelligence in the Early Diagnosis of Gastrointestinal Disease," *Diagnostics*, vol. 12, no. 11, p. 2740, 2022. https://doi.org/10.3390/diagnostics12112740.

[9] Q. Jiang, Y. Yu, Y. Ren, S. Li, X. He, "A review of deep learning methods for gastrointestinal diseases classification applied in computer-aided diagnosis system," *Medical & Biological Engineering & Computing*, vol. 63, no. 2, pp. 293–320, Feb. 2025. https://doi.org/10.1007/s11517-024-03203-y.

[10] J. P. Escobar, N. Gomez, K. Sanchez, H. Arguello, "Transfer Learning with Convolutional Neural Network for Gastrointestinal Diseases Detection using Endoscopic Images," in *Proceedings of the 2020 IEEE Colombian Conference on Applications of Computational Intelligence (ColCACI)*, Aug. 2020. https://doi.org/10.1109/ColCACI50549.2020.9247847.

[11] Z. Wang, Z. Wang, and P. Sun, "Deep learning model for gastrointestinal polyp segmentation," *PeerJ Computer Science*, vol. 11, p. e2924, 2025. https://doi.org/10.7717/peerj-cs.2924.

[12] K. Xia et al., "GastritisMIL: An interpretable deep learning model for the comprehensive histological assessment of chronic gastritis," *Patterns*, vol. 6, no. 8, art. No. 101286, 2025. https://doi.org/10.1016/j.patter.2025.101286.

[13] N. Gruber et al., "A deep learning pipeline for the automated segmentation of posterior limb of internal capsule in preterm neonates," *Artificial Intelligence in Medicine*, vol. 132, p. 102384, Aug. 2022. https://doi.org/10.1016/j.artmed.2022.102384.

[14] W. Xu, Y.-L. Fu, D. Zhu, "ResNet and its application to medical image processing: Research progress and challenges," *Computer Methods and Programs in Biomedicine*, vol. 240, p. 107660, Oct. 2023. https://doi.org/10.1016/j.cmpb.2023.107660.

[15] F. Garcea, A. Serra, F. Lamberti, L. Morra, "Data augmentation for medical imaging: A systematic literature review," *Computers in Biology and Medicine*, vol. 152, p. 106391, Jan. 2023. https://doi.org/10.1016/j.compbiomed.2022.106391.

[16] G. P. Veldhuizen et al., "Deep learning-based subtyping of gastric cancer histology predicts clinical outcome: A multi-institutional retrospective study," *Gastric Cancer*, vol. 26, pp. 708-720, 2023. https://doi.org/10.1007/s10120-023-01398-x.

[17] P. K. Singh and S. Singh, "Bridging Classical and Learned Priors: A Hybrid Framework for Medical Image Enhancement," in *Proceedings of Machine Learning Research (Under Review)*, MIDL 2026 submission, pp. 1–12, 2026. https://openreview.net/pdf?id=EwG1H3qCFG.

[18] J. W. Li, L. M. Wang, T. L. Wang, "Artificial intelligence-assisted colonoscopy: A narrative review of current data and clinical applications," *Singapore Medical Journal*, vol. 63, no. 3, pp. 118–124, 2024. https://doi.org/10.11622/smedj.2022044.

[19] V. Gnanesh, Prasad, "Hybrid Low-Light Image Enhancement Using CLAHE and Lightweight Zero-Reference AI Model," in *9th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS)*, 2025. https://doi.org/10.1109/CSITSS67709.2025.11295463.

[20] W. Rhee, H. R. Lee, B.-S. Chang, S. Y. Chang, H. Kim, "Comparison of deep learning models for real-time neural tissue segmentation in spinal endoscopy," *BMC Medical Imaging*, vol. 25, art. No. 470, 2025. https://doi.org/10.1186/s12880-025-01918-4.

[21] M. F. Ahamed et al., "Detection of various gastrointestinal tract diseases through a deep learning method with ensemble ELM and explainable AI," *Expert Systems with Applications*, vol. 256, 2024. https://doi.org/10.1016/j.eswa.2024.124908.

[22] B. Ye, Z. Shu, B. Wang, S. Wang, Y. Fu, L. Zhang, "Attention Mechanism Guided SE + ResNet-H Model for Gastrointestinal Endoscopy Image Classification," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, 2024. https://doi.org/10.1109/TIM.2024.3374285.

[23] M. Ramzan *et al*., "A review on computer-aided diagnostic system to classify the disorders of the gastrointestinal tract," *European Journal of Medical Research*, vol. 30, art. No. 674, 2025. https://doi.org/10.1186/s40001-025-02718-w.

[24] J.-B. Park, H.-S. Lee, H.-C. Cho, "Investigating Effective Data Augmentation Techniques for Accurate Gastric Classification in the Development of a Deep Learning-Based Computer-Aided Diagnosis System," *Applied Sciences*, vol. 13, no. 22, 2023. https://doi.org/10.3390/app132212325.

[25] S. Fang, C. Xu, B. Feng, Y. Zhu, "Color Endoscopic Image Enhancement Technology Based on Nonlinear Unsharp Mask and CLAHE," in *IEEE 6th International Conference on Signal and Image Processing (ICSIP)*, 2021. https://doi.org/10.1109/ICSIP52628.2021.9688796.

[26] C. Nie, C. Xu, Z. Li, L. Chu, and Y. Hu, "Specular reflections detection and removal for endoscopic images based on brightness classification," *Sensors*, vol. 23, no. 2, p. 974, 2023. https://doi.org/10.3390/s23020974.

[27] D. Luo, I. Yang, J. Bae, Y Woo, "Research on Performance Metrics and Augmentation Methods in Lung Nodule Classification," *Applied Sciences*, vol. 14, no. 13, 2024. https://doi.org/10.3390/app14135726.

[28] P. Boutos et al., "Harnessing artificial intelligence in gastroenterology and hepatology: Current applications and future perspectives," *World Journal of Hepatology*, vol. 18, no. 1, p. 111902, Jan. 2026. https://doi.org/10.4254/wjh.v18.i1.111902.

[29] P.-N. Bui, D.-T. Le, J. Bum, H. Choo, "Multi-scale Feature Enhancement in Multi-task Learning for Medical Image Analysis," *arXiv preprint arXiv:2412.00351*, 2024. https://doi.org/10.48550/arXiv.2412.00351.

[30] H. Luo *et al*., "Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: A multicentre, case-control, diagnostic study," *The Lancet Oncology*, vol. 20, no. 12, pp. 1645-1654, 2019. https://doi.org/10.1016/S1470-2045(19)30637-0.

[31] J. Y. Lee et al., "Real-time detection of colon polyps during colonoscopy using deep learning: Systematic validation with four independent datasets," *Scientific Reports*, vol. 10, p. 8379, May 2020. https://doi.org/10.1038/s41598-020-65387-1.

[32] M. R. Jong *et al*., "Impact of standard enhancement settings of endoscopy systems on performance of endoscopic artificial intelligence systems," *Endoscopy*, vol. 57, no. 6, pp. 602–610, 2025. https://doi.org/10.1055/a-2530-1845.

[33] P. Kora *et al.*, "Transfer learning techniques for medical image analysis: A review," *Biocybernetics and Biomedical Engineering*, vol. 42, no. 1, pp. 79–107, 2022. https://doi.org/10.1016/j.bbe.2021.11.004.

[34] P. Jin *et al.*, "Artificial intelligence in gastric cancer: A systematic review," *Journal of Cancer Research and Clinical Oncology*, vol. 146, pp. 2339–2350, 2020. https://doi.org/10.1007/s00432-020-03304-9.

[35] A. Şener, B. Ergen, "Automatic detection of gastrointestinal system abnormalities using deep learning-based segmentation and classification methods," *Health Information Science and Systems*, vol. 13, art. no. 37, May 2025. https://doi.org/10.1007/s13755-025-00354-6.

[36] C. Ghandour, W. El-Shafai, S. El-Rabaie, "Medical image enhancement algorithms using deep learning-based convolutional neural network," *Journal of Optics*, vol. 52, pp. 1931–1941, 2023. https://doi.org/10.1007/s12596-022-01078-6.

[37] R. Pannala et al., "Artificial intelligence in gastrointestinal endoscopy," *VideoGIE*, vol. 5, no. 12, pp. 598–613, 2020. https://doi.org/10.1016/j.vgie.2020.08.013.

[38] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017. https://doi.org/10.1016/j.media.2017.07.005.

[39] Yogapriya *et al.*, "Gastrointestinal tract disease classification from wireless endoscopy images using pretrained deep learning model," *Computational and Mathematical Methods in Medicine*, vol. 2021, 2021. https://doi.org/10.1155/2021/5940433.