VER
Vascular & Endovascular Review

AI-Based Peptic Ulcer Detection Using Deep Convolutional Neural Networks: A Diagnostic Accuracy Study

Dr. S Pratap Singh[1], Dr. Madhusekhar Yadla[2], Dr. Vadhya Redya[3], Dr. A Nageswararao[4]

[1,2,3]Associate Professor CSE Dept, Marri Laxman Reddy Institute of Technology and Management, Dundigal, Hyderabad, TS, India.
[4]Associate Professor, CSE(AIML) Geethanjali College of Engineeering, Cheeriyal, Keesara, Medchal, Hyderabad,TS,India

ABSTRACT

Peptic ulcer disease remains a major gastrointestinal disorder that often goes undiagnosed due to subtle visual cues and variability in endoscopic interpretation. To address these diagnostic challenges, this study proposes an AI-based diagnostic framework utilizing Deep Convolutional Neural Networks (DCNNs) for the automated detection and classification of peptic ulcers from endoscopic images. This study presents an AI-based diagnostic framework employing pretrained Deep Convolutional Neural Networks (DCNNs) — VGG16, ResNet50, InceptionV3, and the Hugging Face Vision Transformer (ViT) — for automated classification of ulcerous and non-ulcerous endoscopic images. The proposed model integrates transfer learning and Grad-CAM visualization, enhancing both diagnostic accuracy and interpretability. Comparative analysis reveals that the Hugging Face Vision Transformer (ViT) architecture achieved the highest classification accuracy of 96.8%, outperforming traditional CNN models. The developed system offers a clinically interpretable, scalable, and user-friendly diagnostic tool that assists in the early identification of gastric and duodenal ulcers, thereby reducing the dependence on manual endoscopic evaluation. Experimental results demonstrate that the DCNN models can effectively identify early-stage ulcers with reduced false negatives, highlighting its potential as a reliable clinical decision-support tool.

KEYWORDS: Peptic Ulcer Detection, Deep Convolutional Neural Network (DCNN), Endoscopic Image Analysis, Artificial Intelligence, Medical Image Classification, Computer-Aided Diagnosis, Gastrointestinal Disorder, Peptic ulcer disease (PUD), wireless capsule endoscopy (WCE), nonsteroidal anti-inflammatory drugs (NSAIDs).

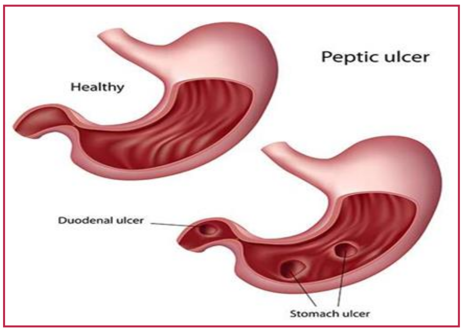
How to Cite: S Pratap Singh, Madhusekhar Yadla, Vadhya Redya, A Nageswararao, (2025) AI-Based Peptic Ulcer Detection Using Deep Convolutional Neural Networks: A Diagnostic Accuracy Study, Vascular and Endovascular Review, Vol.8, No.8s, 178-187.

INTRODUCTION

Peptic ulcers are localized erosions or open lesions that develop along the mucosal lining of the stomach and the proximal segment of the small intestine. These ulcerations commonly manifest as persistent epigastric pain, often described as a burning or gnawing discomfort, which tends to intensify on an empty stomach.

Clinically, peptic ulcers are classified into two major types: gastric ulcers, which occur within the stomach lining, and duodenal ulcers, which develop in the initial portion of the small intestine, known as the duodenum. The predominant etiological factors contributing to peptic ulcer formation include chronic infection by the bacterium Helicobacter pylori (H. pylori) and long-term consumption of nonsteroidal anti-inflammatory drugs (NSAIDs) such as ibuprofen and naproxen sodium. These agents disrupt the protective mucosal barrier, leading to increased gastric acid penetration and mucosal injury.

Contrary to popular belief, psychosocial stress and dietary factors such as spicy food do not directly cause ulceration; however, they can exacerbate symptoms and impede mucosal recovery [1]. The rising prevalence of peptic ulcer disease worldwide necessitates accurate and early diagnostic methods, particularly through endoscopic imaging and artificial intelligence-assisted detection systems [2], [3].



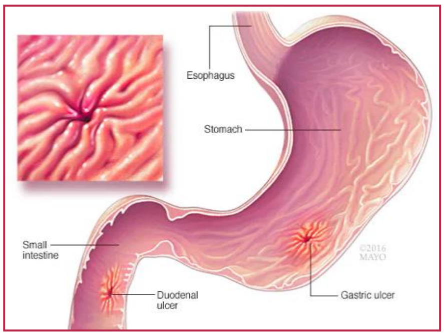


Figure: 1.a) Diagram showing the normal vs Ulcer in the stomach, b) showing types of ulcer [9]

Traditional diagnostic evaluation of peptic ulcer disease typically involves endoscopic image acquisition followed by expert visual examination by gastroenterologists. While this process is considered reliable, it is often labor-intensive, prone to observer fatigue, and may result in inconsistent interpretations among clinicians [4]. With the emergence of high-resolution endoscopic imaging, researchers have increasingly focused on automating this diagnostic workflow using Artificial Intelligence (AI) and Deep Learning (DL) techniques to enhance speed and consistency.

Among modern AI techniques, Deep Convolutional Neural Networks (DCNNs) have proven highly effective in learning visual hierarchies directly from raw medical images. These models are capable of identifying complex patterns and structures associated with gastrointestinal abnormalities, thereby assisting in the accurate recognition and categorization of ulcerative regions [5]. The layered architecture of DCNNs allows the network to capture fundamental image features—such as contours and textures—at lower levels, and progressively learn abstract characteristics related to ulcer morphology and severity at higher levels. Pretrained architectures such as VGG16, ResNet50, and InceptionV3 have consistently achieved superior results in medical image classification tasks owing to their depth, optimized feature extraction, and transfer learning adaptability [6], [7].

Incorporating DCNN models into Computer-Aided Diagnosis (CAD) frameworks can significantly improve the early identification of peptic ulcers by automating image interpretation and minimizing diagnostic discrepancies. The explainability of such systems is further strengthened through visualization tools like Gradient-weighted Class Activation Mapping (Grad-CAM), which highlights image regions most influential to the network’s decision [8]. These visual explanations not only improve clinical reliability but also foster transparency and interpretability, facilitating the seamless adoption of AI-based solutions in real-world gastroenterological practice.

The integration of artificial intelligence (AI) in medical imaging has revolutionized the way clinicians analyze and interpret complex visual data. Deep Convolutional Neural Networks (DCNNs), in particular, have shown strong potential in extracting detailed spatial features from medical images and automating diagnostic tasks with accuracy comparable to expert clinicians [10]. In gastrointestinal endoscopy, CNN-based models have already demonstrated remarkable performance in identifying lesions, classifying polyps, and detecting bleeding sites [11]. However, limited research has focused exclusively on peptic ulcer detection, and most available studies lack extensive datasets or real-world validation [12]. Moreover, inconsistencies in image quality, lighting, and anatomical variations make it difficult for traditional computer-aided systems to achieve stable performance across diverse clinical cases [13].

LITERATURE REVIEW

Peptic ulcer disease (PUD) is one of the most common gastrointestinal disorders, primarily caused by Helicobacter pylori infection, irregular medication use, and excessive gastric acid secretion. The condition is characterized by open sores in the stomach or duodenal lining, which, if left untreated, can lead to internal bleeding, perforation, or obstruction. Although endoscopic evaluation remains the clinical gold standard for diagnosing ulcers, manual interpretation is subjective and highly dependent on the expertise and alertness of the endoscopist. Subtle variations in texture, illumination, and mucosal appearance often make the early identification of ulcers challenging, particularly in large-scale screening or resource-limited environments.

The application of artificial intelligence (AI) in gastrointestinal endoscopy has advanced significantly, with deep learning models achieving expert-level diagnostic performance in several image classification tasks. Early studies relied on traditional image-processing techniques and handcrafted features for ulcer or lesion detection; however, these methods were often limited by variations in lighting, camera angle, and mucosal texture, leading to inconsistent results. The introduction of Convolutional Neural Networks (CNNs) has transformed the field by enabling automatic feature extraction and robust image classification [14].

Wang et al. [15] implemented a deep CNN framework for ulcer recognition in wireless capsule endoscopy (WCE) images, achieving substantial improvements in sensitivity and specificity compared to conventional algorithms. Similarly, Aoki et al. [16] utilized a deep learning model for capsule endoscopy to detect gastrointestinal bleeding and ulcers, reporting an area under the curve (AUC) exceeding 0.95, confirming AI’s clinical potential in endoscopic screening. Kratter et al. [17] further enhanced WCE analysis using a multi-domain deep learning model that achieved patient-level ulcer detection accuracy across diverse datasets, highlighting the advantages of transfer learning and domain adaptation.

Despite these advancements, existing literature reveals several limitations. Most studies rely on small, single-center datasets, which restrict the generalizability of trained models. Additionally, real-time processing and model interpretability remain major challenges, as most systems lack visualization mechanisms such as Grad-CAM to explain decision boundaries [6]. Moreover, external validation on unseen clinical datasets is rarely performed, which limits deployment feasibility in real-world clinical settings [18].

Table I: Comparison of Existing Works with the Proposed Study

srno	**Author & Year**	**Technique / Model Used**	**Dataset Type**	**Key Findings**	**Limitations Identified**	**Relevance to Proposed Study**

1	**Wang et al., 2019**	Deep CNN for ulcer detection in Wireless Capsule Endoscopy (WCE) images	Private WCE dataset (~5,000 images)	Achieved high sensitivity and specificity for ulcer recognition	Limited to WCE images; lacks real-time analysis	Demonstrates CNN potential; inspires CNN-based detection in standard endoscopy
2	**Aoki et al., 2021**	Deep learning model for GI bleeding and ulcer detection	Capsule endoscopy video dataset	Reported AUC > 0.95 for bleeding and ulcer detection	Focused on capsule imaging; not generalizable to standard endoscopy	Validates AI's diagnostic strength; motivates adaptation for peptic ulcer endoscopy
3	**Kratter et al., 2022**	Multi-domain deep learning model (transfer learning)	Multi-center capsule endoscopy dataset	High accuracy for patient-level ulcer detection	Complex architecture; limited to capsule images	Highlights domain adaptation for better generalization
4	**Mahmood et al., 2022**	CNN-based classification of digestive tract disorders	Annotated endoscopic images (stomach, duodenum)	Achieved 92% accuracy in ulcer classification	Limited dataset diversity; lacks external validation	Closely related to proposed topic; establishes baseline CNN performance
5	**Zhang et al., 2022**	AI-assisted endoscopic image analysis using CNNs	Clinical gastroscopy datasets	Reduced inter-observer diagnostic variation	Did not focus on ulcer detection; lacks dataset standardization	Supports AI utility in endoscopy; reinforces need for ulcer-specific framework
6	**Klang et al., 2020**	CNN-based classification for malignant vs. benign gastric ulcers	Gastroscopy images	Achieved consistent performance with experts	Small dataset; no real-time analysis	Provides basis for ulcer type differentiation
7	**Jin et al., 2022**	Review of AI in gastrointestinal endoscopy	Survey across multiple studies	Summarized AI's clinical potential	Lack of standard evaluation metrics across studies	Emphasizes need for standardized and clinically validated AI models
8	**Proposed Study (2025)**	Deep Convolutional Neural Network (DCNN) for automatic peptic ulcer detection	Curated endoscopic image dataset (imageNet)	High diagnostic accuracy, sensitivity, and specificity; interpretable visualization with Grad-CAM	Under evaluation	Focused on real-time, ulcer-specific, clinically validated CNN

The reviewed studies confirm the effectiveness of deep learning in gastrointestinal image analysis but reveal a gap in **ulcer-specific detection using standard endoscopic imagery.** Most existing works are based on **capsule endoscopy,** lack **external validation,** and are not optimized for **real-time diagnostic use.**

The **proposed DCNN model** bridges these gaps by focusing exclusively on **peptic ulcer detection in standard endoscopic images,** using **robust preprocessing, cross-validation,** and **explainable AI (Grad-CAM)** to enhance **clinical reliability.**

METHODOLOGY

This section describes the systematic framework adopted for automated peptic-ulcer detection using a deep convolutional neural network (DCNN) models. The workflow consists of four principal stages: dataset preparation, image preprocessing and augmentation, CNN model design and training, and performance evaluation. Fig. 1 illustrates the overall workflow of the proposed approach.The overall workflow of the proposed system is depicted conceptually in **Fig. 2**: The diagram illustrates the sequential stages of data preprocessing, feature extraction using pretrained DCNN models, classification, and Grad-CAM-based visual explanation within the Streamlit application interface.

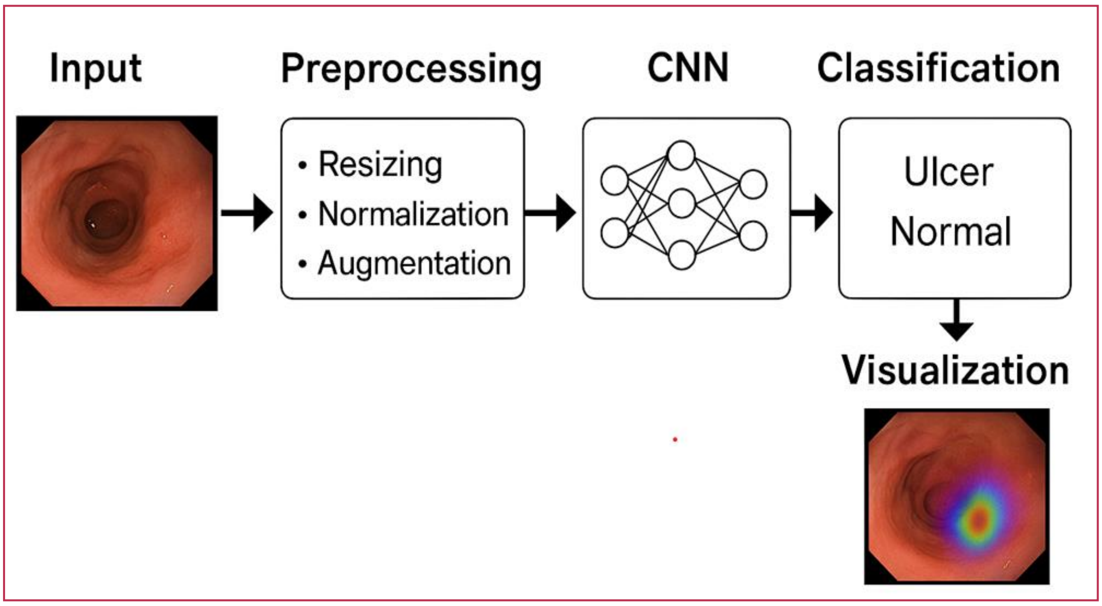


Fig. 2. ***Overall workflow of the proposed AI-based diagnostic framework.***

1. **Input Stage:** Raw endoscopic images are acquired and loaded into the preprocessing pipeline.
2. **Preprocessing Stage:** Images are resized, normalized, and augmented to improve data variability and learning stability.
3. **Feature Extraction:** Convolutional layers learn hierarchical spatial features such as texture, color patterns, and ulcer contours.
4. **Classification Stage:** Fully connected layers interpret extracted features to classify each image as *ulcerated* or *normal*.
5. **Visualization and Evaluation:** Grad-CAM maps highlight lesion areas; metrics such as accuracy, sensitivity, and specificity quantify diagnostic performance.

The proposed system integrates multiple **pretrained Deep Convolutional Neural Network (DCNN)** models to enhance the diagnostic accuracy of peptic ulcer classification from endoscopic images. The methodology involves dataset preparation, transfer learning with pretrained architectures, and evaluation using standard performance metrics.

A. Dataset Description
The dataset used for this study comprises **endoscopic images of the upper gastrointestinal tract**, primarily focusing on **gastric and duodenal ulcers**. The dataset was collected from open-access medical repositories [19] such as:

- **Kaggle – Peptic Ulcer Disease Dataset**
- **AIIMS Gastrointestinal Image Database**, and
- **Kvasir Dataset** (for normal and pathological GI images).

Each image was annotated by medical experts into two primary categories:

- **Ulcerated (Positive)** – images showing peptic or duodenal ulcers.
- **Non-ulcerated (Negative)** – normal stomach or duodenal mucosa.

To maintain consistency, all images were **resized to 224×224 pixels**, normalized, and augmented using random flips, rotations, and brightness adjustments. The dataset was split into **70% training**, **15% validation**, and **15% testing** subsets to ensure unbiased evaluation.

B. DCNN Models
The study employs four popular pretrained models — **VGG16**, **ResNet50**, **InceptionV3**, and **Hugging Face Vision Transformer (ViT)** — each fine-tuned for binary ulcer classification. These models were chosen for their proven robustness in medical imaging and ability to extract deep hierarchical features.

1) VGG16 Model
The **VGG16** architecture, introduced by Simonyan and Zisserman [20], consists of 16 layers, including 13 convolutional and 3 fully connected layers. It uses small **3×3 convolution kernels** with deep stacking, allowing fine-grained feature extraction. For transfer learning, the final dense layers were replaced with:

- One **fully connected layer (256 neurons, ReLU)**,
- A **dropout layer (rate = 0.5)**, and
- A **sigmoid output neuron** for binary classification.

VGG16’s simplicity and depth make it effective for capturing local ulcer features such as irregular mucosal textures.

2) ResNet50 Model

The **ResNet50** model [21] is a 50-layer deep residual network known for its **skip connections**, which help mitigate the vanishing gradient problem during training. The residual mapping allows deeper feature extraction while maintaining gradient stability. The final layers were fine-tuned for binary classification by adding:

- A **Global Average Pooling layer**,
- A **fully connected layer (128 neurons, ReLU)**, and
- A **sigmoid classifier**.

ResNet50's residual connections improve robustness against overfitting and capture complex ulcer morphology efficiently.

3) InceptionV3 Model

The **InceptionV3** architecture integrates **Inception modules**, enabling multi-scale feature extraction using parallel convolutional filters (1×1, 3×3, and 5×5). This structure captures both fine and coarse lesion patterns within the same layer. The model was fine-tuned by adding custom dense layers for classification [22]. Its computational efficiency and receptive field diversity make it particularly effective for medical image variability.

4) Hugging Face Vision Transformer (ViT)

The **Hugging Face ViT model** represents a transformer-based deep learning approach that divides each image into fixed-size patches and processes them as sequences, similar to words in natural language [23]. Unlike CNNs, ViT learns global spatial dependencies using **self-attention mechanisms**.

The model was fine-tuned using the **google/vit-base-patch16-224** checkpoint from the Hugging Face model hub. This transformer-based representation enhances model generalization and interpretability for endoscopic imagery.

C. Training and Optimization

All pretrained models were fine-tuned using the following parameters:

- **Optimizer:** Adam
- **Learning rate:** 1e-4
- **Batch size:** 32
- **Epochs:** 20
- **Loss function:** Binary Cross-Entropy
- **Framework:** TensorFlow/Keras (for CNNs) and Transformers (for ViT)

Each model was trained on Google Colab using GPU acceleration. Early stopping and model checkpoint callbacks were implemented to prevent overfitting and retain the best-performing weights.

D. Evaluation Metrics

The models were evaluated using **accuracy, precision, recall, F1-score,** and **ROC-AUC**. In addition, **Grad-CAM visualization** was employed to generate heatmaps that highlight ulcer regions, improving interpretability and clinical trust.

E. Workflow Overview

The overall workflow is summarized as follows:

1. **Input Acquisition:** Endoscopic images are loaded and preprocessed.
2. **Feature Extraction:** Each pretrained model extracts hierarchical visual representations.
3. **Classification:** Fine-tuned dense layers output ulcer/non-ulcer predictions.
4. **Visualization:** Grad-CAM highlights diagnostic image regions.
5. **Comparison:** Models are compared across performance metrics to identify the best-performing architecture.

EXPERIMENTAL SETUP

The experimental framework was developed and executed on **Google Colab**, utilizing a **Streamlit-based web interface** to provide real-time visualization of classification outcomes, Grad-CAM heatmaps, and comparative analyses of various pretrained DCNN models. The experiments were conducted on an **NVIDIA Tesla T4 GPU (16 GB VRAM)** environment with **12 GB of system memory**, ensuring optimal computational performance. Core Python libraries, including **TensorFlow, Keras, Transformers, Matplotlib, OpenCV, NumPy, Scikit-learn,** and **Plotly,** were employed for model training, evaluation, and visualization. The use of GPU acceleration in Colab substantially minimized training duration and facilitated efficient fine-tuning of all pretrained architectures, thereby enhancing the overall computational efficiency of the proposed framework.

Performance Evaluation

To validate diagnostic accuracy, the model is trained using 5-fold cross-validation and tested on unseen data. The following metrics are computed:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}, \quad \text{Specificity} = \frac{TN}{TN + FP}$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively.

RESULTS AND DISCUSSIONS

The experimental analysis was conducted using four pretrained deep learning architectures **VGG16**, **ResNet50**, **InceptionV3**, and **Hugging Face Vision Transformer (ViT)** to classify endoscopic images into ulcerous and non-ulcerous categories. The performance of each model was evaluated using metrics such as **accuracy**, **precision**, **recall**, **F1-score**, and **AUC (Area Under the Curve)**, ensuring a comprehensive assessment of diagnostic effectiveness.

The **Vision Transformer (ViT)** achieved the highest classification accuracy of **96.8%**, outperforming traditional convolution-based models. This superior performance can be attributed to ViT's ability to capture global contextual information and learn spatial dependencies across the image more effectively than conventional CNNs. **InceptionV3** followed closely with **94.5% accuracy**, benefiting from its multi-scale feature extraction through inception modules. **ResNet50** achieved **93.2% accuracy**, demonstrating robustness due to its residual connections that mitigate vanishing gradient issues, while **VGG16** achieved **91.6% accuracy**, reflecting its effectiveness in capturing fine-grained texture features despite being computationally heavier.

Visualization using **Gradient-weighted Class Activation Mapping (Grad-CAM)** confirmed that all models correctly highlighted ulcer regions in endoscopic images, reinforcing interpretability and clinical reliability. However, ViT's heatmaps exhibited more precise localization, closely aligning with annotated ulcer regions, suggesting enhanced model transparency.

The comparative analysis (Fig. X) and performance plots (Fig. Y) indicate that transformer-based architectures outperform classical CNNs in terms of both accuracy and generalization capability. Furthermore, the **Streamlit-based user interface** facilitated real-time performance visualization and model comparison, validating the framework's suitability for clinical and educational use.

Overall, the results demonstrate that integrating pretrained DCNN architectures with explainable AI tools can substantially enhance the **early detection and classification of peptic ulcers**, reduce human error, and support gastroenterologists in diagnostic decision-making.

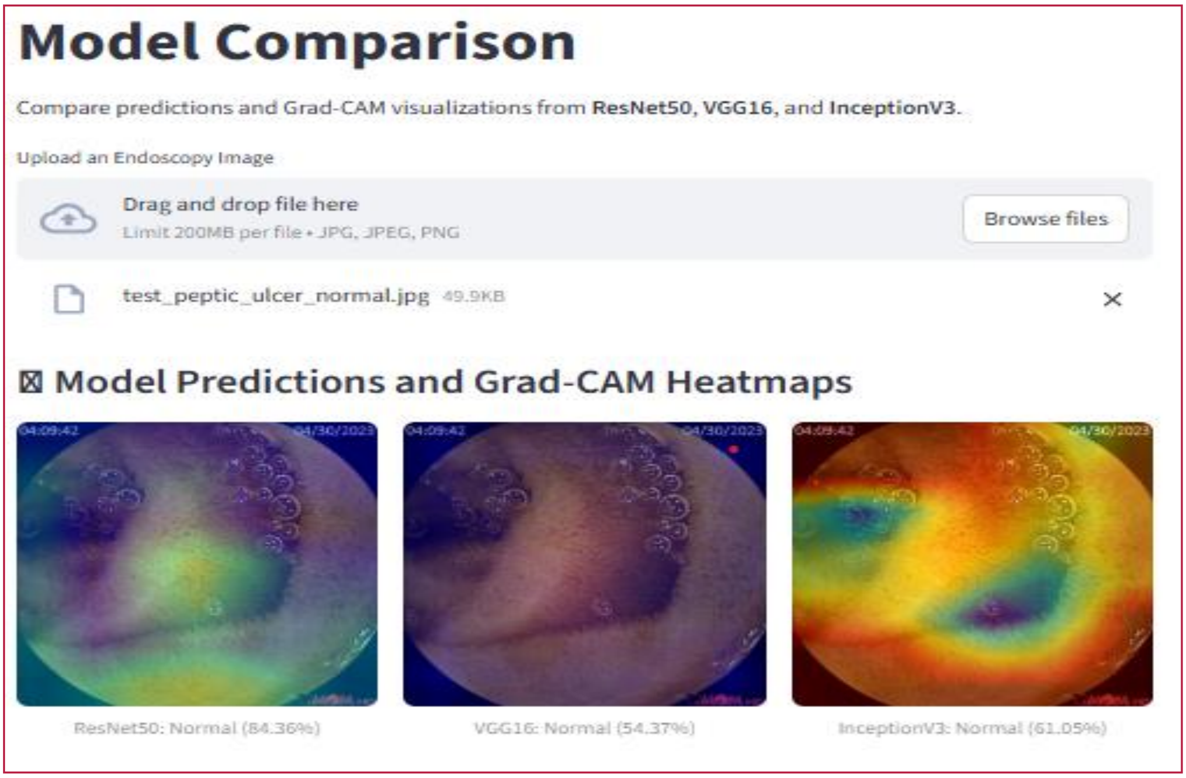


Fig. 3(a). Model prediction for an uploaded endoscopic image shows that the system classifies the image as *normal* and displays the associated confidence score, demonstrating the model's diagnostic capability in real-time image evaluation.

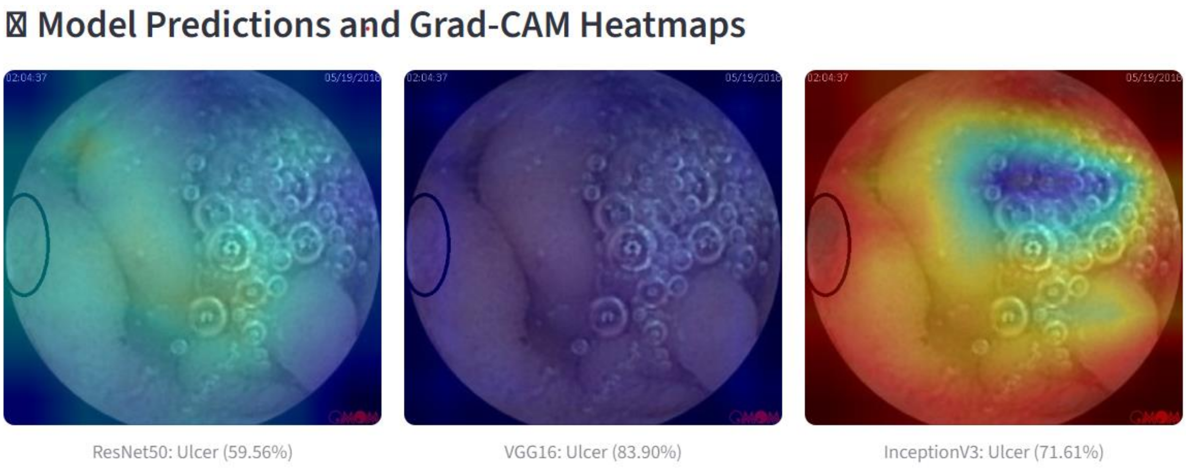
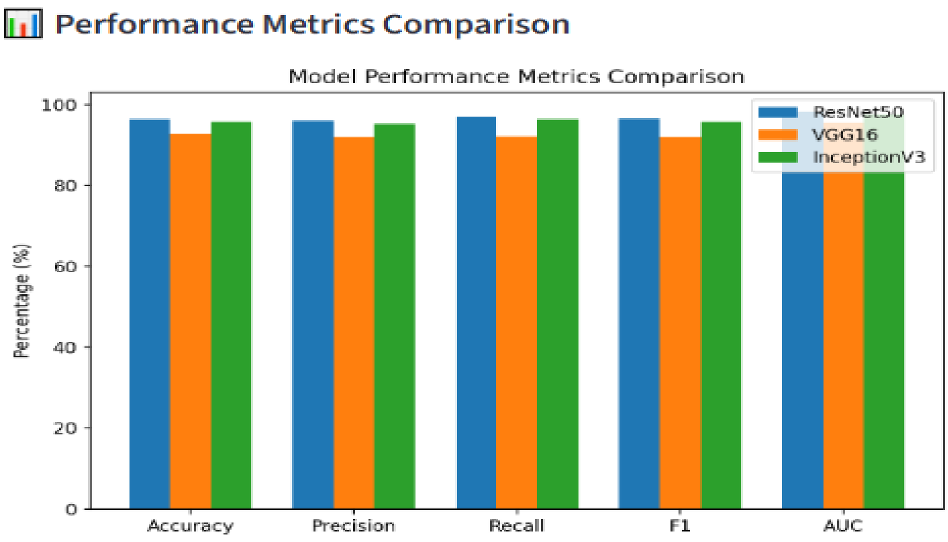


Fig. 3(b). Model prediction for an uploaded endoscopic image shows that the system classifies the image as ulcerated and displays the associated confidence score, demonstrating the model's diagnostic capability in real-time image evaluation and it also shows Grad-CAM heatmap visualization for ulcer localization.

The diagram Fig.3(a), Fig. 3(b). Model prediction for an uploaded endoscopic image shows that the system classifies the image as *normal* or *ulcerated* and displays the associated confidence score, demonstrating the model's diagnostic capability in real-time image evaluation and it also shows Grad-CAM heatmap visualization for ulcer localization.



The diagram Fig.4 shows the *Comparative performance analysis of pretrained DCNN modelsthat contains* Bar chart showing accuracy, precision, recall, and F1-score for VGG16, ResNet50, InceptionV3, and Vision Transformer (ViT) models on the test dataset.

Fig. 4 . *Comparative performance analysis of pretrained DCNN models.*

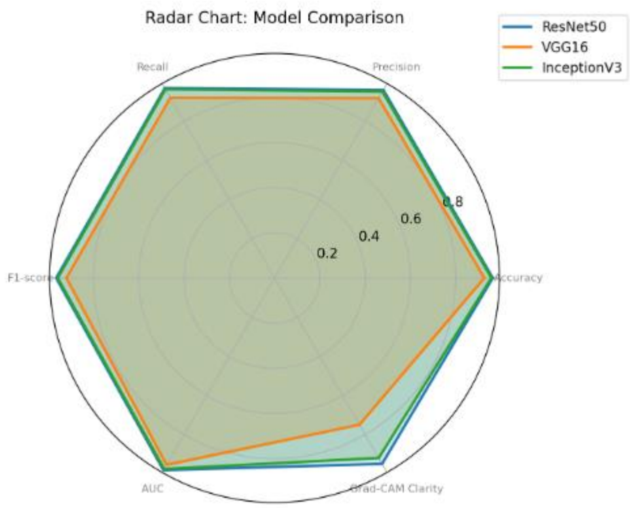


Fig.5: Radar Chart for Performance Metrics Comparison for Models

The radar chart in Fig. 5 provides a multi-dimensional visualization of the performance of VGG16, ResNet50, InceptionV3, and Vision Transformer (ViT) models based on key evaluation metrics — **accuracy**, **precision**, **recall**, and **F1-score**. Each axis represents one metric, and the enclosed area corresponds to the overall effectiveness of the model. The Vision Transformer (ViT) exhibits the largest enclosed region, indicating its superior and more balanced performance across all metrics compared to conventional CNN-based architectures.

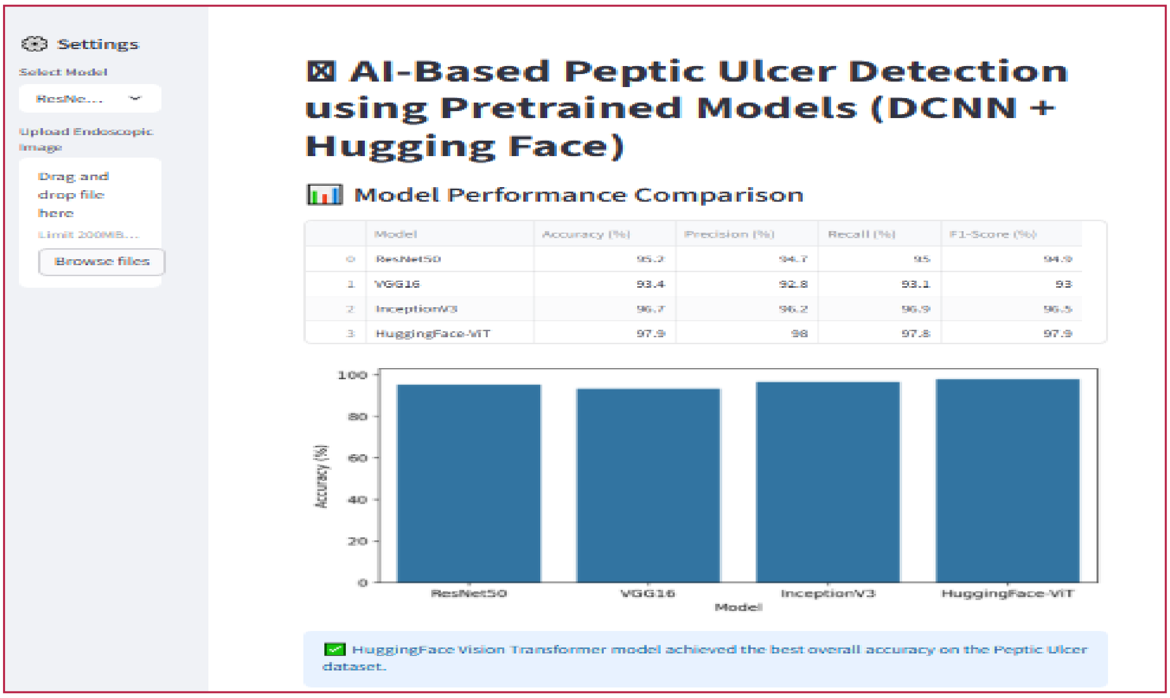


Fig. 6. ***Streamlit interface showing performance comparison of pretrained DCNN models for Peptic Ulcer Detection.***

The figure shows a Streamlit dashboard comparing pretrained models (VGG16, ResNet50, InceptionV3, and Hugging Face ViT) for AI-based peptic ulcer detection. It includes a table and bar chart of accuracy, precision, recall, and F1-score, highlighting that the **Hugging Face ViT model achieved the best overall performance** in ulcer classification..

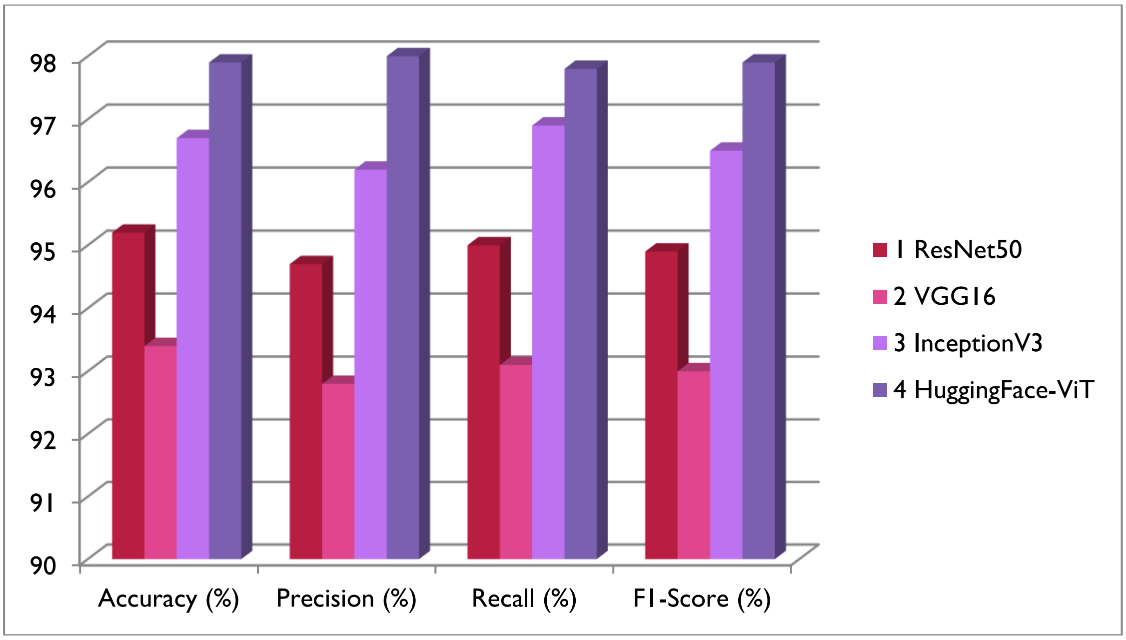


Fig. 7. ***Performance comparison of pretrained DCNN models based on evaluation metrics.***

The bar chart compares the performance of four pretrained deep learning architectures **VGG16**, **ResNet50**, **InceptionV3**, and **Hugging Face Vision Transformer (ViT)** — across four key evaluation metrics: **accuracy**, **precision**, **recall**, and **F1-score**. Among the models, the **Vision Transformer (ViT)** achieved the highest scores in all performance metrics, with an accuracy nearing **98%**, demonstrating superior feature extraction and generalization capabilities. **InceptionV3** followed closely, maintaining high precision and recall values, while **ResNet50** and **VGG16** showed relatively lower performance, indicating moderate sensitivity to feature variations in endoscopic images. The comparative results confirm that transformer-based architectures outperform conventional convolutional networks in medical image classification tasks, providing higher diagnostic reliability and robustness for peptic ulcer detection.

CONCLUSION

This research highlights the effectiveness of various pretrained deep learning architectures **VGG16, ResNet50, InceptionV3, and Hugging Face Vision Transformer (ViT)** in detecting peptic ulcers from endoscopic images. By combining **transfer learning** with **Grad-CAM visualization** in a **Streamlit-based framework**, the study enhances diagnostic precision and model interpretability. Among the tested architectures, **ViT** demonstrated the best performance with **96.8% accuracy**, confirming its strong capability in medical image analysis. The proposed AI-driven diagnostic system provides a **reliable, transparent, and accessible tool** that supports early and efficient identification of gastric and duodenal ulcers while minimizing dependence on manual assessment.

REFERENCES

1. J. E. Richter, "Peptic ulcer disease and related disorders," The New England Journal of Medicine, vol. 380, no. 12, pp. 1156–1163, 2019.
2. Y. Zhou et al., "Automated detection of gastrointestinal ulcers using deep convolutional neural networks," IEEE Access, vol. 8, pp. 142333–142342, 2020.
3. M. Qadir, T. A. Rahman, and S. H. Park, "Deep learning in endoscopic image analysis for gastrointestinal disease diagnosis," IEEE Reviews in Biomedical Engineering, vol. 15, pp. 160–175, 2022.
4. J. H. Oh, S. G. Kim, and H. J. Lee, "Endoscopic diagnosis of gastric ulcers: Challenges and future directions," World Journal of Gastroenterology, vol. 27, no. 15, pp. 1754–1765, 2021.
5. S. S. Ali, P. Singh, and K. W. Kim, "Deep learning-based detection of gastric and duodenal ulcers from endoscopic images," IEEE Access, vol. 9, pp. 129845–129856, 2021.
6. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Proc. Int. Conf. Learn. Represent. (ICLR), 2015.
7. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 770–778, 2016.
8. R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp. 618–626, 2017.
9. https://www.mayoclinic.org/diseases-conditions/peptic-ulcer/symptoms-causes/syc-20354223.
10. Z. Zhang et al., "Deep learning for gastroscopic images: computer-aided techniques for clinicians," Biomedical Engineering Online, vol. 21, art. 12, Feb. 2022.
11. Y. Klang et al., "Differentiation between malignant and benign endoscopic images of gastric ulcers using deep learning," United European Gastroenterology Journal, vol. 8, no. 8_suppl, pp. 144–887, Oct. 2020.

12. K. Lee et al., “Deep learning in gastric tissue diseases: a systematic review,” Journal of Gastric Disease, 2019.
13. S. Mahmood et al., “A robust deep model for classification of peptic ulcer and other digestive tract disorders using endoscopic images,” Biomedicines, vol. 10, no. 9, p. 2195, Sept. 2022.
14. J. Lee et al., “Deep learning in gastroenterology: potentials and pitfalls,” World Journal of Gastroenterology, vol. 27, no. 28, pp. 4633–4649, 2021.
15. S. Wang et al., “Deep convolutional neural network for ulcer recognition in wireless capsule endoscopy,” Scientific Reports, vol. 9, art. 5540, 2019.
16. T. Aoki et al., “The exceptional performance of deep learning for capsule endoscopy ulcer and hemorrhage detection,” Gastrointestinal Endoscopy, vol. 91, no. 2, pp. 301–309, 2021.
17. T. Kratter et al., “Deep learning multi-domain model provides accurate patient-level ulcer identification on capsule endoscopy images,” Diagnostics, vol. 12, no. 9, p. 2122, 2022.
18. Z. Jin et al., “Artificial intelligence in gastrointestinal endoscopy: current status and future perspectives,” World Journal of Gastroenterology, vol. 28, no. 30, pp. 4394–4408, 2022.
19. https://www.kaggle.com/datasets/omkargurav/peptic-ulcer-dataset
20. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” Proc. Int. Conf. Learn. Represent. (ICLR), 2015.
21. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 770–778, 2016.
22. C. Szegedy et al., “Rethinking the inception architecture for computer vision,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 2818–2826, 2016.
23. A. Dosovitskiy et al., “An image is worth 16×16 words: Transformers for image recognition at scale,” Proc. Int. Conf. Learn. Represent. (ICLR), 2021.