International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V71I10P214

Volume 71 Issue 10, 156-162, October 2023

Original Article

SpineResUnet: Classification and Prediction of Spinal Tuberculosis Disease on Exploiting the Structural and Texture Dependencies

K. T. Askarali[1], E. J. Thomson Fredrik[2]

[1]*Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, Tamilnadu, India.*
[2]*Department of Computer Technology, Karpagam Academy of Higher Education, Coimbatore, Tamilnadu, India.*

[1]*Corresponding Author : askaralimcas@gmail.com*

Received: 24 June 2023 Revised: 30 August 2023 Accepted: 15 September 2023 Published: 03 October 2023

Abstract *- Spinal Tuberculosis is a most important and very dangerous extrapulmonary disease type of skeletal tuberculosis, which causes destruction, collapse of vertebrae and angulations of the vertebral column. Hence, detecting the disease early is becoming important to prevent serious complications such as spinal deformity and permanent neurological deficit. Many machine learning and deep learning models have been presented to detect and classify the disease. However, those techniques lead to various challenges in appearance and are complex in fully exploiting the dependency between the structural and texture characteristics of the image. In order to handle the challenges, a new deep learning framework named SpineResUnet was to achieve the classification of spinal tuberculosis disease on volumetric MR Images. SpineGraphNet is composed of automated segmentation and classification of the image. Initially, segmentation of the collapsed region of vertebrate or spinal deformity region is carried out using graph convolution segmentation network as fine-grained segments. The segmented region of the image is processed further towards classifying the segmented region into classes using an adjacency matrix containing semantic features in ResNet and U-Net with Skip Connection. The classes represent kyphosis, gibbus formation and osteomyelitis of the spinal tuberculosis. The proposed model successfully captures the dependency implicitly and explicitly. Experimental analysis of T2 weighted volumetric MR image of 100 subjects has been collected from the KMCH hospital, Coimbatore. In this, 60% of images have been employed to train corresponding experts in Spinal Tuberculosis disease characterization, 20% has been used as testing, and the remaining 20% of the image has been taken for 5-fold cross-validation through a confusion matrix. The proposed model exhibited optimal performance in terms of dice coefficient, specificity, and sensitivity compared against conventional approaches.*

Keywords *- Deep learning, Gibbus formation, Kyphosis, Osteomyelitis, Spinal tuberculosis.*

1. Introduction

Spinal tuberculosis is a manifestation of extrapulmonary tuberculosis with a large frequency of prolonged-term morbidity. It is associated with neurologic deficit due to compression of neurological structures and significant spinal deformity on its slow, insidious progression. In order to prevent persistent neurological dysfunction and to reduce spinal deformity, early diagnosis and classification of the disease have become highly important. Identifying the types of spinal tuberculosis using clinical data has become difficult.

Hence, MRI imaging can discriminate the diseases on the rigid structures of vertebrae and intervertebral discs. Especially lumbar spine is the most common site of the disease, which is followed by the thoracic region. The disease's severity is identified by analysis of features collected using an MRI scan imaging dataset. MRI finding provides the feature of diseases such as spinal deformity on the location of the tuberculosis vertebral lesion, bone destruction in fragments, intervertebral disc destruction, and paravertebral mass or abscess in the region of the lower thoracic and upper lumbar vertebrae.

Existing research using many machine learning models and deep learning models like fully convolution neural is computation efficient in processing vertebrate regions are capable of detecting and classifying the spinal deformities and collapse of vertebrates. Deformable models using CNN are highly exposed to surface errors and fail to explicitly consider dependencies between structures. Further traditional techniques lead to various challenges in appearance and are found to be complex in fully exploiting the dependency between the structural and texture characteristics of the image. In addition, accurately segmenting an intervertebral may lead to failure to compute all subsequent vertebrae.

In order to handle the challenges, a new deep learning framework named SpineResUnet is proposed to achieve vertebral localization, segmentation and classification of the spinal tuberculosis disease on volumetric MR Images. SpineResUnet is composed of automated segmentation and classification of the image on utilization of Generative Adversarial Networks (GAN). GAN implicitly capture dependency between vertebrae. Initially, segmentation of the collapsed region of vertebrate or spinal deformity region

is carried out using graph convolution segmentation network as fine segments. The segmented region of the image is processed further towards classifying the segmented region into classes using an adjacency matrix composed of semantic features in ResUNet. ResUnet is the combination of U-net and ResNet for classification. The classes represent kyphosis, gibbus formation and osteomyelitis of the spinal tuberculosis.

The remaining article has been sectioned as follows: section 2 details the existing deep learning model with spatial and temporal constraints for spinal tuberculosis disease classification. Section 3 provides the proposed new deep learning methodology named SpineResUnet for disease segmentation and classification of disease types and stages. Experimental analysis of the proposed methodology on the MR Dataset has been carried out in section 4, along with performance analysis on various measures like dice coefficient, sensitivity and specificity on the confusion matrix. Finally, section 5 concludes the work with future suggestions.

2. Related work

This part employs a detailed analysis of deep learning techniques to segment and classify Spinal tuberculosis.

2.1. Spinal Tuberculosis Disease using Optimization of CNN using Spatial and Temporal Constraints

In this model, Spatial and Temporal Constrained Convolution Neural Network (SPCCNN) has been designed to detect and classify spinal tuberculosis diseases into classes with respect to their severity on stages. Image Augmentation assists convolution Neural Networks in increasing the training set size without acquiring images. Convolution Neural Network processed with features extracted by PCA on the various layers of the network. The convolution and activation layers have been optimized using the spatial and temporal constraints on the hyperparameter to classify the disease into kyphosis, gibbus formation and osteomyelitis.

3. Proposed Model

This section uses SpineResUnet to segment and classify the spinal tuberculosis disease on volumetric MR Images. It consists of U net and ResNet with skip connection to obtain the semantic image features on utilizing the dependencies between the neighboring vertebra and IVD.

3.1. U Net –ResNet34

The segmentation network aims to generate segmentation feature maps using low-resolution MRI images. The SpineResUnet updates the segmentation outcomes on extracting the high-resolution image using encoder and decoder operation with the pooling process. In [20], M.Nisha *et al*., 2022 presented Deep learning-based Semi-Supervised Generative Adversarial Networks (SSGAN) for Automatic Hippocampus Segmentation with efficient accuracy.

ResUNet is a network model composed of an encoder network, decoder network and convolution layer to achieve segmentation of spinal deformity region and collapse of vertebra regions. Encode Network takes input image in 32 convolution blocks to encode into feature representation at various levels with feature map.

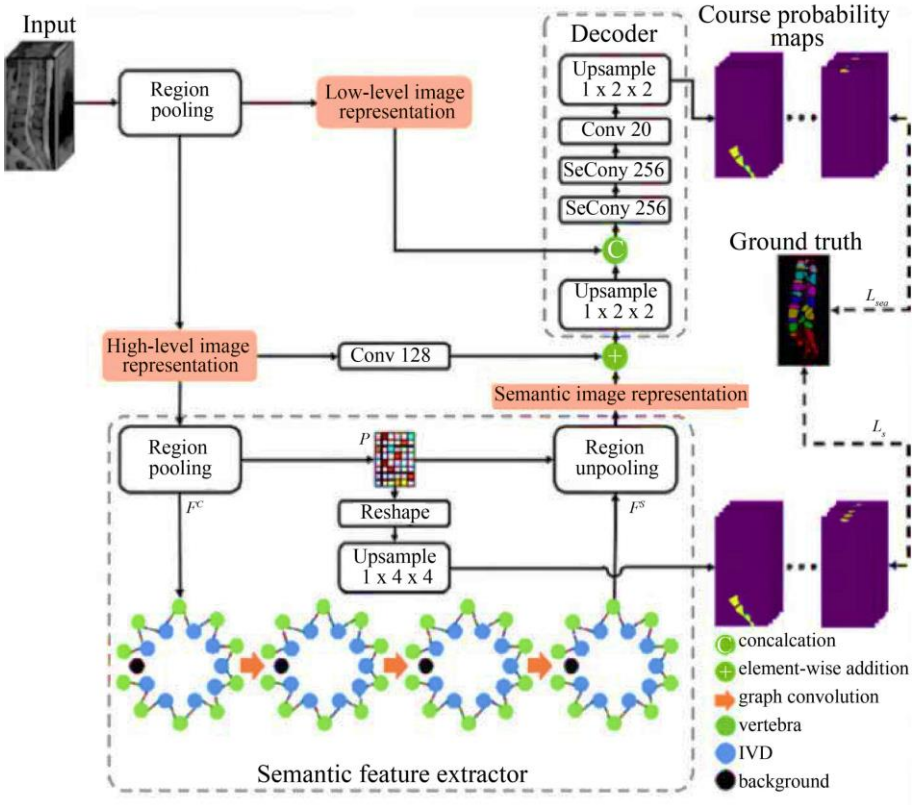

Fig. 1 Segmentation of the spinal tuberculosis vertebrae

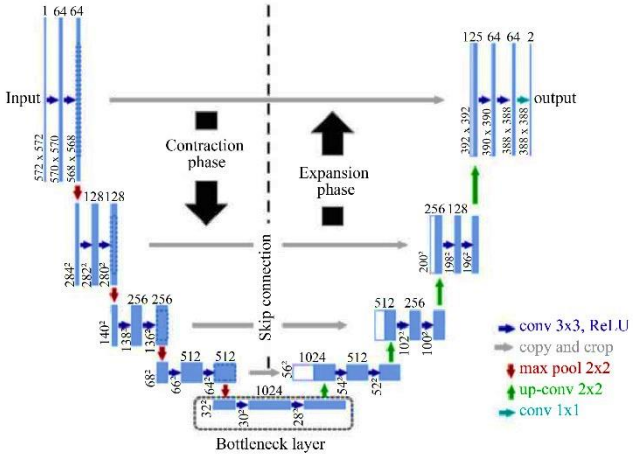

Fig. 2 Spine ResUNet architecture for the disease segmentation

Feature maps will be increased in each convolution block to learn the complex structure of the deformity accurately. The proposed convolution block transforms the high-level features of the MRI image from the encoder network to low-level feature representation in the decoder network. The decoder network semantically transforms the discriminate features of the pixel space to generate the dense network containing the fine-grained segments that represent the collapse of vertebrae and spinal deformity. Figure 2 represents the architecture of the Spinal ResUNet for spinal tuberculosis disease.

3.1.1. Residual Neural Network

Transformation of the dense network will enlarge the receptive layers in the convolution-neutral network by including pooling layers, which can be considered max pooling and average pooling [16]. A neural network becomes deeper due to the occurrence of changes in weight in the layers to the input of the network. In order to degrade the network, the ResNet network has to be employed with a skip connection. ResNet34 is employed as an encoder network to U-Net architecture to enable feature propagation gradient flow and solve the vanishing gradient issues in this work.

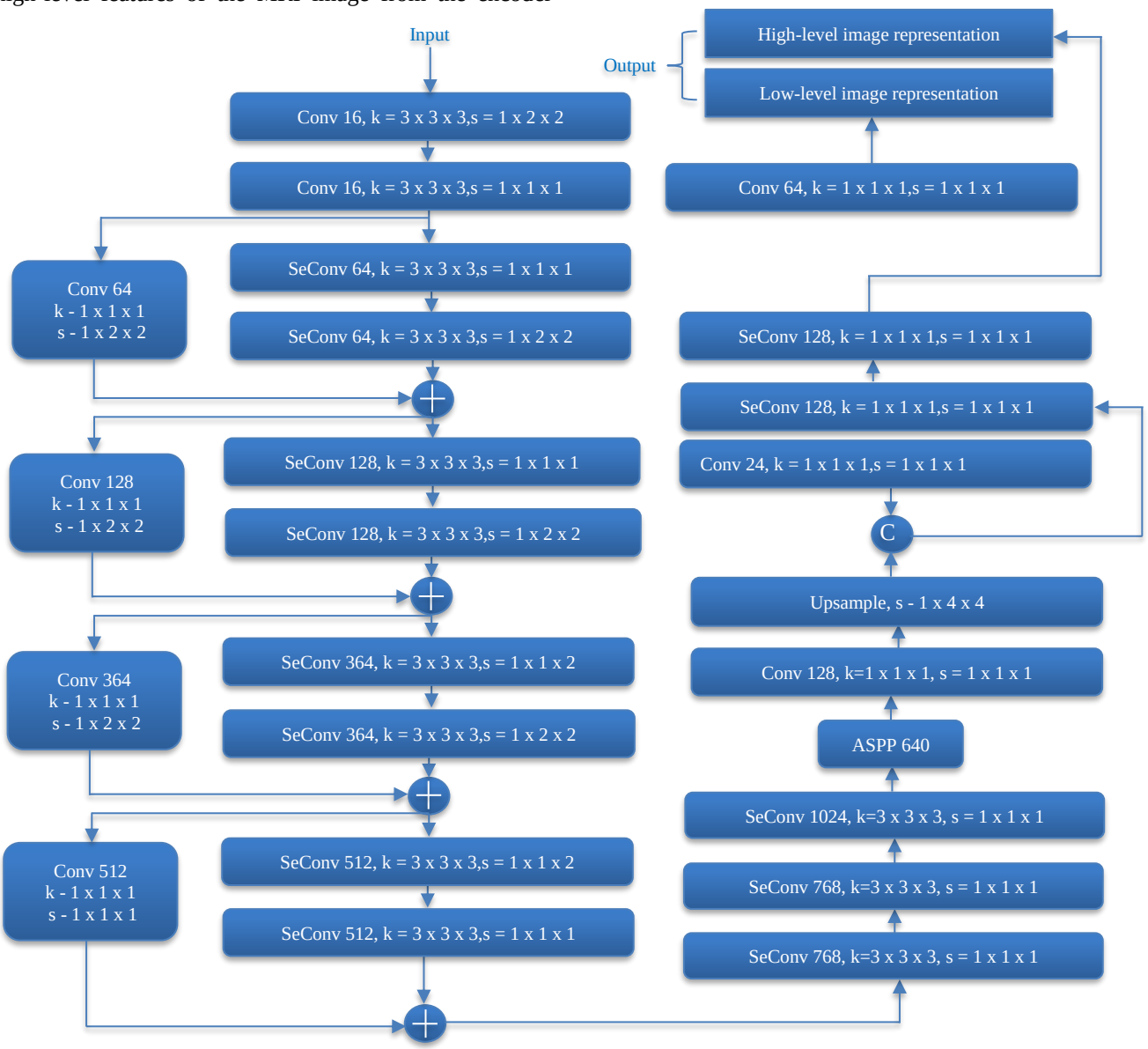

Fig. 3 Encoder and decoder flow of the SpineResUnet architecture with skip connection

Encoder Network of U Net

In this, ResNet34 is a convolution Neural Network composed of 34 layers formed to obtain the details of spinal tuberculosis. Further, the network comprises 16 residual blocks to generate the semantic feature representation. Each block samples the images in each layer using kernel size and stride of convolution. It is considered as a contraction path of the U-net.

Decoder Network of the U Net

A Decoder Network with a decoder block as the residual connection is considered an expansive path incorporating region pooling to up-sample the down-sampled features. The up sample of the low-level image features is carried out in transposed convolutions to expand the size of the convolution layer to obtain high-level features. Features are concatenated from the reducing path to aggregate the object information of the preceding layers to accurately predict spinal deformity region on fine-grained boundaries on the particular spatial locations.

Decoder Block 1

Converted feature convolution is employed with a 3*3 kernel and stride of 2 residual blocks to hold 256 feature maps of different spinal deformities and vertebra collapse. The output of the block is concatenated with the contraction block output containing the image's semantic feature. Skip connection is included with a drop rate of 1.0 to extract the feature map. A 2D Convolution layer is used with a 3*3 kernel size in addition to padding to keep input and output sizes the same. Finally, a sigmoid-based activation function is employed.

Decoder Block 2

Converted convolution is employed with a 3*3 kernel and stride of 2 residual blocks to hold a reduced 128 feature maps. The output of the block is concatenated with the contraction block output containing the image's semantic feature. Skip Connected is included with a drop rate of 1.0. A 2D Convolution layer is used with a 3*3 kernel size in addition to padding to keep input and output sizes the same. Finally, a sigmoid-based activation function is employed.

Decoder Block 3

Converted convolution is employed with a 3*3 kernel and stride of 2 residual blocks to hold a reduced 64 feature maps. The output of the block is concatenated with the contraction block output containing the image's semantic feature. Skip connection is included with a drop rate of 1.0 to obtain the feature map. A 2D Convolution layer is used with a 3*3 kernel size in addition to padding to keep input and output sizes the same. Finally, the sigmoid activation function is used.

Decoder Block 4

Converted convolution is employed with a 3*3 kernel and stride of 2 residual blocks to hold a reduced 32 feature maps. The output of the block is concatenated with the contraction block output containing the image's semantic feature. Skip Connection is included with a drop rate of 1.0 to obtain the feature map. A 2D Convolution layer is used with a 3*3 kernel size in addition to padding to keep input and output sizes the same. Finally, the sigmoid activation function is used.

Decoder Block 5

Converted convolution is employed with a 3*3 kernel and stride of 2 residual blocks to hold a reduced 8 feature maps. The output of the block is concatenated with the contraction block output containing the image's semantic feature. Skip Connection is included with a drop rate of 1.0 to obtain the feature map. A 2D Convolution layer is used with a 3*3 kernel size in addition to padding to keep input and output sizes the same. Finally, the sigmoid activation function is used [6].

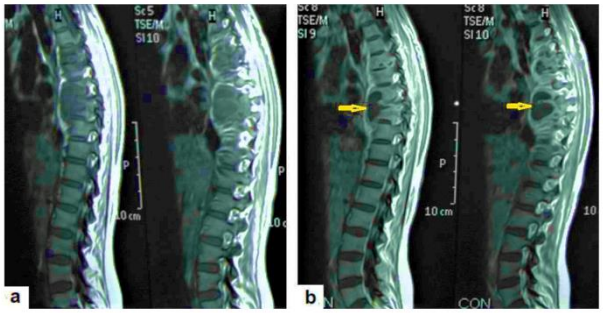

Fig. 4 Prediction of spinal tuberculosis using ResUnet architecture (a) Input Image (b) output Image

Finally, in the output layer, 1*1 convolutions are employed to minimize the resultant feature map to 1, and the sigmoid-based activation function is included in the output layer, which contains the high-level semantic image representation. Figure 4 represents the spinal tuberculosis prediction using ResUNet Architecture.

Convolution Layer

The convolution layer is the vital operation of the Neural Network. The convolution layer computes the kernel to be mapped to segmented regions of spinal deformities using the linear sliding window to determine the object boundaries.

The proposed architecture generates the feature map as semantic features using an adjacency matrix. The Semantic Feature Map of the region of interest is calculated for input image I of the jth convolutional layer of the model.

$$S_{ic} = f(w(c)_i * m)$$

Where S_{ic} convolution operation of adjacency matrix m considers the convolution kernels representing the boundaries and a considers the activation function representing the centroids of the region of interest.

$$W = [W_{i^1}, W_{i^1}, W_{i^1}, \ldots\ldots W_{i^k}]$$

K is the kernel of the layer.

Every kernel of the matrix $W_i^k$is is $M_i * M_j * N_n$ weight of the adjacency matrix M to the segmented region representing the layer's size.

Table 1. Hyperparameter tuning of the ResUnet

Parameter	Value
Tuning method	stochastic gradient descent
The learning rate of the spinalResUnet	0.1
Weight decay	0.002
Momentum	[0,1]
Batch size	10

Activation Layer

The activation layer uses the sigmoid as an activation function. The non-linear function computes the activated class for the disease types with kyphosis, gibbus formation and osteomyelitis.

Loss Function

The loss function of the SpineResUnet is composed of segmentation loss for the result of the decoder and activation loss of the region pooling module. It computes the discrepancy among the determined segmented results and is computed using the Stochastic Gradient Descent (SGD) algorithm.

$$L_s(f) = \frac{1}{n} \sum_{k=0}^{L-1} \{ \sum_x [yseg \log(p(i,j)) + (1 - yseg) \log(1 - c(i,i)) \}$$

Where $L_s(f)$ represents the Convolution layer weighting matrixes and fully connected layers, *n* represents the dataset containing training images, *i* is the index of the training image, and *k* is the kernel of the class index. Loss function *E* is represented as

$$E_k = E_{k-1} - \partial(\partial k(w) / \partial n$$

Where l is the learning rate of the spinalResNet, considered a vital parameter of the semantic image represented with depth, width and height. The *k* is the class index.

Hyper Parameter Tuning

SpinalResUnet models hyperparameters are tuned to obtain high recognition accuracy in this part. It changes the base learning rate. Model optimization is employed using the stochastic gradient descent technique. Further batch training methods will partition the training and testing images into numerous batches. Each fold is composed of 10 training images.

Algorithm: Spinal Tuberculosis Classification using Spinal ResUnet

Input: Segmented Region={S1, S2,…SN}
Output: Spinal Tuberculosis Disease Label T={T1, T2..TN}
Process
For (J= Feature Vector[i], J++ , j<threshold)
Compute convolution () for the segmented region of interest as 3*3 layers for $H_{ic} = f(w_i * x)$
Convolution value = Convolution Window (F)
SpinalResUnet
Compute pooling () for Unet
ResNet (convolution Value)
pooling value = Reduced Convolution Window containing Feature (F)
Dropout () for overfitting
Eliminate the feature with probability Map
Loss function ()
Apply Encoding and Decoding for Semantic feature representation
L(f)= Weighted matrix of the convolution on eliminating error
Hyperparameter tuning ()
Softmax () to generate the classes with Matrix Multiplication and element-wise division
Class Labels = disease region of the image.

4. Experimental Results

Experimental evaluation of the proposed model has been employed using an MRI dataset to classify spinal tuberculosis disease into different classes. Those classes are disease classes of the test images. The proposed architecture is evaluated in MATLAB. In this evaluation, the image dataset is partitioned into train, test and validate partitions such that 60 percent of the image has been utilized for training, 20 percent of the data has been assigned for validation, and the remaining 20 percentages of an image has been assigned for testing. For experimental validation, 5-fold result validations have been employed using a confusion matrix to enhance the performance of the classification and segmentation of the disease region.

SpineResUnet architecture of CNN is used in this work to produce higher disease recognition accuracy on the classification of spinal tuberculosis disease in the segmented regions of spinal deformity and collapse of the vertebra, etc. The proposed model outperforms with good results in computing the deformed spinal classes.

4.1. Performance Evaluation Metrics

The performance of the proposed architecture has been assessed with the Image Dice coefficient, specificity and sensitivity measure against the state-of-the-art approaches for spinal tuberculosis disease detection with volumetric changes of the MR images on various stages of the diseases.

4.1.1. Dice Similarity Coefficient

It is estimated by measuring the variation among the classified outcomes of the segment features. Further, the proposed segmentation outcome can be estimated through true positive, true negative, false positive and false negative values. It is represented as

$$\text{Dice Similarity} = \frac{2TP}{2TP + FP + FN}$$

Table 2. Performance evaluation of spinal tuberculosis classification techniques

MRI images fold	Technique	Dice Coefficient	Sensitivity	Specificity
Fold 1 validation images	Spinal ResUNet-Proposed	0.9998	0.9685	0.9992
	Spatial and temporal constrained convolution Neural Network (SPCCNN)	0.9988	0.9482	0.9992
Fold 2 validation images	Spinal ResUNet-Proposed	0.9997	0.9694	0.9999
	Spatial and temporal constrained convolution Neural Network (SPCCNN)	0.9977	0.9594	0.9998
Fold 3 validation images	Spinal ResUNet-Proposed	0.9999	0.9815	0.9995
	Spatial and temporal constrained convolution Neural Network (SPCCNN)	0.9971	0.9715	0.9985

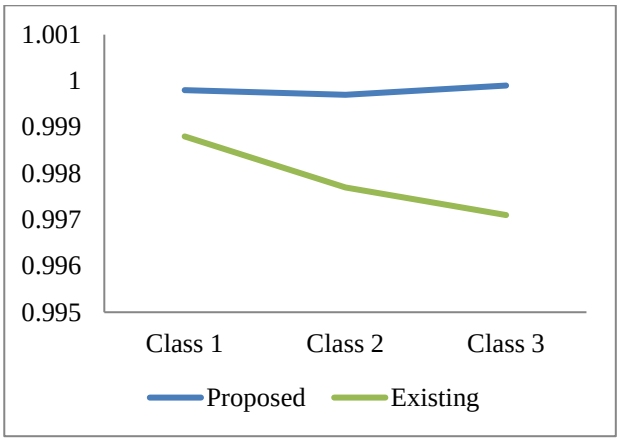

Fig. 5 Performance evaluation of dice coefficient

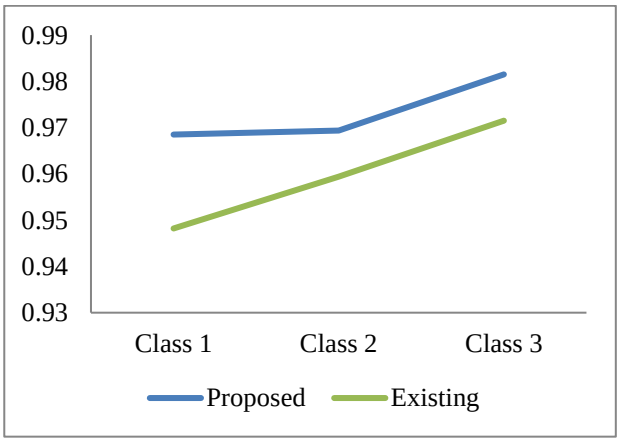

Fig. 6 Performance evaluation of the sensitivity

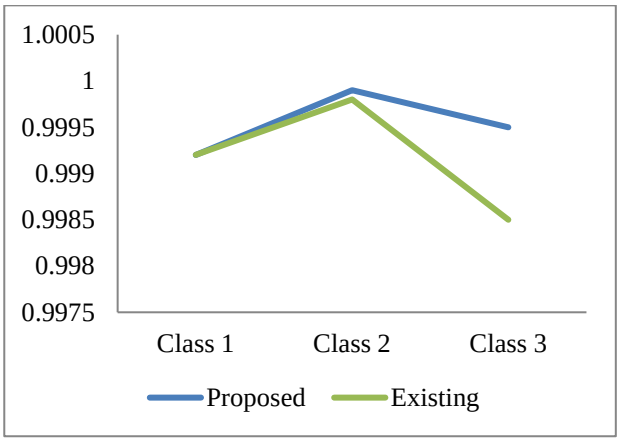

Fig. 7 Performance evaluation of specificity

Sensitivity

It computes the segment result on True Positive of the deformed spinal vertebrae on the semantic features of the segmented region. It is calculated as

$$\text{Sensitivity} = \frac{\text{TP}}{TP+FN}$$

Specificity

It is the computation of the segment result on the True negative of the deformed spinal vertebrae on the semantic features of the segmented region. It is calculated as

$$\text{Specificity} = \frac{\text{TN}}{TN+FN}$$

The analysis of different test images on the MRI dataset on determining the deformity type using SpinalResUnet has been computed with respect to dice coefficient, specificity and sensitivity, depicted in Table 2. The model is capable of segmenting the deformed regions accurately. SpineResUnet performs excellently compared to the existing SPCCNN approach—the proposed architecture results in increased performance [17].

The Dice Coefficient outcomes of the classes with high interclass and intraclass similarity are illustrated in Figure 5. On the measurement of the class sensitivity measure, vertebrae of the deformity are determined accurately. Figure 6 illustrates the outcome of the sensitivity on the collapse of vertebrae identification.

The sensitivity of the class outcome of the proposed Spinal ResUNet is highly capable of differentiating the vertebras. Figure 6 illustrates the performance outcome of the specificity metric.

For improving the model classification accuracy, the proposed architecture is enhanced by changing the size of the kernel using hyperparameter tuning.

5. Conclusion

This work has designed and simulated a novel framework named SpineResUnet towards spinal tuberculosis classification on 5-fold cross-validation. The structural and texture properties of vertebral disk patterns have been extracted as prominent features for abscess volume. The abscess volume is further reconstructed using

an adjacency matrix. Specifically, classification has been carried out on the stretched abscess volume.

The semantic feature has been processed further using the convolution network's ResNet and ReLU activation functions. The output layer generates the small size of the vertebra deformity and the complex appearance of vertebra constituents.

A saliency map has been used in the data visualization step to determine the tuberculosis characteristics and prediction of spinal tuberculosis location accurately.

References

[1] Cheng Chen et al., "Localization and Segmentation of 3D Intervertebral Discs in MR Images by Data Driven Estimation," *IEEE Transactions on Medical Imaging*, vol. 34, no. 8, pp. 1719–1729, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[2] Xiaomeng Li et al., "3D Multi-Scale FCN with Random Modality Voxel Dropout Learning for Intervertebral Disc Localization and Segmentation from Multi-Modality MR Images," *Medical Image Analysis*, vol. 45, pp. 41–54, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[3] Nikolas Lessmann et al., "Iterative Fully Convolutional Neural Networks for Automatic Vertebra Segmentation and Identification," *Medical Image Analysis*, vol. 53, pp. 142–155, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[4] Shuo Li et al., "MRLN: Multi-Task Relational Learning Network for MRI Vertebral Localization, Identification and Segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2902-2911, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Faezeh Fallah et al., "Simultaneous Volumetric Segmentation of Vertebral Bodies and Intervertebral Discs on Fat Water MR Images," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1692–1701, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[6] Sepp Hochreiter, and Jürgen Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997. [CrossRef] [Google Scholar] [Publisher Link]

[7] Zhongyi Han et al., "Spine-GAN: Semantic Segmentation of Multiple Spinal Structures," *Medical Image Analysis*, vol. 50, pp. 23–35, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[8] Tianyang Li et al., "S^3egAnet: 3D Spinal Structures Segmentation via Adversarial Nets," *IEEE Access*, vol. 8, pp. 1892–1901, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Hao Chen et al., "3D Fully Convolutional Networks for Intervertebral Disc Localization and Segmentation," *International Conference on Medical Imaging and Augmented Reality*, pp. 375–382, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[10] Anjany Sekuboyina, "Attention-Driven Deep Learning for Pathological Spine Segmentation," *International Workshop on Computational Methods and Clinical Applications in Musculoskeletal Imaging*, pp. 108–119, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[11] Rens Janssens, Guodong Zeng, and Guoyan Zheng, "Fully Automatic Segmentation of Lumbar Vertebrae from CT Images Using Cascaded 3D Fully Convolutional Networks," *IEEE 15th International Symposium on Biomedical Imaging*, pp. 893–897, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] Chuanbo Wang et al., "Fully Automatic Intervertebral Disc Segmentation using Multimodal 3D U-Net," *2019 IEEE 43rd Annual Computer Software and Applications Conference*, vol. 1, pp. 730–739, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[13] Nassim Guerroumi et al., "Automatic Segmentation of the Scoliotic Spine from Mr Images," *2019 IEEE 16th International Symposium on Biomedical Imaging*, pp. 480–484, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] Jayashree Shedbalkar, K. Prabhushetty, and R.H. Havaldar, "UNet and Transformer-Based Model for Multi-Modality Brain Tumor Segmentation," *SSRG International Journal of Electronics and Communication Engineering*, vol. 10, no. 8, pp. 22-35, 2023. [CrossRef] [Publisher Link]

[15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[16] Liang-Chieh Chen et al., "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," *Proceedings of the European Conference on Computer Vision, 2018*, pp. 801–818, 2018. [Google Scholar] [Publisher Link]

[17] Ke Gong et al., "Graphonomy: Universal Human Parsing via Graph Transfer Learning," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7450–7459, 2019. [Google Scholar] [Publisher Link]

[18] Anjany Sekuboyina et al., "A Localisation-Segmentation Approach for Multi-Label Annotation of Lumbar Vertebrae using Deep Nets," *arXiv*, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[19] Thomas N. Kipf, and Max Welling, "Semi-Supervised Classification with Graph Convolutional Networks," *arXiv*, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[20] M. Nisha et al., "Automatic Hippocampus Segmentation Model for MRI of Human Head through Semi-Supervised Generative Adversarial Networks," *Journal Neuro Quantology*, vol. 20, no. 6, 2022. [Google Scholar]

[21] Kaiming He et al., "Deep Residual Learning for Image Recognition," *In 2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778, 2016. [Google Scholar] [Publisher Link]