Pertanika J. Sci. & Technol. 34 (S3): 1 - 26 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Swin Transfer-Based Auxiliary Classifier Generative Adversarial Network for Colorectal Cancer Analysis

**G. Kanishka[1] and K. Reddy Madhavi[2]***

*[1]Research Scholar, Department of Computer Science and Engineering, School of Computing, Mohan Babu University, Sree Sainath Nagar, 517102 Tirupati, Andhra Pradesh, India*
*[2]Department of Data Sciences, School of Computing, Mohan Babu University, Sree Sainath Nagar, 517102 Tirupati, Andhra Pradesh, India*

## ABSTRACT

Molecular pathways, genetic subtypes, and mutations associated with colorectal cancer (CRC) are important for advancing precision medicine. CRC is an enormous global health issue. Traditional methods of measurement, including gene sequencing, have been shown to be not only costly but also time-intensive, and most deep learning algorithms are inefficient and hard to understand. The present paper proposes a new mechanism to enhance advanced deep learning methods for monitoring biochemical events and predicting significant mutations using cell network information. The goal is to expand the analysis and category of CRC and introduce a distinctive model that integrates an Auxiliary Classifier Generative Adversarial Network (AC-GAN) with Swin Transformer (ST). To address weaknesses commonly encountered in CRC datasets, the AC-GAN component is used to generate high-quality annotated medical images, thereby augmenting the training data and the class distribution. The ST is a high-end vision transformer model that uses these improved datasets to achieve robust organisation by capturing histopathology image-level and global contextual details. The experimental findings from the histopathology CRC dataset show that the AC-GAN-ST combination achieves higher organisation accuracy and sensitivity than a traditional convolutional network. Such a combined approach not only improves the quality of prediction presentation but also enhances the model's overall generalisability, which is why it is a reliable, high-accuracy instrument for detecting CRC and supporting pathologists' clinical decision-making.

ARTICLE INFO

*Article history:*
Received: 03 August 2026
Accepted: 06 October 2026
Published: 20 November 2026

DOI: https://doi.org/10.47836/pjst.34.S3.01

*E-mail addresses*:
kanishkagrk@gmail.com (G. Kanishka)
kreddymadhavi@gmail.com (K. Reddy Madhavi)
* Corresponding author

*Keywords*: Auxiliary Classifier GAN (AC-GAN), colorectal cancer, histopathological image analysis, medical image classification, Swin Transformer

## INTRODUCTION

Cancer is a group of illnesses characterised by abnormal cells proliferating out of

e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

control and forming malignant tumours. Metastasis is the most frequent cause of cancer-related death. It happens when these cells separate, travel, and develop into new tumours elsewhere in the body. Cancer is responsible for one out of every six deaths worldwide, according to the World Health Organisation. The colon, breast, lung, and prostate are among the most affected sites where cancer initially appears. Colorectal cancer (CRC) is the second most deadly and third most common type of cancer (Prezja et al., 2022; Prezja et al., 2023; Skrede et al., 2020). According to data from the American Cancer Society, 56% of patients receive their diagnosis after the cancer has started to spread. Early detection and treatment are crucial. Advances in machine vision, particularly deep neural networks, have improved the automatic classification of diseases, including cancer (Prezja et al., 2022). Despite technological advancements, medical practitioners still need to analyse histologic samples to confirm diagnoses and determine the stage of cancer. Key histological characteristics in these tissues are typically highlighted by hematoxylin and eosin (HE) staining (Khan et al., 2023).

Patients with colorectal cancer are categorised to customise their monitoring and treatment plans. Tumour histopathology, genomic investigations, clinical outcomes, tumour genetics, clinical data, and quantitative biomarkers are among the attributes used for grouping. Molecular and genetic analyses have identified numerous biomarkers. The relevance of the tumour microenvironment to tumour progression has been highlighted by recent studies in tumour immunology. Therefore, identifying new prognostic and predictive biomarkers that accurately characterise cancer traits is crucial (Skrede et al., 2020).

HE-stained whole-slide images have recently enabled the quantification of primary biomarkers using deep learning (DL) methods (Echle et al., 2021; Sirinukunwattana et al., 2021). A measurable indicator generated from medical images, such as histology slides, using deep learning algorithms, is a therapeutically beneficial biomarker produced by convolutional neural networks. This biomarker has the potential to uncover the molecular underpinnings of a disease and predict clinical outcomes such as survival rates and disease progression. To develop a biomarker for colorectal cancer at stages III and IV, Kather and colleagues were among the first teams to use deep learning (Kather et al., 2019). Time and resources may be saved by automatically generating this novel biomarker from images, which performed similarly to the gold standards currently used to forecast the fate of colorectal cancer. It has also been possible to predict known biomarkers using deep learning transformers (Wagner et al., 2023). This strategy reached clinical-level accuracy in colorectal cancer procedures and surpassed existing techniques for detecting microsatellite instability in surgical samples, representing a significant improvement in the area (Ruusuvuori et al., 2023).

In this paper, the AC-GAN collective with the Swin Transformer offers a powerful solution for colorectal cancer analysis by integrating the strengths of generative and

deep learning architectures. The AC-GAN employs a generator and a discriminator: the generator synthesises realistic colorectal tissue images, while the discriminator not only differentiates real from generated images but also predicts class labels, facilitating robust feature extraction. By incorporating Swin Transformers, which excel at capturing long-range dependencies and hierarchical features through shifted windows, the model effectively learns complex spatial patterns in histopathology images, as shown by Khan et al. (2023). The proposed hybrid model leverages the generative nature of the AC-GAN to optimise training data and classification accuracy, and the Swin Transformer's attention model to refine feature representation, enabling accurate detection of cancerous patterns. Subsequently, this architecture overcomes the shortage of annotated data and variability of colorectal cancer images, improves diagnostic accuracy and aids pathologists in the early and accurate detection of cancer.

**Motivation of This Research**

- The accuracy of colorectal cancer detection should be improved to facilitate faster intervention that would greatly benefit patients.
- Through Swin Transformer and Auxiliary Classifier GANs, we can proceed to train the processing of complex medical imaging data with the state-of-the-art AI methods.
- The use of GANs would aid in producing high-quality synthetic data, which is a solution to the problem of the scarcity of annotated medical datasets.
- By adding Swin Transformers, it is possible to have more flexible modelling of image features, resulting in a better performance in recognising tumour patterns.

**The Key Contribution of This Paper**

The major contributions of this study are as follows:

- An integrated AC-GAN and Swin Transformer framework is developed for binary classification of colorectal histopathological images.
- AC-GAN is employed to generate class-conditioned synthetic samples to increase training-data diversity in the limited dataset.
- Transfer learning is used to adapt the pre-trained Swin Transformer to colorectal histopathology images.
- The Swin Transformer captures hierarchical and local-to-global visual representations for classification.
- The proposed framework is evaluated under 90:10, 80:20, and 70:30 training/testing configurations using accuracy, sensitivity, and specificity.
- The study provides a contextual comparison with previously reported approaches while acknowledging differences in datasets and experimental results.

## LITERATURE SURVEY

Analysis of colorectal cancer has attracted immense attention in recent years because of its emerging prevalence and influence on the health of the population. Early detection and diagnosis have been enhanced due to advances in imaging technologies, machine learning (ML) algorithms, and CT and MRI (Wagner et al., 2023). The research also demonstrates the effectiveness of DL methods (particularly convolutional neural networks (CNNs)) in detecting cancerous lesions and predicting patient outcomes. Moreover, studies indicate that clinical and genomic data should be combined to develop better predictive models (Ruusuvuori et al., 2023). In general, the evolving nature of colorectal cancer analysis requires a multidisciplinary approach to enhance diagnostic and treatment accuracy.

Zhang et al. (2020) proposed a multitask learning model. To address the problems of non-increasing survival likelihood and censored data, we utilise sorting loss and semi-supervised loss approaches. In this approach, a supervised learning model for forecasting the risk of familial polyposis colorectal cancer was constructed with five aspects in mind: semi-supervised logistic regression, clinical practicality, interpretability, and discriminative ability. The comparison of approaches improves our comprehension of the generalisation capacity of various models and serves as a guide for creating clinical prediction models. A prognostic model with valuable performance on multicenter real-world medical data was developed with technical assistance, and the strategy's effectiveness was verified using external data.

Sun et al. (2024) created a special loss function known as Label Distribution and Scale Distribution Aware Loss (LSDA-Loss) to address the problem of colorectal lesion imbalance in white light endoscopy. By accounting for the sample distribution and quantifying the degree of imbalance via Bayesian inference, our innovative loss function addresses category imbalance. Additionally, we evaluate the difficulty of classifying each sample based on the proportional distribution. Tests conducted on three datasets demonstrate that the proposed loss function significantly improves detection accuracy over three conventional FPN models, achieving gains of up to 94.56%. Mulenga et al. (2021) proposed a feature intensification technique that improves a dataset's current features by combining data normalisation techniques. The suggested technique improves the CRC organisation's presentation of a DNN model by integrating feature expansion with data augmentation. The model achieved AUC values of 0.89 and 0.96 on two publicly available microbiome datasets.

Hasan et al. (2024) proposed a technique for identifying key genes associated with colorectal cancer (CC) by combining machine learning-based integrative bioinformatics (ML-IB) and weighted gene co-expression network analysis (WGCNA). ML-IB detected differentially expressed genes (DEGs) and then applied a support vector machine to identify DEDGs, taking the common DEDGs across datasets. Hub genes and hub module genes were identified using MCODE scores and connectivity degrees in protein-protein

interaction networks. ML-IB-based important genes were identified by intersecting hub and hub-module genes. The intersection of WGCNA- and ML-IB-derived important genes resulted in four shared genes. Finally, two genes (AURKA and CCNA2) were identified that strongly correlated with CC patient survival, and their discriminative abilities were confirmed on an independent test dataset using AUC analysis.

To identify and categorise colorectal cancer, Ragab et al. (2023) developed the slime mould algorithm with deep transfer learning (SMADTL-CCDC). The proposed SMADTL-CCDC approach aims to accurately identify colorectal cancer by first enhancing the input image quality for the SMADTL-CCDC model. Additionally, a dense-EfficientNet approach was used to extract feature vectors from pre-processed photos. The SMA was then used in conjunction with the DHNN approach to identify and classify colorectal cancer. The SMA enables more effective parameter selection for the DHNN method.

Sarwinda et al. (2021) propose a DL method for picture organisation to identify colorectal cancer (CRC) using the ResNet architecture. The exceptional performance of DL classifiers has prompted researchers to use them in medical imaging. The model was taught to distinguish between benign and malignant colorectal cancer. Mulenga et al. (2021) proposed a feature intensification strategy that increases the number of features in existing data using data normalisation techniques. This method enhances the accuracy of the DNN architecture's CRC classifier by participating in feature extension through information intensification.

Using H&E-stained WSIs (Yamashita et al., 2021) proposed MSINet, a DL-based method for directly forecasting MSI in CRC. The study found that the model outperformed five pathologists in terms of presentation and performed well on both internal and external datasets. Sasmal et al. (2024) developed a semi-supervised generative adversarial network (GAN) to categorise colorectal polyps using histopathology images. In an adversarial manner, our structure learns discriminative features from large unlabelled datasets and incomplete labelled datasets. As in a conventional GAN, the discriminator of the proposed method is trained to distinguish generated from real images in unsupervised mode and to classify genuine histopathological images in supervised mode (Khalid et al., 2024). A robust, generalised model is produced by the discriminator's capacity to identify and extract minute features from unlabelled images during unsupervised training. Using majority voting techniques at 25% and 50%, our approach achieved classification accuracies of 87.50% and 76.25%, respectively, on the UniToPatho dataset.

## Research Gap

Although deep learning has shown promising results in colorectal histopathology analysis, existing studies mainly focus on conventional CNN-based classification, transfer learning, or individual generative approaches. Recent Transformer-based methods can capture

broader spatial relationships in histopathological images, while GAN-based approaches can address limited training data. But their combined use for binary colorectal histopathology classification remains comparatively limited. In particular, the interaction between class-conditioned synthetic-data augmentation and hierarchical Transformer-based feature learning has not been sufficiently explored on small colorectal histopathology datasets. Therefore, this study addresses this gap by integrating AC-GAN with a transfer-learning-based Swin Transformer, in which AC-GAN augments the training data and the Swin Transformer learns hierarchical visual representations for classification. Recent colorectal imaging studies also demonstrate the growing use of Transformer-based architectures in histopathological analysis, supporting the motivation to investigate this integrated framework.

### Problem Identification of the Existing System

- The current imaging methods used, including colonoscopy and CT scans, are susceptible to false negatives because they might fail to detect small polyps or early-stage cancers.
- Most existing systems have high false-positive rates, which subject patients to unneeded stress levels and result in invasive follow-up processes. This may also strain the health facilities.
- The current systems are not very effective in incorporating multimodal data, which restricts the ability to analyse the information thoroughly and to plan the treatment individually.
- Traditional analysis methods may require significant time and resources, leading to delays in diagnosis and treatment. Automated systems may not be widely implemented due to high costs or complexity.

## PROPOSED SYSTEM

### Overall Architecture of Colorectal Cancer Analysis

The proposed framework integrates AC-GAN, transfer learning, and Swin Transformer to address the limited size and visual variability of the colorectal histopathology dataset. AC-GAN is used to generate class-conditioned synthetic training samples, thereby increasing the diversity of the training data. Transfer learning provides the Swin Transformer with pre-trained visual representations (Zidan et al., 2023), which is particularly useful when the target medical dataset is relatively small. The Swin Transformer is then fine-tuned on colorectal histopathology images to learn hierarchical visual features via shifted-window attention, capturing both local tissue patterns and broader spatial relationships. Thus, AC-GAN addresses training-data limitations, transfer learning improves the initialisation of the feature extractor, and Swin Transformer performs the final hierarchical feature learning and classification. The overall architecture is illustrated in Figure 1.

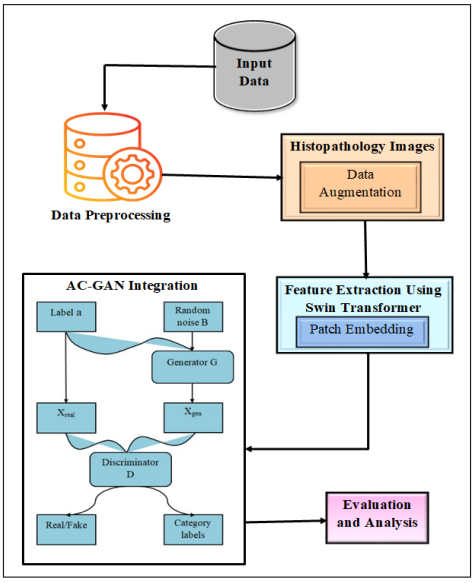



*Figure 1.* Block diagram of proposed AC-GAN-ST

## Dataset

The recently released Warwick-QU dataset, which was the subject of the GlaS@ MICCAI2015 challenge, is used to test our proposed method, AC-GAN-ST, and is publicly available on the contest website (https://www.kaggle.com/code/kerneler/starter-warwick-qu-changed-dataset-d429aecc-a) (Sujatha et al., 2024). It contains 171 images categorised as either cancerous or normal. Figure 2 shows several example images of a colorectal tumour.

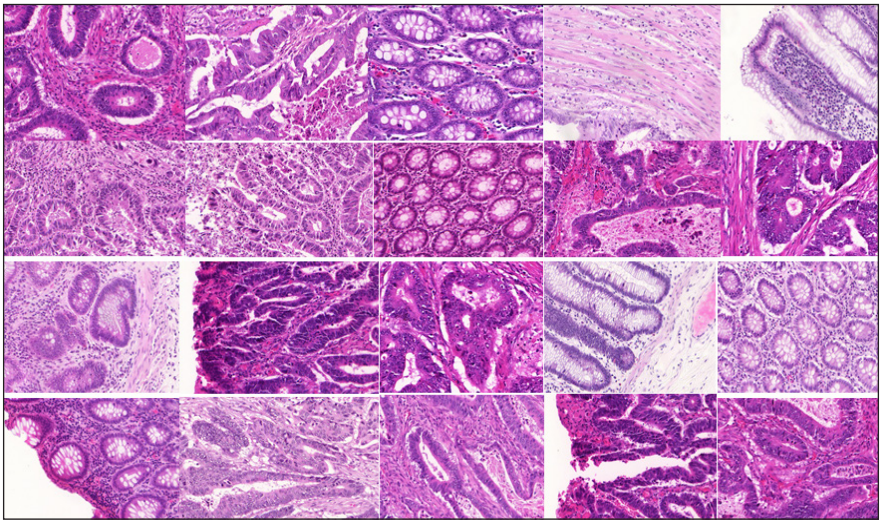


*Figure 2.* Sample images of colorectal cancer

## Data Pre-Processing and Augmentation

Histopathological images may exhibit noise and high visual variability, thereby complicating automatic classification. In the current study, the images have been pre-processed using the proposed Wiener-based Midpoint Weighted Non-Local Means Filter (WMW-NLM) to remove noise while preserving diagnostically significant structures. The geometric augmentation approach, including translation, scaling, and rotation, was used to increase the variability of the training images.

It should be noted that data augmentation was considered a training-stage procedure. Images reserved for testing were left unaltered and were considered evaluation samples. Therefore, the goal of data augmentation in the current study was to expand data diversity during training rather than to enlarge the test set. It helped prevent the artificial improvement in test performance that would otherwise be caused by augmented evaluation samples.

For the AC-GAN part, synthetically generated images conditioned on the class were used to expand the training set. Generated images were used in the training procedure and were not considered additional test cases. As the current study is oriented towards classification accuracy assessment, the synthetically generated images are described as synthetic training samples rather than clinically validated substitutes for real histopathological images. (Khalid et al., 2024).

## WMW-NLM Filter for Denoising

The middle weight (MW) in the NLM photo denoising algorithm represents the central pixel's weight. The efficiency of NLM is significantly impacted by the MW. In this work, a Wiener-like estimator is used to determine the MW. The following is the statement model for the proposed WMW-NLM (Equation 1):

$$I_N(mw,n) = I_{NF}(mw,n) + \eta(mw,n) \quad [1]$$

$I_N$ (mw, n) represents a noisy image, $I_{NF}$ (mw, n) signifies a noise-free image, and η (mw, n) specifies Gaussian noise with a variance of $\sigma^2$. The spatially adaptive NLM filter transforms the experimental data $I_N$ (mw, n) into U (mw, n) as follows (Equation 2):

$$U(mw,n) = \frac{\sum(mw',n')\in R I_{NF}(mw',n')W(mw,n,mw',n')}{\sum(mw',n')W(mw,n,mw',n')} \quad [2]$$

The weights are represented as W (mw, n, mw′, n′), and R is a search window centred at (mw, n). Two patches, ρ (mw, n), n')) containing pixels at (mw, n) and (mw′, n′), separately, are compared. The midway weight W (mw, n, mw′, n′) is defined in Equation 3:

$$W(mw,n,mw',n') = exp\left(-\frac{||I_N(\rho(mw,n))-||I_N\left(\rho(mw',n')\right)||_{2,\sigma}^2}{f^2}\right) \quad [3]$$

Where f is the filtering level-determining smooth out variable. The weighted Euclidean reserve of image patch ρ (mw, n) is designated by $|\left|I_N(\rho(mw,n)) - -I_N\left(\rho(mw',n')\right)\right||_{2,\sigma}^2$ since patch ρ (mw′, n′) has a size of $2(MW+1)\times 2(MW_\rho+1)$. The following can be shown mathematically using Equation 4:

$$||I_N(\rho(mw,n)) - I_N\left(\rho(mw',n')\right)||_{2,x}^2 = \sum_{i=-MW_\rho}^{MW_\rho}\sum_{j=-M_\rho}^{M_\rho}\zeta_\sigma(i,j)\left(I_N(mw'+i,n'+j) - I_N(mw+i,n+j)\right)^2 \quad [4]$$

In this case, $\zeta_\sigma(i,j)$ stands for the Gaussian Kernel with σ > 0. The (mw, n) whims of the related window R′ are shown here. In accordance with Equation 3, the centre weight is always 1. A Wiener filter is used in this study to build the MWs, thereby achieving high picture-dispensing quality. Let $mw_n = (mw,n), mw_{n'}' = (mw',n'), u_{mw_n}^1 = U(mw,n), I_{mw_n} = I_N(mw,n)$, $I_{mw_n} = I_N(mw,n)$ and $W_{mw_n mw_{n'}'} = W(mw,n,mw',n')\times\lambda_{mw_n}$ is used to represent the MW. Consider $R_0$ = R − (mw, n) and (2) is simplified as Equation 5:

$$u_{mw_n}^1 - I_{mw_n} = \frac{1}{\lambda_{mw} + \sum_{mw_{n'}',\in R_0} W_{mw_n mw_{n'}'}}\left(\sum_{mw_{n'}',\in R_0} W_{mw_n mw_{n'}'}\left(I_{mw_{n'}'} - I_{mw_n}\right)\right) \quad [5]$$

Equation 5 designates that the noise $u_{mw_n}^1 - I_{mw_n}$ is removed through the NLM filter procedure. The noise that has stood removed is considered by reproducing the two terms: $1/\left(\lambda_{mw_n} + \sum_{mw_{n'}',\in R_0} W_{mw_n mw_{n'}'}\right)$ and $\sum_{mw_{n'}',\in R_0} W_{mw_n mw_{n'}'}\left(I_{mw_{n'}'} - I_{mw_n}\right)$. Here, $\lambda_{m_n}$ is crucial for the denoising procedure. The quantity of noise removed decreases $\lambda_{m_n}$. As a result, excessive smoothness is easily produced. It's conceivable that some visual characteristics will be diminished or eliminated. For instance, NLM will produce an oversmoothed output with a centre weight of 0. The removed noise diminishes as $\lambda_{m_n}$ enhance. It demonstrates that denoising alone is not sufficient. For example, when the midpoint weight is set to one, NLM produces the desired result. Nevertheless, there is still a sizable amount of residual noise in the denoised image. As a result, the middle weight ought to fluctuate suitably between 0 and 1. This is the foundation for figuring out the intermediate weight. Additionally, smaller midway weights are used to minimise noise while maintaining greater structure in the denoised image. In the frequency domain, higher midway weights

are used to minimise noise without over-smoothing. The calculation of midpoint weights is considerably improved by this rule. The proposed midpoint weight can be expressed using these parameters (Equation 6):

$$\lambda_{mw_n} = \frac{\tau f^2}{V_{mw_n} + \tau f^2} \quad [6]$$

It is comparable to the Wiener filter. In both the gradient and wavelet domains, the proposed $V_{m_n}$. It is considered better than the previous Wiener filters, as shown in Equation 7.

$$V_{mw_n} = max\left(q_{mw_n} - \tau f^2, 0\right) \quad [7]$$

Here, the image's features are calculated using the feature descriptor $q_{m_n}$, which is shown in Equation 8:

$$q_{mw_n} = \frac{1}{|R_0|} \sum_{(mw',n') \in R_0} ||I_N\left(\rho(mw,n)\right) - I_N\left(\rho\left(mw',n'\right)\right)||^2_{2,\sigma} \quad [8]$$

Where the continuous τ is used for tuning the demonising boundary in the local Wiener filter. Equation 6 indicates that $\lambda_{m_n}$ would be large for the section with reduced $V_{m_n}$. As a result, Eq. 6 is consistent with the additional rule. Consequently, the WMW-NLM container produces images of higher value than NLM filters with varying midpoint weightiness.

**Feature Extraction Using Swin Transfer**

The paper "Swin Transformer for Colorectal Cancer Analysis" utilises the Swin Transformer model to identify and analyse colorectal cancer in medical imaging. The Swin (Shifted Window) Transformer is a variation of the Transformer architecture specifically designed for vision tasks. Unlike standard vision transformers, which operate globally, Swin Transformers use attention mechanisms within shifted windows, allowing them to capture both local and global context at a lower computational cost. This makes them ideal for high-resolution medical images, such as those used in colorectal cancer analysis. The Swin Transformer can evaluate colorectal tissue slides by learning intricate spatial relationships within the images. The model identifies key patterns and characteristics that indicate malignant tissue. high-quality diagnostic and analytical accuracy. To enhance the identification of cancer without. This method usually uses transfer learning, since it requires a larger labelled dataset. Here, colorectal cancer images are used to fine-tune a Swin Transformer pre-trained on a large dataset.

- **Self-Attention Mechanism in a Window:** The self-attention within each window is calculated as follows, given a window W with query Q, key K, and charge V conditions as in Equation 9:

$$Attention(Q, K, T) = soft\,max\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad [9]$$

Where $d_k$ is the dimension of the key vectors.

- **Shifted Window Mechanism:** The Swin Transformer alternates between regular window-based attention and shifted window attention. This shift is achieved by moving the window by a predetermined number of pixels, enabling patches to interact at window boundaries.

- **Patch Embedding:** The image patches are embedded in a high-dimensional space as shown in Equation 10:

$$X = Conv(X_{input}) \quad [10]$$

Conv is a convolution operation that converts input patches into fixed-dimensional vectors.

- **MLP Layer:** The output is passed to a multi-layer perceptron (MLP) for feature transformation, following the attention layers, as given in Equation 11.

$$MLP(X) = \sigma(W_2.Re\,L\,U(W_1.X + b_1) + b_2) \quad [11]$$

Where σ is a motivation function like ReLU, $b_1$ and $b_2$ are biases, and $W_1$ and $W_2$ are weight matrices.

- **Loss Function:** Colorectal cancer classification can be performed using a basic cross-entropy loss (Equation 12).

$$L_{CE} = -\sum_{i=1}^{N} y_i\, log(\hat{y}_i) \quad [12]$$

Where $y_i$ is the true label and $\hat{y}i$ is the forecast possibility for each class.

This Swin Transformer-based technique enhances the accuracy and robustness of colorectal cancer analysis by utilising sophisticated attention mechanisms that capture both micro-level characteristics within windows and macro-level features across the entire image.

**Model the Architecture of AC-GAN**

ACGAN is a modified GAN, as illustrated in Figure 3. Unlike GAN, ACGAN can generate samples from specific categories while also recognising and categorising input data. The generator G uses random noise and label b to create new samples G (a,b). After analysing the input samples, discriminator D decides if they are authentic or fake. Through adversarial training, ACGAN's ability to generate and identify samples is continually improved. Ultimately, the model can generate new examples with accurate labels.

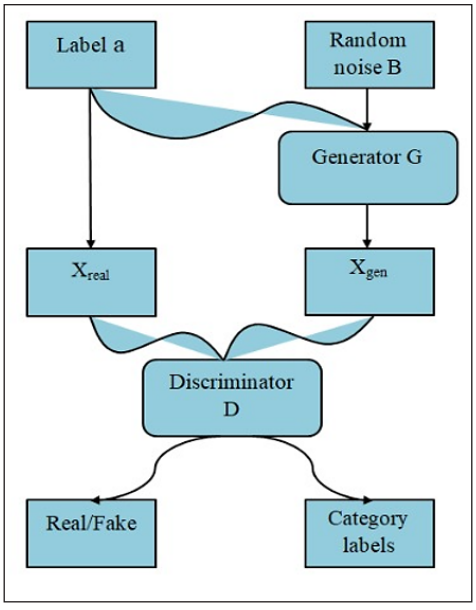



*Figure 3.* Block diagram of ACGAN

ACGAN's loss function is divided into two halves since it needs to assess the input samples' source and class label information (Equations 13 to 16):

$$L_{source} = E_{X\sim P_{data}}[log\, D\,(x)] + E_{B\sim P_b}\left[log\left(1 - D\big(G(b,a)\big)\right)\right] \quad [13]$$

$$L_{class} = E_{X\sim P_{data}}[log\, P\,(A = a|X_{real})] + E_{B\sim P_b}\left[log\, P\,(A = a|X_{generated})\right] \quad [14]$$

In this case, $P(A = a|X_{real})$ represents the provisional probability distribution of the genuine example, $P(A = a|X_{generated})$ represents the provisional probability distribution of the produced example, and G (a, b) represents the generated example when the generator inputs are random noise and example label b.

Making the most of $L_{source} + L_{class}$ is the impartial function of the D, and maximising $L_{source} - L_{class}$ is the impartial function of the G, as illustrated in the folllowing:

$$L_G = E_{X\sim P_{data}}[log\, D\,(x)] + E_{B\sim P_b}\left[log\left(1 - D(G(b,a))\right)\right] + E_{X\sim P_{data}}[log\, P\,(A = a|X_{real})] + E_{B\sim P_b}[log\, P\,(A = a|X_{generated})] \quad [15]$$

$$L_G = E_{X\sim P_{data}}[log\, D\,(x)] + E_{B\sim P_b}\left[log\left(1 - D(G(b,a))\right)\right] + E_{X\sim P_{data}}[log\, P\,(A = a|X_{real})] - E_{B\sim P_b}[log\, P\,(A = a|X_{generated})] \quad [16]$$

**Advantages of the Proposed Model**

- The Swin Transformer gathers multi-scale features to improve lesion detection and classification accuracy.
- The Auxiliary Classifier GAN creates high-quality synthetic samples to overcome data scarcity and enhance model resilience.
- The auxiliary classifier in the AC-GAN provides class-conditioned information during generation, allowing synthetic samples to be associated with the corresponding classification categories. In this study, the classification task is binary and should not be interpreted as cancer-stage prediction.

## RESULT AND DISCUSSIONS

### Experimental Setup

The computer used for the simulation has an Intel Core i7 processor. Additionally, Python 3.8 and PyTorch 1.10 were utilised to develop the proposed AC-GAN-ST model. To investigate the impact of the train-test split on classification performance, three splits were employed: 90:10, 80:20, and 70:30. The validation data used in model development were held separate from the test data. The samples generated by the AC-GAN were utilised only during training to increase the number of training samples. The original test data were not utilised in the image generation process or training process. Selecting and tuning hyperparameters is necessary for optimal model performance. Kather et al. (2019). The suggested network employs a random search methodology, with the main components being batch size, optimisation method, number of epochs, and learning rate. The suggested AC-GAN-ST hyperparameter combinations are listed in Table 1.

Table 1
*Hyperparameters setting*

| Parameter | Value |
|---|---|
| Epochs | 100 |
| Spatial patch sizes | [9 × 9], [11 × 11], [13 × 13], [15 × 15], [17 × 17] |
| Optimiser | ADAM |
| Learning rate | 0.001 |
| Batch size | 64 |

**Sensitivity**

Sensitivity, sometimes referred to as recall, is the percentage of true positives a test correctly identifies. It is a significant metric for evaluating the presentation of diagnostic tests, especially in medical contexts, as in Equation 17.

$$Sensitivity = \frac{TP}{TP + FN} \quad [17]$$

**Specificity**

Specificity is an arithmetical term that estimates the accuracy of a binary diagnostic test. It signifies the fraction of true negatives properly identified among all actual negatives. A height specificity value indicates that the test is effective at recognising negative suitcases (Equation 18).

$$Specificity = \frac{TN}{TN + FP} \quad [18]$$

**Accuracy**

Accuracy is a statistical indicator that is used to measure the presentation of organisational models. It estimates the fraction of correctly predicted cases among all examples in the dataset. The formula for calculating accuracy is as follows (Equation 19):

$$Accuracy = \frac{Number\ of Correct\ Pr\ e\ dictions}{Total\ Number\ of\ Pr\ e\ dictions} \times 100\% \quad [19]$$

**Root Mean Square Error (RMSE)**

RMSE is a commonly utilised statistic for calculating the difference between predicted and observed values in regression models. It indicates how far the predictions deviate from the actual data, with lower values indicating better model performance. It can be calculated using Equation 20.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad [20]$$

In the present study, accuracy, sensitivity, and specificity are treated as the primary measures for evaluating the binary classification task. RMSE is retained only as an additional numerical measure reported for consistency with the comparative results included in the manuscript. Since RMSE is more commonly associated with continuous-valued prediction, it is not used as the primary criterion for evaluating classification performance or for establishing superiority over other methods.

## Execution Time

Execution time refers to the total duration essential to run a specific procedure or process, from initiation to accomplishment. It is a crucial measure in performance evaluation, typically expressed in seconds or milliseconds, as in Equation 21.

$$T = T_{start} - T_{end} \quad [21]$$

The collection of confusion matrices for the AC-GAN-ST representation across various training/testing (TR/TS) data sizes is shown in Figure 4. Figure 4 presents the confusion matrices obtained under the three training/testing configurations considered in the study. The matrices are shown separately for the 90:10, 80:20, and 70:30 configurations to illustrate the class-wise prediction behaviour of the proposed AC-GAN-ST model. The values in the matrices represent the corresponding classification outcomes for the evaluated samples and should be interpreted together with the quantitative performance measures reported in Tables 2 to 4.

Because the three panels correspond to different training/testing configurations, their sample counts are not expected to be identical. The confusion matrices are therefore intended to provide a visual representation of the classification outcomes for each experimental setting, whereas Tables 2 to 4 provide the summarised accuracy, specificity, and sensitivity values.

Table 2
*Result analysis of the AC-GAN-ST technique on TR/TS data of 90:10*

| | TR/TS (90:10) | | |
|---|---|---|---|
| **Class Label** | **Accuracy** | **Specificity** | **Sensitivity** |
| **Training Phase** | | | |
| Benign | 97.31 | 92.35 | 95.47 |
| Malignant | 97.31 | 95.47 | 92.35 |
| Average | 97.31 | 93.91 | 93.91 |

Table 2 (*continued*)

| TR/TS (90:10) | | | |
|---|---|---|---|
| **Class Label** | **Accuracy** | **Specificity** | **Sensitivity** |
| **Testing Phase** | | | |
| Benign | 94.46 | 90.00 | 95.47 |
| Malignant | 94.46 | 95.47 | 90.00 |
| Average | 94.46 | 92.735 | 92.735 |

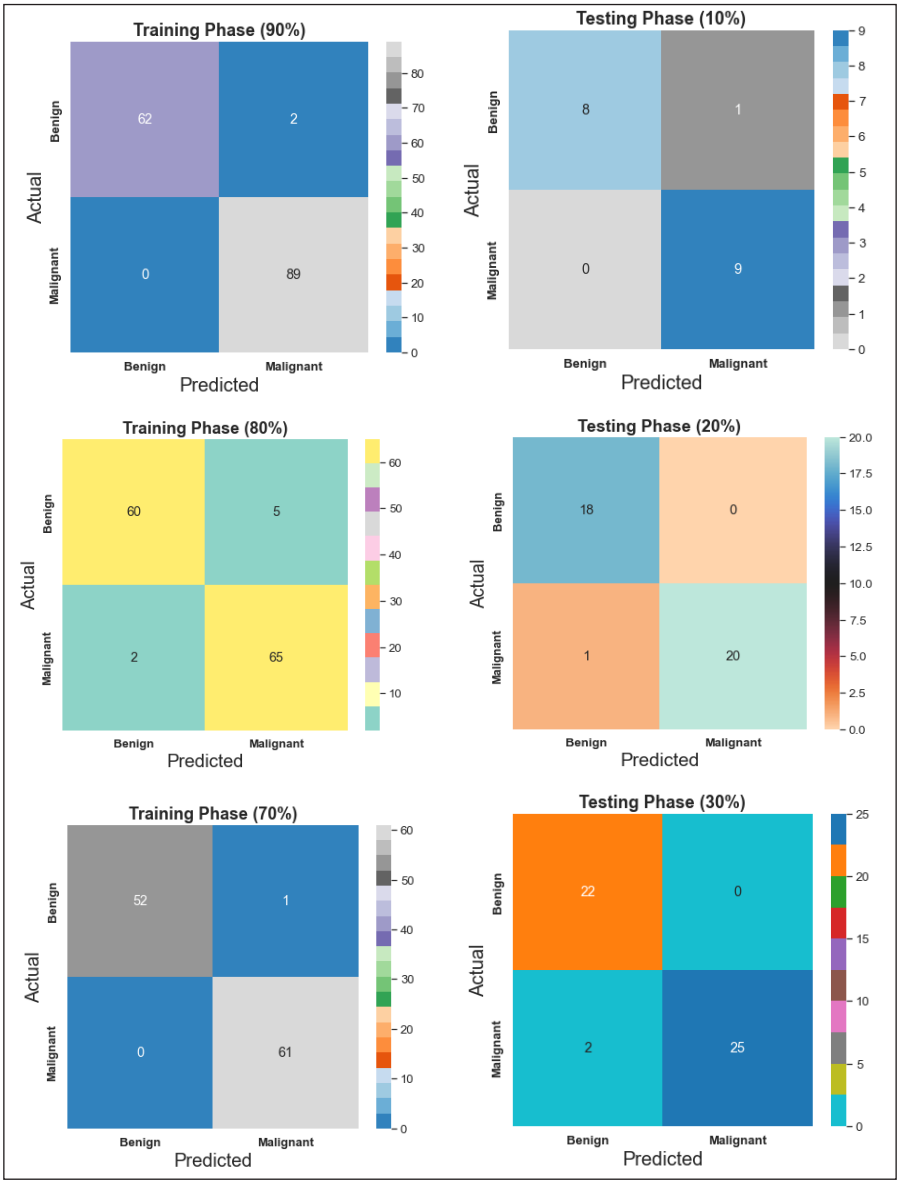



*Figure 4.* Confusion matrix of the AC-GAN-ST method on distinct sizes of TR/TS data

Table 2 summarises the performance of AC-GAN-ST under the 90:10 training/testing configuration. The model achieved an average training accuracy of 97.31%, with average specificity and sensitivity of 93.91% each. On the testing data, the corresponding average accuracy, specificity, and sensitivity were 94.46%, 92.74%, and 92.74%. These values represent the performance obtained under the 90:10 configurations and are interpreted as a dataset-specific evaluation of the proposed framework. The average metrics for the testing phase were accuracy 94.46%, specificity 92.74%, and sensitivity 92.74%.

Using training/testing (TR/TS) data with a 90:10 ratio, Figure 5 presents the training accuracy (TA), validation accuracy (VA), training loss (TL), and validation loss (VL) for this configuration. The plotted curves show the progression of training and validation metrics across training epochs. The validation and training curves follow a similar overall trend, while the corresponding loss curves decrease during training. These observations provide a qualitative indication of the model's learning behaviour; no statistical significance test was performed on the differences between the curves. The AC-GAN-ST model exhibited the lowest TL and VL values, with VL significantly below TL, as indicated by the experimental data.

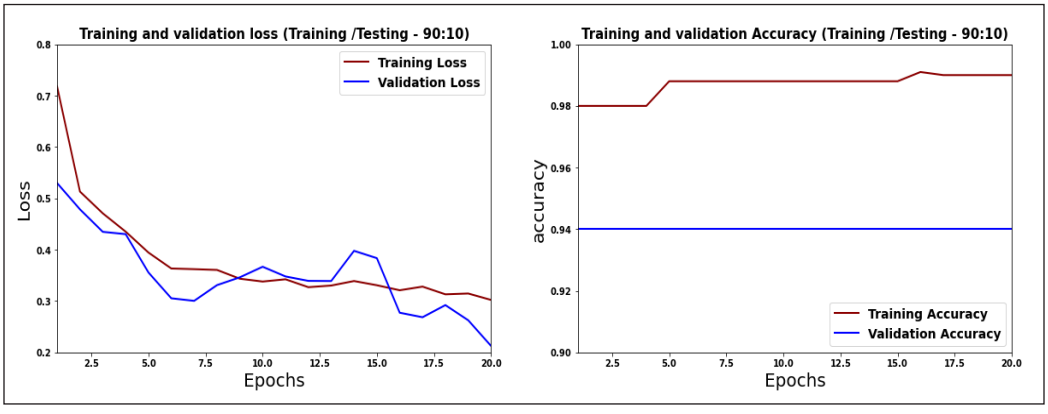



*Figure 5.* TA/TL and VA/VL analysis of the AC-GAN-ST model on the TR/TS data of 90:10

Table 3 reports the results obtained using the 80:20 training/testing configuration. The model achieved an average training accuracy of 95.23%, with average specificity and sensitivity of 93.05% each. For the test data, the average accuracy was 95.97%, and the average specificity and sensitivity were 92.78% each. The results demonstrate that the reported performance remains relatively consistent as the proportion of testing data increases under this experimental configuration. This resulted in an overall average accuracy of 95.23%, specificity of 93.05%, and sensitivity of 93.05%. In testing, the model maintained high performance, achieving 95.97% accuracy across both classes, with the benign and malignant classes showing specificities of 90.24% and 95.33% and sensitivities

of 95.33% and 90.24%, respectively. The average test results were 95.97% accuracy, 92.78% specificity, and 92.78% sensitivity.

Table 3
*Result analysis of the AC-GAN-ST technique on the TR/TS data of 80:20*

| TR/TS (80:20) | | | |
|---|---|---|---|
| **Class Label** | **Accuracy** | **Specificity** | **Sensitivity** |
| **Training Phase** | | | |
| Benign | 95.23 | 91.73 | 94.38 |
| Malignant | 95.23 | 94.38 | 91.73 |
| Average | 95.23 | 93.05 | 93.05 |
| **TR/TS (80:20)** | | | |
| **Class Label** | **Accuracy** | **Specificity** | **Sensitivity** |
| **Testing Phase** | | | |
| Benign | 95.97 | 90.24 | 95.33 |
| Malignant | 95.97 | 95.33 | 90.24 |
| Average | 95.97 | 92.78 | 92.78 |

Figure 6 displays the TA and VA produced by the AC-GAN-ST method using 80:20 TR/TS data. According to the experimental data, the AC-GAN-ST model obtained the highest TA and VA values, with VA seemingly exceeding TA. Figure 6 also displays the TL and VL derived from 80:20 TR/TS data using the AC-GAN-ST technique. The AC-GAN-ST model exhibited the lowest TL and VL values in the experimental data, with VL appearing to be lower than TL.

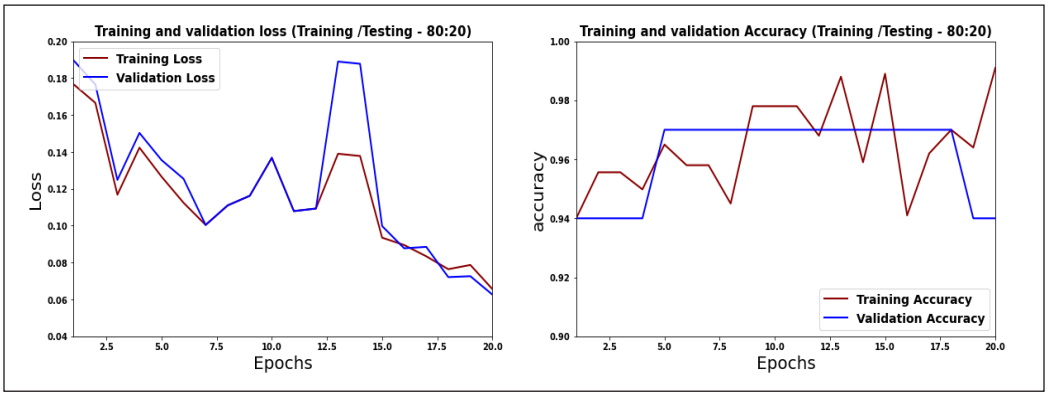



*Figure 6.* TA/TL and VA/VL analysis of the AC-GAN-ST model on the TR/TS data of 80:20

Table 4 presents an analysis of the AC-GAN-ST method's results using various performance measures on a 70:30 training-to-testing split (TR/TS). In the training stage,

the model achieved 97.13% accuracy across both benign and malignant classes, with specificity and sensitivity averaging 97.19% and 95.73%, respectively. In the testing phase, the model maintained high accuracy at 96.45% for both classes, with an average specificity of 96.38% and sensitivity of 93.45%. These results indicate robust model performance in distinguishing between benign and malignant classes in both training and testing phases.

Table 4
*Result analysis of the AC-GAN-ST technique on the TR/TS data of 70:30*

| TR/TS (70:30) | | | |
|---|---|---|---|
| **Class Label** | **Accuracy** | **Specificity** | **Sensitivity** |
| **Training Phase** | | | |
| Benign | 97.13 | 96.39 | 96.23 |
| Malignant | 97.13 | 98.00 | 95.23 |
| Average | 97.13 | 97.19 | 95.73 |
| **TR/TS (70:30)** | | | |
| **Class Label** | **Accuracy** | **Specificity** | **Sensitivity** |
| **Testing Phase** | | | |
| Benign | 96.45 | 95.76 | 94.18 |
| Malignant | 96.45 | 97.00 | 93.32 |
| Average | 96.45 | 94.76 | 93.45 |

The TA and VA produced by the AC-GAN-ST algorithm utilising 70:30 TR/TS data are displayed in Figure 7. The AC-GAN-ST model produced the highest values for both TA and VA, according to the experimental results, with VA seeming to be higher than TA. Figure 7 also displays the TL and VL derivatives since the AC-GAN-ST model on 70:30 TR/TS data. The AC-GAN-ST method yielded the lowest TL and VL values, with VL apparently lower than TL, consistent with the experimental results.

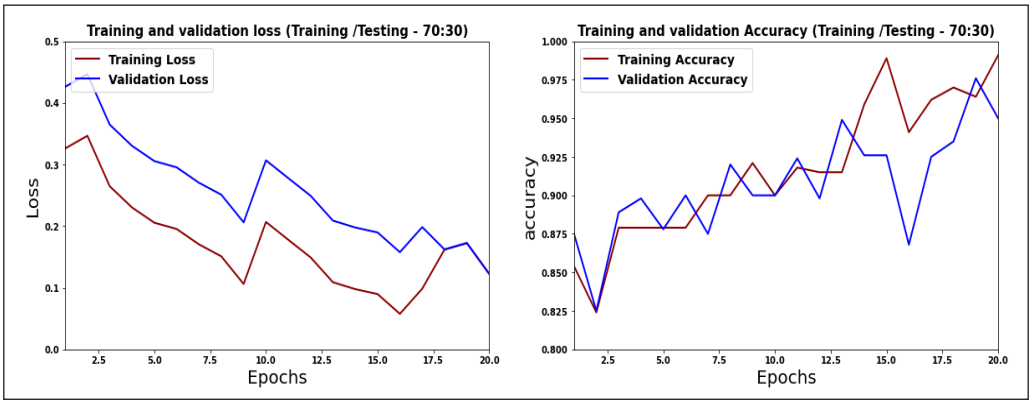



*Figure 7.* TA/TL and VA/VL analysis of the AC-GAN-ST method on the TR/TS data of 70:30

The presentation was proposed for comparative analysis in Table 5 and Figures 8, 9, 10, 11, and 12. comparisons between metrics of different models, such as the proposed AC-GAN-ST approach and current ones. models from recent research. The table has identified major presentation measures: Sensitivity, Specificity, Accuracy, RMSE, and Time of Execution. The proposed AC-GAN-ST algorithm. has a better presentation of most metrics and has achieved the highest values of Sensitivity (93.45%), Specificity (94.76%), and Accuracy (96.45%). It also shows the lowest RMSE (15.87) and the minimum execution time (1.876 seconds), both of which improve prediction accuracy and the effectiveness of predictive models relative to individuals (Kather et al., 2019; Wang et al., 2021; Yang et al., 2021). This suggests that AC-GAN-ST would be a more effective and valuable solution to the problem under study.

Table 5
*Performance measures of the proposed model*

| Research Work | Sensitivity | Specificity | Accuracy | RMSE | Execution Time |
|---|---|---|---|---|---|
| Yang et al. (2021) | 76.46 | 81.23 | 65.47 | 31.11 | 15.224 |
| Wang et al. (2021) | 80.44 | 86.23 | 80.22 | 24.55 | 7.332 |
| Shen et al. (2022) | 76.45 | 80.11 | 85.13 | 19.33 | 8.654 |
| Peng et al. (2019) | 74.56 | 87.56 | 89.11 | 17.11 | 12.117 |
| Kavitha et al. (2022) | 79.43 | 90.45 | 91.34 | 22.43 | 9.223 |
| Neelakandan et al. (2023) | 87.99 | 91.11 | 93.76 | 20.55 | 5.887 |
| Prezja et al. (2023) | 92.13 | 91.45 | 95.77 | 21.76 | 3.876 |
| Proposed AC-GAN-ST | 93.45 | 94.76 | 96.45 | 15.87 | 1.876 |

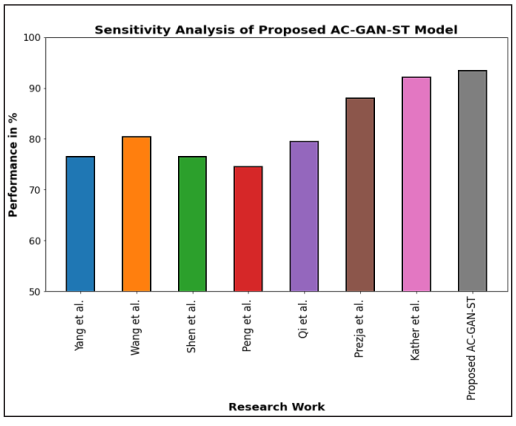



*Figure 8.* Sensitivity analysis for the proposed model

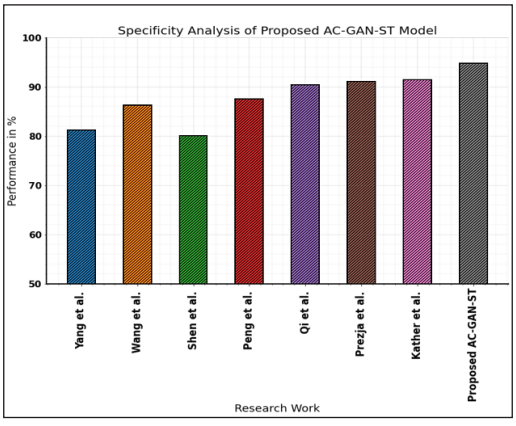



*Figure 9.* Specificity analysis for the proposed model

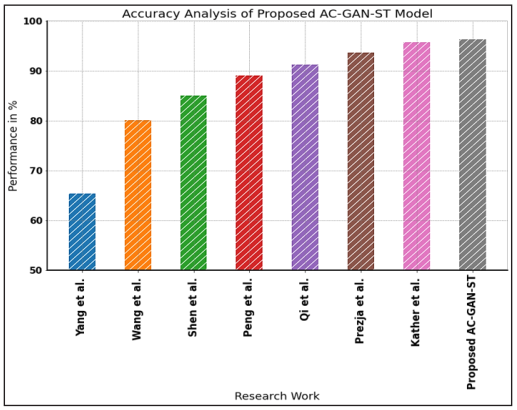



*Figure 10.* Accuracy analysis for the proposed model

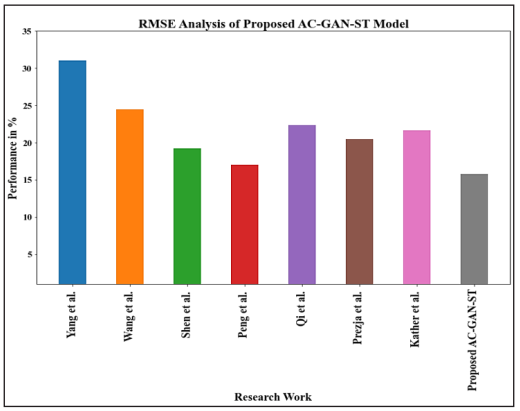



*Figure 11.* RMSE analysis for the proposed model

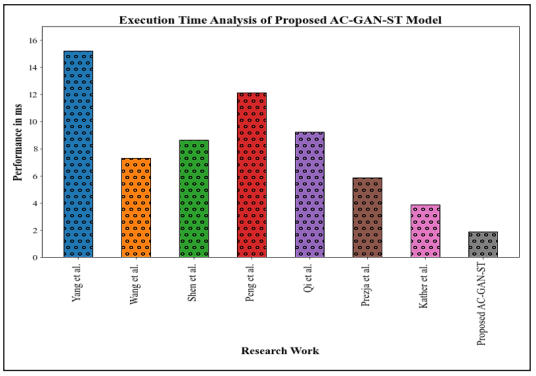



*Figure 12.* Execution Time analysis for the proposed model

## ROC Curve

The ROC curve is an important method of evaluating the efficiency of colorectal cancer detection. models work. The sensitivity-specificity trade-off between the true positive rate and the false positive rate. positive rate) by the ROC curve that relates the true positive rate with the False positive rates of different threshold values as specified in Figure 13. The accuracy of models is measured by the area under the curve. ROC curve (AUC): an AUC of approximately

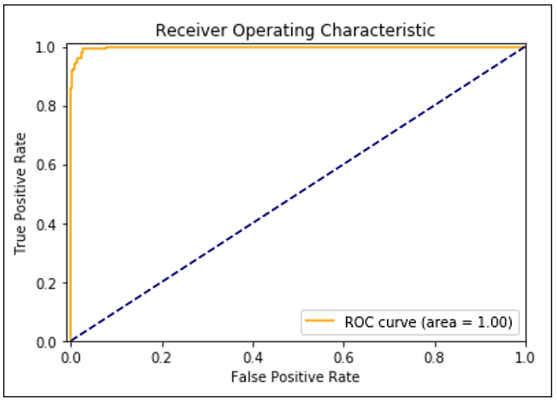



*Figure 13.* ROC curve analysis for the proposed AC-GAN-ST model

1 indicates that the procedure is highly accurate. In colorectal cancer screening, a high AUC suggests that the diagnostic instrument is useful in differentiating. between malignant and non-cancerous, in favour of early detection and evolution.

## Comparative Analysis of the Proposed Framework

Each module in the proposed model serves a significant role. In this section, we present a series of ablation studies on the Warwick-QU colorectal cancer dataset to compare the proposed AC-GAN with other models, including GAN, DNN, ANN-LR, and GRNN, to investigate performance improvements and demonstrate the motivations behind our proposed AC-GAN.

## Influence of the GAN

The AC-GAN has shown promising potential for colorectal cancer research by improving diagnostic accuracy and enabling earlier detection. Researchers can create high-quality, realistic images using the generative property of ACGAN with limited information. Simulate various colorectal cancer scenarios to help train and perfect ML models. models used for detection. This approach not only minimises the use of big data, as these are known to be. present in small amounts but similarly increase model strength by producing images that represent different stages and varieties of colorectal cancer, ultimately leading to more precise and timely results. diagnoses. Lastly, the AC-GAN+ST model gave a result. accuracy = 96.45 per cent of our input data. In the evaluation, the current studies by Yang et al. (2021) are utilised. The accuracy performances are 65.44%, 80.26%, 85.17%, 89.17%, 91.34%, 93.76%, 95.77%, and 93.67%, respectively (Peng et al., 2019; Qi et al.; Prezja et al., 2023; Shen et al., 2022; Wang et al., 2021).

## Challenges and Limitations

The Combination of Auxiliary Classifier Generative Adversarial Network (AC-GAN) with. Swin Transformers for analysing colorectal cancer have limitations. First, because of the model architecture's sophistication, additional computation can be introduced. It is not so practical in real-time applications because it requires demands. In addition, the performance of the AC-GAN is highly dependent on the quality and quantity of the training dataset, which is limited in clinical settings where annotated images are scarce. There is also a risk of overfitting, especially with small datasets, which can impair the model's generalisability to unseen data. Moreover, it is possible to analyse GAN outputs. be difficult because the generated images might not conform to clinical expectations, raising doubts about their usefulness in diagnostic workflows. Lastly, integrating modern models such as Swin Transformers can be highly education-intensive and time-consuming, limiting their

widespread use in medical institutions. An additional limitation concerns the evaluation of the generated AC-GAN samples. The present experimental protocol focused primarily on downstream classification performance and did not include dedicated synthetic-image quality measures such as Fréchet Inception Distance, Kernel Inception Distance, or an independent expert-based assessment. Consequently, the contribution of the AC-GAN should be interpreted as a training-data augmentation mechanism rather than as evidence of clinically validated realism in synthetic images. Future work should incorporate quantitative and expert-based evaluations of the generated histopathological samples, together with controlled component-wise experiments.

## CONCLUSION

Finally, the combination of ACGAN and Swin Transformers for colorectal cancer. The analysis shows great improvement in medical imaging and diagnostics. technology. This hybrid approach will utilise the strengths of both designs to improve the system. The production of high-quality synthetic images and the effective classification of cancerous lesions. The findings suggest enhanced accuracy and strength in detecting and analysing colorectal. cancer, which preconditions more reliable and efficient diagnostic tools. This approach has potential for clinical applications and adds to the growing body of studies on the applications of DL in oncology. There should be further studies on perfecting this in the future. model and discussing its possible integration with the actual workflow in the clinical setting, and the final objective of enhancing patient outcomes by detecting and treating them early.

## ACKNOWLEDGEMENT

We are very grateful to the management and senior officials of Mohan Babu University for providing research and infrastructure facilities to carry out this research work.

## REFERENCES

Echle, A., Laleh, N. G., Schrammen, P. L., West, N. P., Trautwein, C., Brinker, T. J., Gruber, S. B., Buelow, R. D., Boor, P., Grabsch, H. I., Quirke, P., & Kather, J. N. (2021). Deep learning for the detection of microsatellite instability from histology images in colorectal cancer: A systematic literature review. *ImmunoInformatics, 3-4*, Article 100008. https://doi.org/10.1016/j.immuno.2021.100008

Hasan, M. A. M., Maniruzzaman, M., & Shin, J. (2024). WGCNA and machine learning-based integrative bioinformatics analysis for identifying key genes of colorectal cancer. *IEEE Access, 12*, 144350-144363. https://doi.org/10.1109/ACCESS.2024.3472688

Kather, J. N., Krisam, J., Charoentong, P., Luedde, T., Herpel, E., Weis, C.-A., Gaiser, T., Marx, A., Valous, N. A., Ferber, D., Jansen, L., Reyes-Aldasoro, C. C., Zörnig, I., Jäger, D., Brenner, H., Chang-Claude, J., Hoffmeister, M., & Halama, N. (2019). Predicting survival from colorectal cancer histology slides using

deep learning: A retrospective multicenter study. *PLOS Medicine, 16*(1), Article e1002730. https://doi.org/10.1371/journal.pmed.1002730

Kavitha, T., Mathai, P. P., Karthikeyan, C., Ashok, M., Kohar, R., Avanija, J., & Neelakandan, S. (2022). Deep learning based capsule neural network model for breast cancer diagnosis using mammogram images. *Interdisciplinary Sciences: Computational Life Sciences, 14*(1), 113-129. https://doi.org/10.1007/s12539-021-00467-y

Khalid, M., Deivasigamani, S., V, S., & Rajendran, S. (2024). An efficient colorectal cancer detection network using atrous convolution with coordinate attention transformer and histopathological images. *Scientific Reports, 14*(1), Article 19109. https://doi.org/10.1038/s41598-024-70117-y

Khan, U., Koivukoski, S., Valkonen, M., Latonen, L., & Ruusuvuori, P. (2023). The effect of neural network architecture on virtual H&E staining: Systematic assessment of histological feasibility. *Patterns, 4*(5), Article 100725. https://doi.org/10.1016/j.patter.2023.100725

Mulenga, M., Abdul Kareem, S., Qalid Md Sabri, A., Seera, M., Govind, S., Samudi, C., & Mohamad, S. B. (2021). Feature extension of gut microbiome data for deep neural network-based colorectal cancer classification. *IEEE Access, 9*, 23565-23578. https://doi.org/10.1109/ACCESS.2021.3050838

Neelakandan, S., Rajesh, P., Mahesh, P., Karthik, P., & Mukesh, M. (2023). An efficient WBC cancer cells prediction using dense net algorithm. In *2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICAISC58445.2023.10199502

Peng, T., Boxberg, M., Weichert, W., Navab, N., & Marr, C. (2019). Multi-task learning of a deep k-nearest neighbour network for histopathological image classification and retrieval. In D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, & A. Khan (Eds.), *Medical image computing and computer assisted intervention – MICCAI 2019* (Lecture Notes in Computer Science, Vol. 11764, pp. 676-684). Springer. https://doi.org/10.1007/978-3-030-32239-7_75

Prezja, F., Annala, L., Kiiskinen, S., Lahtinen, S., & Ojala, T. (2023). *Adaptive variance thresholding: A novel approach to improve existing deep transfer vision models and advance automatic knee-joint osteoarthritis classification* [Preprint]. arXiv. https://arxiv.org/abs/2311.05799v1

Prezja, F., Äyrämö, S., Pölönen, I., Ojala, T., Lahtinen, S., Ruusuvuori, P., & Kuopio, T. (2023). Improved accuracy in colorectal cancer tissue decomposition through refinement of established deep learning solutions. *Scientific Reports, 13*(1), Article 15879. https://doi.org/10.1038/s41598-023-42357-x

Prezja, F., Paloneva, J., Pölönen, I., Niinimäki, E., & Äyrämö, S. (2022). Deepfake knee osteoarthritis x-rays from generative adversarial neural networks deceive medical experts and offer augmentation potential to automatic classification. *Scientific Reports, 12*(1), Article 18573. https://doi.org/10.1038/s41598-022-23081-4

Prezja, F., Pölönen, I., Äyrämö, S., Ruusuvuori, P., & Kuopio, T. (2022). H&E multi-laboratory staining variance exploration with machine learning. *Applied Sciences, 12*(15), Article 7511. https://doi.org/10.3390/app12157511

Ragab, M., Mahmoud, M. M., Asseri, A. H., Choudhry, H., & Yacoub, H. A. (2023). Optimal deep transfer learning based colorectal cancer detection and classification model. *Computers, Materials & Continua, 74*(2), 3279-3295. https://doi.org/10.32604/cmc.2023.031037

Ruusuvuori, P., Valkonen, M., & Latonen, L. (2023). Deep learning transforms colorectal cancer biomarker prediction from histopathology images. *Cancer Cell, 41*(9), 1543-1545. https://doi.org/10.1016/j.ccell.2023.08.006

Sarwinda, D., Paradisa, R. H., Bustamam, A., & Anggia, P. (2021). Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. *Procedia Computer Science, 179*, 423-431. https://doi.org/10.1016/j.procs.2021.01.025

Sasmal, P., Sharma, V., Prakash, A. J., Bhuyan, M.-K., Patro, K. K., Samee, N. A., Alamro, H., Iwahori, Y., Tadeusiewicz, R., Acharya, U. R., & Pławiak, P. (2024). Semi-supervised generative adversarial networks for improved colorectal polyp classification using histopathological images. *Information Sciences, 658*, Article 120033. https://doi.org/10.1016/j.ins.2023.120033

Shen, Y., Luo, Y., Shen, D., & Ke, J. (2022). RandStainNA: Learning stain-agnostic features from histology slides by bridging stain augmentation and normalization. In L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, & S. Li (Eds.), *Medical image computing and computer assisted intervention – MICCAI 2022* (Lecture Notes in Computer Science, Vol. 13432, pp. 212-221). Springer. https://doi.org/10.1007/978-3-031-16434-7_21

Sirinukunwattana, K., Domingo, E., Richman, S. D., Redmond, K. L., Blake, A., Verrill, C., Leedham, S. J., Chatzipli, A., Hardy, C., Whalley, C. M., Wu, C.-H., Beggs, A. D., McDermott, U., Dunne, P. D., Meade, A., Walker, S. M., Murray, G. I., Samuel, L., Seymour, M., … Koelzer, V. H. (2021). Image-based consensus molecular subtype (imCMS) classification of colorectal cancer using deep learning. *Gut, 70*(3), 544-554. https://doi.org/10.1136/gutjnl-2019-319866

Skrede, O.-J., De Raedt, S., Kleppe, A., Hveem, T. S., Liestøl, K., Maddison, J., Askautrud, H. A., Pradhan, M., Nesheim, J. A., Albregtsen, F., Farstad, I. N., Domingo, E., Church, D. N., Nesbakken, A., Shepherd, N. A., Tomlinson, I., Kerr, R., Novelli, M., Kerr, D. J., & Danielsen, H. E. (2020). Deep learning for prediction of colorectal cancer outcome: A discovery and validation study. *The Lancet, 395*(10221), 350-360. https://doi.org/10.1016/S0140-6736(19)32998-8

Sujatha, R., Mahalakshmi, K., & Yoosuf, M. S. (2024). Colorectal cancer prediction via histopathology segmentation using DC-GAN and VAE-GAN. *EAI Endorsed Transactions on Pervasive Health and Technology, 10*. https://doi.org/10.4108/eetpht.10.5395

Sun, W., Zhao, R., Zhang, K., Gao, J., & Qu, G. (2024). Distribution-aware loss for lesions detection using white-light endoscopy in colorectal region. *IEEE Access, 12*, 46138-46148. https://doi.org/10.1109/ACCESS.2024.3381614

Wagner, S. J., Reisenbüchler, D., West, N. P., Niehues, J. M., Zhu, J., Foersch, S., Veldhuizen, G. P., Quirke, P., Grabsch, H. I., van den Brandt, P. A., Hutchins, G. G.-A., Richman, S. D., Yuan, T., Langer, R., Jenniskens, J. C.-A., Offermans, K., Mueller, W., Gray, R., Gruber, S. B., … Kather, J. N. (2023). Transformer-based biomarker prediction from colorectal cancer histology: A large-scale multicentric study. *Cancer Cell, 41*(9), 1650-1661.e4. https://doi.org/10.1016/j.ccell.2023.08.002

Wang, K. S., Yu, G., Xu, C., Meng, X. H., Zhou, J., Zheng, C., Deng, Z., Shang, L., Liu, R., Su, S., Zhou, X., Li, Q., Li, J., Wang, J., Ma, K., Qi, J., Hu, Z., Tang, P., Deng, J., … Deng, H. W. (2021). Accurate diagnosis of colorectal cancer based on histopathology images using artificial intelligence. *BMC Medicine, 19*(1), Article 76. https://doi.org/10.1186/s12916-021-01942-5

Yamashita, R., Long, J., Longacre, T., Peng, L., Berry, G., Martin, B., Higgins, J., Rubin, D. L., & Shen, J. (2021). Deep learning model for the prediction of microsatellite instability in colorectal cancer: A diagnostic study. *The Lancet Oncology, 22*(1), 132-141. https://doi.org/10.1016/S1470-2045(20)30535-0

Yang, J., Shi, R., & Ni, B. (2021). MedMNIST classification decathlon: A lightweight AutoML benchmark for medical image analysis. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)* (pp. 191-195). IEEE. https://doi.org/10.1109/ISBI48211.2021.9434062

Zhang, F., Jiang, S., & Li, Y. (2020). Application of massive parallel deep learning algorithm in the prediction of colorectal carcinogenesis of familial polyposis. *IEEE Access, 8*, 129432-129445. https://doi.org/10.1109/ACCESS.2020.3009412

Zidan, U., Gaber, M. M., & Abdelsamea, M. M. (2023). SwinCup: Cascaded Swin transformer for histopathological structures segmentation in colorectal cancer. *Expert Systems with Applications, 216*, Article 119452. https://doi.org/10.1016/j.eswa.2022.119452