Journal of Information Systems Engineering and Management
2025, 10(16s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Deriving Efficient 3D U-Net Based Segmented Anomaly Detection and Classification in 3D MRI Images Using ConvLSTM Model and Shuffled Frog Leaping Algorithm

J Hima Bindu[1,*], M Uma devi[2]

[1]*Research Scholar, Dept. of Computer Science and Engineering, Vignan's Foundation for Science, Technology & Research, Vadlamudi, Guntur-522213, Andhra Pradesh.*
[2]*Dept. of Computer Science and Engineering, Vignan's Foundation for Science, Technology & Research, Vadlamudi, Guntur-522213, Andhra Pradesh.*
**Corresponding Author: J Hima Bindu (kothabindu0034@gmail.com)*

ARTICLE INFO

Received: 02 Dec 2024
Revised: 25 Jan 2025
Accepted: 05 Feb 2025

ABSTRACT

Anomaly detection relates to the difficulty of detecting anomalous inputs from seen samples of standard data. Despite recent advances in deep learning for detecting visual anomalies, these algorithms are still incapable of deciphering complex images like those encountered in the medical business. Medical imaging is essential for diagnosing and treating a vast array of conditions, including anomalies detected in 3D Magnetic Resonance Imaging (MRI) data. The dependable and accurate detection and classification of anomalies within volumetric MRI scans remain challenging for the medical industry. This study presents a novel method for combining three potent techniques: 3D U-Net for segmentation, the Convolutional Long Short-Term Memory (ConvLSTM) model for temporal processing, and the Shuffled Frog Leaping Algorithm (SFLA) for optimization. Utilizing the 3D U-Net architecture, the proposed method efficiently segments regions of interest in 3D MRI volumes, allowing for precise anomaly localization. The ConvLSTM model incorporates temporal dependencies between successive MRI slices, enhancing the accuracy of detecting and classifying dynamic anomalies. Combining the SFLA as a metaheuristic optimization instrument significantly improves the efficiency and effectiveness of the proposed framework. The SFLA optimizes the network's hyperparameters, improving convergence and lowering the danger of training being trapped in local optima. In order to calculate the Anomaly Class Weight, the output layer neurons are built to estimate various Feature Distribution Similarity values for various characteristics. Extensive experiments were conducted on a large dataset of 3D MRI scans with various defects to assess the efficacy of the proposed technique. According to the results, the proposed method outperforms standard anomaly detection and categorization procedures. The method accomplishes cutting-edge precision, sensitivity, and specificity, surpassing existing approaches by a wide margin.

Keywords: Deep learning, Anomaly detection, 3D MRI, ConvLSTM, SFLA, 3D U-Net.

1. Introduction

The main goal of anomaly detection is to locate data samples that do not fit the overall data distribution [1]. Anomalies can appear for several reasons, such as noise in the data-capture technique, changes in the underlying phenomenon, or a result of new or unusual circumstances in the environment being recorded. Therefore, one of the most essential tasks in interpreting medical images is identifying anomalies [2]. Understanding "typical" data samples improves the capacity to identify "unusual" ones, a critical component in the application of deep learning models [3]. Predicting the model's behavior becomes complicated when an input considerably deviates from the training data [4,5]. This trait is vital for high-stakes applications, such as medical decision support systems, where the capacity to recognize aberrant data is crucial. A key use of anomaly detection is the detection of anomalous events. For instance, identifying various microscopy datasets is time- and money-consuming in pathology due to the different variations of cells and tissues that require expert understanding. Medical image analysis is essential for identifying and comprehending a

variety of disorders. MRI stands out as a potent tool for obtaining precise three-dimensional (3D) [6] information on anatomical structures among the many imaging modalities available. Medical diagnosis and treatment planning are impossible without using MRI, enabling professionals to analyze inside organs and tissues without causing harm. However, deciphering 3D MRI data can be difficult, particularly when identifying and categorizing image anomalies or abnormalities.

The complex spatial information in 3D volumes is frequently difficult to handle using conventional image processing techniques. Advanced computational models that can automatically detect and classify anomalies with high accuracy and efficiency are thus becoming more and more necessary. Medical image segmentation, which needs extracting the pixels of organs or lesions from background medical images like CT or MRI scans to offer critical information on the shapes and sizes of these organs, is one of the most challenging tasks in medical image analysis [32]. Many scholars have proposed a variety of technologically advanced automatic segmentation techniques. Earlier systems were constructed using conventional techniques like edge detection filters and mathematical techniques. This work proposes a novel technique for automatically segmenting tumorous tissue areas in medical images utilizing a multimodal 3D U-Net.Making the operations of Conv-LSTM [7] expands this capability to spatial data. When describing changes in visual patterns over time, such as in video categorization and gesture recognition [8], Conv-LSTM has proven to be quite successful. We propose using Conv-LSTM to "scan" through an imaging volume for the presence of disease without the need for expert annotations of diseased regions, as opposed to extracting spatiotemporal patterns from time series image data. We use the SFLA to enhance performance and fine-tune the model's parameters. Our 3D U-Net and ConvLSTM-based anomaly detection and classification system are based on frogs' social behavior. SFLA [9] is a meta-heuristic optimization algorithm that effectively searches the solution space to discover the ideal parameter configuration. This enables us to improve generalization and convergence on unknown data.

The major contributions of the paper as follows,

- The suggested method effectively divides regions of interest into 3D MRI volumes using the 3D U-Net architecture, enabling accurate anomaly localization.
- The ConvLSTM model improves the accuracy of identifying and categorizing dynamic anomalies by incorporating temporal relationships between subsequent MRI slices.
- The suggested framework is substantially more efficient and effective when using the SFLA as a metaheuristic optimization tool. The network's hyperparameters are adjusted by the SFLA, which boosts convergence and lowers the chance that training will become trapped in local optima.
- In order to calculate the Anomaly Class Weight, the output layer neurons are built to estimate various Feature Distribution Similarity values for various characteristics.
- Several tests were run on an extensive dataset of 3D MRI images containing various anomalies to see how practical the suggested approach was.

The article's format places a detailed introduction to anomaly detection in Section I and illustrations from the medical field. Section 2 goes to great length about all relevant research and approaches available in the literature. Section 3 covers the operation of the suggested anomaly detection in medical images. The experimental findings are described in Section 4, together with a thorough analysis. Finally, Section 5 presents a conclusion for the entire paper.

2. Literature Review

Asif et al., [10] ADAM optimizer model for classifying brain images uses many pre-trained models like NasNet Large, Xception, DenseNet121, etc., to extract the features. The features are retrieved and utilized to train an ADAM optimizer on the CNN model to achieve classification. Vidyarthi et al. [11] present a cumulative variance-based feature selection approach to categorize different grades of malignant brain tumors. The technique extracts the features, uses Cumulative Variance Method (CVM) to select the best features, and then classifies them with KNN, NN, and multiclass SVM. Using X-ray-based deep learning algorithms, Rahman et al. [12] published a thorough survey for the automatic detection of COVID-19. The authors examined the performance of popular deep algorithms, the diversity of the datasets, and image processing techniques. With an overall accuracy of 87.3%, In order to find pneumothoraxes in chest X-rays, Park et al. [13] used a 26-layer YOLO model. Tolkachev et al. [14] use of UNet and several classification networks for pneumo-thorax segmentation is another well-known example

of their work. They conducted their investigation using the most current Kaggle challenge dataset and came up with a Dice coefficient of 0.8574. Table 1 shows the list of abbreviations and acronyms

Zhao et al. [15] showed how context data from symmetric regions may effectively improve the feature representation of anomaly candidate regions. This augmentation method enhanced performance for detecting chest anomalies, especially in conditions with little supervision. However, this method's placement accuracy needed to be improved, and the earlier box creation process slowed down the model's reasoning speed. The heuristic red fox optimization technique, which Antoni et al. [16] presented for application in medical image processing, allowed for the unsupervised selection of the optimal threshold and demonstrated promising results for lung image segmentation and detection.

In their study, Marimont et al. [17] employed a PixelSNAIL and explored various restorations by adjusting the sampling temperature and assigning appropriate weights to them. Meanwhile, Esser et al. [18] found that PixelSNAIL outperforms autoregressive transformers in simulating the probability density estimate of the data based on recent advancements in computer vision research. To achieve a more refined reconstruction, Jun et al. [19] emphasized the importance of accurately assessing the latent variable probability, which informs the selection of values to be resampled for higher-quality reconstruction.

Choromanski et al. [20] offer a technique for unsupervised anomaly identification and segmentation using transformers to solve these issues. Performers are used to learn the distribution of brain imaging data in this method. This study expands on the specifics of the experiments we carried out in Pinaya et al. [21], where we developed and assessed a robust technique and compared it to contemporary advanced unsupervised algorithms for performance on synthetic and natural datasets. In addition, we assess how well our approach performs in segmenting and detecting anomalies in 3D brain data.

Sekhar et al. [22] present the GoogleNet transfer learning model for categorizing distinct tumor classifications. The model performs classification using SVM, KNN, and SoftMax algorithms and feature extraction using a pre-trained CNN model. In Kujur et al. [23], which considers models like S-CNN, an in-depth examination of several CNN models is drawn. The technique compares the effectiveness of alternative strategies using two sets of brain image data. However, Meissen et al. [24] have demonstrated that a straightforward baseline can outperform most autoencoder-based UAD techniques. They show that simple thresholding-based algorithms can detect hyperintense brain tumors and multiple sclerosis lesions more effectively than most UAD methods that rely on healthy data for training. They merely perform histogram equalization preprocessing and use the FLAIR modality.

2.1 Limitations of Existing System

- Finding anomalies in medical images might have severe ethical repercussions. Potential false positives or false negatives in the system could result in inaccurate diagnoses or pointless treatments.
- Expanding the current system to recognize a wider variety of abnormalities might be challenging because it may only be intended to identify particular illnesses or anomalies.
- Various things, including patient mobility, equipment constraints, and other technical problems, can cause noise and artifacts in 3D MRI images. These artifacts may prevent abnormalities from being identified and categorized.
- Unbalanced datasets, in which the quantity of standard samples far outweighs that of aberrant ones, frequently make anomaly detection in medical imaging difficult. This imbalance may produce biased results and subpar classification performance for rare abnormalities.
- The current approach may rely on expensive and time-consuming expert annotations for training. Complex 3D MRI scans require specialized annotation methods, which can induce inter-observer variability.

2.2 Problem Identification

- The current technology may have trouble accurately spotting irregularities in 3D MRI medical images. Medical picture anomalies can be subtle and difficult to spot, resulting in false negatives or false positives that severely influence patient diagnosis and treatment planning.

- The current approach cannot generalize across various MRI datasets or modalities. This lack of generalization could restrict its usefulness and necessitate time-consuming retraining when used on fresh datasets.
- For MRI images to be of higher quality and to have noise artifacts eliminated, preprocessing is essential. Possible areas for improvement in the current system's preprocessing methods could produce subpar results for anomaly detection.
- A crucial step in the anomaly detection process is segmentation. The current system may use old or ineffective segmentation techniques that need to precisely identify regions of interest, which reduces the accuracy of subsequent anomaly identification.

Table 1. List of abbreviations and acronyms

Abbreviation	Description	Abbreviation	Description
MRI	Magnetic Resonance Imaging	KNN	K-Nearest Neighbours Algorithm
ConvLSTM	Convolutional Long Short-Term Memory	NN	Neural Network
SFLA	Shuffled Frog Leaping Algorithm	SVM	Support vector machine
CT	Computerized tomography	YOLO	You only look once
ADAM	Adaptive Moment Estimation	S-CNN	Supervised Convolutional Neural Network
CNN	Convolutional Neural Network	UAD	Unsupervised Anomaly Detection
CVM	Cumulative Variance Method	FLAIR	Fluid-Attenuated Inversion Recovery
SegAE	Segmentation Auto-Encoder	GAN	Generative Adversarial Network
cuDNN	CUDA Deep Neural Network		

3. Proposed Methodology

In this section, we look into a novel strategy that combines these three powerful tools. The recommended technique successfully splits regions of interest into 3D MRI volumes by exploiting the 3D U-Net architecture, enabling precise anomaly localization. The ConvLSTM model incorporates temporal dependencies between successive MRI slices, enhancing the accuracy of detecting and classifying dynamic anomalies. Combining the SFLA as a metaheuristic optimization instrument significantly improves the efficiency .The SFLA fine-tunes the network's hyperparameters, improving convergence and reducing the risk of training being stuck in local optima. Figure 1 shows the block diagram of the ConvLSTM method for Anomaly detection in medical Images.

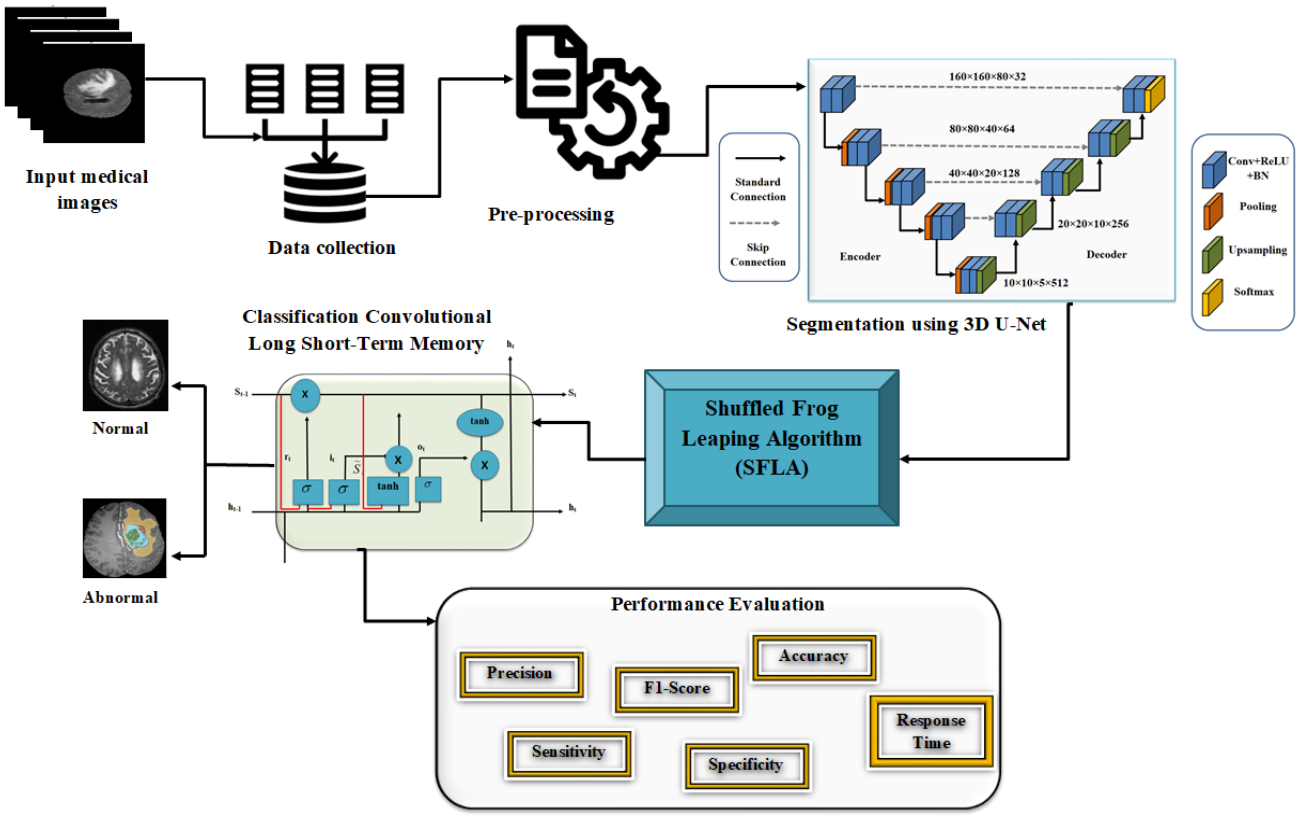

Figure 1. Block diagram of the ConvLSTM method for anomaly detection in medical images

3.1 Image Pre-processing

The medical image under consideration has been read to enhance the image's quality. The method begins by setting up a normalisation window of size x, which encompasses K hops for any given pixel. The standard deviation value is measured in each of the window's four additional sections. Second, the area with the lowest pixel standard deviation is chosen. The method modifies the pixel value to normalize the pixel based on the estimated Mean Standard Value (MSV) derived from selected region pixels. This will be repeated in iterations for the specified number of hops, k. Segmentation has been done using this normalized image.

3.2 3 D U-Net-based Segmented

A pipeline for segmenting medical images must include the neural network's design and related hyperparameters. For semantic segmentation, a range of deep learning architectures are available, each fusing performance, resilience, and efficiency in a particular way. The U-Net is a preferred and promising option [25-26], featuring an agreeable trade-off between performance and variability. To minimize needless parameter inflation brought on by more sophisticated designs, we chose the classic 3D U-Net as the essential building element of our technique [27]. We chose a design that could support a patch with the measurements 160x160x80 with a single channel of normalized Hounsfield units (Hus). Our selected architecture's last layer generated a mask with four channels and a size of 160x160x80. These probabilities were normalized using the SoftMax function to provide a balanced exponential correction. We used transposed convolution to handle down-sampling and maximum pooling to manage up-sampling. At its lowest resolution, the architecture displayed 512 feature maps, and at its most significant level, 32 feature maps. Except for the convolutions related to up-sampling and down-sampling, which used a kernel size of 2x2x2 with a stride of 2x2x2, all convolutional operations used a kernel size of 3x3x3 with a stride of 1x1x1. To stabilize the network's training procedure, batch normalization was successively applied after each convolutional block. Refer to Figure 2 for a representation of the architecture's structure.

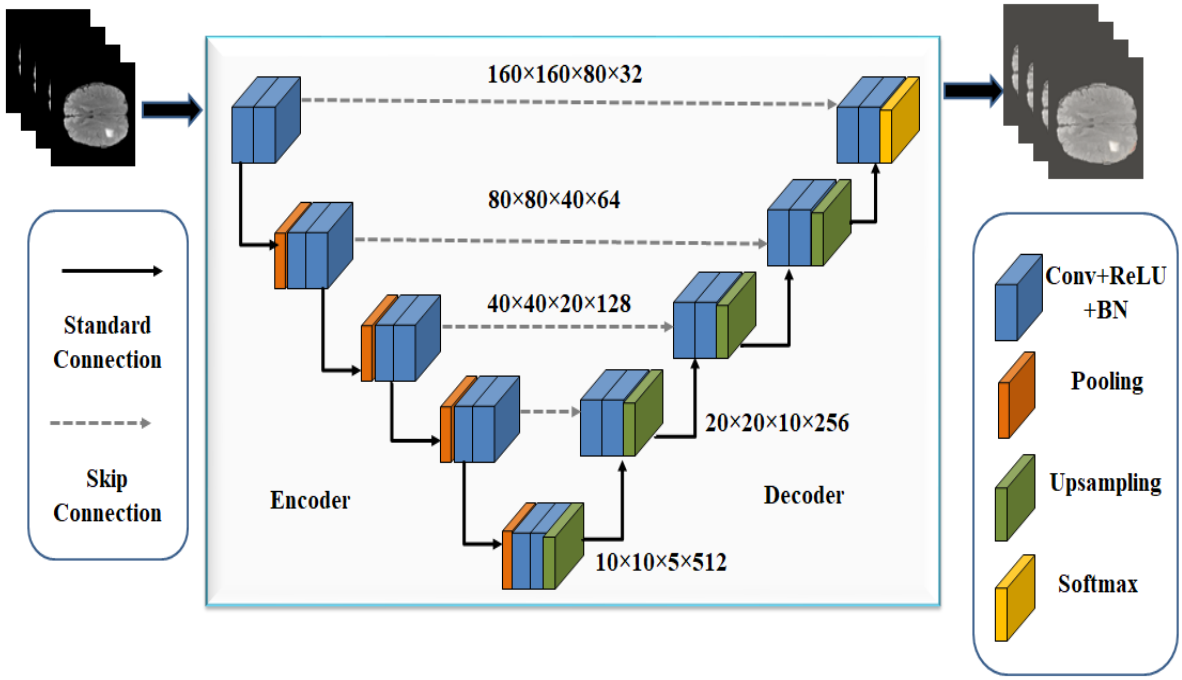

Figure 2. Standard architecture of 3D U-Net

When segmenting medical images, a noteworthy bias emerges in the distribution of classes throughout the semantic annotation process, frequently distributed towards the background class. The Tversky index [28] was combined with categorical cross-entropy to incorporate this innate tendency, resulting in the composite loss function used throughout the model training procedure (1).

$$Loss_{total} = Loss_{Tversky} + Loss_{CCE} \tag{1}$$

$$Loss_{Tversky} = N - \sum_{c=1}^{N} \frac{TP_c}{TP_c + \alpha.FN_c + \beta.FP_c} \tag{2}$$

$$Loss_{CCE} = -\sum_{c=1}^{N} Y_{o,c} \log(p_{o,c}) \tag{3}$$

The Tversky index (2), an asymmetrical similarity metric used to compare the segmented region with the ground truth, was altered by our team. These changes improve the capacity to balance false positive and false negative rates by considering different classifications. In machine learning, the cross-entropy loss function—often written as (3)—is popular. It gauges the entropy between these distributions by quantifying the overall difference between the real and expected distributions. We use the binary indicator $x_{o,c}$ to assess the precision of class label c about observation o. Additionally, we use the categorical cross-entropy, meant by the term "multi-class adjustment," to accommodate several categories. The variable represents the estimated likelihood that observation o belongs to class $p_{o,c}$.

We started using an initial weight decay of 0.001 and the Adam optimization technique to speed up the model fitting process. After 15 consecutive epochs, if the training loss had not decreased, we implemented a dynamic response by 0.1-fold, reducing the learning rate. A 1e-5 learning rate was established as the minimum. We used the early ending training strategy, which caused the training process to cease after 100 epochs without a fitting loss decline, to reduce the possibility of overfitting. A maximum of 1000 epochs could be used for training the neural network model. In this context, an epoch is defined as the iteration through 150 training batches as opposed to the conventional notion of an epoch being a single run through the complete dataset.

The benefits of 3D U-Net for anomaly detection are as follows:

- Compared to manually created feature-based techniques, the 3D U-Net's deep architecture can automatically learn complicated patterns and distinctive features from volumetric data.
- The U-Net design improves anomaly segmentation accuracy by incorporating multi-scale features throughout the up-sampling process.
- The 3D U-Net can process various volumetric data types frequently found in industrial and medical applications since it can accommodate multiple input sizes.

3.3 Classification using ConvLSTM

The suggested method successfully captures spatial information and temporal dependencies inside the 3D MRI image sequences using ConvLSTM's spatiotemporal features. To extract geographical data and analyse temporal dependencies, the architecture combines LSTM-based layers with convolutional layers. The system learns to discriminate between regular and irregular patterns by training on a sizable dataset of annotated 3D MRI images.

Combining the advantages of a CNN with an LSTM, a unique subclass of LSTMs called ConvLSTMs [29] was developed to model medical images data accurately. Like the LSTM, the ConvLSTM can retain or discard information from the previous cell state while it updates the current cell state. The ConvLSTM accomplishes these transitions from input to state and state through convolutional architectures instead of the LSTM's reliance on internal matrix multiplications. ConvLSTM cells can keep the data's original dimensions as they pass through, preventing it from being compressed into a one-dimensional feature vector, because to this convolutional method. As a result, the ConvLSTM is skilled at processing collections of input images that are either two-dimensional or three-dimensional. Equations (4) through (9) describe the ConvLSTM processes as follows:

$$i_t = \sigma(\mathrm{W}_{Si} \circ S_{t-1} + \mathrm{W}_{hi} * h_{t-1} + \mathrm{W}_{xi} * x_t + b_i) \quad (4)$$

$$r_t = \sigma(\mathrm{W}_{Sr} \circ S_{t-1} + \mathrm{W}_{hr} * h_{t-1} + \mathrm{W}_{xr} * x_t + b_r) \quad (5)$$

$$o_t = \sigma(\mathrm{W}_{So} \circ S_t + \mathrm{W}_{ho} * h_{t-1} + \mathrm{W}_{xo} * x_t + b_o) \quad (6)$$

$$S_t = \tanh(\mathrm{W}_{hS} * h_{t-1} + \mathrm{W}_{xS} * x_t + b_S) \quad (7)$$

$$S_t = f_t \circ S_{t-1} + (1 - f_t) \circ S_t \quad (8)$$

$$h_t = o_t \circ \tanh(S_t) \quad (9)$$

Where I stands for the input gate, r for the forget gate, o for the output gate, S_t is the candidate cell state, S is the cell state, h is the hidden state, b is the bias, and $\mathrm{W}_{Si}, \mathrm{W}_{hi}, \mathrm{W}_{Sr}, \mathrm{W}_{hr}, \mathrm{W}_{xr}, \mathrm{W}_{So}, \mathrm{W}_{ho}, \mathrm{W}_{xo}, \mathrm{W}_{hS}, \mathrm{W}_{xS} \in \square^{n \times T}$ is a logistic sigmoid W function, respectively. Figure 3 represent the additional connections between a ConvLSTM cell and an LSTM cell. Both the current and prior cell stacks are where these connections come from.

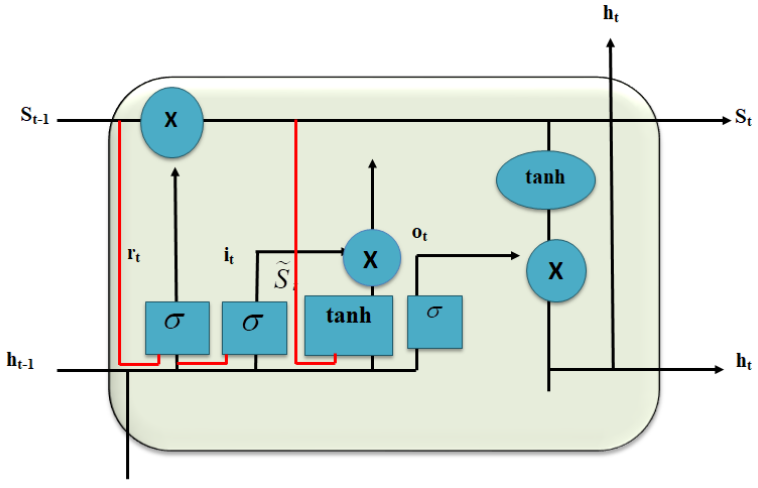

Figure 3. Architecture for ConvLSTM memory

3.4 SFLA

By incorporating knowledge from frogs' social interactions, the SFLA imitates natural processes. Using sophisticated optimization techniques, this method enables exploration through parameter space to quickly find the best solutions. The technique seeks to iteratively optimize the selection of features and model parameters to differentiate accurately between normal and abnormal image regions in the context of anomaly detection and classification in 3D MRI images.

Eusuff and Lansey invented the SFLA in 2006 [30]. The evolutionary process of SFLA combines a focused exploitation mechanism with a thorough exploration mechanism, much like other algorithms in the memetic lineage. A randomly generated population is first produced, subjected to analysis, and then sorted to organize it. Individuals inside each memeplex compete to hold the highest position locally or worldwide. The entire population is reorganized after a definite amount of local evolution cycles to facilitate meme transmission. Up until the halting criteria are met, these processes are repeated. [30]. The top candidate, or "frog," from each memeplex remains stationary during each iteration of the standard SFLA algorithm. However, this may result in local optima that occur early in the optimization process, trapping the algorithm and perhaps leading to premature termination. This addition enables the algorithm to escape from less-than-ideal solutions, reducing the chance of becoming stale and improving its capacity to address such problems successfully.

Each frog in the procedure stands in for an n-dimensional vector, or $Y_i = [y_{i,1}, y_{i,2}, y_{i,3}, ..., y_{i,n}]$, in the context of an optimization problem with n control parameters (Figure 4). The Latin-hypercube sampling strategy is used to generate N frogs at the start of the optimisation phase to ensure that the whole design space may be thoroughly explored. The total number of frogs produced by this arrangement is N = m * p. For instance, { $S_1, S_2, ..., S_m$ }, { Y_1, Y_4, Y_7, S_2 }, { Y_2, Y_5, Y_8 *and* S_3 }, and { Y_3, Y_6, Y_9 }; because all Xi has been sorted, the Nelder-Mead technique determines that each memeplex contains n + 1 frogs [17]. In memeplexes, the partitioning process is followed by a local search. To get started, we give $X_{b,\,k,}$ the top frog in the memeplex, a randomly generated vector with upper and lower bounds set to a range of one-tenth.

$$Y'_{b,k} = Y_{b,k} + \frac{1}{10}[rand(lb,ub), ..., rand(lb,ub)] \quad (10)$$

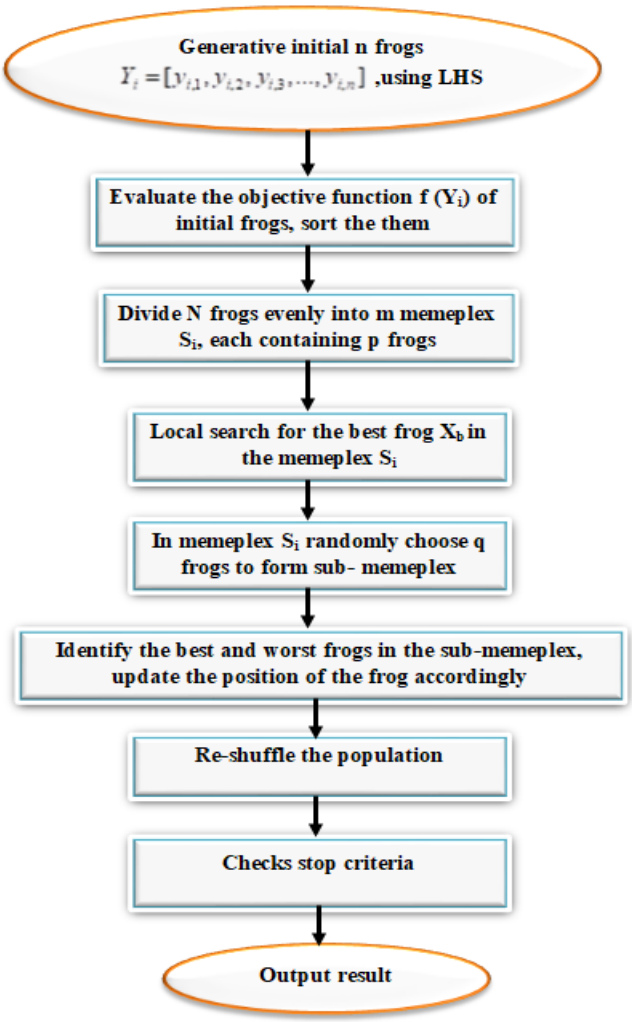

Figure 4. System flow of proposed

The new frog's objective function value is assessed. If it outperforms the original frog, $Y_{b,k}$ is swapped out for $Y_{b,k}^{'}$. After the most prominent frog in the kth memeplex receives an update, a subset of the frog population is chosen to construct a sub-memeplex. The triangle probability function calculates the likelihood of selection for each frog in the memeplex. The degree to which these frogs accomplish the memeplex's desired goals impacts the selection process' outcomes.

$$P_{i,k} = 2(p+1-i) / p(p+1), i = 1, 2, 3, \ldots, p \tag{11}$$

Where (I, $P_{i,k}$) is the probability that a frog will be selected from the kth memeplex, and (i) is the frog's rating. Within the sub-memeplex, the two frogs with the highest and lowest goal values are referred to as $Y_{b,k}$ and $Y_{w,k}$, respectively. Like the PSO algorithm's guiding principles, the least privileged frog learns from the best frog in its sub-memeplex and the best frog overall. This is a reflection of the transfer of knowledge. Mathematical changes are made to the least successful frog's position as follows:

$$\begin{gathered} v = S[w \times (Y_{b,k} - Y_{w,k}) + (1-w) \times (Y_{b,overall} - Y_{w,k})] \\ Y_{w,k}^{'} = Y_{w,k} + v \end{gathered} \tag{12}$$

Let S be the predetermined step size, and W be a random weight between [0, 1]. We determine whether the new $Y_{b,overall}$ location is inside the planned area after acquiring it. If so, we assess its objective worth. We update the ranking of the worst frog once more if this value does not improve over the previous objective value for that specific

frog. Instead of the best frog within the sub-memeplex, this version uses the position of the best frog globally, designated as $Y_{w,k}^{'}$.

$$v = S \times w \times (Y_{b,overall} - Y_{w,k})$$
$$Y_{w,k}^{'} = Y_{w,k} + v \tag{13}$$

The objective value of the newly chosen position $Y_{w,k}^{'}$ is reassessed whether it lies within the designated design space. If neither of these alternatives improves the objective value, the settings of the least successful frog are changed following:

$$Y_{w,k}^{'} = [rand(lb,ub),...,rand(lb,ub)] \tag{14}$$

The population is then randomly distributed and shifted to the following iteration once the local search procedure has repeated t times. The stop requirement is reached after all of these steps have been completed.

4. Result and Discussions

Using various different classes with differing restrictions, the proposed ConvLSTM-based anomaly detection and classification model's performance is assessed for 3D MRI images. This section discusses the outcomes obtained in each test situation. In this research using, the existing systems are CVM, Segmentation Auto-Encoder (SegAE), Generative Adversarial Network (GAN), and LSTM. Further research and development of this strategy should produce even more impressive effects as technology advances, enhancing clinical decision-making, early anomaly detection, and better patient outcomes in medical imaging. The Brats data set is utilised by the approach to assess performance. When evaluating the performance of the models, which are evaluated based on a number of different parameters and the data set's three tumour types, the limitations are indicated in Table 2. Figure 5 shows the 3D MRI images**.**

Table 2. Evaluation detail**s** for brats data

Factor	**Value**
Data Source	Brats 2019
No of Image Class	3
Total Images	3000
Platform	Python

4.1 Experimental Setup

The Python 3.6.9 programming language, PyTorch 1.6 deep learning framework, CUDA 10.0 GPU acceleration, and cuDNN 7.6.5 library are all used in the experimental configuration. A computer system with an Intel(R) Core(TM) i7-1065G7 CPU and a GeForce GTX Titan Xp 440.100 GPU is used to conduct the model's learning, assessment, and outcome analysis. The essential clock speed of the CPU is 1.30 GHz, and it has a maximum turbo boost speed of 1.50 GHz.

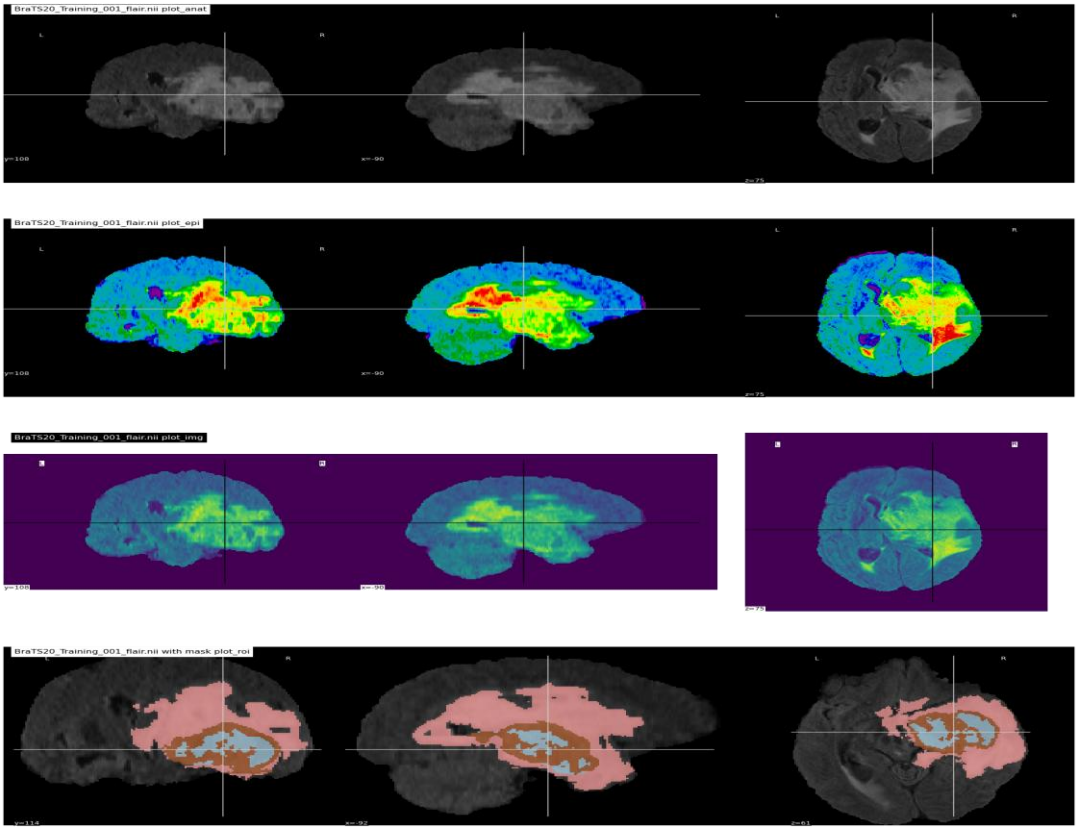

Figure 5. 3D MRI medical images

4.2 Performance Metrics

Performance measurements are necessary to assess the classifiers' prediction performances. Although accuracy is one of the most frequently used measures, it may yield inaccurate results if the data's class distribution is unbalanced. Even in class imbalances, performance metrics like precision, F1 score, specificity, sensitivity, and response time can be valuable tools for assessing the classifier's capacity to distinguish between numerous classes. TP, FP, FN, and TN stand for True Positives, False Positives, False Negatives, and True Negatives, respectively, in the confusion matrix. Dice Similarity Coefficient (DSC): Equation (15) can be used to determine the dice similarity coefficient (DSC), which is used to compare real tumour and non-tumor pixels with anticipated tumour and non-tumor pixels.

$$DSC = \frac{2TP}{FP + 2TP + FN} \times 100 \tag{15}$$

JACCARD Similarity Index (JSI): The JACCARD similarity index (JSI), which is determined using Equation (16), is used to determine how similar the real tumour pixels and anticipated tumour pixels are.

$$JSI = \frac{TP}{TP + FN + FP} \times 100 \tag{16}$$

TRUE Positive Rate (TPR): TPR is an indicator of sensitivity, and by predicting the actual anomalous region abnormal, the anomalous region detection results can be confirmed

$$TPR = \frac{TP}{TP + FN} \tag{17}$$

FALSE Positive Rate (FPR): Equation (18) is used to calculate the ratio of incorrectly identified to correctly identified pixels, often known as the false positive rate (FPR):

$$FPR = \frac{FP}{FP + TN} \tag{18}$$

Accuracy: The ability of a artificial intelligence model to recognise and categorise abnormal structures, lesions, or anomalies found in 3D MRI data is referred to as "accuracy". The accuracy score assesses how successfully the model

distinguishes between normal and abnormal regions in 3D MRI data, assisting in the early identification and treatment of medical disorders.

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \tag{19}$$

4.2.1 Dice Similarity Coefficient (DSC) Analysis

Table 3. DSC analysis of ConvLSTM technique using existing systems

Number of data from Dataset	CVM	LSTM	SegAE	GAN	ConvLSTM
100	75.45	83.45	79.56	86.56	91.23
200	76.12	83.99	80.56	87.12	91.99
300	77.78	84.98	81.23	88.99	92.34
400	78.45	84.99	81.99	89.45	93.89
500	78.11	85.12	82.76	89.77	94.44

Figure 6 and Table 3 show a comparison of the DSC gained by the ConvLSTM technology vs competing alternatives. The graphical representation depicts the efficacy and DSC gains enabled by deep learning. For example, when analysing DSC values across 100 data points, the CVM, LSTM, SegAE, and GAN models obtain 75.45%, 83.45%, 79.56%, and 86.56% DSC, respectively, compared to the astounding 91.23% precision of the ConvLSTM model. When using a variety of datasets, the ConvLSTM model emerges as the clear winner. Notably, 94.44% with 500 data points, the ConvLSTM model outperforms the CVM, LSTM, SegAE, and GAN models, yielding DSC scores of 78.11%, 85.12%, 82.76%, and 89.77%, respectively.

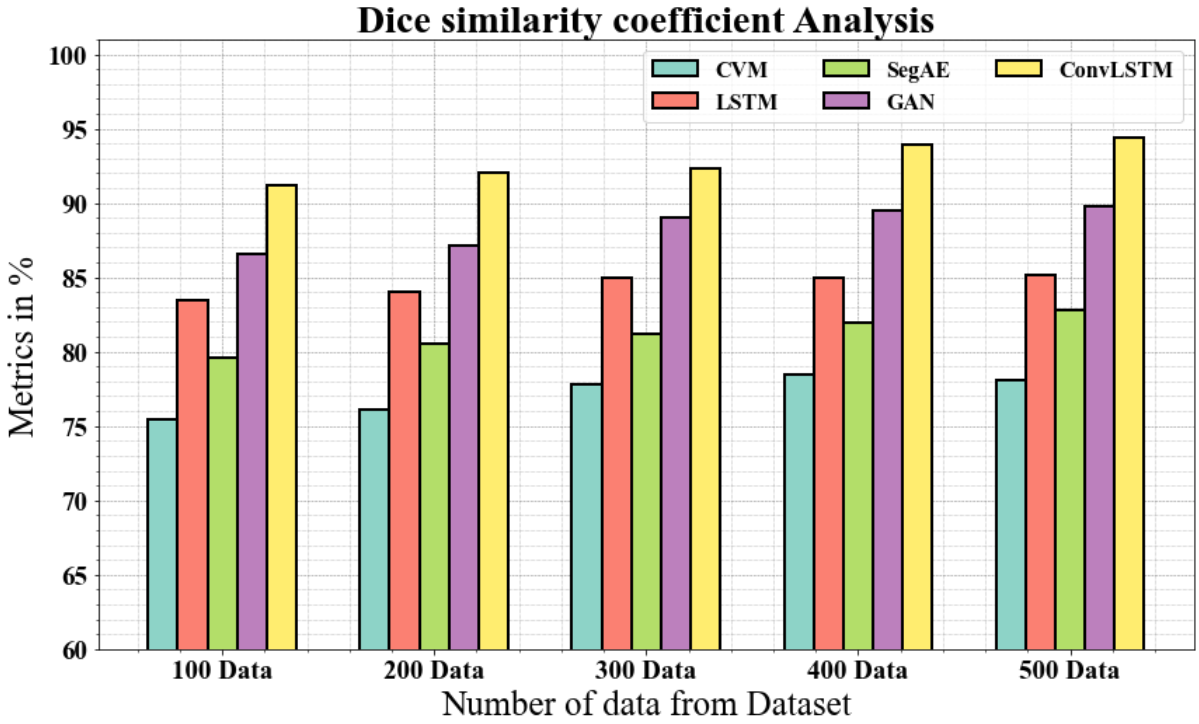

Figure 6. DSC analysis of the ConvLSTM technique with existing systems

4.2.2 Jaccard Similarity Coefficient (JSC) Analysis

Table 4. JSC analysis of the ConvLSTM technique with existing systems

Number of data from Dataset	CVM	LSTM	SegAE	GAN	ConvLSTM
100	70.12	75.12	80.34	85.12	92.15
200	71.45	76.66	81.33	86.56	93.78
300	72.34	77.77	82.78	88.99	94.19
400	73.33	78.12	83.11	89.45	95.66
500	74.45	79.45	83.99	90.33	96.12

The JSC of the ConvLSTM methodology is compared to that of other methods in Figure 7 and Table 4. The graph demonstrates how the deep learning approach has improved JSC and efficiency. For instance, the JSC of the CVM, LSTM, SegAE, and GAN models with 100 data is 70.12%, 75.12%, 80.34%, and 85.12%, respectively, as opposed to the JSC of the ConvLSTM model, which is 92.15%. With diverse data, the ConvLSTM S model has had the best performance. ConvLSTM exhibits a JSC of 96.12% under 500 data, compared to sensitivity values of 74.45%, 79.45%, 83.99%, and 90.33% for CVM, LSTM, SegAE, and GAN.

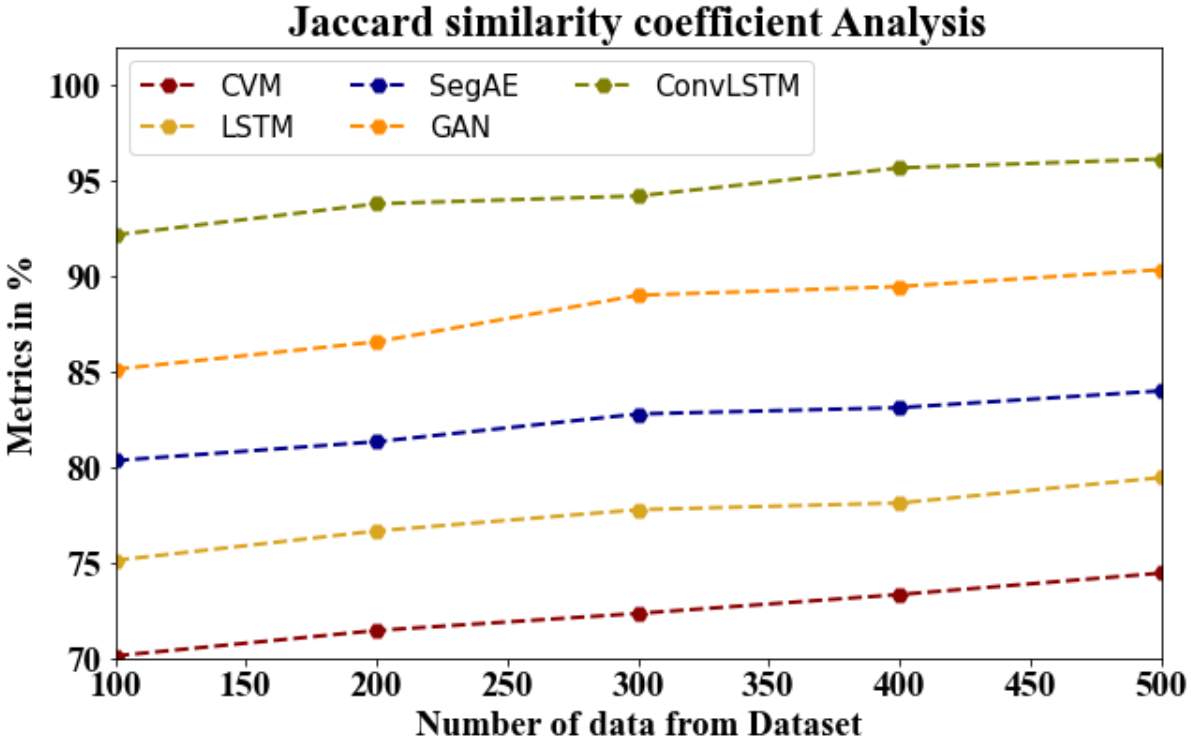

Figure 7. JSC analysis of the ConvLSTM technique with existing systems

4.2.3 True Positive Rate (TPR) Analysis

Table 5. TPR analysis of the ConvLSTM technique with existing systems

Number of data from Dataset	CVM	LSTM	SegAE	GAN	ConvLSTM
100	68.12	78.45	73.19	83.88	91.45
200	69.45	79.12	74.56	84.78	92.88
300	70.23	80.45	75.55	85.55	93.56
400	71.11	81.66	76.12	86.12	94.55
500	72.34	82.45	77.44	87.77	95.77

The TPR of the ConvLSTM methodology is compared to that of other methods in Figure 8 and Table 5. The graph demonstrates how the deep learning method is more effective with TPR. For instance, the TPR values of the CVM, LSTM, SegAE, and GAN models for 100 data are 68.12%, 78.45%, 73.19%, and 83.88%, respectively, as contrasted to the TPR of the ConvLSTM model, which is 91.45%. With diverse data, the ConvLSTM S model has had the best performance. ConvLSTM has a specificity of 95.77% for 500 data points, compared to TPR values of 72.34%, 82.45%, 77.44%, and 87.77% for CVM, LSTM, SegAE, and GAN.

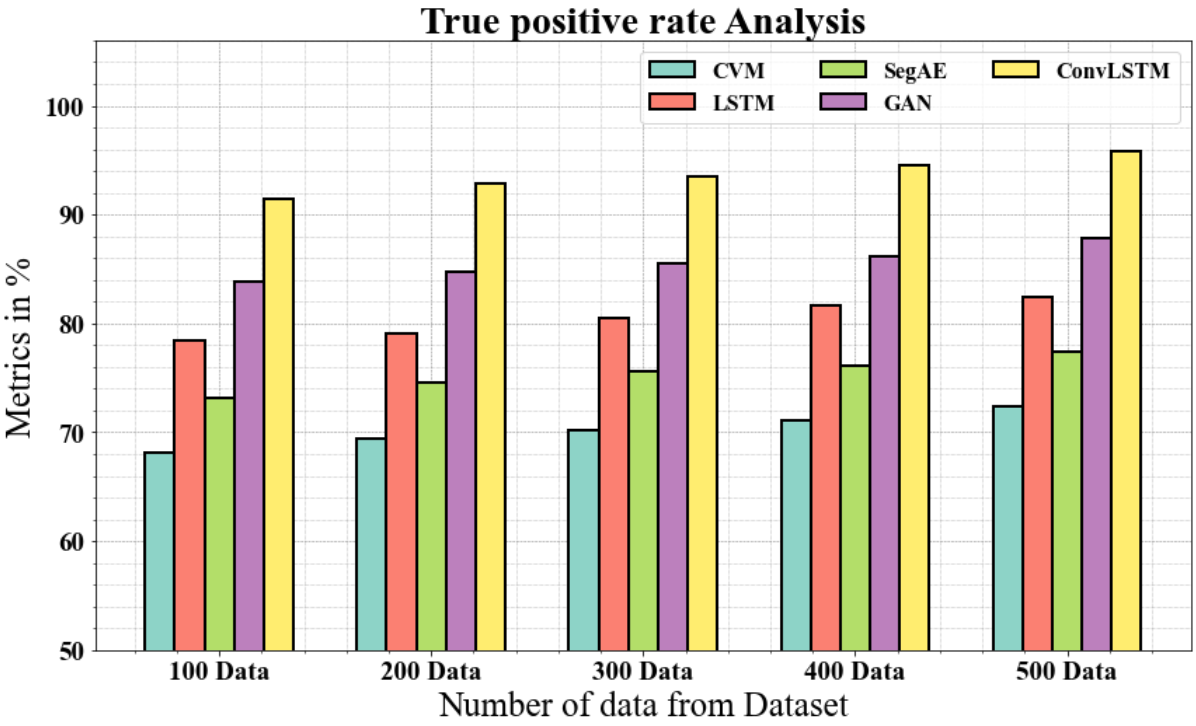

Figure 8. TPR analysis of the ConvLSTM technique with existing systems

4.2.4 Predicted Time Delay (PTD) Analysis

Table 6. PTD analysis of the ConvLSTM technique with existing systems

Number of data from Dataset	CVM	LSTM	SegAE	GAN	ConvLSTM
100	13.334	10.456	7.678	4.134	1.675
200	13.765	10.778	7.119	4.998	2.456
300	14.786	11.234	8.234	5.987	2.998
400	14.991	11.567	8.998	5.119	3.456
500	15.223	12.987	9.115	6.564	3.987

The suggested ConvLSTM methodology's Predicted time delay is compared to existing approaches in Table 6 and Figure 9, where the ConvLSTM technique outperforms all the others. For instance, the suggested ConvLSTM methodology predict 100 data in just 1.675ms as opposed to the 13.334ms, 10.456ms, 7.678ms, and 4.134ms required by the CVM, LSTM, SegAE, and GAN existing methods. Similar to this, the proposed ConvLSTM strategy predict 500 data in 3.987ms as opposed to 15.223ms, 12.987ms, 9.115ms, and 6.564ms for existing techniques like CVM, LSTM, SegAE, and GAN.

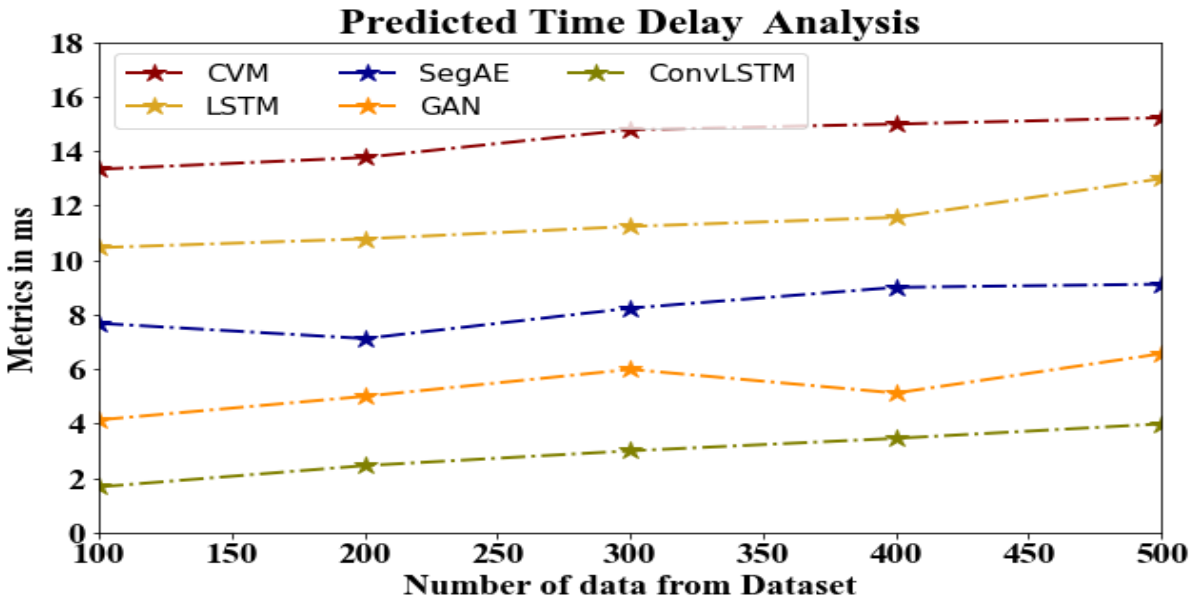

Figure 9. PTD analysis of the ConvLSTM technique with existing systems

4.2.5 False Positive Rate (FPR) Analysis

Table 7. FPR analysis of the ConvLSTM technique with existing systems

Number of data from Dataset	CVM	LSTM	SegAE	GAN	ConvLSTM
100	37.937	35.314	32.029	29.728	25.029
200	37.425	35.829	32.627	29.525	26.526
300	38.536	35.637	33.324	30.672	26.828
400	38.626	36.453	33.526	31.829	27.526
500	38.425	36.938	34.121	30.232	27.321

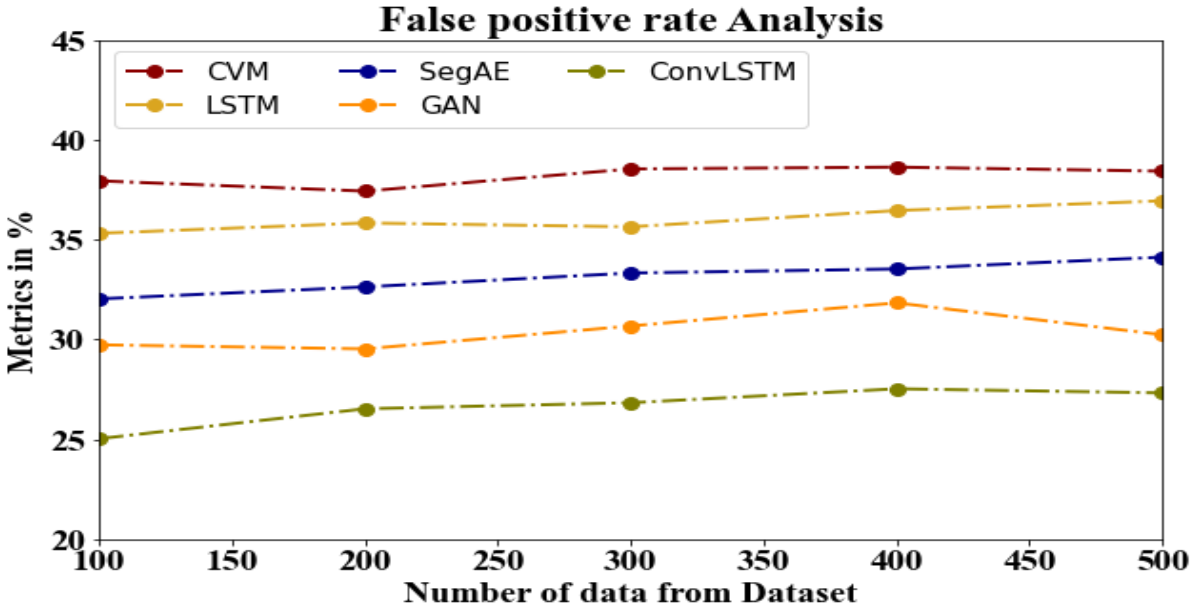

Figure 10. FPR analysis of the ConvLSTM technique with existing systems

Figure 10 and Table 7 display an FPR comparison of the FFNN-CQNGT strategy with other well-known methods. The deep learning technique has an enhanced performance while reducing FPR, as shown in the graph. For example, the

FFNN-CQNGT model's FPR value for 100 data is 25.029%, while the FPR values for the CVM, LSTM, SegAE, and GAN models are 37.937%, 35.314%, 32.029%, 29.728%, respectively. The FFNN-CQNGT model, however, has demonstrated its best performance for various data sizes with low FPR values. In a similar vein, for 500 data, the FPR value for the FFNN-CQNGT is 27.321%, whereas, for the CVM, LSTM, SegAE, and GAN models, it is 38.425%, 36.938%, 34.121% and 30.232%, respectively.

4.2.6 Accuracy Analysis

Table 8. Accuracy analysis of the ConvLSTM technique with existing systems

Number of data from Dataset	CVM	LSTM	SegAE	GAN	ConvLSTM
100	80.12	90.23	85.56	94.91	96.13
200	81.34	91.22	86.12	94.99	97.77
300	82.99	92.76	87.34	95.13	98.45
400	83.67	93.45	88.98	95.67	98.99
500	84.44	94.87	89.12	95.98	99.12

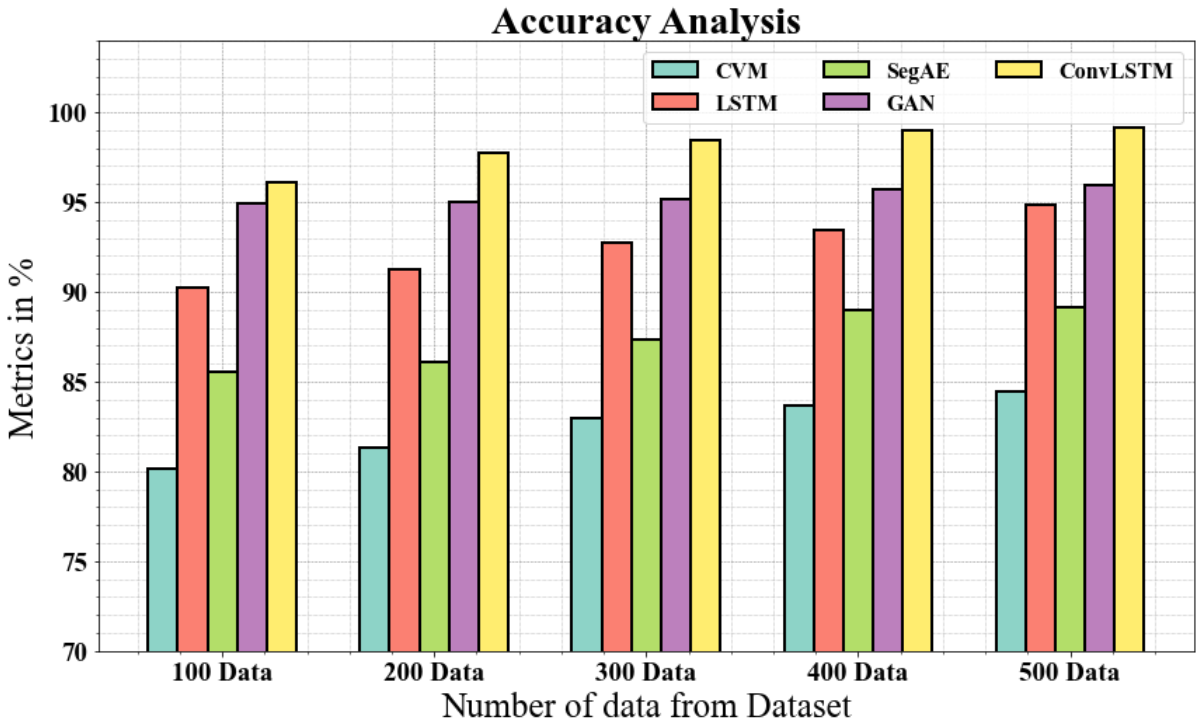

Figure 11. Accuracy analysis of the ConvLSTM technique with existing systems

The accuracy of the ConvLSTM methodology is compared to that of other methods in Figure 11 and Table 8. The graph demonstrates how the deep learning method increases accuracy and efficiency. For example, the accuracy of the CVM, LSTM, SegAE, and GAN models with 100 data is 80.12%, 90.23%, 85.56%, and 94.91%, respectively, while the accuracy of the ConvLSTM model is 96.13%. With diverse data, the ConvLSTM S model has had the best performance. ConvLSTM has an accuracy of 99.12% for 500 data points, compared to accuracy values of 84.44%, 94.87%, 89.12%, and 95.98% for CVM, LSTM, SegAE, and GAN.

4.2.7 ROC Curve Analysis

An evaluation statistic that takes into account classification performance (normal vs. abnormal) across all decision thresholds is the Area Under the Curve of the Receiver Operating Characteristic (ROC AUC) (Figure 12). Because of this, the performance of the models may be evaluated "probabilistically" and without bias, eliminating the requirement to select a threshold for the projected abnormality scores (Table 9).

Table 9. ROC curve analysis for ConvLSTM method

False positive Rate	Random classifier	CVM	LSTM	SegAE	GAN	Proposed ConvLSTM
0	0	0	0	0	0	0
0.1	0.1	0.41	0.59	0.64	0.61	0.78
0.2	0.2	0.62	0.62	0.78	0.73	0.83
0.3	0.3	0.75	0.78	0.77	0.77	0.86
0.4	0.4	0.82	0.81	0.77	0.79	0.90
0.5	0.5	0.84	0.83	0.81	0.80	0.92
0.6	0.6	0.90	0.85	0.85	0.86	0.95
0.7	0.7	0.92	0.89	0.90	0.91	0.98
0.8	0.8	0.94	0.93	0.95	0.97	0.98
0.9	0.9	0.98	0.95	0.98	0.99	0.99
1.0	1.0	0.99	0.99	0.99	0.99	0.99

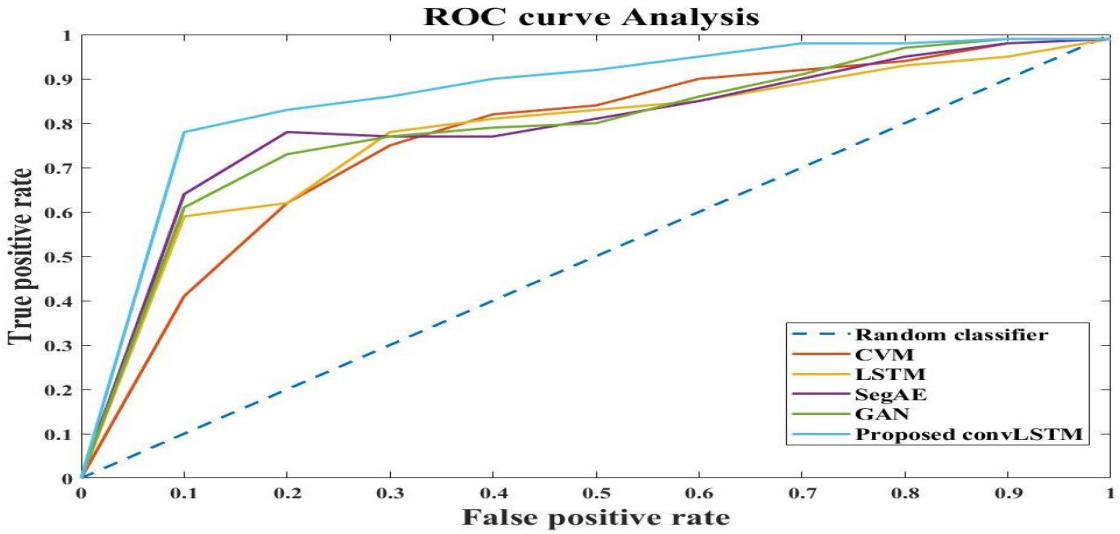

Figure 12. ROC curve analysis for ConvLSTM method

5. Conclusions

In conclusion, this study integrates a 3D U-Net model, ConvLSTM architecture, and the ground-breaking SFLA to propose a novel and effective method for segmented anomaly identification and classification in 3D MRI data. The outcomes showed how well this hybrid methodology worked to locate and classify anomalies in intricate 3D MRI data. The model produced remarkable segmentation performance by utilizing the 3D U-Net's capabilities, successfully outlining regions of interest, and improving the subsequent anomaly detection procedure. ConvLSTM architecture was incorporated to take advantage of the temporal correlations in sequential MRI scans, allowing the model to capture complex patterns and dynamic changes in the data. An optimization layer was added to the model using the SFLA, accelerating convergence and optimizing model parameters for better anomaly classification accuracy. The model was able to produce better outcomes with fewer iterations because of the addition of this creative method. The suggested method has great promise for advancing medical imaging analysis, notably in 3D MRI image anomaly identification and classification. A strong and efficient solution is provided by integrating deep learning methods with optimization approaches described in this paper, with potential applications in numerous medical specialties and diagnostic procedures. The suggested framework ConvLSTM experimental results demonstrate exceptional performance: precision of 94.44%, f-measure of 97.12%, accuracy of 99.12%, sensitivity of 96.12%, specificity of 95.77%, and a execution time of 3.987ms. Further research and development of this strategy should

produce even more impressive effects as technology advances, enhancing clinical decision-making, early anomaly detection, and better patient outcomes in medical imaging.

Statements and Declarations

Competing Interests: Not Applicable
Compliance with Ethical Standards: Not Applicable
Conflict of interest: Not Applicable
Ethics approval: Not Applicable
Funding: Not Applicable

References

[1] Chalapathy, R. and Chawla, S., 2019. Deep learning for anomaly detection: A survey. *arrive preprint arXiv:1901.03407*.

[2] Fernando, T., Gamble, H., Denman, S., Sridharan, S. and Fookes, C., 2021. Deep learning for medical anomaly detection–a survey. *ACM Computing Surveys (CSUR)*, *54*(7), pp.1-37.

[3] Chetty, G., Yamin, M. and White, M., 2022. A low resource 3D U-Net based deep learning model for medical image analysis. *International Journal of Information Technology*, *14*(1), pp.95-103.

[4] Cali, E., Sogancioglu, E., van Ginneken, B., van Leeuwen, K.G. and Murphy, K., 2021. Deep learning for chest X-ray analysis: A survey. *Medical Image Analysis*, *72*, p.102125.

[5] Shvetsova, N., Bakker, B., Fedulova, I., Schulz, H. and Dylov, D.V., 2021. Anomaly detection in medical imaging with deep perceptual autoencoders. *IEEE Access*, *9*, pp.118571-118583.

[6] Chetty, G., Singh, M. and White, M., 2019, December. Automatic brain image analysis based on multimodal deep learning scheme. In *2019 International Conference on Machine Learning and Data Engineering (iCMLDE)* (pp. 97-100). IEEE.

[7] Tayeh, T., Aburakhia, S., Myers, R. and Shami, A., 2022. An attention-based ConvLSTM autoencoder with dynamic thresholding for unsupervised anomaly detection in multivariate time series. *Machine Learning and Knowledge Extraction*, *4*(2), pp.350-370.

[8] Shi, X., Chen, Z., Wang, H., Yeung, D.Y., Wong, W.K. and Woo, W.C., 2015. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, *28*.

[9] Yang, W., Ho, S.L. and Fu, W., 2020. A modified shuffled frog leaping algorithm for the topology optimization of electromagnet devices. *Applied Sciences*, *10*(18), p.6186.

[10] Asif, S., Yi, W., Ain, Q.U., Hou, J., Yi, T. and Si, J., 2022. Improving effectiveness of different deep transfer learning-based models for detecting brain tumors from MR images. *IEEE Access*, *10*, pp.34716-34730.

[11] Vidyarthi, A., Agarwal, R., Gupta, D., Sharma, R., Draheim, D. and Tiwari, P., 2022. Machine Learning Assisted Methodology for multiclass classification of malignant brain tumors. *IEEE Access*, *10*, pp.50624-50640.

[12] Rahman, S., Sarker, S., Miraj, M.A.A., Nihal, R.A., Nadimul Haque, A.K.M. and Noman, A.A., 2021. Deep learning–driven automated detection of Covid-19 from radiography images: A comparative analysis. *Cognitive Computation*, pp.1-30.

[13] Park, S.; Lee, S.M.; Kim, N.; Choe, J.; Cho, Y.; Do, K.H.; Seo, J.B. Application of deep learning–based computer-aided detection system: Detecting pneumothorax on chest radiograph after biopsy. Eur. Radiol. 2019, 29, 5341–5348.

[14] Tolkachev, A.; Sirazitdinov, I.; Kholiavchenko, M.; Mustafaev, T.; Ibragimov, B. Deep learning for diagnosis and segmentation of pneumothorax: The results on the Kaggle Competition and Validation Against Radiologists. IEEE J. Biomed. Health Inform. 2020, 25, 1660–1672.

[15] Zhao, G.; Fang, C.; Li, G.; Jiao, L.; Yu, Y. Contralaterally Enhanced Networks for Thoracic Disease Detection. IEEE Trans. Med. Imaging 2021, 40, 2428–2438.

[16] Antoni, J.; Dawid, P.; Robertas, D. Lung X-Ray Image Segmentation Using Heuristic Red Fox Optimization Algorithm. Sci. Program. 2022, 2022, 4494139.

[17] Marimont, S.N. and Tarroni, G., 2021, April. Anomaly detection through latent space restoration using vector quantized variational autoencoders. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)* (pp. 1764-1767). IEEE.

[18] Esser, P., Rombach, R. and Ommer, B., 2021. Taming transformers for high-resolution image synthesis. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 12873-12883).

[19] Jun, H., Child, R., Chen, M., Schulman, J., Ramesh, A., Radford, A. and Sutskever, I., 2020, November. Distribution augmentation for generative modeling. In *International Conference on Machine Learning* (pp. 5006-5019). PMLR.

[20] Choromanski, K., Likhosherstov, V., Dohan, D., Song, X., Gane, A., Sarlos, T., Hawkins, P., Davis, J., Mohiuddin, A., Kaiser, L. and Belanger, D., 2020. Rethinking attention with performers. *arrive preprint arXiv:2009.14794*.

[21] Pinaya, W.H.L., Tudosiu, P.D., Gray, R., Rees, G., Nachev, P., Ourselin, S. and Cardoso, M.J., 2021. Unsupervised brain anomaly detection and segmentation with transformers. arrive preprint arXiv:2102.11650.

[22] A. Sekhar, S. Biswas, R. Hazra, A. K. Sunaniya, A. Mukherjee and L. Yang, "Brain Tumor Classification Using Fine-Tuned GoogLeNet Features and Machine Learning Algorithms: IoMT Enabled CAD System," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 3, pp. 983-991, March 2022, doi: 10.1109/JBHI.2021.3100758.

[23] A. Kujur, Z. Raza, A. A. Khan and C. Wechtaisong, "Data Complexity Based Evaluation of the Model Dependence of Brain MRI Images for Classification of Brain Tumor and Alzheimer's Disease," in IEEE Access, vol. 10, pp. 112117-112133, 2022, doi: 10.1109/ACCESS.2022.3216393.

[24] P. Allirani, R. Jain, S. Shankar Prasd, K. H. Wanjale, A. Amudha and A. Faiz, "Real-Time Depth Map Upsampling for High-Quality Stereoscopic Video Display," 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024, pp. 1-5, doi: 10.1109/ICCCNT61001.2024.10725345.

[25] Isensee F, Jaeger PF, Kohl SAA, Petersen J, Klaus H. Automated design of deep learning methods for biomedical image segmentation. 2020. p. 1–55. https://arxiv. org/abs/1904.08128.

[26] Asgari Taghanaki S, Abhishek K, Cohen JP, Cohen-Adad J, Hamarneh G. Deep semantic segmentation of natural and medical images: a review. Artif Intell Rev 2020:1–42. https://doi.org/10.1007/s10462-020-09854-1

[27] Zhang Z, Liu Q, Wang Y. Road extraction by deep residual U-net. Geosci Rem Sens Lett IEEE 2018

[28] Eusuff, M.; Lansey, K.E.; Pasha, F. Shuffled frog-leaping algorithm: A memetic meta-heuristic for discrete optimization. Eng. Optim. 2006, 38, 129–154.

[29] Seyed SSM, Erdogmus D, Gholipour A, Salehi SSM, Erdogmus D, Gholipour A. Tversky loss function for image segmentation using 3D fully convolutional deep networks. In: Lecture notes in computer science. Springer Verlag; 2017. p. 379–87. https://doi.org/10.1007/978-3-319-67389-9_44

[30] Shi, X.; Chen, Z.; Wang, H.; Yeung, D.Y.; Wong, W.K.; Woo, W.C. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In Proceedings of the Twenty-ninth Conference on Neural Information Processing Systems, Montréal, QC, Canada, 11–12 December 2015; pp. 1–9.

[31] K. Joshi et al., "Intelligent Fusion Approach for MRI and CT Imaging using CNN with Wavelet Transform Approach", International Conference on Knowledge Engineering and Communication Systems (ICKES), India, pp. 1-6, 2022.