Original Article

DOI: 10.7860/JCDR/2026/87356.23274

Internal Medicine Section

Automated Detection of Acute Appendicitis from Contrast-enhanced CT Images using CNN with GAN-based Data Augmentation: A Retrospective Observational Study

G CHINTHANA[1], S APPAVU ALIAS BALAMURUGAN[2], AKSHAYA SARAVANAN[3], S MOHAMMED FAIZAL[4], S ARULMOZHIVARMAN[5]

ABSTRACT

Introduction: Acute appendicitis is a common cause of acute abdominal pain, which should be appropriately diagnosed, so that the subsequent perforation, or unjustified surgeries, may be prevented. The use of Artificial Intelligence (AI) powered structures can combine clinical and imaging data to enable faster, more precise detection.

Aim: To develop an automated system for detecting acute appendicitis using Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GAN)- based data augmentation.

Materials and Methods: A retrospective observational study has been conducted at the General Medicine Unit, Vinodhagan Memorial Hospital, Thanjavur, Tamil Nadu, India from January 2022 to June 2025, to analyse clinical data from 500 patients, of whom 100 underwent Contrast-enhanced Computed Tomography (CECT) scans according to clinical indications. The training input included 70 original CT images, which were enhanced using a GAN to produce 140 synthetic images, for a total of 210 training images. The validation set (10 images) and Independent test set (20 images) consisted solely of original CT-acquired images to avoid bias in evaluating the model. A Feedforward Neural Network (FNN) was used to process the clinical symptoms (body temperature, abdominal pain, nausea, and appetite loss) and provide the probability of appendicitis. The CT Images of high probability patients were collected for further analysis. The analysis of CT images was processed using a three-dimensional Residual Network (3D ResNet)-based CNN. GANs were used to generate synthetic images to improve model robustness through data augmentation. The system was evaluated based on accuracy, sensitivity, specificity, and F1 score, and the results were compared with surgical outcomes.

Results: The image-based CNN with GAN augmentation produced the best diagnostic results with an accuracy of 90%, precision of 0.91, recall of 0.91, F1 score of 0.91, specificity of 0.89, and AUC-ROC of 0.90. The FNN model, which uses symptoms as the independent variable, achieved an accuracy of 80% and an AUC-ROC of 0.81. The CNN baseline with no augmentation got the 85% accuracy and an AUC-ROC of 0.85.

Conclusion: GAN-based augmentation improves model generalisation on small datasets. The model has great potential for clinical implementation, thereby preventing misdiagnosis and unnecessary surgeries.

Keywords: Augmented classification, Clinical decision support system, Convolutional neural network, Computed tomography, Generative adversarial networks, Medical imaging, Machine Learning

INTRODUCTION

One of the most frequent causes of acute abdominal pain that needs emergency surgery is acute appendicitis. Otherwise, it may cause severe complications like perforation, abscess, and peritonitis, which are life-threatening unless diagnosed and treated promptly. Despite being a relatively widespread medical condition, appendicitis is a condition that is difficult to diagnose as the clinical manifestations of the condition are often confused with other diseases such as gastrointestinal disorders, urinary tract infections, and gynaecological diseases. Such overlap often leads to diagnostic and treatment delays, especially in emergency medical facilities [1].

The diagnosis of acute appendicitis is traditionally based on the integration of clinical examination, laboratory investigations, and medical imaging. Clinicians normally check symptoms like abdominal pain, nausea, vomiting, fever, and leukocytosis, which can help determine whether it is probable to have appendicitis or not. Imaging modalities, including ultrasound and Computed Tomography (CT) scans, are then performed to confirm the diagnosis. Of these methods, CECT imaging is widely regarded as the most effective technique for diagnosing appendicitis because it offers detailed visualisation of the appendix and the surrounding tissues. Nonetheless, interpreting CT images can be difficult and requires substantial radiological experience and may be influenced by workload, varying work experience among radiologists, and individual anatomy [1].

Since fast, precise diagnosis in an emergency clinical setting requires an immediate response, the demand for automated diagnostic systems that help medical practitioners detect appendicitis more efficiently and reliably is rising. Recent progress in Artificial Intelligence (AI) and Deep Learning (DL) has shown that these technologies can be applied to medical imaging to yield meaningful results [2].

In particular, CNNs and other DL approaches have been widely used in medical image analysis, including disease detection, segmentation, and classification. CNN-based models have been trained to discover hierarchical representations of imaging data, making them describe the presence of subtle abnormalities in more complex medical images. Although there are many applications, recent research has shown that DL models can be more effective than expert radiologists on specific diagnostic tasks [2].

The proposed work uses a deep CNN architecture, named 3DResNet, to automatically identify appendicitis in CECT scans. A three-dimensional CNN architecture is used to enable the model to learn spatial relations among multiple CT slices, thereby enabling more accurate analysis of anatomical structures in volumetric medical imaging scans [2].

Besides an analysis based on using images, clinical symptoms are also significant in the process of early diagnosis of appendicitis. As such, the proposed work combines a symptom-based prediction system with FNNs to further increase diagnostic accuracy. There are clinical symptoms, including abdominal pain, nausea, fever, and other corresponding symptoms, that were analysed through the FNN model to calculate the likelihood of a patient having appendicitis. The system initially scores the symptoms a user reports and generates a probability score indicating how likely the person is to have appendicitis. When the predicted probability exceeds a set threshold, the user is asked to post images from a CT scan, which will be analysed further. This two-step diagnostic methodology is useful for prioritising high-risk cases, thereby directing imaging resources to the most likely patients and improving diagnostic efficiency in any given healthcare setting [3].

The other major issue with designing AI-based medical diagnostic systems is the lack of large, well-labelled medical imaging datasets. The difficulty posed by access to training datasets that are large enough to comprehensively cover privacy regulations and be dealt with ethically, and the prohibitive cost of expert annotation, is likely to restrict access to suitable large training data. The scarcity of data can reduce the potential of DL models to generalise and increase overfitting. To address this issue, this work employs GANs for data augmentation [4].

The GAN models consist of two rival neural networks, one being a generator and the other being a discriminator that are co-trained and adversarially trained. The generator aims to produce synthetic images similar to real CT scans, while the discriminator aims to distinguish real from fake images. In this antagonistic training method, GANs can produce highly realistic synthetic medical images that closely match the properties of real clinical images. Images generated by these are then utilised to enlarge the training data, improving the robustness of the DL models and their generalisation ability. GAN-based augmentation is beneficial because it helps the model learn a wider range of anatomical variations [4].

The main objective of this work is to develop and test a DL- based diagnostic system that can help medical workers to effectively diagnose appendicitis through CECT scans and patient symptom data. The effectiveness of the proposed solution is evaluated using conventional classification metrics, including accuracy, precision, recall, F1 Score, and AUC. To investigate this hypothesis, the study focuses on the following research questions:

- How effective is the 3DResNet model in detecting appendicitis from CT images?
- Does model generalisation perform better when GAN-generated images are added as well?
- How does the performance of the proposed system compare with that of radiologists and conventional diagnostic methods?

MATERIALS AND METHODS

A retrospective observational study was conducted at the General Medicine Unit, Vinodhagan Memorial Hospital, Thanjavur, Tamil Nadu, India, from January 2022 to June 2025. The study received the ethical clearance of the Institutional Ethics Committee (VMH/TN/IEC/01/2025). Informed consent was waived by the Institutional Ethics Committee (Approval No: VMH/TN/IEC/01/2025) due to retrospective design and anonymised data usage.

The dataset used in the present study consists of two main types of data: clinical symptoms such as body temperature, nausea, loss of appetite, and pain location, and CT scan images from patients diagnosed with appendicitis. The clinical symptom data provide vital signs that help diagnose appendicitis, while the CECT images are utilised to train the CNN to visually detect appendicitis.

Sample size: A retrospective study of clinical data from 500 patients, including 100 who underwent CECT scans per clinical indications, was conducted. From these cases, 70 original CT images were selected to train a model. The small dataset is also a limitation, as we used GANs to augment the data, generating 140 synthetic CT images in total, which were included in the training set. To aid model testing, the validation set of 10 original CT images were used and the independent test set of 20. The validation and testing did not involve synthetic images to avoid evaluation bias. Thus, the imaging dataset used for model development comprised 210 training images, with separate validation and test sets (only real CT images).

Inclusion and Exclusion criteria: The inclusion criteria included patients aged 18-60 years who had CECT imaging due to suspected appendicitis and had full clinical records. As exclusion criteria, individuals older than 60 years, those with a previous appendectomy, pregnant individuals, and those with missing records were excluded.

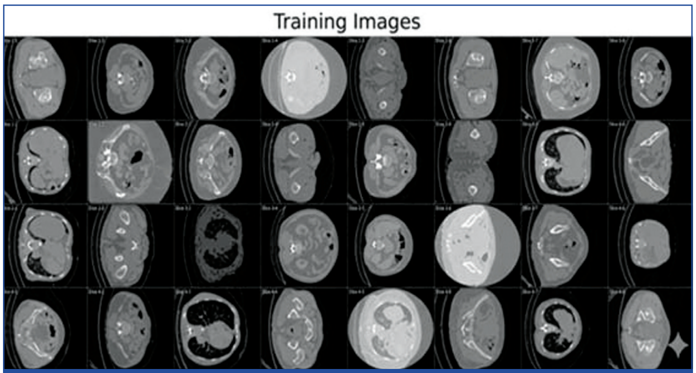
Study Procedure

A retrospective observational study using previously recorded clinical and imaging data collected between January 2022 and June 2025. The cases were categorised as cases of suspected appendicitis according to the clinical and imaging data and the cases of confirmed appendicitis according to the pathology reports after surgery. The patients who were diagnosed with other causes of acute abdomen, like gastroenteritis, diverticulitis, or ovarian cysts, were kept under negative control.

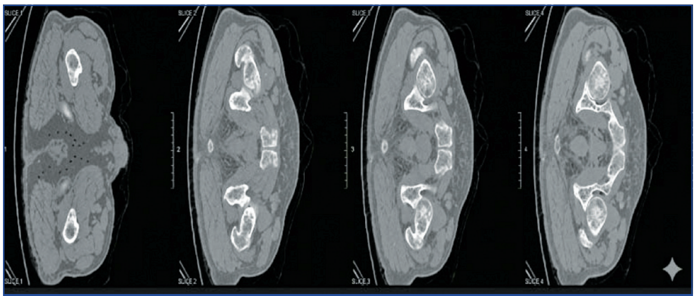
Data preprocessing and augmentation with generative adversarial networks (GANs): The present work uses both a GAN and a CNN to automatically detect acute appendicitis based on CECT. A Deep Convolutional Generative Adversarial Network (DCGAN) architecture was adopted to generate realistic synthetic CT images that supplement the original dataset and address class imbalance. The CNN comprised numerous convolutional and pooling layers, followed by final fully connected layers to categorise the images as either appendicitis or non appendicitis.

Initial processing: A GAN was used in the study to generate synthetic images to supplement the dataset used to classify acute appendicitis with the CNN. No anomaly detection methodology was employed, and any mention of latent space-based anomaly detection has been removed to avoid confusion. The generator consists of four transposed convolutional blocks, using ReLU and batch normalisation, with a 100-dimensional noise image as input and an output of 256x256 grayscale CT images. The discriminator was trained for 200 epochs with a batch size of 32 using the Adam optimiser (learning rate = 0.0002, $\beta1 = 0.5$). Both the discriminator and the generator were trained on binary cross-entropy loss. The convergences were tracked by monitoring the generator and discriminative model curves until convergence. The quality of synthetic images was assessed using the Fréchet Inception Distance (FID), which was low (25), indicating that the GAN-generated images resemble real CECT images and can be effectively used to supplement the dataset during model training. Further, the anatomical plausibility and pathological realism of a random sampling of 50 synthetic images were verified by a board-certified radiologist and were found to be consistent with known cases of appendicitis. [Table/Fig-1,2] show high-resolution real and artificial CT images; all images are anonymised to preserve patient confidentiality.

Synthetic data generation using GANs: The GANs were used to produce high-quality synthetic CT images to overcome data scarcity and increase the model's diversity of pathological features. It was a technique that helped not only increases the number of entries in the database but also add comprehensive depictions of rare medical conditions, which are frequently underrepresented in publicly available datasets.



[Table/Fig-1]: Actual CT scan images from the dataset illustrate various patient conditions and imaging situations that appeared during training.



[Table/Fig-2]: Synthetic CT images produced by the GAN, showcasing the model's ability to generate detailed and clinically relevant imaging data.

The GAN Set-up with a Dual-component Architecture

- **Generator:** The training process was gradually developed to the point where the generator can create synthetic CT images very similar to those produced by real clinical CT scans. To accomplish this, the generator performed this step by continuously optimising its parameters by minimising its loss function.
- **Generator Loss:**

$$L_G = -\log(D(G(z)))$$

The mathematical formulation demonstrates how the generator optimises the discriminator's classification error by creating increasingly realistic synthesised images.

- **Discriminator:** The discriminator serves as the authentic image verifier and provides feedback to the generator, thereby enhancing generator performance. In this pursuit, the discriminator has to establish the vision and accuracy in real image identification as well as generated image classification and to evaluate its performance, we use this loss function:
- **Discriminator loss:**

$$L_D = -\{\log(D(x)) + \log(1-D(G(z)))\}$$

In this evaluation process, the discriminator determines two probabilities: the probability that the real image x is authentic, D(x), and the probability that an image generated by G(z) is authentic, D(G(z)). The algorithm optimises the discriminator function to correctly recognise between real images and synthetic images and provides important feedback to the generator.

Model development:

- **FNN for symptom-based prediction:** The FNN had three hidden layers with 64, 32, and 16 neurons, using ReLU activations, and a sigmoid output layer, with the output probability indicating the presence of appendicitis. Input variables were based on the standardised clinical variables: the pain description, duration, nausea, vomiting, fever, WBC, and C-reactive protein. The median (continuous) and mode (categorical) imputation were used to fill in missing values. It was a feature selection in which the variables were highly collinear. The FNN was trained on a dataset of 500 patients (70% training, 15% validation, 15% test), utilising backpropagation, the Adam optimiser, and binary cross-entropy loss.
- **3D CNN for CT image classification:** Volumetric CT classification was performed using a 3D ResNet-18 (R3D18) network. The input volumes were formed by running 64 successive 2D slices (reduced to 256x256 pixels) of the input volume (interpolated to 1 mm slice thickness). Preprocessing was done with pixel normalisation and contrast enhancement. The network was used to do binary classification (appendicitis vs. non appendicitis).
- **Integration and scoring system:** The results of the FNN and 3D CNN were integrated to produce an appendicitis probability score. The average of the two probabilities yielded a final likelihood score used to determine the classification, with a threshold of 0.5.

The processing of custom-developed machine learning algorithms that use both symptom data and CT scans as core inputs is also included in the methodology. The output side of the FNN is trained on clinical symptom data to produce likely appendicitis outcomes. In the network, several hidden-layer neurons extract features from the data in successive layers until they produce a final output: the predicted probability.

Training is performed on the FNN using backpropagation and gradient descent to maximise prediction accuracy. It carries out its training process by making simultaneous weight adjustments and reducing the prediction errors step by step. The weight update rule influences the gradient descent method to update to the weights using the gradient descent update.

- **Weight update:** $W = W - \eta \cdot \frac{\partial E}{\partial W}$

Where, η is the learning rate, and $\frac{\partial E}{\partial W}$ denotes the gradient of the loss function E with respect to the weights W. Repeating this model training procedure helps the system learn more from the training examples, improving its appendicitis predictions with each round.

The CNN were trained to predict which patterns or features of appendicitis are present in the images. For example, the dimensions of the data are reduced using a series of convolutional and pooling layers, which progressively reduce spatial dimensionality through pooling operations. The CNN were trained on CT scan images, and after training, it performed binary classification to determine whether the scan showed evidence of appendicitis. To improve the robustness of the CNN, the CT scan dataset was enhanced with GANs. The CNN was trained on a combination of real and synthetic data, with synthetic images generated to help it learn more generalised features, thereby improving its ability to detect appendicitis across diverse CT images.

Uncertainty quantification using bayesian CNN: Uncertainty quantification is crucial in medical imaging, as it helps assess the reliability of model predictions, particularly when the consequences of errors are severe. A Bayesian CNN offers a principled way to quantify uncertainty by modelling the distribution over weights and predictions, rather than providing point estimates.

Approach

- **Bayesian CNN model:** A Bayesian CNN is a variation of traditional CNNs, where the network weights are treated as distributions rather than fixed values. This allows the network to produce a distribution of outputs rather than a single deterministic output [5].
- **Uncertainty estimation:** Using techniques such as Monte Carlo dropout or variational Inference, we approximate the posterior distribution over the model weights. The variance in predictions can then be interpreted as the uncertainty in the model's decision [5].
- **Evaluation and validation:** The CNN and FNN were trained with a batch size of 32 to balance computation time and steady gradient updates. The maximum number of epochs was 200,

and early stopping occurred when 10 consecutive epochs without improving the validation loss were reached. Data on the symptoms and CT images were split into 70% for training, 15% for validation, and 15% for testing. Cross-validation was performed on the FNN at a five-fold level to ensure generalisation. GAN-based synthetic image augmentation was applied alongside classical data augmentation, i.e., random rotations (±15 degrees), horizontal/vertical, and scaling, in the case of CT images (±.10%)

The data comprised 100 patients who had undergone CECT, with 70 in the training group, 10 in the validation group, and 20 in the Independent test set. The test set was independent, comprising 20 patients (20% of the total dataset). Of these, 11 were positive for appendicitis, and nine were diagnosed with no appendicitis, according to the surgical results. No data on test patients have been utilised in the training and validation of the model.

An independent dataset of 40 patients collected in Vinodhagan Memorial Hospital, Thanjavur, was used in external validation and was not included in model development. The external dataset performance scores were Accuracy=88%, Precision=0.88, Recall=0.87, F1 Score=0.87 and AUC=0.89. This confirms the model generalisation across hospitals.

The performance of the proposed model was compared with that of two board-certified radiologists, each with over five years of work experience, who were also asked to independently assess the same CT test dataset. The radiologists obtained a diagnostic accuracy of 82% and 80%, respectively, and the external validation dataset achieved an accuracy of 88%, whereas the primary test dataset showed an accuracy of 90%.

The AI model was compared with the standard clinical Alvarado scoring system on an independent test set of 20 patients. The Alvarado score had an AUC of 0.71, whereas the integrated AI model had an AUC of 0.90, indicating better predictive performance.

The DeLong ROC comparison test showed that the AI model was significantly more effective than Radiologist 1 (p=0.014) and Radiologist 2 (p=0.032), and more effective than the clinical scoring system (p=0.048). The p-values that were considered significant were below 0.05. Cohen's kappa (0.82) was high, indicating strong inter-rater reliability between the two radiologists.

The CNN+GAN model on 20 test cases gave one false positive, which was a patient with diverticulitis or an ovarian cyst that appeared as appendicitis on imaging, and one false negative, including a patient who had early-stage appendicitis with mild signs of inflammation.

Ablation study: Contribution of each module: The model was tested on varying configurations in the Ablation Study to test the effect of each component:

Ablation variants:

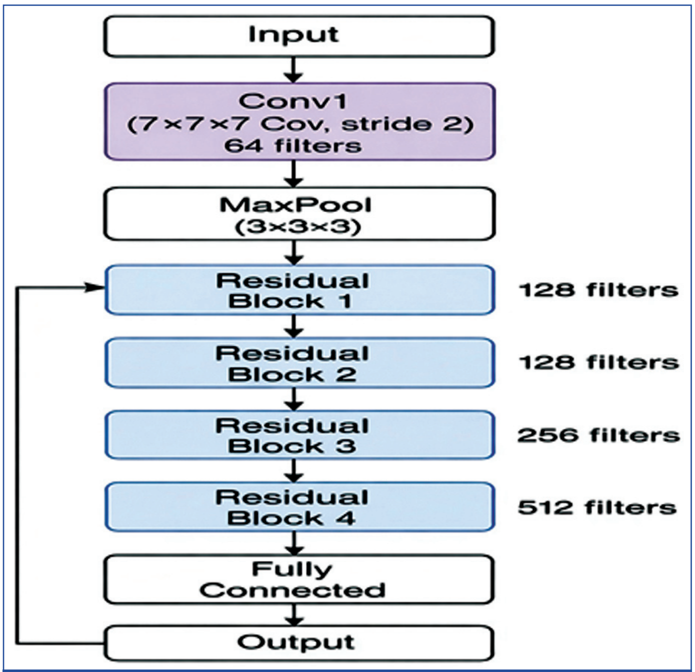
- **Without augmentation:** The CNN is trained on the initial dataset without data augmentation, using the GAN.
- **With augmentation**, the CNN is trained on a GAN-based dataset of images that produce synthetic images similar to real CT images of appendicitis.
- **Model architecture:** The model uses a modified version of 3D ResNet-18 (R3D18) or 3D MC3-18 (MC3D18), selected dynamically during initialisation. The architecture includes 3D convolutions in the ResNet or MC3 model, with a fully connected layer for binary classification (Appendicitis vs. Non-Appendicitis). The step-by-step system architecture of the ResNet-18 model are shown in [Table/Fig-3].

Training parameters:

Loss function: Binary Cross-Entropy with Logits

Optimiser: AdamW optimiser with a learning rate of 0.001.

Changed



[Table/Fig-3]: System architecture of ResNet-18.

Batch size: 32

Epochs: The model is trained for 200 epochs, with the option to resume from previous checkpoints if needed.

Additional details:

Validation: Validation accuracy is calculated at the end of each epoch.

Model saving: Model weights are saved periodically after each epoch to prevent loss of progress.

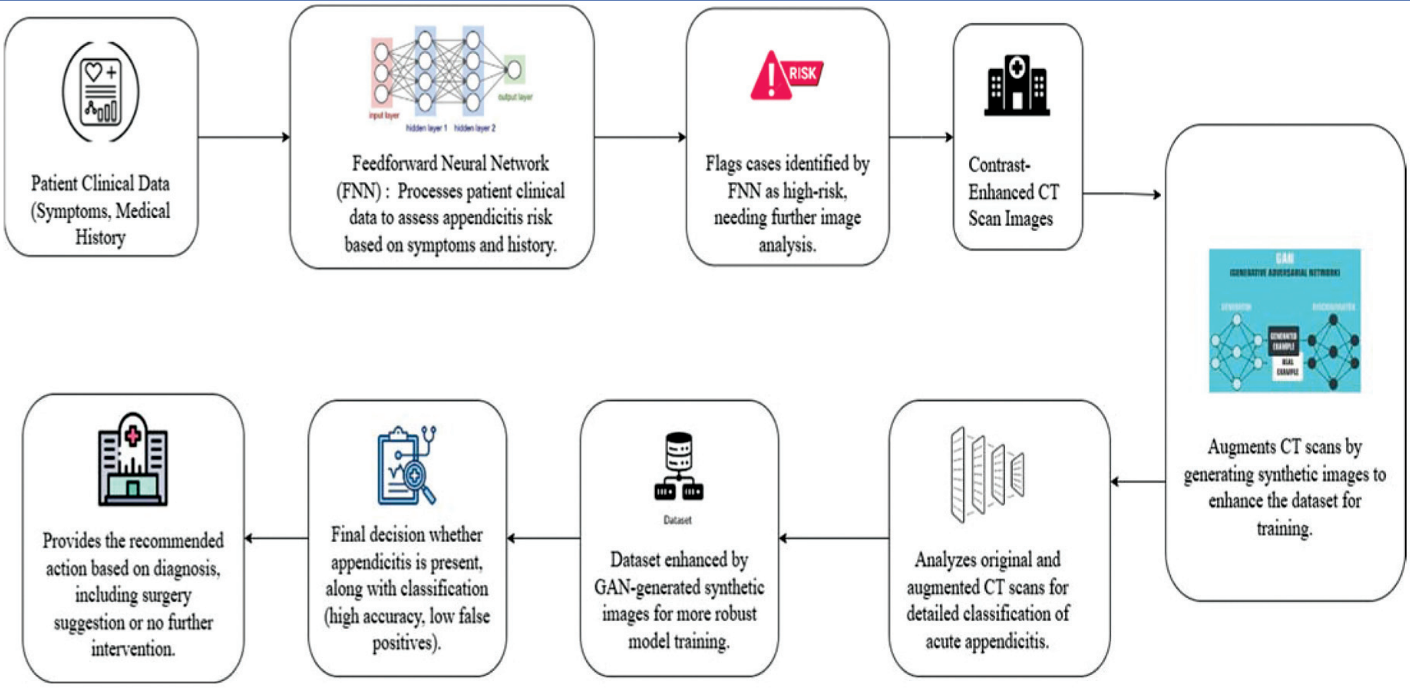
Memory monitoring: Memory usage is tracked and displayed during training.

Training environment: The model was trained on a GPU (specifically, an MPS GPU on a Mac), which accelerated training. The total training time per epoch was monitored and optimised to improve efficiency. The various hyperparameters used for the ResNet-18 model are given in [Table/Fig-4].

Hyperparameter	Value	Description
Model Architecture	Modified 3D ResNet-18 (R3D18)	A modified version of ResNet-18 with 3D convolutions was used for appendicitis detection.
Optimiser	AdamW	AdamW optimiser used with weight decay for regularisation.
Learning rate	0.001	The learning rate for the AdamW optimiser.
Batch size	32	The batch size used during training.
Epochs	200	The number of epochs for training.
Loss function	Binary Cross-Entropy with Logits	Loss function used for binary classification (Appendicitis vs non-appendicitis).
Dropout rate	0.5	Dropout rate applied in the fully connected layers to prevent overfitting (if used).
Weight decay	1e-5	L2 regularisation term (weight decay) to prevent overfitting.
Activation function	ReLU	ReLU activation function used in convolutional layers and residual blocks.
Validation accuracy	Calculated during training	Accuracy is calculated during each validation phase and printed after every epoch.

[Table/Fig-4]: Hyperparameter tuning.

The proposed AI-based framework for diagnosing appendicitis using clinical data and improved CT images is presented in [Table/Fig-5]. The framework involves patient clinical data, along with improved



[Table/Fig-5]: Proposed system architecture.

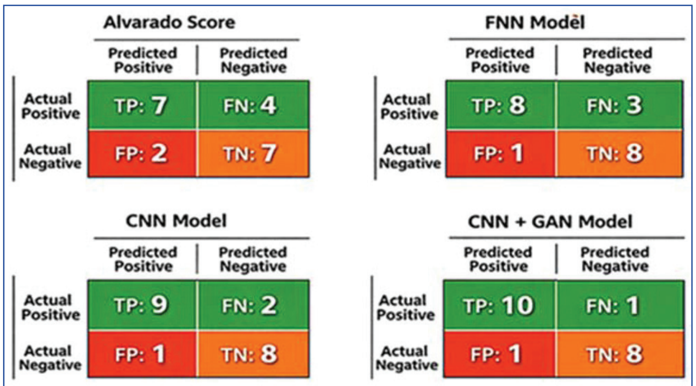
CT, to enable automatic diagnosis of appendicitis. In a FNN, a high-risk case would be evaluated using clinical data, and such cases could be excluded or additional imaging performed. The CT scans with contrast enhancement, which are organised alongside GAN-generated synthetic images, are subject to detailed analysis to improve the model's robustness. The system ends up giving a diagnostic recommendation, which in turn assists clinicians in making a decision on whether to perform surgery or non surgical treatment.

STATISTICAL ANALYSIS

To compare the predictive ability of the Traditional Clinical Scoring (Alvarado), FNN, CNN, and CNN+GAN models for differentiating between positive and negative cases of appendicitis, performance metrics such as accuracy, precision, recall, F1 Score, and AUC-ROC were analysed. The training dataset was used to train the model, and an independent test set of 20 patients was used to evaluate it. The standard measures used to evaluate model performance included accuracy, sensitivity, specificity, and Area Under the ROC Curve (AUC).

RESULTS

The sample size used was 100 patients undergoing CECT scans. The overall performance of the GAN-augmented CNN was the best, and the values are: Accuracy: 90%, Sensitivity: 0.91, Specificity: 0.89, AUC-ROC: 0.90. The values of the confusion matrix (TP, FP, FN, TN) of each of the models on the test set, which provides a detailed picture of how well each model was working is presented in [Table/Fig-6].



[Table/Fig-6]: Detailed confusion matrix for all models on the test dataset.

The performance of different models were compared in [Table/Fig-7]. It compares clinical and imaging-based diagnostic procedures for appendicitis. Patient symptoms and laboratory data were used to create clinical models, such as the Alvarado Score and a symptom-based FNN. The CECT images were trained on imaging models, such as a CNN and a CNN augmented with a GAN. A matched subset of patients with clinical and imaging data was utilised, and all models were tested on the same independent test set, ensuring an impartial and balanced comparison. GAN-based data augmentation increased recall from 0.82 to 0.91 and reduced the number of false-negative cases in the test group from 2 to 1. This is a significant reduction, which is clinically significant because the missed diagnosis of appendicitis can result in perforation and greater morbidity. False negatives (n=1) in the GAN-augmented CNN model were related to cases of appendicitis in an early stage that showed slight inflammatory alterations. False positives (n=1) were mostly cases in which the patients had the inflammatory bowel disorder, in which CT imaging made it appear to be appendicitis.

Model	Accuracy	Precision	Recall	F1 Score	Specificity	AUC-ROC
Traditional Clinical Scoring (Alvarado)	70%	0.78	0.64	0.70	0.78	0.71
FNN	80%	0.89	0.73	0.80	0.89	0.81
Baseline CNN	85%	0.90	0.82	0.86	0.89	0.85
CNN + GAN Augmentation	90%	0.91	0.91	0.91	0.89	0.90

[Table/Fig-7]: Performance comparison of different models.

There is a clear performance gradient, and CNN+GAN proves to be the most resilient model, achieving the best balance across all evaluation metrics. The implications of these findings are that, in terms of diagnostic accuracy, DL models are more effective than traditional models.

The data-specific modalities of the symptom-based FNN and the image-based CNN, focusing on the role of clinical and imaging data in diagnostic accuracy are compared in [Table/Fig-8]. The FNN model, which was built on symptoms, performed remarkably well with only structured clinical data, achieving an AUC of 0.81. Nonetheless, the CNN based on images showed better sensitivity and discriminative power overall (AUC=0.85) for the definite diagnosis of appendicitis.

Model type	Accuracy	Precision	Recall	F1 score	Specificity	AUC-ROC
Symptom-Based FNN	80%	0.89	0.73	0.80	0.89	0.81
Image-Based CNN	85%	0.90	0.82	0.86	0.89	0.85

[Table/Fig-8]: Clinical (FNN) image-based (CNN) model comparison.

Training loss curve are shown for the first 30 epochs (early convergence phase) out of total of 200 epochs [Table/Fig-9]. The training loss consistently decreases, indicating that the model is improving at fitting the training data. However, the validation loss shows some fluctuations, reflecting the model's generalisation capability. Over time, the gap between training and validation losses narrows, suggesting that the model is beginning to generalise better to unseen data, though some minor overfitting can be observed in later epochs. This behaviour is typical in DL models and suggests the need for potential regularisation or early stopping techniques to further optimise model performance.

[Table/Fig-9]: Training vs validation loss.

DISCUSSION

The results of the current research confirm the significance of computer-assisted analysis of CECT images for identifying acute appendicitis. The findings also show that combining a CNN with GAN data augmentation improves diagnostic quality and enhances the detection model's resilience. A comparison of previous studies on AI applications in medical imaging and appendicitis prediction are provided in [Table/Fig-10] [1-19].

In previous studies, it was noted that clinical and image-based diagnostic methods, including CT and ultrasound, play a critical role in detecting appendicitis and its related complications [1,9]. A number of studies have examined machine learning models, such as ANN, logistic regression, and XGBoost, for predicting appendicitis and its severity [10,13,14,17]. Also, DL and CNN-based models have been used in CT imaging to automatically detect appendicitis [5], and [12], with improved diagnostic performance.

In addition to improving diagnostic accuracy, this hybrid approach enables healthcare professionals to make decisions that can help avoid misdiagnosis and needlessly performing surgeries. Together,

S. No.	Author and year	Method/Model	Dataset/Data Type	Performance/Outcome	Key limitation
1	Moris D et al., 2021 [1]	CNN-based detection	CECT images	Improved appendicitis detection accuracy	Limited dataset size affects generalisation.
2	Park JJ et al., 2020 [2]	CT imaging analysis	CT scan data	Improved diagnosis of appendicitis using CT imaging.	No automated AI mode
3	Schipper A et al., 2024 [3]	Machine learning prediction model	Emergency department patient data	Diagnostic accuracy comparable to that of physicians.	Limited imaging features for appendicitis detection
4	Yang Y et al., 2024 [4]	Organ-aware GAN	Medical imaging datasets	Generated anatomically realistic synthetic images for appendicitis	Focus mainly on image synthesis
5	Takaishi T et al., 2025 [5]	Deep Learning (DL) with 3D bounding boxes	CT scans	Automated appendicitis localisation	Requires high-resolution CT data
6	Ye Z et al., 2022 [6]	CNN-based Deep Learning (DL)	CT scans	Improved detection of contrast-sensitive regions	Requires large annotated datasets
7	Kim SW et al., 2021 [7]	Deep learning (DL) contrast synthesis	CT abdominal scans	Generated synthetic contrast-enhanced images	No anatomical constraint modeling
8	Issaiy M et al, 2023 [8]	AI diagnostic models	Clinical datasets	Improved prediction of appendicitis outcomes	Limited imaging integration
9	Terasawa T et al., 2004 [9]	CT and ultrasound review	Imaging studies	Evaluated imaging methods for appendicitis detection	Pre-AI diagnostic techniques
10	Chen S et al., 2025 [10]	XGBoost predictive model	Clinical datasets	Accurate prediction of complicated appendicitis	Imaging data not used
11	Maleš I et al., 2025 [11]	AI applications review	Multiple medical datasets	AI models useful in triage and diagnosis of appendicitis	Limited practical deployment
12	Dogan K et al., 2024 [12]	Deep Learning (DL) model	CT imaging datasets	Automated appendicitis diagnosis	Requires a large training dataset
13	Wei W et al., 2024 [13]	Machine learning model	Clinical patient data	Predicted complicated appendicitis risk	No CT imaging integration
14	Shahmoradi L et al., 2019 [14]	ANN vs logistic regression	Clinical datasets	ANN achieved better prediction accuracy	Limited dataset
15	Park SY et al., 2015 [15]	Artificial Neural Network	Clinical data	Developed early ANN diagnostic system	Limited computational power
16	Bianchi V et al., 2024 [16]	AI-based decision support	Clinical datasets	AI proposed for diagnosis and treatment planning	Conceptual model
17	Lin HA et al., 2023 [17]	Multilayer perceptron	Clinical datasets	Classified complicated vs uncomplicated appendicitis	Limited dataset size
18	Hamilton A, 2024 [18]	AI healthcare simulation	Medical education datasets	Demonstrated AI use in training environments	Not diagnostic-focused
19	Li J et al., 2025 [19]	ML + CT imaging review	Radiology studies	Summarised ML progress in appendicitis diagnosis.	Identified challenges like data scarcity

[Table/Fig-10]: Comparison of previous studies on AI applications in medical imaging and appendicitis prediction [1-19].

the system provides a complete and robust solution for early-stage detection of appendicitis by combining clinical data with the latest advances in image analysis. This work suggests that future work could introduce real-time data processing and cloud-based storage for CT scan images, thereby increasing system mobility and scalability.

Limitation(s)

This retrospective single-centre design limits generalisability and introduces potential selection bias. Although GAN-based augmentation was employed, synthetic images may not fully capture true biological variability and could introduce artefacts, increasing the risk of overfitting.

CONCLUSION(S)

The proposed appendicitis detection system utilises a dual approach integrating symptom-based prediction and image-based analysis, and has demonstrated better diagnostic accuracy. The system can offer accurate, timely predictions for appendicitis scenarios using FNN for symptom analysis and CNN for CT scan image classification. The system is further strengthened by data augmentation with GANs, enabling efficient analysis of volumetric CT data using multiple slices per patient through a 3D CNN framework, allowing the CNN to generalise across different imaging conditions and patient demographics.

REFERENCES

[1] Moris D, Paulson EK, Pappas TN. Diagnosis and management of acute appendicitis in adults: A review. JAMA. 2021;326(22):2299-311. Doi: 10.1001/jama.2021.20502.

[2] Park JJ, Kim KA, Nam Y, Choi MH, Choi SY, Rhie J. Convolutional-neural-network-based diagnosis of appendicitis via CT scans in patients with acute abdominal pain presenting in the emergency department. Sci Rep. 2020;10(1):9556. Doi: 10.1038/s41598-020-66674-7.

[3] Schipper A, Belgers P, O'Connor R, Jie KE, Dooijes R, Bosma JS, et al. Machine-learning based prediction of appendicitis for patients presenting with acute abdominal pain at the emergency department. World J Emerg Surg. 2024;19(1):40. Doi: 10.1186/s13017-024-00570-7. PMID: 39716296; PMCID: PMC11664873.

[4] Yang Y, Liu J, Zhan G, Chen Q, Wang F, Li Y, et al. OA-GAN: Organ-aware generative adversarial network for synthesizing contrast-enhanced medical images. Biomed Phys Eng Express. 2024;10(3):035012.

[5] Takaishi T, Kawai T, Kokubo Y, Fujinaga T, Ojio Y, Yamamoto T et al. Deep learning for appendicitis: Development of a three-dimensional localization model on CT. Jpn J Radiol. 2025;43(11):1859-69. Doi: 10.1007/s11604-025-01834-1. Epub 2025 Jul 16. PMID: 40668351; PMCID: PMC12575542.

[6] Ye Z, Qian JM, Hosny A, Zeleznik R, Plana D, Likitlersuang J, et al. Deep learning-based detection of intravenous contrast enhancement on CT scans. Radiol Artif Intell. 2022;4(3):e210285. Doi: 10.1148/ryai.210285. PMID: 35652117; PMCID: PMC9152686.

[7] Kim SW, Kim JH, Kwak S, Seo M, Ryoo C, Shin CI, et al. The feasibility of deep learning-based synthetic contrast-enhanced CT from nonenhanced CT in emergency department patients with acute abdominal pain. Sci Rep. 2021;11(1):20390. Doi: 10.1038/s41598-021-99896-4.

[8] Issaiy M, Zarei D, Saghazadeh A. Artificial intelligence and acute appendicitis: A systematic review of diagnostic and prognostic models. World Journal of Emergency Surgery. 2023;18(1):59.

[9] Terasawa T, Blackmore CC, Bent S, Kohlwes RJ. Systematic review: Computed tomography and ultrasonography to detect acute appendicitis in adults and adolescents. Ann Intern Med. 2004;141(7):537-46.

[10] Chen S, Xia J, Xu B, Huang Y, Teng M, Pan J. Risk prediction and effect evaluation of complicated appendicitis based on XGBoost modeling. BMC Gastroenterol. 2025;25:295.

[11] Maleš I, Kumrić M, Huić Maleš A, Cvitković I, Šantić R, Pogorelić Z, et al. A Systematic integration of artificial intelligence models in appendicitis management: A comprehensive review. Diagnostics (Basel). 2025;15:866.

[12] Dogan K, Selcuk T. A novel deep learning approach for the automatic diagnosis of acute appendicitis. Journal of Clinical Medicine. 2024;13:4949.

[13] Wei W, Tongping S, Jiaming W. Construction of a clinical prediction model for complicated appendicitis based on machine learning techniques. Scientific Reports. 2024;14:16473.

[14] Shahmoradi L, Safdari R, Mirhosseini MM, Arji G, Jannat B, Abdar M. Predicting risk of acute appendicitis. A comparison of artificial neural network and logistic regression models. Acta Med Iran. 2019;56:784-95.

[15] Park SY, Kim SM. Acute appendicitis diagnosis using artificial neural networks. Technol Health Care. 2015;23 Suppl 2:S559-S565.

[16] Bianchi V, Giambusso M, De Iacob A, Chiarello MM, Brisinda G. Artificial intelligence in the diagnosis and treatment of acute appendicitis: A narrative review. Updates Surg. 2024;76:783-92.

[17] Lin HA, Lin LT, Lin SF. Application of Artificial Neural Network Models to Differentiate Between Complicated and Uncomplicated Acute Appendicitis. J Med Syst. 2023;47:38.

[18] Hamilton A. Artificial intelligence and healthcare simulation: The shifting landscape of medical education. Cureus. 2024;16:e59747.

[19] Li J, Ye J, Luo Y, Xu T, Jia Z. Progress in the application of machine learning in CT diagnosis of acute appendicitis. Abdom Radiol (NY). 2025;50(9):4040-49.

PARTICULARS OF CONTRIBUTORS:

1. Assistant Professor, Department of Pharmacology, Thanjavur Medical College, Thanjavur, Tamil Nadu, India.
2. Professor, Department of Computer Science and Engineering, Periyar Maniammai Institute of Science and Technology, Thanjavur, Tamil Nadu, India.
3. Undergraduate Student, Department of Computer Science and Engineering, Periyar Maniammai Institute of Science and Technology, Thanjavur, Tamil Nadu, India.
4. Undergraduate Student, Department of Computer Science and Engineering, Periyar Maniammai Institute of Science and Technology, Thanjavur, Tamil Nadu, India.
5. Undergraduate Student, Department of Computer Science and Engineering, Periyar Maniammai Institute of Science and Technology, Thanjavur, Tamil Nadu, India.

NAME, ADDRESS, E-MAIL ID OF THE CORRESPONDING AUTHOR:

Dr. G Chinthana,
Assistant Professor, Department of Pharmacology, Thanjavur Medical College,
Thanjavur-613004, Tamil Nadu, India.
E-mail: chinbala2006@gmail.com

PLAGIARISM CHECKING METHODS: [Jain H et al.]

- Plagiarism X-checker: Jan 12, 2026
- Manual Googling: Mar 07, 2026
- iThenticate Software: Mar 11, 2026 (1%)

ETYMOLOGY: Author Origin

EMENDATIONS: 8

AUTHOR DECLARATION:

- Financial or Other Competing Interests: None
- Was Ethics Committee Approval obtained for this study? Yes
- Was informed consent obtained from the subjects involved in the study? No
- For any images presented appropriate consent has been obtained from the subjects. NA

Date of Submission: **Jan 11, 2026**
Date of Peer Review: **Feb 25, 2026**
Date of Acceptance: **Mar 15, 2026**
Date of Publishing: **May 01, 2026**