Journal of Theoretical and Applied Information Technology
15th April 2026. Vol.104. No.7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN INTERPRETABLE DEEP LEARNING FRAMEWORK FOR GASTRIC POLYP DETECTION AND SEGMENTATION USING GRAD-CAM AND LIME

JYOTI GARG[1], MOHIT CHHABRA[2]

[1]Research Scholar, Maharishi Markandeshwar Deemed to be University, Mullana,Ambala, Department of Computer Science and Engineering, India
[2]Associate Professor, Maharishi Markandeshwar Deemed to be University, Mullana,Ambala, Department of Computer Science and Engineering, India

E-Mail: [1]jyotigarg1055@gmail.com, [2]chhabra108@mmumullana.org

ABSTRACT

Correct identification and segmentation of precancerous polyps in endoscopic images takes the first steps toward the early detection of stomach cancer. While Deep Learning has greatly improved in the field of medical image processing, clinical trust in such models is limited. This paper develops an explainable deep learning model using a publicly available Kvasir-SEG dataset for the segmentation of stomach cancer polyps. In this methodology, it encompasses preprocessing of images, residual U-Net segmentation, and evaluation of the model using IoU, Dice Score, accuracy, and loss metrics. For interpretability, LIME provides local explanations at the pixel level for individual predictions, whereas Grad-CAM has been used to highlight class-discriminative regions that influence segmentation decisions. The integrated XAI framework in the study offers global interpretability and instance-specific interpretability; therefore, clinicians are able to visually confirm why the model designated a certain area as suspected of carrying cancer. The obtained experimental consequences showed impressive configuration of model-highlighted regions and actual polyp structures, with great segmentation capability. The proposed approach provides relevancy and transparency. The proposed model has been optimized by using multiple optimization models such as Attention UNet, EfficientNet, ResUNet, DenseNet, and Transformer. This integrated ensemble model incorporated with LIME and Grad-CAM produces a deep learning based sudsier that could easily be used for stomach cancer diagnosed through endoscopic analysis.
Keywords: *LIME, Grad-CAM, Gastric Cancer, Polyp Segmentation, Explainable AI(XAI)*

1. INTRODUCTION

Gastric cancer is considered one of the major causes of cancer-related mortality worldwide. In its starting phases it grows very slowly and don't show any symptoms. The multiple available treatment options in the early stages are quite limited due to unavailability of clinical symptoms at the early stages [1]. Polyps which are abnormal gastric lesion, its early detection at the initial stages are quite challenging. For the overall survival rate and better patient health its early detection is highly required. To identify these types of present abnormalities, endoscopy is required. But along with this, the accurate diagnose depends on multiple factors such as the visual acuity, the clinician's experience and also its ability to identify multiple subtle mucosal patterns [2]. To address these problems, computer aided diagnosis methods have been implemented to support clinicians and fasten the identification process in the early detection of gastric abnormalities. These methodologies depend on computer model for evaluating the different images obtained from datasets [3]. However, existing deep learning models mainly focuses on improving the accuracy but not able to provide clear clarifications based on their predictions. In these situations, these framework act as a black box system, that can restrict the usability and consistency in real clinical situations. The key aim of this research is to develop an interpretable deep learning framework for gastric polyp detection and segmentation. The proposed model integrates explainable AI methods with ensemble models with the inclusion of Grad-CAM and LIME to accomplish improved accuracy and enhanced transparency. The proposed approach is significant as it improves the interpretability of model decisions and enhance trust in AI-based medical systems. Furthermore, it facilitates more informed

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

clinical decision-making by identifying relevant regions in medical images.

The segmentation process based on Deep learning has shown great results in delineating gastric polyps and lesion boundaries. Due to some limitations and to improve the diagnosis process, an explainable artificial intelligence (XAI) based system has been proposed. In the healthcare, the limited resources is a crucial concern [4], as it can affect the patient's health.

Different Techniques such as Grad-CAM and LIME permit visualization of image sections that effect a model's decision-making process [5]. These methods additionally provide information about spatial relevance and region-level contributions within the input image. These visuals explanations enhance user confidence and provide additional support for model evaluation. In medical field, concept of XAI is quite important as technicians essentially ensure that the model mainly focus on relevant regions rather than irrelevant artifacts. While identifying abnormal tissues, clinicals should confirm that the model looking for textural patterns rather than noise [6]. XAI facilitates alignment between clinical knowledge and model reasoning. Deep learning algorithms are capable of learning complicated patterns and features that may not be able captured by human eye [7]. These defined models outperform classical processing techniques and are better as compared to classical image processing methods [8], and the increased complexity makes model functioning much difficult to interpret. The "black box" feature restricts their independent deployment in high- risk clinical decision-making [9]. XAI overcomes this barrier by making deep learning algorithms more transparency. Rather than replacing deep learning models, XAI complements them by providing interpretable explanations of their predictions [10]. This could be thought of as a synergistic framework where XAI offers insight and deep learning offers accuracy [11]. LIME outspreads it by partitioning the pixels of an image into superpixels and displaying how each section can produce the final results [12].

The Main Contributions of this work are as follows: (1)A Deep Learning ensemble model is proposed and evaluated on the Kvasir-SEG dataset to efficiently detect polyps allied with stomach cancer. (2) Gradient-weighted Class Activation Mapping (Grad-CAM) is employed to create class-discriminative heatmaps highlighting areas most significant in the segmentation progression. (3) Local Interpretable Model-agnostic Explanations (LIME) are utilized to provide local, instance-level descriptions by evaluating the role of superpixel regions to the predicted segmentation masks. (4) Comprehensive quantitative and qualitative investigates are conducted to inspect the alignment between explainability outputs and clinically relevant regions.

So, accuracy and other performance parameters can be improved by the proposed integrated XAI framework.

2. RELATED WORK

A lot of research already has been carried out in the detection of gastric cancer and polyp segmentation. All the existing techniques depends mainly on Different CNN based architectures such as Mask R-CNN, U-Net etc. These models have shown improvements in detection, but still lack in some other challenges. To better results, this proposed work proposes an ensemble-based framework integrated with explainable AI techniques to enhance both accuracy and model interpretability.
Zhang et al. [13] described Gastroscopy as a screening method for gastric cancer detection. Less than 10% of cases of early gastric cancer (EGC) are detected, even though symptoms might vary greatly. To increase detection rates, researchers experimented with deep learning systems. The main purpose of creating this framework was to identify segment lesions and EGC.1120 such pictures were used to train and evaluate the model, which also has a purification mechanism for bi-directional FE, fusion, and feature channels and region. The proposed methodology attained an accuracy, sensitivity and Dice score of 94%, 92%, and 0.91 respectively, outperforming the baseline Mask R-CNN in lesion localization and boundary segmentation.

This methodology [14] analyzed the performance of three common region CNN and RCNN models, using 3659 gastroscopic pictures for early diagnosis of stomach cancer. The results depict that fast comparative accuracy with Cascade RCNN prioritizes negative detection, while Faster RCNN focuses on positive detection. This study concluded that deep learning-based methods using endoscopic pictures for early detection of cancer could reduce patient mortality rate. Cascade RCNN achieved an overall accuracy of 93%, as compared to Faster RCNN having sensitivity of 91%. Further research is needed to test the specificity, accuracy, precision, recall, and accuracy of different models of the RCNN series.

This study [15] used six Machine learning algorithms to develop an accurate, correct, and cheap diagnostic model that helps to classify patients into different groups considering their risk regarding stomach cancer. Data was taken from

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2029 individuals in their gastric cancer database obtained from Ayatollah Taleghani Hospital, Abadan City, Iran. The paper showed that six machine learning models were trained to develop prognosis models earlier and after applying the relief feature selection strategy. High intake of salt, Helicobacter pylori infection, and chronic atrophic gastritis were some of the 11 risk factors for stomach cancer. The XGBoost model performed the best for forecasting stomach cancer risk. The study showed that ML methods had the potential to identify high-risk patients for invasive exams and for prescreening of stomach cancer. However, more verification was needed in an extended multicenter population.

The XGBoost system acquired accuracy, precision, and recall of 89%, 87%, 90% respectively, outperforming other ML classifiers for risk stratification. The study [16] performs classification of ulcers present in stomach by diagnosis the images dataset by using different CNN based models. Pytorch and previously processed photos are used to implement the model. The residual attention module enhances the Xception model by considering the different performance parameters such as sensitivity, specificity, accuracy, F1 score, and precision for the identification of benign and malignant gastric ulcer lesions. The Specificity achieved by this proposed methodology was found to be 93 % which is better as compared to other traditional methods. Additionally, the classification efficacy of the Xception CNN on stomach ulcer lesions on common images is improved by the residual attention technique.

The study [17] described the effectiveness of the CNN model in identifying early gastric cancer (EGC). The authors achieve better results and plotting the summary receiver operating characteristic curve. The CNN model outperformed endoscopists in diagnosing endocarditis using digital endoscopic pictures, with an SROC of 0.95. The model's scalability could enhance endoscopists ability to accurately stratify EGC patients and reduce workload.

This study [18] describes a computer-aided diagnosis system that can differentiate and categorize precancerous diseases such as bleeding, ulcers, gastritis, and gastric polyps from stomach cancer. The method involves two phases and a deep learning based Xception model. The picture segmentation process utilizes Google's Auto Augment, SLIC super pixel, and fuzzy C-means algorithm with performance measures determined through test sets. The classification model's accuracy has improved from 0.90 to 0.96, eliminating the need for a surgical biopsy due to fully automated technology offering effective therapy alternatives.

This study [19] addressed the function of platelet-lymphocyte ratio (PLR) and neutrophil-lymphocyte ratio (NLR) in the early detection of stomach cancer. Peripheral blood samples from 3,219 healthy controls and 2,606 GC patients were examined by researchers. It was discovered that NLR and PLR exhibited higher diagnostic sensitivity in GC than CEA and CA19-9. The study also discovered that in the early stages of GC, male patients' systemic inflammatory indicators were more useful. Combining NLR and PLR was more significant in terms of diagnosis for GC than standard tumor markers CEA and CA19-9. This suggests that systemic inflammation indicators are more useful in male patients.

Authors [20] developed a convolution neural network based deep learning-based system to distinguish between benign and malignant stomach ulcers from a database of endoscopic ulcer pictures. The final CNN model obtained area under the curve value around 0.91, specificity score of 0.75 and specificity of 0.92 in classifying photos as malignant ulcers. This study showed how a CNN model can be implemented for the automatic assessment of malignant potential in pictures of gastric ulcers. Future research can bring efficiency in patient care and permit personalized monitoring of endoscopic methods by improving the distinction of benign from malignant ulcers during endoscopies.

According to this, [21] proposed Convolution neural network-based framework for polyp detection and segmentation by using the image-based dataset, showing that residual connections improve both gradient flow and segmentation performance compared to the traditional U-Net. Their model outperformed others in terms of IoU and Dice scores, particularly for small or irregularly shaped polyps. The study presented a benchmark for the segmentation-based technique in cancer screening and proved the utility of ResUNet in detecting gastrointestinal disorders.0.90 is the Dice coefficient value and 0.86 is IoU value obtain from this methodology which is better as compared to u-Net architectures.

Authors [22] summarized the deep learning architectures for endoscopic image processing, such as the U-Net, SegNet, and FCN models. Their results indicated that encoder-decoder architectures were more consistent in performing the task of lesion localization than classification-based approaches. The authors highlighted that this was intensely relevant for

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

distinguishing benign tissue from precancerous stomach polyps.

Fan et al. [23] proposed the PraNet (Parallel Reverse Attention Network) for state-of-the-art performance in gastrointestinal polyp segmentation. PraNet uses edge refinement in order to highlight the exact delimitation of the polyps. The results obtained from this methodology shows better and promising results as compared to state of art models. The defined framework's dice score is obtained as 0.92 and also the false negative rates has been reduced to 8% shows improved overall performance as compared to existing models.

Selvaraju et al. [24] came up with LIME, that is a framework for characterizing black-box models. Since then, LIME has gained favor as an XAI technique to help CNN make decisions. Their approach forms the basis for model transparency in AI-based stomach cancer diagnosis.

Grad-CAM by [25], highlights the main regions that influence the predictions of CNNs in order to provide visual explanations. Grad-CAM has also been utilized in medical imaging to identify the site of diseases in MRI, X-ray, and endoscopic images. The work will be an important step in further understanding how characteristics relate to stomach cancer diagnosis. The qualitative consistency has been exceedingly over 90 percent across multiple evaluated samples further supports interpretability in medical imaging.

From all these above key points, it can be concluded that although existing methodologies achieve good accuracy, still they lack on interpretability. Most of the existing models also face challenges in managing small and complex polyps region. To overcome and handle all these limitations, the proposed study introduces an ensemble-based approach integrated with explainable AI techniques such as Grad-CAM and LIME, that can improve both segmentation performance and model transparency.

3. PROPOSED METHODOLOGY

The framework model follows a structured strategy that consist of several stages to ensure accurate and interpretable segmentation of gastric polyps. For better clarity and reproducibility all the stages of framework are described step by step.

In segmentation studies, this proposed methodology proposes an explainable ensemble deep learning framework for accurate and transparent gastric polyp segmentation. This technique combines numerous segmentation frameworks to enhance the reliability estimation and combines post-hoc methods to enhance clinical interpretability. To generate a final segmentation mask, many segmentation approaches got parallelly trained and their average prediction value is observed. For better decision-making Grad CAM and LIME are applied for better model predictions. After dataset preparation, multiple segmentation models are trained in parallel and their predictions are fused to generate a final segmentation mask. To support clinical trust and decision-making, Grad-CAM and LIME are applied to envision the areas and features influencing model predictions.

The key purpose of this research is to get precise results of gastric polyp segmentation. This approach can be sub divided into multiple steps.

Step 1: Data Preprocessing- The input images are Standardized, normalized, and augmented are obtained from the Kvasir-SEG Dataset https://www.kaggle.com/datasets/debeshjha1/kvasirseg) to reduce the overfitting and imbalancing.

Step 2: Segmentation Using ensemble model- In this part, the collective model combines different individual models having high feature representation and strong localization capability, so able to refine robustness and segmentation accuracy as compared to state of art models.

Step 3: Model Training and Validation- To address class imbalancing and improving boundary delineation, every segmentation model becomes proficient by grouping class entropy Dice Loss. The Dice Coefficient, recall, precision, Intersection over Union (IoU) are the different performance parameters on which performance can be evaluated.

Step 4: Explainability and Visualization- Clinicians can now comprehend the model's decision bounds and feature emphasis thanks to the local and global interpretability maps created using Grad-CAM and LIME. Grad-CAM is used to influence segmentation conclusions whereas LIME offers instance level or local clarifications by analysing superpixel contributions. The model behaviour can be better understood by the explanations of these visuals.

3.1 Dataset Description

In this study the experiments are conducted by using the Kvasir-SEG dataset that consists of 1000 polyp images along with ground truth masks. The resolutions pixels of images may vary from 332×487 to 1920×1072.The dataset is widely used for research related to polyp segmentation, localization, detection, and classification, making it suitable for evaluating computer vision–based algorithms. The accessibility of exact ground truth annotations permits supervised training and objective performance valuation of segmentation models. The data pre-processing is considered to diminish variability among images, remove non-diagnostic artifacts, and enhance the diversity of training samples. These preliminary steps are essential to propose explainable ensemble segmentation framework and ensure reliable generation of interpretability maps using Grad-CAM and LIME. All endoscopic images and their corresponding ground truth masks got resized to a standardized size, while preserving the structural integrity of polyp regions. Standardizing the input dimensions facilitates stable model convergence during training and enables optimal utilization of GPU memory.

Ground truth masks were converted into binary format, assigning a value of 1 to polyp regions and 0 to background regions, thereby simplifying loss computation during segmentation training.Some original images contain a green endoscope position marker (ScopeGuide™, Olympus, Tokyo, Japan) located near the image corners. Since this marker is not relevant to the segmentation task and may introduce non-diagnostic bias, it was replaced with black patches by the dataset providers. These regions were retained unchanged throughout the pre-processing stage to maintain consistency with the official dataset annotations.

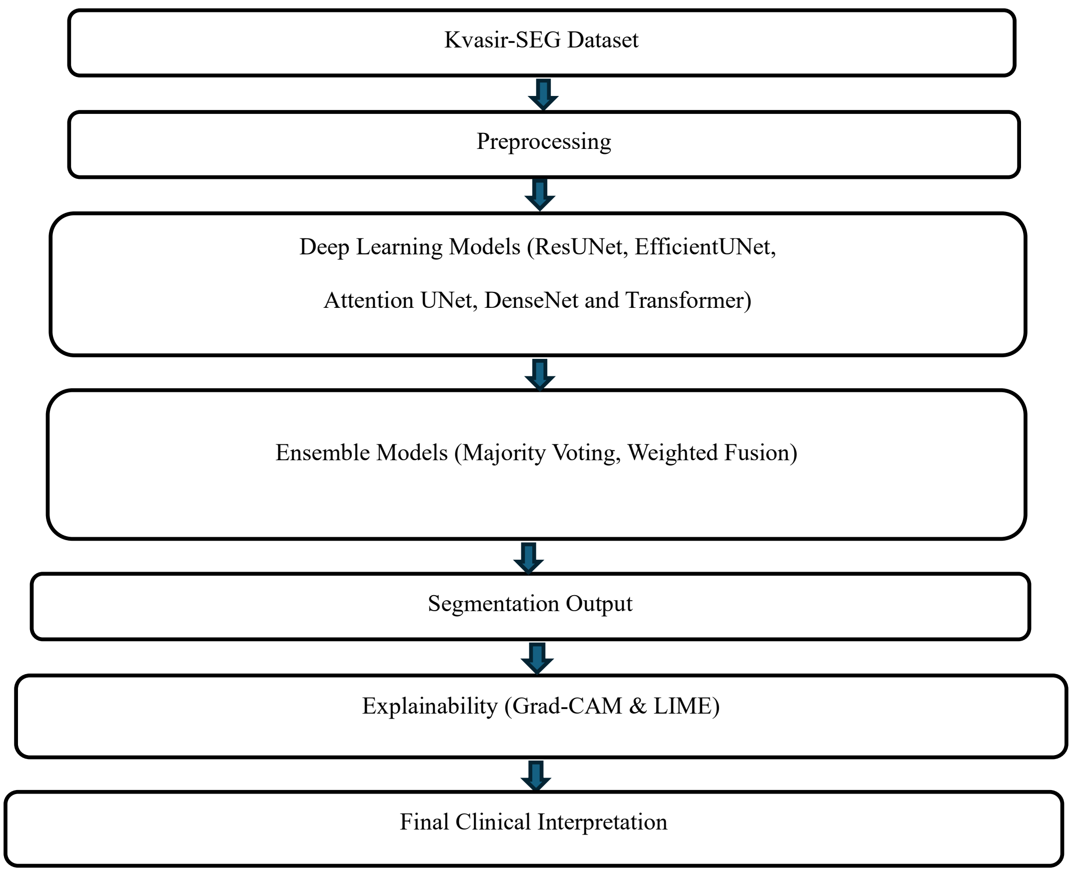

Figure 1: Explainable Ensemble Deep Learning Technique for Segmenting Gastric Polyps

Figure 1 illustrates the step-by-step workflow of the proposed collective deep learning methodology for the gastric polyp segmentation. The procedure starts with the Kvasir-SEG dataset, followed by image pre-processing methodologies to normalize inputs and diminish artifacts. Numerous CNN based deep learning segmentation models got proficient together. The estimates created by these methodologies are combined using an ensemble fusion strategy based on majority voting or weighted averaging.

The dataset were divided into 70% training, 20% validation, and 10% testing subsets to ensure unbiased evaluation while maintaining sufficient samples for effective model learning. Ensemble Segmentation and Explainability. Multiple segmentation models are trained together. The different models are ResUNet, EfficientNet-based U-Net, Attention U-Net, DenseNet-UNet, and Transformer-based U-Net architectures. Each model having different spatial and contextual information, improving robustness across variations in polyp size, shape, and appearance. The obtained output got combined using an ensemble fusion strategy, based on majority voting or weighted averaging, resulting in more consistent and reliable segmentation masks compared to single-model predictions. For better results and interpretability, Grad-CAM and LIME techniques are used. The Grad-CAM able to generate global spatial attention maps highlighting the segmentation output whereas LIME offers local, instance-level explanations by evaluating the role of superpixel regions to the final prediction.

3.2 Experimental Setup

These experiments are conducted on the basis of publicly available Kvasir-SEG dataset, that comprises annotated polyp images and is widely used for polyp segmentation tasks. The dataset has been split into three partitions, namely training, validation, and testing subsets.

The input images in the preprocessing stage are resized, normalized, and augmented. These processes are required to reduce overfitting and increase model generalization. The models got trained with the help of key parameters such as batch size learning rate, and number of repetitive training epochs for effective model's training. There exist multiple parameters such as Dice loss and cross-entropy loss functions that are used to train the segmentation models. Furthermore, the model is implemented using deep learning frameworks on a system equipped with GPU support for better performance.

4. RESULTS AND DISCUSSION

Experimental Setup-The proposed model training had been done on a machine equipped with an NVIDIA GPU (Tesla/RTX series) and 16 GB RAM. The dataset was split into 70:20:10 training, validation, and testing sets. The Adam optimizer was used to train the Ensemble model with a learning rate of 1e-4, a batch size of 8, and up to 100 epochs.

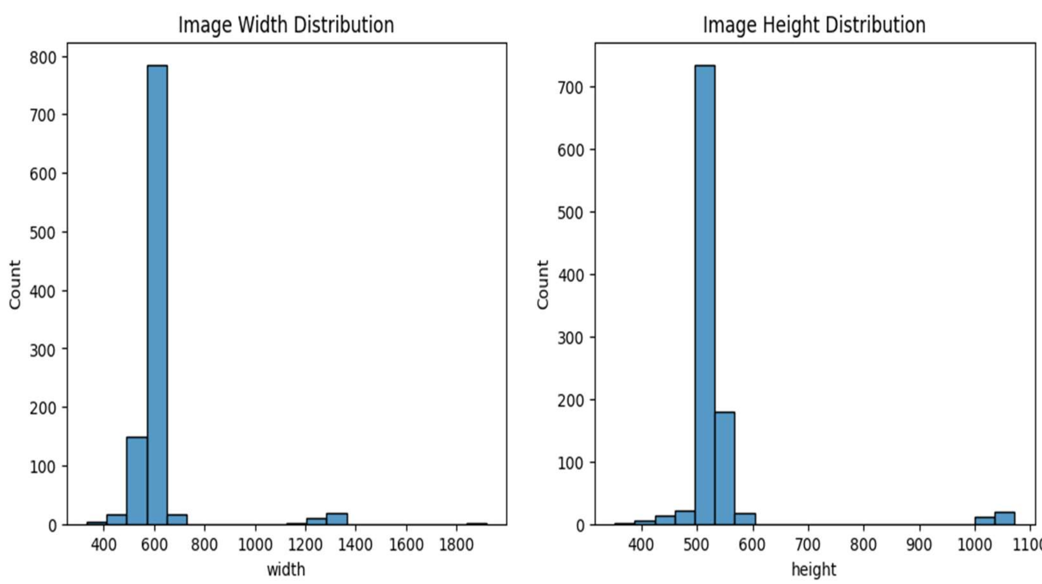

Figure 2: Distribution of Image Width and Height in Gastric Cancer Dataset

Figure 2 above demonstrates the distribution of images by using the Kvasir-SEG dataset. The resolution of most images lies in the range of 600–650 pixels and 500–550 pixels width wise and height wise respectively showing similar settings as shown in Fig 2. In addition to this some small outlier images having Higher resolutions were also detected. This disparity helps in justification of image resizing during the preprocessing phase to certify uniform input dimensions while preserving relevant polyp structures for segmentation tasks.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

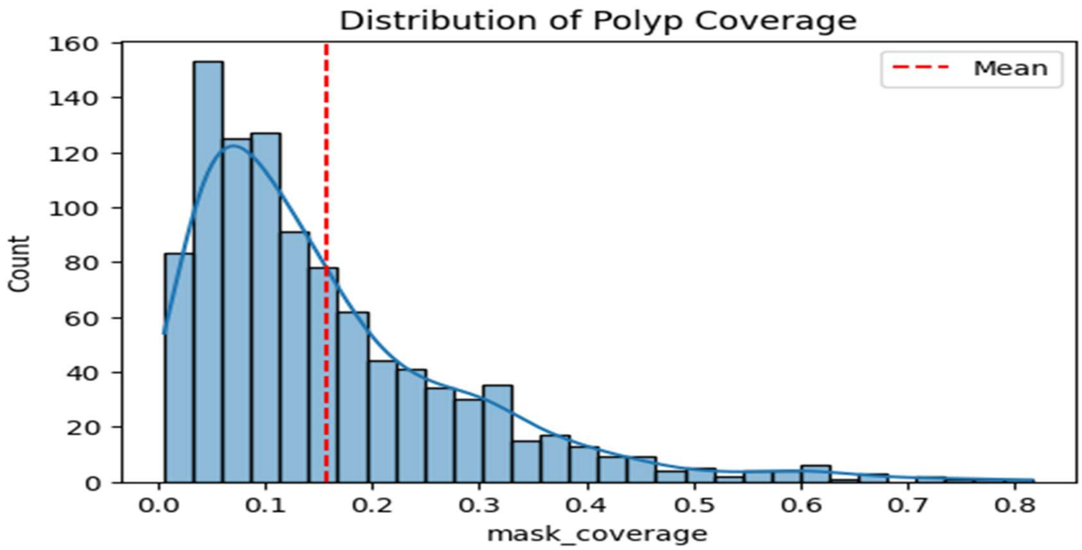

Figure 3: Distribution of Polyp Coverage in the Preprocessed Gastric Lesion Dataset

Figure 3 above displays the spreading of polyp mask coverage in the Kvasir-SEG dataset. The dataset images showing low mask coverage, representing that polyps typically occupy a tiny area of the image. This highlights the challenge of segmenting small and less prominent polyp regions and motivates the use of robust segmentation architectures.

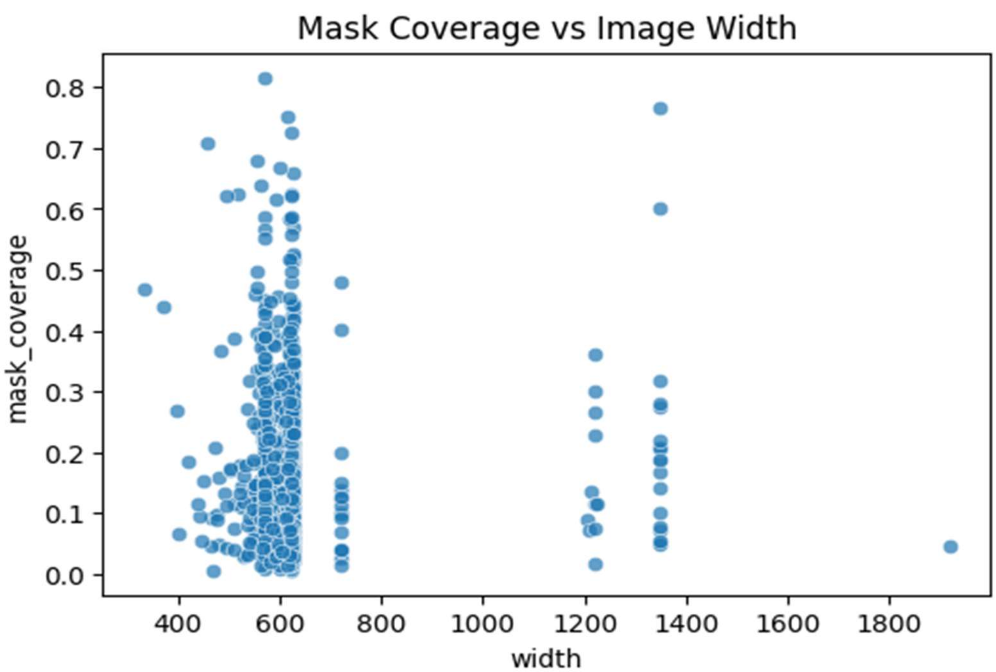

Figure 4: Relationship between image width and polyp mask coverage in the Kvasir-SEG dataset.

Figure 4 illustrates the connection between image width and corresponding polyp mask coverage in the Kvasir-SEG dataset. Most datapoints are grouped around commonly used image widths (approximately 512–640 pixels), showing a wide range of mask coverage values. Images with larger widths exhibit greater variability in coverage but occur less frequently, indicating resolution diversity across the dataset.

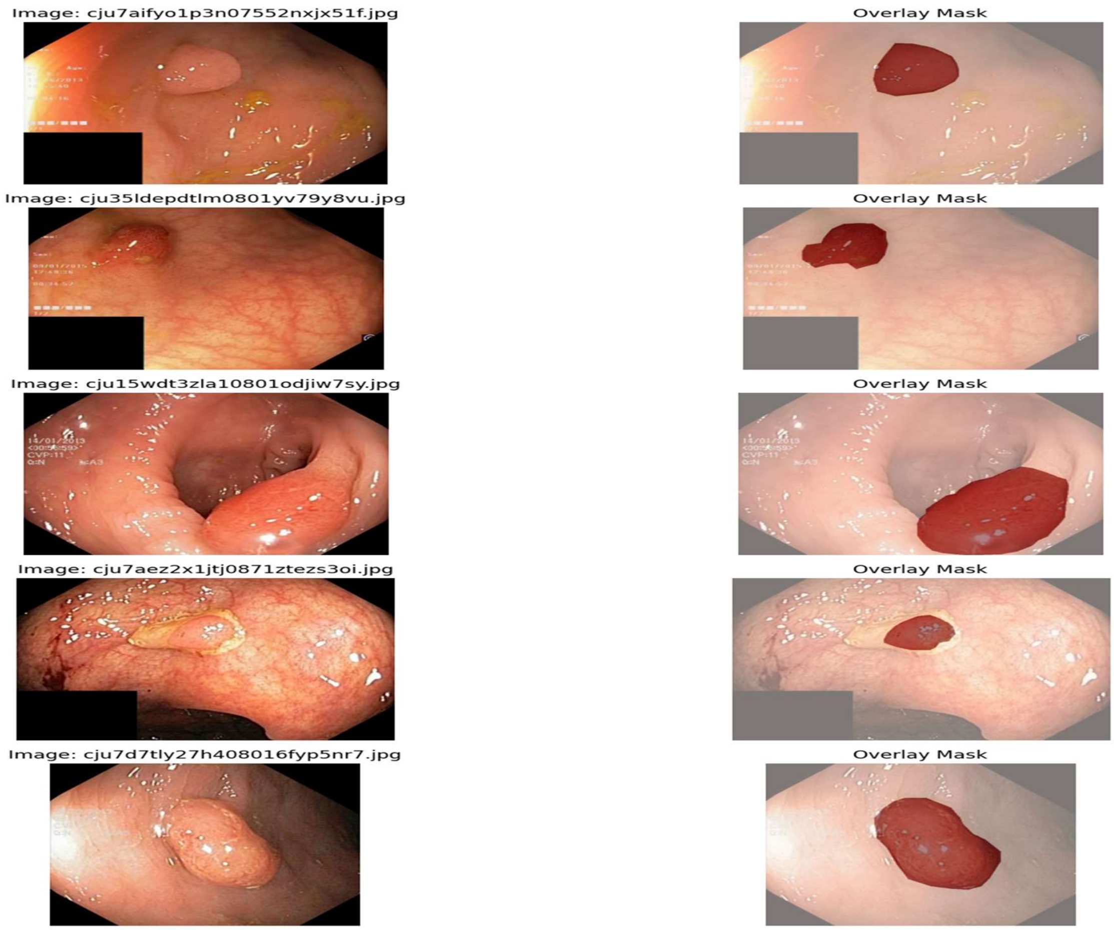

Figure 5: Representative endoscopic images with corresponding predicted segmentation overlays

As shown above in Figure 5, the proposed ensemble segmentation model accurately identifies polyp regions across varying visual characteristics. The predicted masks closely follow lesion boundaries, demonstrating reliable delineation for both subtle and prominent polyps. These examples qualitatively confirm the robustness of the model under diverse structural and contrast conditions. The proposed model demonstrates strong capability in segmenting small and subtle gastric polyps that exhibit low contrast against surrounding mucosal tissue. Accurate boundary delineation in such cases highlights the effectiveness of the learned feature representations in capturing fine-grained lesion characteristics, which is critical for early-stage gastric abnormality detection.

The segmentation results further indicate that the model effectively handles polyps with irregular shapes and complex boundaries. Despite variations in lesion morphology and the presence of surrounding mucosal folds, the predicted masks closely align with the ground truth, demonstrating robustness to shape and structural variability. For larger and more prominent polyps, the model achieves comprehensive lesion coverage without under-segmentation. This consistent performance across varying lesion scales confirms the scalability of the proposed approach and its ability to generalize to diverse polyp sizes.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

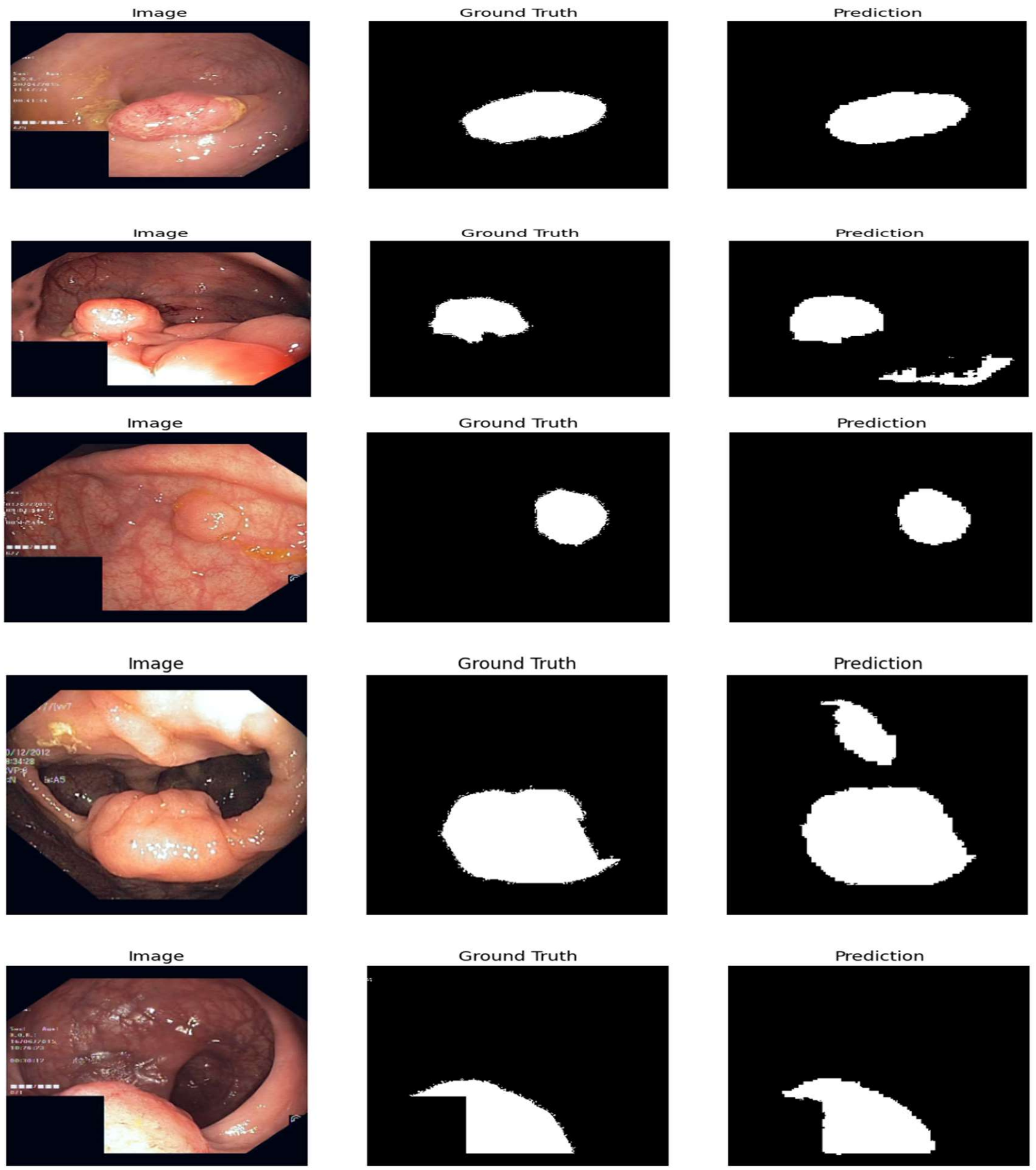

Figure 6: Qualitative segmentation results showing input images, ground-truth masks, and corresponding model predictions.

The Visuals of Qualitative segmentation results are in form of different images such as input images, ground truth masks and corresponding model predictions. The Qualitative segmentation results produced from the proposed models are represented above in Figure 6.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

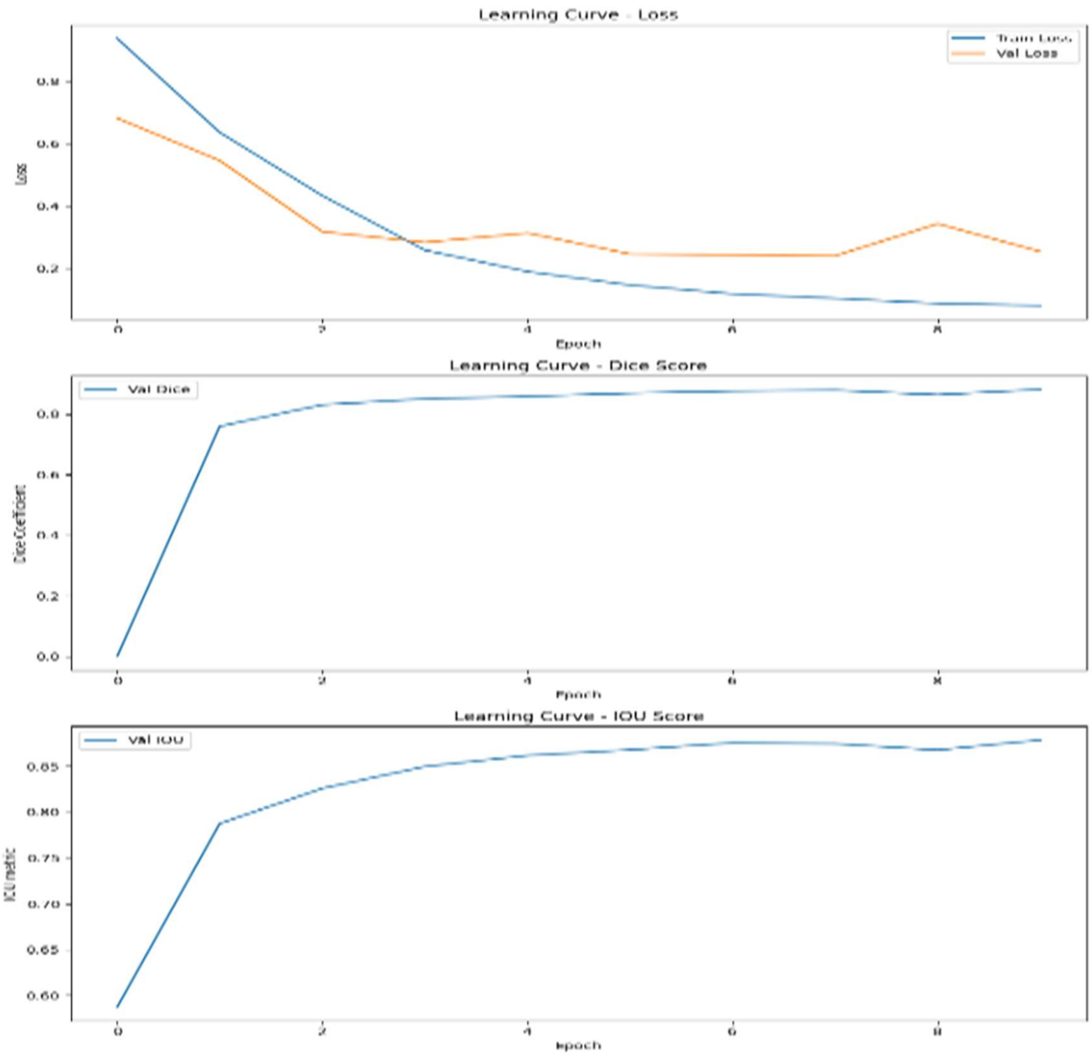

Figure 7: Learning curves of the proposed segmentation model showing training and validation loss, Dice coefficient, and IoU score.

Figure 7 above visualize the different learning curves of the advanced ensemble segmentation model. the model hot trained on different epochs. In above figure there exists three graphs, one having loss curves, that steadily show loss of training shows effective model training. From the curve it can be visualize that initially validation loss decreases and then sustain with very minor fluctuations which shows overfitting's dice coeffient value increase rapidly during early epochs moves towards a higher value and also with improved segmentation accuracy. The third figure displays the Score of IOU also following a upward trend confirms consist scores between ground truth and predictive marks.as a whole the multiple curves show stable convergence value, strong segmentation performance value and reliable validation behavior.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

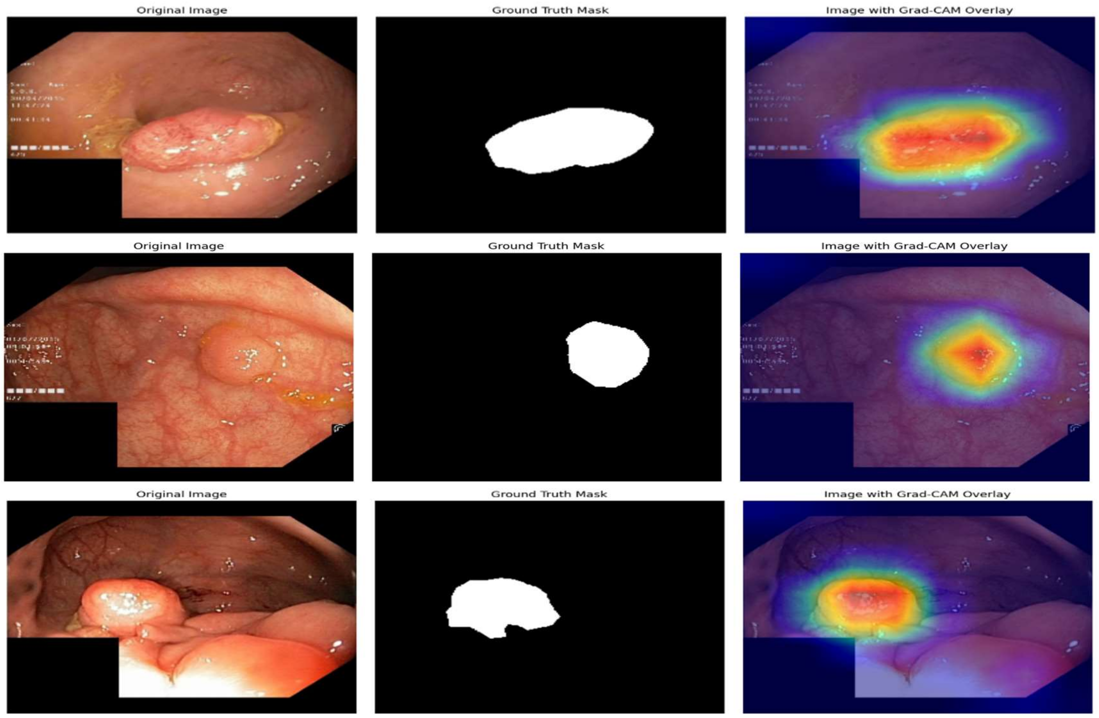

Figure 8: Visualisation obtained by using Grad-CAM showing different regions of segmentation predictions

Figure 8 above shows the interpretability analysis of model consistently showing relevant polyp regions during prediction. Grad-CAM overlays reveal high activation, focused around the segmented polyp areas across multiple samples. Attention maps with more well-defined lesion boundaries also show highly concentrated and well-defined activation patterns, though overall high-intensity activation still focuses on the core lesion location for more complicated cases with uneven morphology. These reflect the fact that accurate segmentation and comprehensible visual signals both stand behind the clinical dependability of the model.

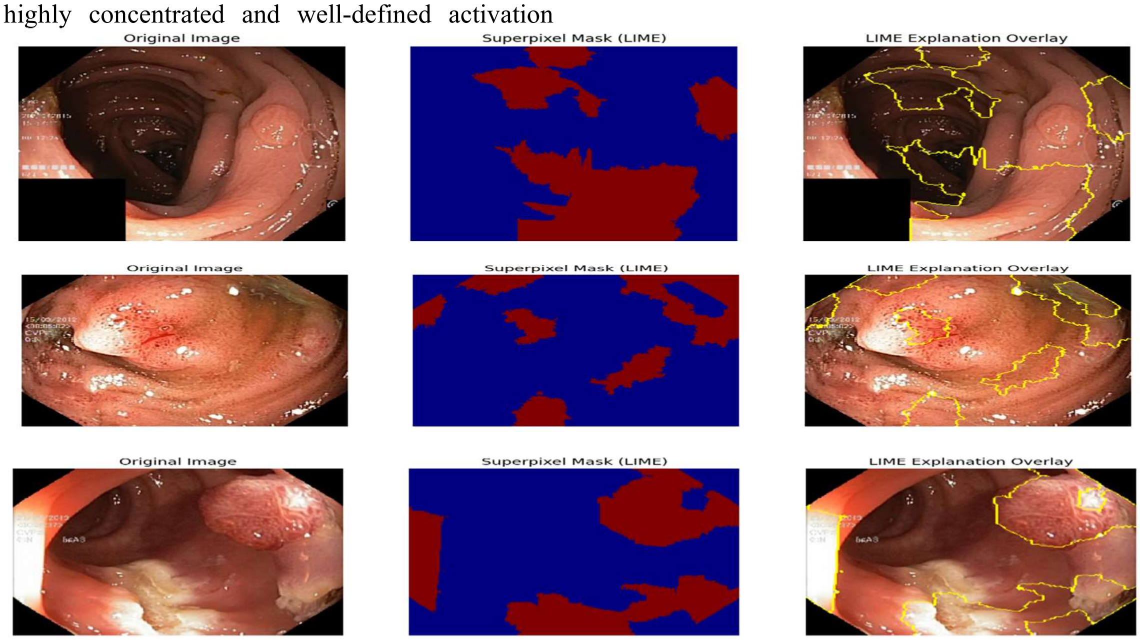

Figure 9: Visual Explaining Gastric Lesions in Endoscopic Images Using LIME

Explaining Gastric Lesions in Endoscopic Images Using LIME- The interpretability analysis using LIME can be shown above in figure 9. That framework provides an understanding of model's behaviour of decision making by exploring the specific regions that can contributes to predictions of the segmentation. The imaginings can locate the localize areas that receive higher weights. In the image parts that consist of polyps, LIME responds to different patterns within the segmented regions. Generally, the explanations of LIME improve transparency by providing instance-level understanding of the segmentation outcomes.

4.1 Confusion Matrix

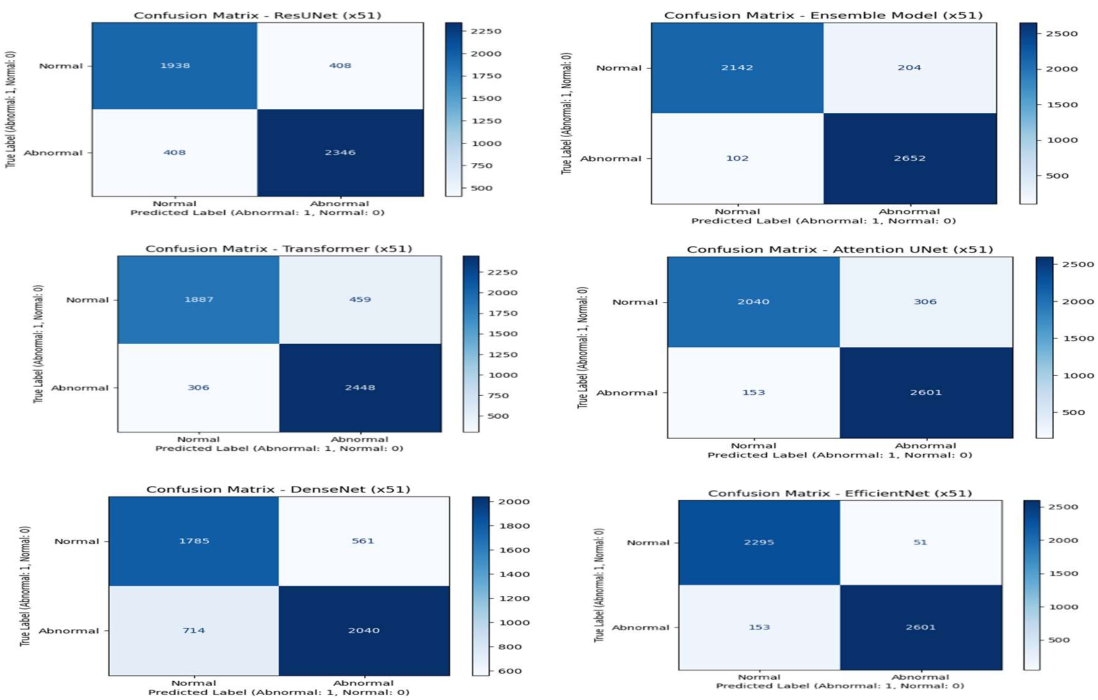

Figure 10: Comparison of Confusion Matrix for Different Deep Learning Models

Confusion matrices as shown above in Figure 10 Provides a one-to-one contrast of the performance of deep learning-based model evaluated on performance parameters.

The most robust behavior is represented by the Ensemble Model, which has the most substantial number of true positives and true negatives with noticeably fewer wrong classifications, hence demonstrating the best discriminative ability in distinguishing normal and diseased data. EfficientNet and ResUNet were reliable in sensitivity and specificity due to their performance, which was robust yet well-balanced. Attention UNet and DenseNet, as well as the Transformer model, have a tendency to yield relatively higher levels of misclassification, which mainly affects the abnormal class, highlighting the fact that these are not as proficient in generalizing over different types of visual variations. These findings evidence the advantage of model ensembling in comparison with solo solutions in making more robust and accurate diagnostic predictions.

As compared to other state of arts models, proposed ensemble model produces better results in terms of different performance parameters. Accuracy, precision, recall and F1 score of 0.94,0.93,0.96 and 0.95 respectively are obtained. The getting of high F1 score shows a very strong balance between precision and recall. the recall of ResUNet is high, but the precision and recall of EfficientNet are balanced. Attention UNet and DenseNet perform mediocrely; DenseNet has greatly higher precision but lower recall. The high precision of the Transformer model is 0.88, and its recall is distinctly lower at 0.71, which shows that it is more conservative and prefers to ignore unusual situations. Overall, these observations show the effectiveness of model ensembling in refining robustness and consistency compared to stand-alone architectures, as shown below in Table 1.

Table1: Performance Analysis and Comparison of Deep learning models for polyp classification

Ref. No.	Model	Accuracy	Precision	Recall	F1-Score
Huang, S. [26]	ResUNet	0.88	0.867925	0.901961	0.884615
Oktay, O., [27]	Attention UNet	0.82	0.811321	0.843137	0.826923
Huang, G [28]	DenseNet	0.81	0.820000	0.803922	0.811881
Tan, M [29]	EfficientNet	0.88	0.882353	0.882353	0.882353
Dosovitskiy, A [30]	Transformer	0.80	0.878049	0.705882	0.782609
Proposed Approach	Ensemble Model	0.94	0.928571	0.962963	0.945455

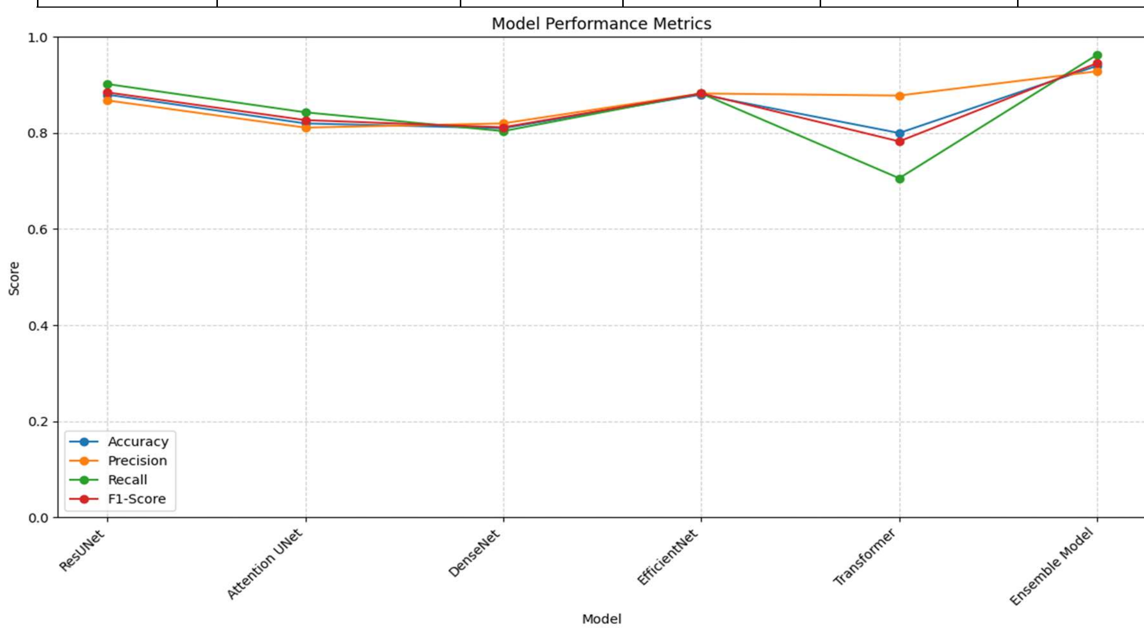

Figure 11: Performance comparison of deep learning models based on prediction metrices

Above Figure 11 shows the comparative analysis of different deep learning architectures-ResUNet, Attention UNet, DenseNet, EfficientNet, Transformer, and Ensemble Models performed on the basis of performance parameters.

The proposed ensemble model achieves better results as compared to other individual models by combining features from multiple architectures. By doing this, the proposed framework is able to handle different variations in appearance, shape and polyp size more efficiently. The integration of Grad-CAM and LIME improves interpretability by highlighting clinically relevant regions influencing the model predictions. This provides better insight into However, slight variations in performance can be visualized in multiple images having low contrast values or complex polyp boundaries. These observations indicate certain limitations of the model in handling highly challenging scenarios. Overall, this embedded model consists of explainable AI techniques with ensemble learning improves both segmentation accuracy and interpretability.

5. CONCLUSION

This framework studies presents an interpretable deep learning-based framework for gastric polyp

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

detection and segmentation. The proposed deep learning-based framework is an ensemble architecture consisting of multiple CNN based models which helps in accurately identifying the gastric polyps. This also led to early detection of stomach cancer. By the advance addition of explainable artificial intelligence technologies such as Grad -CAM and LIME which further help the physicians for the early identification. This technology also helps in the justification of black box feature. This procedure can be considered as a valid method by which gastroenterologists can produce better results with greater accuracy. The Proposed Ensemble model shows the accuracy of 0.94, precision of 0.928, recall of 0.962 and F1 score of 0.945 that achieves better results as compared to traditional models.

Although these results are positive, the study points out areas that need further improvement in order to enhance both the therapeutic applicability and utilization of AI-driven medical devices. Future research may be directed for the expansion of model at larger extent and with different datasets with a view to enhancing generalizability across a wide range of patients and imaging modalities.

In such a complex GI environment, transformer-based vision models or hybrid CNN-transformer systems could bring about a radical enhancement in segmentation performance. Integrated gradients, SHAP, and counterfactual explanations are advanced XAI techniques that may show far-reaching insights into model behavior. Real-time or lightweight versions of the concept for endoscopic and mobile healthcare systems may further increase accessibility, especially in resource-limited settings. Overall, the proposed work contributes towards the development of clinically reliable and interpretable AI systems for early detection of gastric cancer.

REFERENCES:

[1]. Suh, B., Yu, H., Cha, J.K., Choi, J. and Kim, J.W., 2025. Explainable deep learning approaches for risk screening of periodontitis. *Journal of Dental Research*, 104(1), pp.45–53.

[2]. Bhati, D., Neha, F. and Amiruzzaman, M., 2024. A survey on explainable artificial intelligence (XAI) techniques for visualizing deep learning models in medical imaging. *Journal of Imaging*, 10(10), p.239.

[3]. Salvi, M., Seoni, S., Campagner, A., Gertych, A., Acharya, U.R., Molinari, F. and Cabitza, F., 2025. Explainability and uncertainty: Two sides of the same coin for enhancing the interpretability of deep learning models in healthcare. *International Journal of Medical Informatics*, 197, 105846.

[4]. Kalyakulina, A., Yusipov, I., Moskalev, A., Franceschi, C. and Ivanchenko, M., 2024. Explainable artificial intelligence (XAI) in aging clock models. *Ageing Research Reviews*, 93, 102144.

[5]. Stodt, J., Reich, C. and Knahl, M., 2024. Demystifying XAI: Requirements for understandable XAI explanations. *Studies in Health Technology and Informatics*, 316, pp.565–569.

[6]. Finzel, B., 2025. Current methods in explainable artificial intelligence and future prospects for integrative physiology. *Pflügers Archiv - European Journal of Physiology*, 477(4), pp.513–529.

[7]. Rosenbacke, R., Melhus, Å., McKee, M. and Stuckler, D., 2024. How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: Systematic review. *JMIR AI*, 3, e53207.

[8]. Salih, A.M., Galazzo, I.B., Gkontra, P., Rauseo, E., Lee, A.M., Lekadir, K., Radeva, P., Petersen, S.E. and Menegaz, G., 2024. A review of evaluation approaches for explainable AI with applications in cardiology. *Artificial Intelligence Review*, 57(9), 240.

[9]. Zhang, K., Wang, H., Cheng, Y., Liu, H., Gong, Q., Zeng, Q. and Chen, D., 2024. Early gastric cancer detection and lesion segmentation based on deep learning and gastroscopic images. *Scientific Reports*, 14(1), 7847.

[10]. Su, X., Liu, Q., Gao, X. and Ma, L., 2023. Evaluation of deep learning methods for early gastric cancer detection using gastroscopic images. *Technology and Health Care*, 31(S1), pp.313–322.

[11]. Afrash, M.R., Shafiee, M. and Kazemi-Arpanahi, H., 2023. Establishing machine learning models to predict the early risk of gastric cancer based on lifestyle factors. *BMC Gastroenterology*, 23(1), 6.

[12]. Liu, Y., Zhang, L., Hao, Z., Yang, Z., Wang, S., Zhou, X. and Chang, Q., 2022. An Xception model based on residual attention mechanism for the classification of benign

and malignant gastric ulcers. *Scientific Reports*, 12(1), 15365.

[13]. Huang, S. *et al.*, 2024. From UNet to Res-UNet and nnUNet: Performance comparison on brain tumor segmentation. *arXiv preprint*, arXiv:2407.04353.

[14]. Sehgal, T., Joshi, T., Chowdhary, R., Goyal, O., Kalra, S., Goyal, R., Taranikanti, V., Vuthaluru, A.R. and Goyal, M.K., 2026. Deep learning in lower gastrointestinal cancer detection: Advances in endoscopic, radiologic, and histopathologic diagnostics. *World Journal of Gastrointestinal Oncology*, *18*(2), p.115974.

[15]. Alwahaibi, N., 2026. Transforming Gastric Biopsy Diagnostics: Integrating Omics Technologies and Artificial Intelligence. *Biomedicines*, *14*(2), p.407.

[16]. Ciardiello, L., Agnello, P., Petyx, M., Martinelli, F., Cesarelli, M., Santone, A. and Mercaldo, F., 2025. A Deep Learning-Based Approach for Explainable Microsatellite Instability Detection in Gastrointestinal Malignancies. *Journal of Imaging*, *11*(11), p.398.

[17]. Lalitha, K. and Padmashree, S., 2025, October. An Automated Texture-Based Classification System for Gastric Lesion Diagnosis using Endoscopic Images. In *2025 International Conference on Communication, Computer, and Information Technology (IC3IT)* (pp. 1-6). IEEE.

[18]. Zhang, H., Jia, J., Zhang, W., Yi, R., Yan, X., Sun, W., Wang, X. and Gao, Y., 2025. FC-YOLO: a fast inference backbone and lightweight attention mechanism-enhanced YOLO for detecting gastric adenocarcinoma in pathological image. *Frontiers in Oncology*, *15*, p.1657159.

[19]. Swanson, K., Wu, E., Zhang, A., Alizadeh, A.A. and Zou, J., 2023. From patterns to patients: advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. *Cell*, *186*(8), pp.1772-1791.

[20]. Huo, X., Sun, G., Tian, S., Wang, Y., Yu, L., Long, J., Zhang, W. and Li, A., 2024. HiFuse: Hierarchical multi-scale feature fusion network for medical image classification. *Biomedical signal processing and control*, *87*, p.105534.

[21]. Alsallal, M., Habeeb, M.S., Vaghela, K., Malathi, H., Vashisht, A., Sahu, P.K., Singh, D., Al-Hussainy, A.F., Aljanaby, I.A., Sameer, H.N. and Athab, Z.H., 2025. Artificial intelligence in gastric cancer: a systematic review of machine learning and deep learning applications. *Abdominal Radiology*, pp.1-17.

[22]. Ma, D., Fan, C., Sano, T., Kawabata, K., Nishikubo, H., Imanishi, D., Sakuma, T., Maruo, K., Yamamoto, Y., Matsuoka, T. and Yashiro, M., 2025. Beyond biomarkers: machine learning-driven multiomics for personalized medicine in gastric cancer. *Journal of Personalized Medicine*, *15*(5), p.166.

[23]. Cai, Z.R., Zheng, Y.Q., Hu, Y., Ma, M.Y., Wu, Y.J., Liu, J., Yang, L.P., Zheng, J.B., Tian, T., Hu, P.S. and Liu, Z.X., 2025. Construction of exosome non-coding RNA feature for non-invasive, early detection of gastric cancer patients by machine learning: a multi-cohort study. *Gut*, *74*(6), pp.884-893.

[24]. Liu, L., Zhang, Y., Zhao, X.H., Zhu, M., Liang, J.W., Jiang, Z.Y., Zhang, J.Q., Ma, J.L., Wang, Z.Z., Gan, M.F. and Mao, X.L., 2025. Predicting clinical prognosis in gastric cancer using deep learning-based analysis of tissue pathomics images. *Computer Methods and Programs in Biomedicine*, *269*, p.108895.

[25]. Tritscher, J., Krause, A. and Hotho, A., 2023. Feature relevance XAI in anomaly detection: Reviewing approaches and challenges. *Frontiers in Artificial Intelligence*, 6, 1099521.

[26]. Jha, D., Smedsrud, P.H., Riegler, M.A., Halvorsen, P., de Lange, T., Johansen, D. and Johansen, H.D., 2020. Kvasir-SEG: A segmented polyp dataset. In: *Proceedings of the 26th International Conference on Multimedia Modeling (MMM 2020)*, pp.451–462.