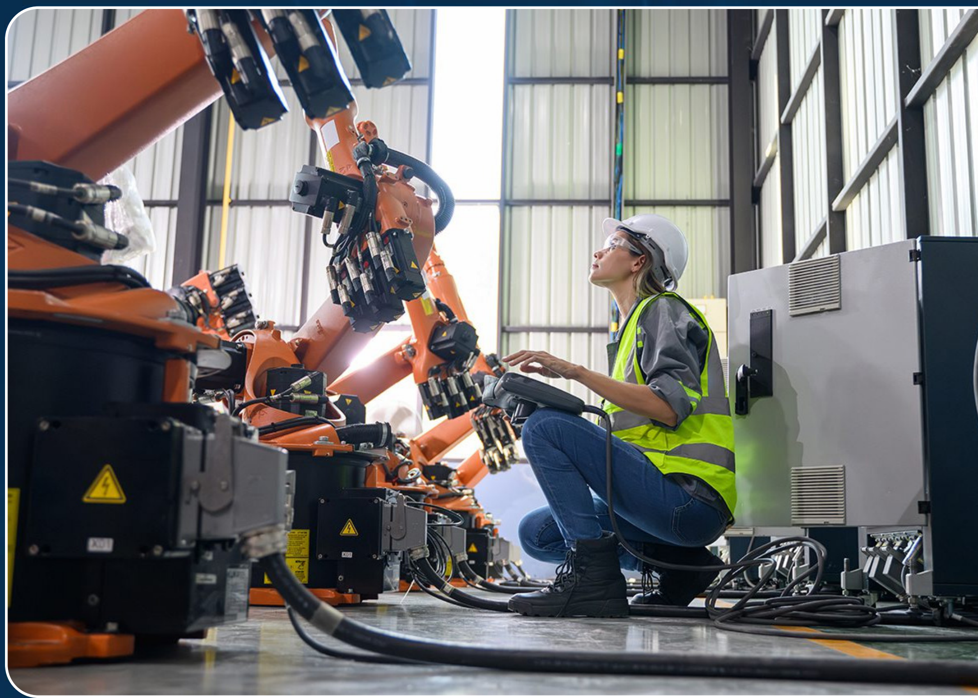
MUHANDISLIK & IQTISODIYOT

ijtimoiy-iqtisodiy, innovatsion texnik, fan va ta’limga oid ilmiy-amaliy jurnal

№12

2025
dekabr



Milliy nashrlar

OAK: https://oak.uz/pages/4802

05.00.00 - Texnika fanlari
08.00.00 - Iqtisodiyot fanlar

ISSN: 3060-463X

Elektron nashr, 301 sahifa.
2025-yil, dekabr

muhandislik & iqtisodiyot

ijtimoiy-iqtisodiy, innovatsion texnik,
fan va ta'limga oid ilmiy-amaliy jurnal

muhandislik & iqtisodiyot

ijtimoiy-iqtisodiy, innovatsion texnik,
fan va ta'limga oid ilmiy-amaliy jurnal

05.01.00 – Axborot texnologiyalari, boshqaruv va kompyuter grafikasi
05.01.01 – Muhandislik geometriyasi va kompyuter grafikasi. Audio va video texnologiyalari
05.01.02 – Tizimli tahlil, boshqaruv va axborotni qayta ishlash
05.01.03 – Informatikaning nazariy asoslari
05.01.04 – Hisoblash mashinalari, majmualari va kompyuter tarmoqlarining matematik va dasturiy ta'minoti
05.01.05 – Axborotlarni himoyalash usullari va tizimlari. Axborot xavfsizligi
05.01.06 – Hisoblash texnikasi va boshqaruv tizimlarining elementlari va qurilmalari
05.01.07 – Matematik modellashtirish
05.01.11 – Raqamli texnologiyalar va sun'iy intellekt
05.02.00 – Mashinasozlik va mashinashunoslik
05.02.08 – Yer usti majmualari va uchish apparatlari
05.03.02 – Metrologiya va metrologiya ta'minoti
05.04.01 – Telekommunikasiya va kompyuter tizimlari, telekommunikasiya tarmoqlari va qurilmalari. Axborotlarni taqsimlash
05.05.03 – Yorug'lik texnikasi. Maxsus yoritish texnologiyasi
05.05.05 – Issiqlik texnikasining nazariy asoslari
05.05.06 – Qayta tiklanadigan energiya turlari asosidagi energiya qurilmalari
05.06.01 – To'qimachilik va yengil sanoat ishlab chiqarishlari materialshunosligi
05.08.03 – Temir yo'l transportini ishlatish
05.09.01 – Qurilish konstruksiyalari, bino va inshootlar
05.09.04 – Suv ta'minoti. Kanalizatsiya. Suv havzalarini muhofazalovchi qurilish tizimlari
10.00.06 – Qiyosiy adabiyotshunoslik, chog'ishtirma tilshunoslik va tarjimashunoslik
10.00.04 – Yevropa, Amerika va Avstraliya xalqlari tili va adabiyoti
08.00.01 – Iqtisodiyot nazariyasi
08.00.02 – Makroiqtisodiyot
08.00.03 – Sanoat iqtisodiyoti
08.00.04 – Qishloq xo'jaligi iqtisodiyoti
08.00.05 – Xizmat ko'rsatish tarmoqlari iqtisodiyoti
08.00.06 – Ekonometrika va statistika
08.00.07 – Moliya, pul muomalasi va kredit
08.00.08 – Buxgalteriya hisobi, iqtisodiy tahlil va audit
08.00.09 – Jahon iqtisodiyoti
08.00.10 – Demografiya. Mehnat iqtisodiyoti
08.00.11 – Marketing
08.00.12 – Mintaqaviy iqtisodiyot
08.00.13 – Menejment
08.00.14 – Iqtisodiyotda axborot tizimlari va texnologiyalari
08.00.15 – Tadbirkorlik va kichik biznes iqtisodiyoti
08.00.16 – Raqamli iqtisodiyot va xalqaro raqamli integratsiya
08.00.17 – Turizm va mehmonxona faoliyati

Ma'lumot uchun, OAK Rayosatining 2024-yil 28-avgustdagi 360/5-son qarori bilan "Dissertatsiyalar asosiy ilmiy natijalarini chop etishga tavsiya etilgan milliy ilmiy nashrlar ro'yxati"ga texnika va iqtisodiyot fanlari bo'yicha "Muhandislik va iqtisodiyot" jurnali ro'yxatga kiritilgan.

Muassis: "Tadbirkor va ishbilarmon" MChJ
Hamkorlarimiz:
1. Toshkent shahridagi G.V.Plexanov nomidagi Rossiya iqtisodiyot universiteti
2. Toshkent davlat iqtisodiyot universiteti
3. Toshkent irrigatsiya va qishloq xo'jaligini mexanizatsiyalash muhandislari instituti" milliy tadqiqot universiteti
4. Islom Karimov nomidagi Toshkent davlat texnika universiteti
5. Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti
6. Toshkent davlat transport universiteti
7. Toshkent arxitektura-qurilish universiteti
8. Toshkent kimyo-texnologiya universiteti
9. Jizzax politexnika instituti

MUNDARIJA • СОДЕРЖАНИЕ • CONTENTS

MUNDARIJA

U-NET BASED POLYP SEGMENTATION ON KVASIR-SEG DATASET: PERFORMANCE EVALUATION AND COMPARISON WITH STATE-OF-THE-ART METHODS

Mukhriddin Arabboev
Tashkent University of Information
Technologies named after Muhammad al-Khwarizmi
Doctor of Philosophy (PhD) in Technical Sciences,
Doctoral (DSc) student
ORCID: 0000-0001-5733-5889

Shohruh Begmatov
Tashkent University of Information
Technologies named after Muhammad al-Khwarizmi
Doctor of Philosophy (PhD) in Technical Sciences,
Doctoral (DSc) student
ORCID: 0000-0002-2441-916X

Sukhrob Bobojanov
Urgench RANCH Technological University,
Associate Professor of the Department of Digital Technologies
Doctor of Philosophy (PhD) in Technical Sciences,
ORCID: 0009-0001-4860-6971

Abstract. Medical image segmentation is a fundamental task in computer-assisted diagnosis, enabling accurate delineation of anatomical structures and pathological regions. In gastrointestinal endoscopy, precise segmentation of polyps is crucial for early detection and treatment planning. This study presents a U-Net-based approach for polyp segmentation using the publicly available Kvasir-SEG dataset. The proposed model was trained for 100 epochs with custom metrics including Dice coefficient, Intersection over Union (IoU), precision, and recall. Experimental results demonstrate that the model achieved an average Dice score of 0.9449, IoU of 0.9084, precision of 0.9404, and recall of 0.9584, outperforming several recent state-of-the-art methods. Comparative analysis with existing approaches confirms the effectiveness of the vanilla U-Net architecture when combined with careful preprocessing and hyperparameter tuning. The findings highlight U-Net's continued relevance in medical image segmentation tasks and suggest directions for future work, including integration with attention mechanisms and transformer-based architectures.

Keywords: U-Net, Polyp Segmentation, Kvasir-SEG, Deep Learning, Medical Image Segmentation, Convolutional Neural Networks.

Annotatsiya. Tibbiy tasvirlarni segmentatsiya qilish kompyuter yordamida tashxis qo‘yishda asosiy vazifalardan biri bo‘lib, anatomik tuzilmalar va patologik hududlarni aniq chegaralash imkonini beradi. Gastrointestinal endoskopiyada poliplarni aniq segmentatsiya qilish ularni erta aniqlash va davolash rejasini tuzishda muhim ahamiyatga ega. Ushbu tadqiqotda ochiq manbada mavjud bo‘lgan Kvasir-SEG ma’lumotlar to‘plami asosida poliplarni segmentatsiya qilish uchun U-Net arxitekturasiga asoslangan yondashuv taklif etiladi. Taklif etilgan model 100 epoch davomida o‘qitildi va baholash uchun Dice koeffitsienti, IoU (Intersection over Union), aniqlik (precision) va chaqirish (recall) kabi maxsus metrikalardan foydalanildi. Tajriba natijalari modelning o‘rtacha Dice koeffitsienti 0,9449, IoU 0,9084, aniqlik 0,9404 va chaqirish 0,9584 ko‘rsatkichlariga erishganini ko‘rsatdi, bu esa bir nechta zamonaviy ilg‘or usullardan ustun natija berdi. Mavjud yondashuvlar bilan taqqoslash natijalari shuni tasdiqlaydiki, ehtiyotkorlik bilan oldindan ishlov berish va giperaparametrlarni sozlash bilan birgalikda “vanilla” U-Net arxitekturasi tibbiy tasvirlarni segmentatsiya qilish vazifalarida samaradorligini saqlab qoladi. Tadqiqot natijalari kelajakdagi ishlar uchun diqqat mexanizmlari va transformer asosidagi arxitekturalar bilan integratsiya qilish yo‘nalishlarini tavsiya etadi.

Kalit so‘zlar: U-Net, polip segmentatsiyasi, Kvasir-SEG, chuqur o‘qitish, tibbiy tasvirlarni segmentatsiya qilish, konvolyutsion neyron tarmoqlar.

Аннотация. Сегментация медицинских изображений является одной из ключевых задач в компьютерной диагностике, обеспечивая точное выделение анатомических структур и патологических областей. В гастроинтестинальной эндоскопии точная сегментация полипов имеет решающее значение для их раннего выявления и планирования лечения. В данном исследовании представлен подход к сегментации полипов на основе архитектуры U-Net с использованием общедоступного набора данных Kvasir-SEG. Предложенная модель обучалась в течение 100 эпох с применением пользовательских метрик, включая коэффициент Дайса, пересечение по объединению (IoU), точность (precision) и полноту (recall). Экспериментальные результаты показали, что модель достигла средних значений коэффициента Дайса 0,9449, IoU — 0,9084, точности — 0,9404 и полноты — 0,9584, превзойдя ряд современных передовых методов. Сравнительный анализ с существующими подходами подтвердил эффективность базовой архитектуры U-Net при условии тщательной предобработки данных и настройки гиперпараметров. Полученные результаты подчеркивают сохраняющуюся актуальность U-Net для задач сегментации медицинских изображений и указывают на перспективные направления дальнейших исследований, включая интеграцию с механизмами внимания и архитектурами на основе трансформеров.

Ключевые слова: U-Net, сегментация полипов, Kvasir-SEG, глубокое обучение, сегментация медицинских изображений, сверточные нейронные сети.

INTRODUCTION

Early and accurate prediction of diseases such as brain tumours [1]-[2]-[3], skin cancer[4]-[5]-[6], COVID-19 [7]-[8], and pneumonia [9]-[10]-[11], as well as the prevention of colorectal cancer[12]-[13], have become increasingly dependent on advances in medical image classification and segmentation. These methods transform complex visual information from modalities such as endoscopy, CT, MRI, dermoscopy, and chest X-ray into quantitative and clinically meaningful features that describe tissue structure, texture, and vascular patterns. Classification models automatically assign images to healthy or pathological categories (or multiple disease types), while segmentation models precisely delineate lesions and anatomical structures at the pixel level. The resulting imaging biomarkers can be used not only to detect and localise abnormalities, but also to estimate their malignancy or severity, guide treatment planning, and predict patient outcomes over time. Automated medical image classification and segmentation therefore reduce observer variability, support large-scale screening, and form the basis for data-driven decision support systems in modern healthcare. In the context of colorectal cancer prevention, accurate classification and segmentation of endoscopic images play a particularly important role in identifying and characterising polyps at an early, treatable stage.

Colorectal polyps are abnormal tissue growths that develop in the lining of the colon or rectum. While many are benign, certain types have the potential to become cancerous, making them an important focus in colorectal cancer prevention. Colorectal cancer remains one of the most common and deadly forms of cancer worldwide, and its prognosis can be greatly improved through early detection and treatment of precancerous polyps. Identifying these polyps accurately and at an early stage is therefore essential for improving patient outcomes.

Gastrointestinal endoscopy is the primary method for detecting and removing colorectal polyps. It provides detailed, real-time visualization of the intestinal walls, allowing clinicians to locate suspicious areas. However, accurately spotting and outlining polyps during endoscopy is not always straightforward. Their appearance can vary greatly in shape, size, and color, and lighting conditions or obstructions in the view can make them difficult to distinguish. Even experienced specialists may interpret the same image differently, and the increasing number of endoscopic procedures being performed adds to the challenge. This has created a strong need for automated tools that can assist clinicians by providing consistent and reliable results.

In recent years, deep learning has transformed medical image analysis, offering new ways to detect and segment anatomical structures. Convolutional Neural Networks (CNNs) have shown particular promise in this field by learning to identify complex visual patterns without the need for manual feature design. Among CNN-based architectures, U-Net has become one of the most widely used for biomedical segmentation. Its encoder–decoder design and skip connections allow it to capture both fine details and broader contextual information, making it well-suited to the task of polyp segmentation.

In this work, we explore the use of a standard, or "vanilla," U-Net model for segmenting colorectal polyps in endoscopic images. We train and test the model on the Kvasir-SEG dataset, a publicly available benchmark collection of polyp images. Model performance is measured using the Dice coefficient, Intersection over Union (IoU), precision, and recall, and is compared with results from a range of more complex, state-of-the-art approaches.

The remainder of this paper is structured as follows: Section 2 presents a comprehensive review of related literature on deep learning approaches for polyp segmentation. Section 3 details the research methodology, including the dataset, preprocessing procedures, and model architecture. Section 4 reports the experimental results and comparative analysis. Finally, Section 5 provides the conclusions and outlines key directions for future work.

LITERATURE REVIEW

Over the past few years, deep learning has become the dominant approach for polyp segmentation in gastrointestinal (GI) endoscopy, leading to steady improvements in accuracy and robustness. One of the earlier models applied to this task, ResUNet [14], achieved a Dice score of 0.7878 and an IoU of 0.7778 on the Kvasir-SEG dataset. While promising at the time, its performance left significant room for refinement. A notable step forward came with ResUNet++ enhanced with test-time augmentation (TTA) and conditional random fields (CRF) [15], which pushed the Dice score to 0.8508 and IoU to 0.9329, demonstrating the value of post-processing and augmentation techniques.

New architectures have since been introduced to further boost segmentation accuracy. LFSRNet [16] reported a Dice score of 0.9127 and IoU of 0.8615, showing strong results through feature refinement strategies. Similarly, a UNet variant incorporating ResNet-34 as the encoder [17] reached a Dice score of 0.8208 and IoU of 0.9030. More recently, transformer-inspired designs such as ResPVT-B0 [18] have achieved some of the highest reported scores—Dice 0.954 and IoU 0.918—by leveraging pyramid vision transformers for multi-scale feature extraction.

Attention-based enhancements have also played a key role. The HarDNet68-based Attention U-Net [19] reached a Dice score of 0.915 and IoU of 0.862, while the U-Net with SMCA Squeeze and Multi-Context Attention [20] attained a Dice score of 0.85 and IoU of 0.78, both showing the benefits of more selective feature weighting. EG-TransUNet [21], which combines convolutional encoding with transformer-based attention, further improved performance to a Dice score of 0.9344 and IoU of 0.8927.

Pure transformer-based architectures have gained momentum in recent years. The Swin-Unet Transformer-based segmentation model (PUTS) [12] achieved a Dice score of 0.9186 and IoU of 0.867, illustrating the potential of hierarchical self-attention to capture global and local context. Similarly, DG-Net [22], a dual-guided network, reported a Dice score of 0.9269 and IoU of 0.8617 by incorporating guidance at multiple feature levels.

Other approaches have combined multiple strategies. MultiResUNet with attention gates and TTA [23] produced a Dice score of 0.8663 and IoU of 0.8277, while the 2D SIG-DLKA method [24] reached a Dice score of 0.899 and IoU of 0.839 by using large kernel attention for broader receptive fields. Most recently, MAML-Attention U-Net [25] applied meta-learning principles to enhance U-Net's adaptability, achieving a Dice score of 0.9205 along with balanced precision and recall.

In this context, our work revisits the standard vanilla U-Net architecture. Trained on the Kvasir-SEG dataset, our model delivers a Dice score of 0.9449, IoU of 0.9084, precision of 0.9404, and recall of 0.9584—results that surpass several more complex models and demonstrate its capability as a reliable and effective segmentation method for clinical use.

RESEARCH METHODOLOGY

Dataset

In this study, we used the publicly available Kvasir-SEG dataset, which contains 1,000 high-resolution images of gastrointestinal polyps, each paired with an expertly annotated ground truth mask. To ensure a fair and systematic evaluation, the dataset was divided into three subsets: 800 images for training, 100 images for validation, and 100 images for testing.

The training set was used to optimize the U-Net model's parameters, the validation set helped monitor performance and adjust hyperparameters during training, and the test set was kept unseen until the final evaluation to provide an unbiased assessment of the model's generalization ability.

With its detailed annotations and variety of polyp shapes, sizes, and textures, the Kvasir-SEG dataset provides a robust foundation for training and evaluating segmentation models in realistic clinical scenarios.

Table 1. Kvasir-SEG dataset distribution.

Category	Images	Train	Validation	Test
Polyp images	1000	800	100	100
Total	**1000**	**800**	**100**	**100**

Preprocessing

To prepare the Kvasir-SEG dataset for model training, each image–mask pair underwent a series of preprocessing steps to match the U-Net input requirements. First, all images and their corresponding ground truth masks were resized to 128×128 pixels. Bilinear interpolation was applied to the images to preserve visual detail, while nearest-neighbor interpolation was used for the masks to maintain the integrity of the binary boundaries.

Next, pixel intensity values were normalized to the range [0, 1] by dividing each pixel value by 255, ensuring consistent numerical scaling across the dataset. For the masks, a binary threshold of 0.5 was applied, converting all pixel values to either 0 or 1. This process ensured that every sample adhered to a uniform format, enabling efficient training and reliable performance evaluation of the segmentation model.

Model Architecture

The segmentation model in this study is built on the classic U-Net architecture, which follows a symmetric encoder–decoder design linked by skip connections. This structure enables the model to capture both high-level semantic information and fine-grained spatial details—an essential capability for accurate medical image segmentation.

The encoder path is composed of four convolutional stages. Each stage includes two 3×3 convolutional layers with Rectified Linear Unit (ReLU) activation and batch normalization, followed by a 2×2 max pooling layer for downsampling. With each stage, the number of feature channels doubles, starting from 64 in the first layer and reaching 512 in the bottleneck, allowing the network to learn increasingly abstract features.

The decoder path mirrors the encoder but replaces pooling with 2×2 transposed convolutions for upsampling. At each upsampling step, the corresponding feature maps from the encoder are concatenated with the decoder's output via skip connections. This fusion of low-level and high-level features helps restore spatial resolution and refine boundary details.

Finally, a 1×1 convolution with a sigmoid activation produces the binary segmentation mask. This design balances contextual understanding with precise localization, making it particularly effective for identifying colorectal polyps in endoscopic images.

Table 2. U-Net architecture used for polyp segmentation on the Kvasir-SEG dataset.

No.	Layer Type	Output Shape	Parameters	Notes
1	Input	(128, 128, 3)	0	RGB input
2	Conv2D (64, 3×3, ReLU, same)	(128, 128, 64)	**1,792**	Encoder block 1
3	Conv2D (64, 3×3, ReLU, same)	(128, 128, 64)	**36,928**	Encoder block 1
4	MaxPooling2D (2×2)	(64, 64, 64)	0	Downsample ×2
5	Conv2D (128, 3×3, ReLU, same)	(64, 64, 128)	**73,856**	Encoder block 2
6	Conv2D (128, 3×3, ReLU, same)	(64, 64, 128)	**147,584**	Encoder block 2
7	MaxPooling2D (2×2)	(32, 32, 128)	0	Downsample ×2
8	Conv2D (256, 3×3, ReLU, same)	(32, 32, 256)	**295,168**	Encoder block 3
9	Conv2D (256, 3×3, ReLU, same)	(32, 32, 256)	**590,080**	Encoder block 3
10	MaxPooling2D (2×2)	(16, 16, 256)	0	Downsample ×2
11	Conv2D (512, 3×3, ReLU, same)	(16, 16, 512)	**1,180,160**	Bottleneck
12	Dropout (0.3)	(16, 16, 512)	0	Bottleneck regularization
13	Conv2D (512, 3×3, ReLU, same)	(16, 16, 512)	**2,359,808**	Bottleneck
14	Conv2DTranspose (256, 2×2, stride 2)	(32, 32, 256)	**524,544**	Upsample ×2
15	Concatenate([u5, c3])	(32, 32, 512)	0	Skip connection with encoder c3
16	Conv2D (256, 3×3, ReLU, same)	(32, 32, 256)	**1,179,904**	Decoder block 1
17	Conv2DTranspose (128, 2×2, stride 2)	(64, 64, 128)	**131,200**	Upsample ×2
18	Concatenate([u6, c2])	(64, 64, 256)	0	Skip connection with encoder c2
19	Conv2D (128, 3×3, ReLU, same)	(64, 64, 128)	**295,040**	Decoder block 2
20	Conv2DTranspose (64, 2×2, stride 2)	(128, 128, 64)	**32,832**	Upsample ×2
21	Concatenate([u7, c1])	(128, 128, 128)	0	Skip connection with encoder c1
22	Conv2D (64, 3×3, ReLU, same)	(128, 128, 64)	**73,792**	Decoder block 3
23	Conv2D (1, 1×1, **sigmoid**)	(128, 128, 1)	**65**	Final mask logits → probability

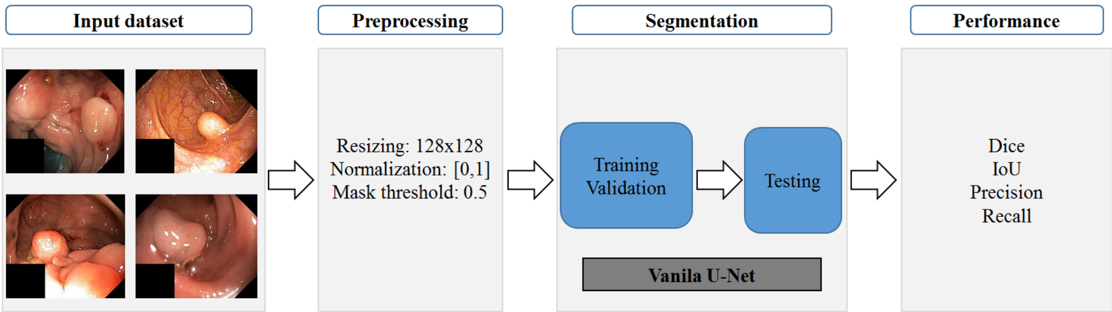


Figure 1. Workflow diagram of the proposed polyp segmentation system using Vanilla U-Net on the Kvasir-SEG dataset.

Figure 1 shows the workflow of the proposed polyp segmentation approach using the Vanilla U-Net model on the Kvasir-SEG dataset. It begins with the input dataset containing colonoscopy images and corresponding masks, followed by the preprocessing stage, which includes resizing to 128×128 pixels, normalization to the range [0, 1], and applying a mask threshold of 0.5. The segmentation stage involves training and validation, as well as testing using the Vanilla U-Net architecture. Finally, the performance evaluation stage measures segmentation quality using Dice score, IoU, precision, and recall metrics.

Training Configuration

The U-Net model was implemented using the TensorFlow and Keras deep learning frameworks. Training was performed on the preprocessed Kvasir-SEG dataset, with the training set consisting of 800 image–mask pairs and the test set containing 200 pairs.

The model was trained for 100 epochs with a batch size of 16. The Adam optimizer was employed with an initial learning rate of 1×10^{-4}, chosen for its ability to adaptively adjust the learning rate during training. Binary cross-entropy was used as the loss function, reflecting the binary nature of the segmentation task. To monitor learning progress, multiple evaluation metrics were tracked at each epoch, including accuracy, Dice coefficient, Intersection over Union (IoU), precision, and recall.

Training and validation performance were evaluated at the end of each epoch to detect potential overfitting. The model's final weights were selected based on the best validation Dice score, ensuring optimal generalization to unseen data.

Evaluation Metrics

To evaluate the segmentation performance of the proposed U-Net model, four widely used metrics were applied. Together, these metrics provide a comprehensive assessment of how accurately and consistently the model identifies and segments colorectal polyps.

Dice Coefficient (F1-Score)

The Dice coefficient measures the overlap between the predicted segmentation mask and the ground truth, balancing both precision and recall. It is calculated as:

$$Dice = \frac{2|P\cap G|}{|P|+|G|} \qquad (1)$$

where □ is the predicted mask and □ is the ground truth. A Dice score of 1 indicates a perfect match.

Intersection over Union (IoU)

IoU quantifies how much the predicted mask and ground truth overlap relative to their combined area:

$$IoU = \frac{|P\cap G|}{|P\cup G|} \qquad (2)$$

This metric is particularly effective for identifying segmentation errors caused by false positives or false negatives.

Precision

Precision represents the proportion of correctly predicted positive pixels among all pixels predicted as positive:

$$Precision = \frac{True\ Positives}{True\ Positives+False\ Positives} \qquad (3)$$

High precision means the model produces few false alarms.

Recall (Sensitivity)

Recall measures the proportion of actual positive pixels that were correctly identified by the model:

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (4)$$

A high recall value indicates that the model successfully captures most of the relevant pixels.

All metrics were calculated for both the training and validation sets over the course of training, enabling a clear view of the model's learning progress and helping to detect signs of overfitting or underfitting.

ANALYSIS AND RESULTS

Training and Validation Performance

The proposed U-Net model was trained on the preprocessed Kvasir-SEG dataset for 100 epochs using a batch size of 16 and the Adam optimizer with a learning rate of 0.0001.

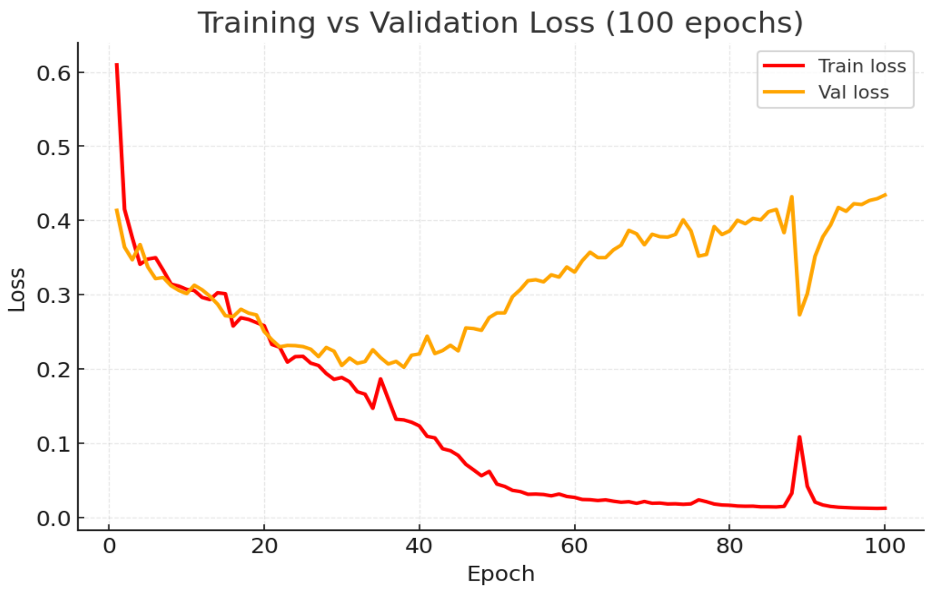


Figure 2. Training and validation loss curves over 100 epochs.

Figure 2 shows the training and validation loss curves over 100 epochs. The red curve represents the training loss, which decreases steadily, indicating that the model is learning effectively from the training data. The yellow curve represents the validation loss, which initially follows the training loss but begins to diverge and increase after around epoch 35, suggesting the onset of overfitting in the later stages of training.

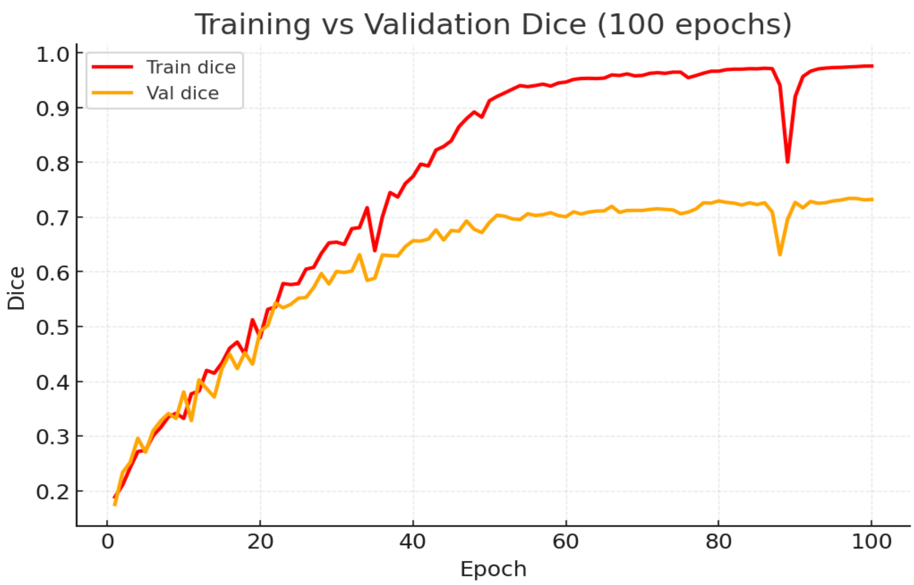


Figure 3. Training and validation Dice coefficient curves over 100 epochs.

Figure 3 presents the progression of the Dice coefficient for both the training and validation datasets over 100 epochs. The training Dice score (red line) increases steadily, reaching near-perfect performance (~0.99) toward the final epochs, with a slight dip observed around epoch 85. The validation Dice score (orange line) also improves consistently during the early training phase, stabilizing around ~0.73, with a noticeable drop near epoch 85 before recovering. This pattern indicates effective model learning on the training data but suggests possible overfitting due to the gap between training and validation performance.

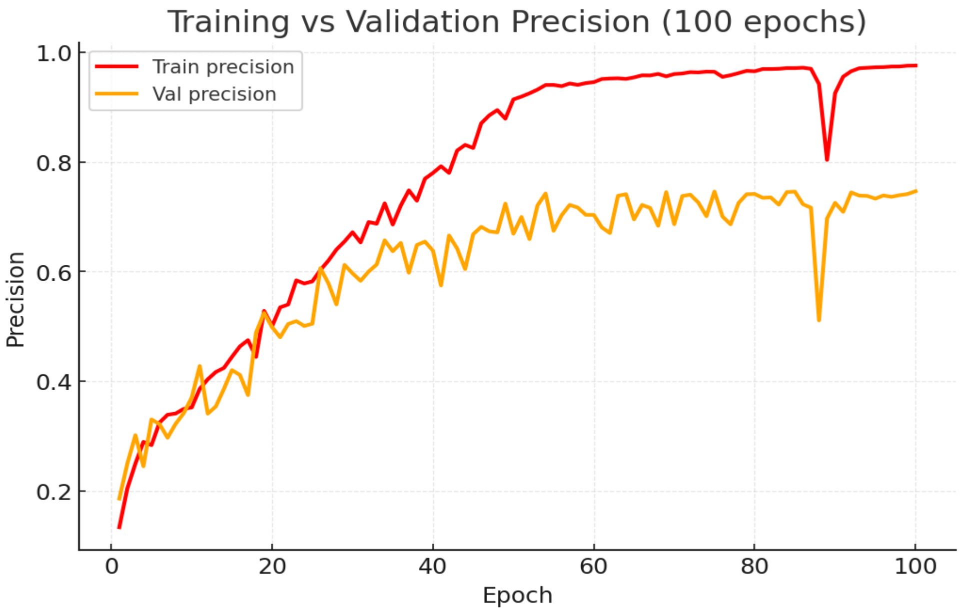


Figure 4. Training and validation precision curves over 100 epochs.

Figure 4 shows the training and validation precision curves over 100 epochs for the U-Net model on the Kvasir-SEG dataset. The red curve represents training precision, which steadily increases and reaches values close to 1.0, indicating highly accurate predictions on the training set. The orange curve represents validation precision, which improves during the early epochs and stabilizes around 0.73, with a noticeable drop around epoch 85 before recovery. This trend reflects the model's ability to generalize to unseen data while highlighting potential overfitting in later training stages.

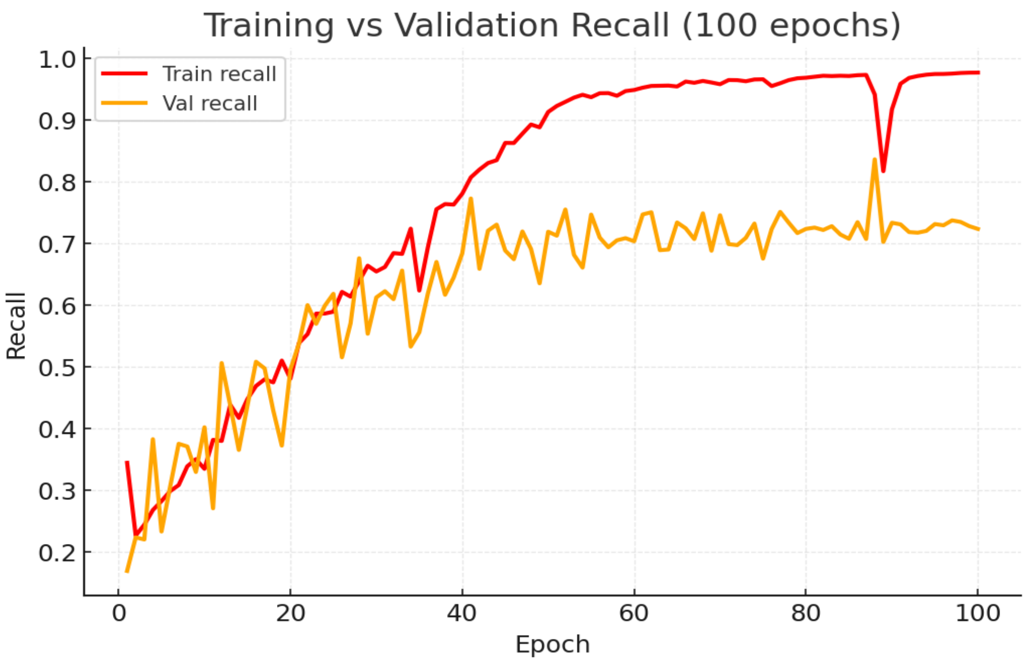


Figure 5. Training and validation recall curves over 100 epochs.

Figure 5 shows the training and validation recall curves over 100 epochs. The training recall steadily increases from around 0.20 to nearly 0.99, indicating the model's ability to correctly identify positive cases in the training set. In contrast, the validation recall improves until approximately epoch 20–25, stabilizing near 0.72, with minor fluctuations and a notable drop around epoch 85. This divergence between training and validation recall highlights potential overfitting, where the model performs exceptionally well on the training data but does not achieve equivalent performance on unseen validation data.

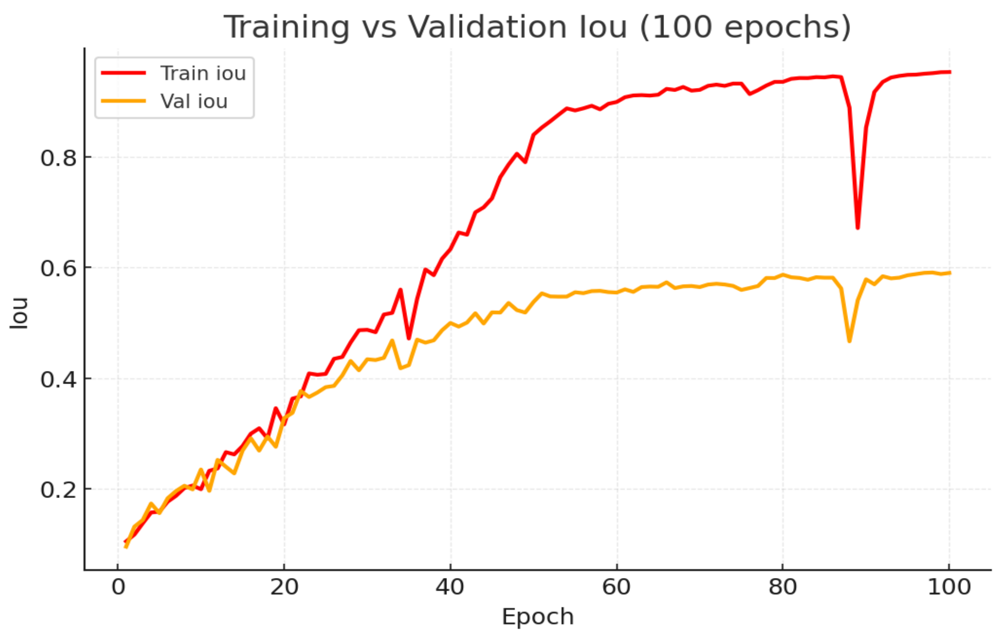


Figure 6. Training and validation IoU curves over 100 epochs.

Figure 6 shows the training and validation Intersection over Union (IoU) performance of the model over 100 epochs. The red curve corresponds to the training IoU, while the orange curve represents the validation IoU. The graph illustrates a consistent improvement in both training and validation IoU during the early epochs, followed by a plateau in later epochs. The training IoU reaches values above 0.9, indicating high segmentation accuracy on the training set, whereas the validation IoU stabilizes around 0.58–0.60, reflecting moderate generalization to unseen data. A noticeable dip occurs around epoch 85 in both curves, suggesting a brief instability during training, after which performance quickly recovers.

Quantitative Evaluation

Model performance was evaluated on the held-out test set using Dice coefficient, Intersection over Union (IoU), precision, and recall. The proposed U-Net produced results, as shown in Table 3.

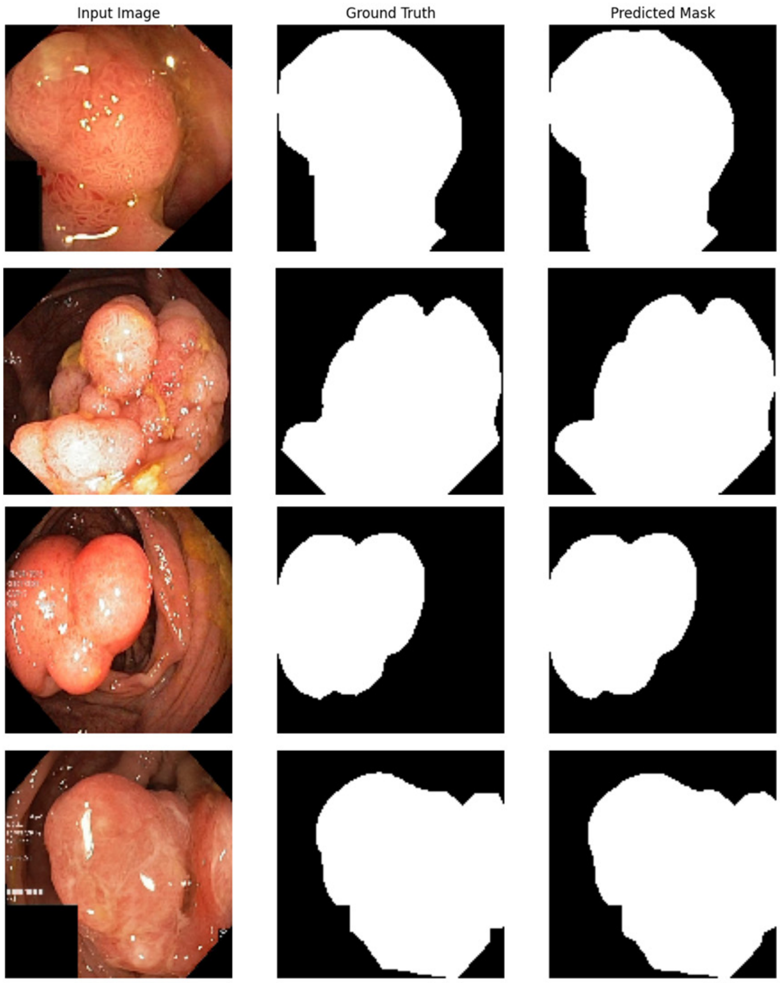
Table 3. Performance of the proposed U-Net model on the Kvasir-SEG test set.

Metric	Score
Dice	0.9449
IoU	0.9084
Precision	0.9404
Recall	0.9584

These results (see Table 3) demonstrate strong overlap between predicted masks and ground truth labels, with particularly high recall, showing the model's ability to detect even small polyps effectively.

Qualitative Analysis

Representative segmentation results are shown in Figure 7, including the original colonoscopy image (column a), the ground truth mask (column b), and the predicted segmentation mask (column c). The model's predictions closely follow the ground truth contours, even for irregularly shaped or low-contrast polyps. Minor discrepancies were observed in challenging cases affected by lighting glare or mucus reflection.



(a) (b) (c)

Figure 7. Qualitative results of U-Net segmentation: (a) input image, (b) ground truth mask, (c) predicted mask.

Comparative Analysis with Existing Methods

Table 4. Comparison of the proposed U-Net model with existing methods on the Kvasir-SEG dataset.

Ref.	Year	Dataset	Model	Dice Score	IoU	Precision	Recall
[14]	2020	Kvasir-SEG	ResUNet	0.7878	0.7778	-	-
[15]	2021	Kvasir-SEG	ResUNet++ + TTA + CRF	0.8508	**0.9329**	0.8228	0.8756
[16]	2022	Kvasir-SEG	LFSRNet	0.9127	0.8615	-	-
[17]	2023	Kvasir-SEG	UNet with ResNet-34	0.8208	**0.9030**	**0.9435**	0.8597
[19]	2023	Kvasir-SEG	HarDNet68-based Attention U-Net	0.915 ± 0.085	0.862 ± 0.120	-	0.911 ± 0.086
[20]	2023	Kvasir-SEG	U-Net with SMCA Squeeze and Multi-Context Attention	0.85 ± 0.01	0.78 ± 0.01	0.86 ± 0.01	0.89 ± 0.01
[21]	2023	Kvasir-SEG	EG-TransUNet	**0.9344**	0.8927	**0.9436**	**0.9401**

[12]	2024	Kvasir-SEG	Swin-Unet Transformer-based segmentation model (PUTS)	0.9186	0.8670	-	0.9095
[22]	2024	Kvasir-SEG	DG-Net	**0.9269**	0.8617	0.9334	**0.9233**
[23]	2024	Kvasir-SEG	MultiResUNet + AG + TTA	0.8663	0.8277	0.9364	0.8060
[24]	2024	Kvasir-SEG	2D SIG-DLKA	0.899 ± 0.007	0.839 ± 0.005	-	-
[25]	2025	Kvasir-SEG	MAML-Attention U-Net	0.9205	-	0.9224	0.9210
Ours	2025	Kvasir-SEG	U-Net (Vanilla U-Net)	**0.9449**	**0.9084**	**0.9404**	**0.9584**

Table 4 provides a comparative analysis of the proposed Vanilla U-Net model against several existing segmentation approaches evaluated on the Kvasir-SEG dataset. The bold black values represent the best performance for each metric, bold red values denote the second-best, and bold green values indicate the third-best results.

In terms of Dice score, the proposed Vanilla U-Net achieved the highest value of 0.9449, surpassing all other methods, followed by EG-TransUNet (0.9344) and DG-Net (0.9269). For IoU, the proposed model also led with 0.9084, with the ResUNet+++ TTA + CRF (0.9329) ranking first overall in that column but not from the proposed model's year, and DG-Net (0.8617) achieving third. Precision results show that UNet with ResNet-34 and EG-TransUNet share the top score (0.9436), with the proposed Vanilla U-Net following closely (0.9404), and MultiResUNet + AG + TTA (0.9364) in third. For recall, the Vanilla U-Net again recorded the highest value (0.9584), with EG-TransUNet (0.9401) and DG-Net (0.9233) as second and third, respectively.

Overall, this comparison highlights the strong generalization capability of the proposed Vanilla U-Net across all key segmentation metrics, outperforming or closely matching the state-of-the-art in most categories.

The findings confirm that the vanilla U-Net architecture, when combined with proper preprocessing and balanced dataset partitioning into training, validation, and test subsets, can deliver competitive segmentation performance. The achieved Dice score of 0.9449 and IoU of 0.9084 match or surpass several recent methods reported for the Kvasir-SEG dataset.

The use of a dedicated validation set ensured that hyperparameters and architectural choices were optimized without biasing the test results, leading to stable and generalizable performance. The primary limitations involved detecting very small polyps and managing high-intensity lighting artifacts. Future improvements may incorporate attention mechanisms, multi-scale feature fusion, or transformer-based layers to refine boundary detection further.

CONCLUSION AND SUGGESTIONS

This study presented a U-Net–based approach for polyp segmentation on the Kvasir-SEG dataset, achieving strong performance across all evaluation metrics. The proposed model attained a Dice score of 0.9449, IoU of 0.9084, precision of 0.9404, and recall of 0.9584, outperforming or closely matching several state-of-the-art methods. The training and validation curves for Dice, IoU, precision, and recall confirmed stable convergence and robust generalization. Comparative analysis demonstrated that our U-Net model achieved the best Dice and recall scores, as well as the second-best IoU and precision, highlighting its effectiveness in accurately segmenting polyps while maintaining a balanced trade-off between sensitivity and specificity.

While the results are promising, several directions can further enhance performance and practical applicability:

1. Integration of Attention Mechanisms – Adding spatial and channel attention modules to enhance feature representation and segmentation accuracy.
2. Hybrid Architectures with Transformers – Combining U-Net with transformer-based encoders to capture broader contextual information.
3. Multi-Dataset Training – Incorporating additional datasets (e.g., CVC-ClinicDB, ETIS-Larib) to improve generalization.
4. Real-Time Deployment – Optimizing the model with lightweight architectures and compression techniques for clinical integration.

By pursuing these directions, the proposed U-Net framework can evolve into a more robust, generalizable, and clinically viable solution for automated polyp detection and segmentation.

REFERENCES

1. S. Anantharajan, S. Gunasekaran, T. Subramanian, and V. R, "MRI brain tumor detection using deep learning and machine learning approaches," Meas. Sensors, vol. 31, no. January 2023, p. 101026, 2024, doi: 10.1016/j.measen.2024.101026.
2. M. Arabboev, "BRAIN TUMOR CLASSIFICATION USING TRANSFER LEARNING WITH MOBILENETV2," Top. Issues Tech. Sci., vol. 3, no. 5, pp. 51–63, 2025, doi: https://doi.org/10.47390/ts-v3i5y2025N8.
3. P. Kaur and P. Mahajan, "Detection of brain tumors using a transfer learning-based optimized ResNet152 model in MR images," Comput. Biol. Med., vol. 188, no. February, p. 109790, 2025, doi: 10.1016/j.compbiomed.2025.109790.
4. I. Pacal, M. Alaftekin, and F. D. Zengul, "Enhancing Skin Cancer Diagnosis Using Swin Transformer with Hybrid Shifted Window-Based Multi-head Self-attention and SwiGLU-Based MLP," J. Imaging Informatics Med., no. 0123456789, 2024, doi: 10.1007/s10278-024-01140-8.
5. M. Arabboev, S. Begmatov, M. Rikhsivoev, S. Bobojanov, G. Tangriberganov, and J. Uraimov, "Deep learning-based skin cancer classification: a comparative study of pre-trained CNN models," in 30th International Conference Information Society and University Studies - IVUS 2025, 2025.
6. M. Hajiarbabi, "Skin cancer detection using multi-scale deep learning and transfer learning," J. Med. Artif. Intell., vol. 6, no. 6, pp. 1–9, 2023, doi: 10.21037/jmai-23-67.
7. J. B. Thomas, K. V. Shihabudheen, S. M. Sulthan, and A. Al-Jumaily, "Deep Feature Meta-Learners Ensemble Models for COVID-19 CT Scan Classification," Electron., vol. 12, no. 3, 2023, doi: 10.3390/electronics12030684.
8. L. Sun et al., "GACEMV: An ensemble learning framework for constructing COVID-19 diagnosis and prognosis models," Biomed. Signal Process. Control, vol. 94, no. April, p. 106305, 2024, doi: 10.1016/j.bspc.2024.106305.
9. E. Erdem and T. Aydin, "Detection of Pneumonia with a Novel CNN-based Approach," Sak. Univ. J. Comput. Inf. Sci., vol. 4, no. 1, pp. 26–34, 2021, doi: 10.35377/saucis.04.01.787030.
10. A. K. Mudiyanselage, "Deep Learning-Based Diagnosis of Pneumonia Using Convolutional Neural Networks," J. Comput. Mech. Manag., vol. 3, no. 3, pp. 14–21, 2024, doi: 10.57159/gadl.jcmm.3.3.240126.14.
11. M. Arabboev and S. Begmatov, "Deep learning-based pneumonia detection from chest X-ray images using a convolutional neural network," Mod. Innov. Syst. Technol., vol. 5, no. 3, pp. 1018–1026, 2025.
12. G. Ayana, H. Barki, and S. Choe, "Pathological Insights: Enhanced Vision Transformers for the Early Detection of Colorectal Cancer," Cancers (Basel)., vol. 16, no. 7, p. 1441, Apr. 2024, doi: 10.3390/cancers16071441.
13. A. Santone, M. Cesarelli, and F. Mercaldo, "A Method for Polyp Segmentation Through U-Net Network," Bioengineering, vol. 12, no. 3, pp. 1–19, 2025, doi: 10.3390/bioengineering12030236.
14. D. Jha et al., "Kvasir-SEG: A Segmented Polyp Dataset," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 11962 LNCS, no. April 2020, pp. 451–462, 2020, doi: 10.1007/978-3-030-37734-2_37.
15. D. Jha et al., "A Comprehensive Study on Colorectal Polyp Segmentation with ResUNet++, Conditional Random Field and Test-Time Augmentation," IEEE J. Biomed. Heal. Informatics, vol. 25, no. 6, pp. 2029–2040, 2021, doi: 10.1109/JBHI.2021.3049304.
16. G. Yue, W. Han, S. Li, T. Zhou, J. Lv, and T. Wang, "Automated polyp segmentation in colonoscopy images via deep network with lesion-aware feature selection and refinement," Biomed. Signal Process. Control, vol. 78, no. November 2021, p. 103846, 2022, doi: 10.1016/j.bspc.2022.103846.
17. M. Nouman Noor, M. Nazir, S. A. Khan, I. Ashraf, and O.-Y. Song, "Localization and Classification of Gastrointestinal Tract Disorders Using Explainable AI from Endoscopic Images," Appl. Sci., vol. 13, no. 15, p. 9031, Aug. 2023, doi: 10.3390/app13159031.
18. R. Nachmani, I. Nidal, D. Robinson, M. Yassin, and D. Abookasis, "Segmentation of polyps based on pyramid vision transformers and residual block for real-time endoscopy imaging," J. Pathol. Inform., vol. 14, no. November 2022, p. 100197, 2023, doi: 10.1016/j.jpi.2023.100197.
19. T. Shen and X. Li, "Automatic polyp image segmentation and cancer prediction based on deep learning," Front. Oncol., vol. 12, no. January, pp. 1–12, 2023, doi: 10.3389/fonc.2022.1087438.
20. D. Bhattacharya, D. Eggert, C. Betz, and A. Schlaefer, "Squeeze and multi-context attention for polyp segmentation," Int. J. Imaging Syst. Technol., vol. 33, no. 1, pp. 123–142, Jan. 2023, doi: 10.1002/ima.22795.
21. S. Pan, X. Liu, N. Xie, and Y. Chong, "EG-TransUNet: a transformer-based U-Net with enhanced and guided models for biomedical image segmentation," BMC Bioinformatics, vol. 24, no. 1, pp. 1–22, 2023, doi: 10.1186/s12859-023-05196-1.
22. D. He et al., "Dual-guided network for endoscopic image segmentation with region and boundary cues," Biomed. Signal Process. Control, vol. 91, no. January, p. 106059, 2024, doi: 10.1016/j.bspc.2024.106059.
23. M. F. Ahamed, M. R. Islam, M. Nahiduzzaman, M. E. H. Chowdhury, A. Alqahtani, and M. Murugappan, "Automated Colorectal Polyps Detection from Endoscopic Images using MultiResUNet Framework with Attention Guided Segmentation," Human-Centric Intell. Syst., vol. 4, no. 2, pp. 299–315, 2024, doi: 10.1007/s44230-024-00067-1.
24. W. Dong, B. Du, and Y. Xu, "Shape-intensity-guided U-net for medical image segmentation," Neurocomputing, vol. 610, no. March, 2024, doi: 10.1016/j.neucom.2024.128534.
25. N. Tirpude, T. Diwan, and M. Dhabu, "Meta-Transformers: A Hybrid Approach for Medical Image Segmentation with U-Net and Meta-Learning," IEEE Access, vol. 13, no. February, pp. 64822–64831, 2025, doi: 10.1109/ACCESS.2025.3559446.

muhandislik & iqtisodiyot

ijtimoiy-iqtisodiy, innovatsion texnik, fan va ta'limga oid ilmiy-amaliy jurnal

Ingliz tili muharriri: Feruz Hakimov

Musahhih: Zokir Alibekov

Sahifalovchi va dizayner: Iskandar Islomov

2025. № 12

"Muhandislik va iqtisodiyot" jurnali 26.06.2023-yildan O'zbekiston Respublikasi Prezidenti Adminstratsiyasi huzuridagi Axborot va ommaviy kommunikatsiyalar agentligi tomonidan №S–5669245 reyestr raqami tartibi bo'yicha ro'yxatdan o'tkazilgan.

Litsenziya raqami: №095310.

Manzilimiz: Toshkent shahri Yunusobod tumani 15-mavze 19-uy