IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 April 2024, accepted 19 July 2024, date of publication 29 July 2024, date of current version 6 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3434592

RESEARCH ARTICLE

Dimension Reduction Using Dual-Featured Auto-Encoder for the Histological Classification of Human Lungs Tissues

AMNA ASHRAF[1,2], NAZRI MOHD NAWI[2], TARIQ SHAHZAD[3], MUHAMMAD AAMIR[4], MUHAMMAD ADNAN KHAN[5], (Senior Member, IEEE), AND KHMAIES OUAHADA[3], (Senior Member, IEEE)

[1]Department of Artificial Intelligence, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan
[2]Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), Parit Raja, Johor 86400, Malaysia
[3]Department of Electrical and Electronic Engineering Science, University of Johannesburg, Johannesburg 2006, South Africa
[4]School of Electronics, Computing and Mathematics, University of Derby, DE22 3AW Derby, U.K.
[5]Department of Software, Faculty of Artificial Intelligence and Software, Gachon University, Seongnam-si 13557, Republic of Korea

Corresponding authors: Tariq Shahzad (tariqshahzadd@gmail.com) and Muhammad Adnan Khan (adnan@gachon.ac.kr)

This work was supported by the Research Fund from the University of Johannesburg, Johannesburg, South Africa.

ABSTRACT Histopathology images are visual representations of tissue samples that have been processed and examined under a microscope in order to establish diagnoses for various disorders. These images are categorized by deep transfer learning due to the absence of big annotated datasets. There are some classifiers such as softmax and Support Vector Machine (SVM) used to perform multiple and binary classification respectively. Feature reduction for high dimensional images, is an emerging technique which can meet two basic criteria's of classification i.e. it deals with over-fitting issue and it can also incredibly increase the classification accuracy. As disease diagnosis requires accurate histopathological image classification, so the proposed Dual Featured Auto-encoder (DFAE) based transfer learning is introduced with Triple Layered Convolutional Architecture. The Histological CIMA dataset is used after pre-processing by PHAT, a mathematical and computational framework to get spatial features as well as spectral features. In order to achieve the two objectives, the proposed integrated methodology uses reduced informative features from DFAE and fed them to Triple Layered Convolutional Architecture (TLCA). The conventional Convolutional Neural Network (CNN), ResNet50, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are also tested against reduced dimensional image data but we found moderate or even low accuracies i.e. 25% for DFAE-ResNet50, 66% for DFAE-LSTM, 33% for DFAE-GRU and 67% for DFAE-CNN. While the accuracy of our proposed architecture Dual Featured Auto-encoder with TLCA (DFAE-TLCA) is better i.e. 96.07%. The proposed methodology has the potential to revolutionize the medical research.

INDEX TERMS Feature extraction, high-resolution image, histological slide image classification, triple layered convolutional architecture.

I. INTRODUCTION

In digital pathology, glass slides are scanned using a whole-slide image scanner, and the digital images are then examined using an image viewer, generally on a computer screen or mobile device [1]. A crucial step in identifying and comprehending various lung disorders, such as lung cancer [2] and pulmonary fibrosis; Idiopathic Pulmonary Fibrosis (IPF) and Fibrotic Hypersensitivity Pneumonitis (FHP), are the histological classification of human lung

The associate editor coordinating the review of this manuscript and approving it for publication was Yudong Zhang.

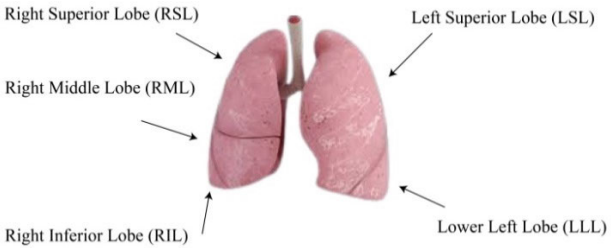

FIGURE 1. Lungs division into five lobes [6].

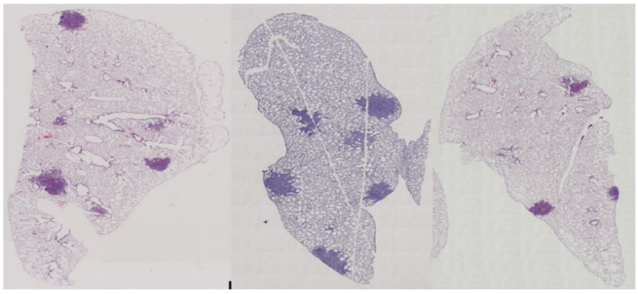

FIGURE 2. Internal morphology of lung's lobes [12].

tissues [3]. The proper classification of lung tissues is essential for enhancing patient outcomes and directing effective treatment plans. Computer-aided diagnosis systems have become important tools to help pathologists in their diagnostic decision-making process and image analysis techniques. The structure of the right and left lungs differs in structure as shown in Fig. 1. Right Superior Lobe (RSL), Right Middle Lobe (RML), and Right Inferior Lobe (RIL) are the three lobes that make up the right lung [4]. The Left Superior Lobe (LSL) and the Left Lower Lobe (LLL) are the two lobes that make up the left lung.

Lung tissue samples' inherent complexity and heterogeneity represent one of the major difficulties in histological classification. These samples frequently display variances in tissue architecture, cellular morphology, and staining patterns, making it challenging to extract useful information for precise classification [5]. Internal morphological complexity can be easily perceived thorough Fig. 2 taken from dataset ''Histology (CIMA)''. These tissue slices are stained with several different stains such that Clara cell 10 protein, Prosurfactant protein C, Hematoxylin and Eosin, Antigen KI-67, Platelet endothelial cell adhesion molecule, Human epidermal growth factor receptor 2, Estrogen receptor and Progesterone receptor. Additionally, pathologists' manual evaluation of histological slides requires a lot of time and is subject to inter-observer variability.

Researchers have automated the classification process using machine learning and deep learning techniques to overcome these difficulties. Convolutional Neural Networks (CNNs) have demonstrated astounding success in image classification tasks in recent years [7], [8]. However, when training deep learning models, the high dimensionality of histological images places a large computational strain on the system and raises the possibility of over-fitting. Techniques for dimensionality reduction present a potential answer to these problems. These methods enable effective processing and analysis while maintaining the critical discriminatory information by converting the high-dimensional histology pictures into a lower-dimensional representation [9]. Dimension Reduction based CNN is one such method that combines the advantages of auto-encoders and deep learning classification methods. Similar strategies such as CNN based auto-encoder for breast cancer detection [10] and for COVID-19 patient survival chances [11] have been proposed in 2021.

In this study, we present histological classification of human lung tissues using dimension reduction technique based on an Image Super-Resolution Auto-encoder (ISRAE). By identifying useful features from a lower-dimensional space, our method attempts to address the issues raised by high-dimensional histology pictures. We predict that this method will increase the classification of lung tissue's precision, effectiveness, and interpretability and will be applicable to other medical researches of histological image classification problem. The major objectives of this research are as follows; to address spectral and spatial both type of features while performing image classification for better accuracy. Second is to apply an ISRAE for successful dimension reduction of histological pictures while preserving crucial structural and textural information which leads data to best fit for classification.

The rest of the paper is structured as follows. We give a thorough overview of related work in dimensionality reduction and histological image analysis in Section II of this article. The technique, including the design of the image dual featured auto-encoder and the training process, are covered in Section III. We outline the experimental design and go over the findings in Section IV. Finally, we review the results, evaluate their implications, and offer suggestions for further research in Section V.

II. METHODOLOGY

High dimensional image datasets are usually considered as challenging due to their high resolution, visual discrepancies and a dearth of objects with a distinguishable appearance [13]. These challenges affect the classification performance of images. The work done is this research is basically categorized in the two main steps. In the first step, Dual featured auto-encoder is used to reduce the dimensions of data. While in the second step, classification is performed by state of the art algorithms. The workflow for this research is depicted by Fig. 3 which reveals that after parameters initialization, spectral features obtained from data preprocessing and the spatial features obtained by Persistent Homology Algorithm Toolbox (PHAT) are passed to ISRAE. The functionality of auto-encoder is based on Image-Super-Resolution-Auto encoder. The compressed images obtained are then classified by traditional models of classification such as Triple Layered Convolutional Architecture (TLCA),

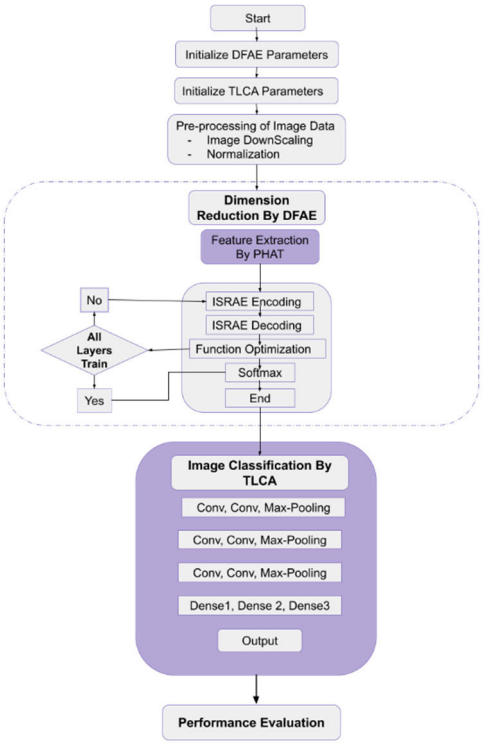

FIGURE 3. Proposed framework.

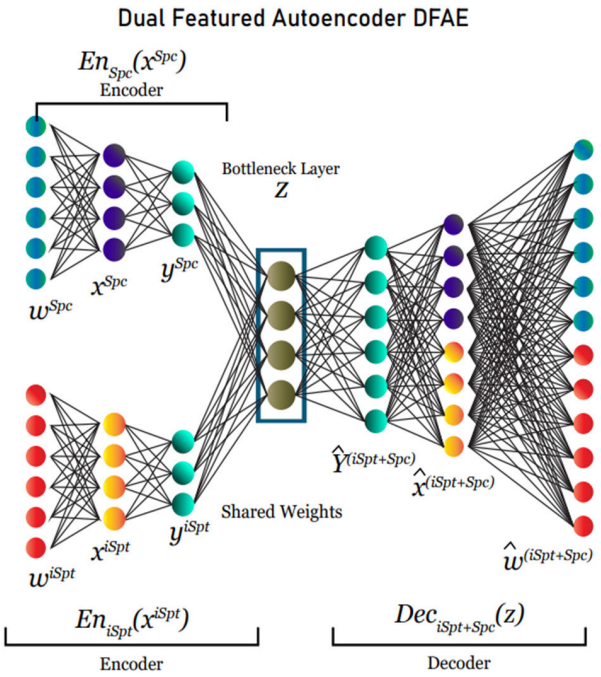

FIGURE 4. Dual featured auto-encoder (DFAE).

CNN, Capsule Network (CapsNet), Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU).

A. DUAL FEATURED AUTO-ENCODER

To accomplish unsupervised learning and dimensionality reduction, neural networks such as auto-encoders [14], [15] are frequently employed. An encoder network and a decoder network learn a compressed version of the input data [16] by coordinating their operations. Input to the auto-encoder, consists of both spectral data and spatial data as it is fed by both spectral 'X_{Spc}' and spatial features 'X_{iSpt}' as shown in the Fig 4. Spectral features are the information about a signal's or an image's frequency content, whereas information about its spatial organization or structure is known as spatial features.

PHAT is a mathematical and computational technique; that can be used to examine the topological characteristics of a dataset. When standard methods, such as principal component analysis [17] or t-SNE [18], fail to adequately capture the underlying structure of the data, this method performed well to gain insight into its form and organization. The topological characteristics of data at different scales are the primary focus of persistent homology. The utilization of PHAT can be particularly advantageous in extracting and analyzing intricate topological characteristics, particularly in the case of high-resolution photographs. Persistent homology-based high-resolution image analysis finds utility in several domains such as medical imaging e.g., tumor identification [19], materials science e.g., microstructure analysis [20], environmental science e.g., land cover classification [21], and others. Conventional image processing techniques may struggle to identify the intricate structures and tiny features present in high-resolution photos. Persistent homology enables the acquisition and examination of the spatial and topological attributes of these images.

Topological features extracted with PHAT can be employed in the context of spatial data analysis, such as image analysis or computer vision, to reveal the underlying spatial relationships and structures. These capabilities may prove useful when trying to comprehend complicated data sets with sophisticated spatial characteristics.

In order for an auto-encoder to make use of both spectral and spatial properties, the input data must be formatted in a way that maintains the data's spatial arrangement while also retaining spectrum information. Input data from multiple channels, each of which represents a distinct spectral band or feature, is used for this purpose. The ISRAE [22] is then trained to encode and decode this multi-channel input, compressing it while preserving its spectral and spatial features. An auto-encoder is a neural network specifically designed for unsupervised learning and the task of reducing the dimensionality of data. The system basically comprises of two primary components: an encoder and a decoder. The weights of the encoder and decoder are changed iteratively until the model can effectively encode and decode data to get output 'Z' with minimized information loss. Fig. 4 shows the transformation of weights from '$w^{spc} + w_i^{spc}$' to '$x^{spc} + y^{spc}$' i.e. the sum of spectral and improved spatial features for encoder section. Similarly, weights transformation for spectral and improved spatial features for decoder section can also be seen. The encoder reduces the dimensionality of the input data by consecutive layers of convolutional and

TABLE 1. Pseudo code proposed DFAE.

Algorithm 1 (DFAE)

Parameters initialization
- DataFrame : *df*
- DownScaled size of Image : *input_downsize*
- PHAT : *phat*

Image data processing loop

For each image in *df*
- Open and resize: *input_downsize*
- Convert to NumpyArray: *X-data*

End

Extract Features using PHAT
- Convert *X-data* to point clouds: *point_clouds*
- Compute persistent homology and extract persistent diagram : *X_Diagram*

ISRAE (input: *X_Diagram***)**

Encoding
- Compute Convolutional encoded input $f_1(w)$ by all feature map values f[i, j]
- Compute Convolutional encoded input $f_2(w)$ using $f_1(w)$ as input
- Compute Maxpooling encoded input $f_3(w)$ using $f_2(w)$
- Compute Convolutional encoded input $f_4(w)$ by all feature map values f[i, j]
- Compute Convolutional encoded input $f_5(w)$ using $f_4(w)$ as input
- Compute Maxpooling encoded input $f_6(w)$ using $f_5(w)$
- Compute Convolutional encoded input: *encoded* by $f_6(w)$

Decoding
- Compute UpSampling output $f_6(w^o)$ by *encoded*
- Compute Convolutional decoded output $f_5(w^o)$ using $f_6(w^o)$ as input
- Compute Convolutional decoded output $f_4(w^o)$ using $f_5(w^o)$ as input
- Perform addition operation to $f_4(w^o)$*and* $f_5(w)$: $f_{add\text{-}1}(w^o)$
- Compute UpSampling output $f_3(w^o)$ by $f_{add\text{-}1}(w^o)$
- Compute Convolutional decoded output $f_2(w^o)$ using $f_3(w^o)$ as input
- Compute Convolutional decoded output $f_1(w^o)$ using $f_2(w^o)$ as input
- Perform addition operation to $f_2(^o)$*and* $f_1(w^o)$: $f_{add\text{-}2}(w^o)$
- Compute Convolutional decoded output: *decoded* by $f_{add\text{-}2}(w^o)$

Optimization
- Optimize error value

While (All layers trained)

return *output*

max-pooling, while the decoder tries to recreate the original input data by alternative layers of up-sampling and convolutional, based on this reduced representation 'Z'. The network is trained to limit the reconstruction error, thereby acquiring a compressed representation of the data. Below is the pseudo code for DFAE.

The target to get reduced image size is (224, 224, 3), which is accomplished by ISRAE. Where '224 × 224' represents the reduced dimensions while 3 is number of channels used to denote a color image consists of Red, Green, and Blue (RGB). The pseudo code of the model DFAE is given in the Table 1 which reflects feature extraction by PHAT and the 6 layers of ISRAE encoder, encoder's output layer i.e. the bottle neck 'Z' or the reduced output and then there are the 6 layers of decoder.

The different layers of ISRAE are illustrated in the Fig. 5. Encoder part of auto-encoder represented by left side starts from Layer 0 and ends by Layer 7. Likewise, decoder part of auto-encoder represented by right side starts from Layer 8, by taking Layer 7 as an input and ends at Layer 16. The very first layer is the input layer of resized image having 224 by 224 dimensions and 3 channels. The three channels are the Red, Green and Blue (RGB) channels in the context of images. The intensity of these color channels is represented by three values in each pixel in the input image. The generated model summary authenticates the channel information and the number of parameters used. As we can see in the summary, the model can adjust to different input sizes by using 'None' in the spatial dimensions (height and width). Because of its adaptability, input photos of various sizes or resolutions can be handled without the need for intentional data reshaping or resizing. The next two layers: Layer 1 and Layer 2 are the 2D-convolutional layers having 64 filters and 262144 neurons. Filter size i.e. height and width of filter is 3 and 3. These 3 × 3 filters glide across the input data during the convolution process to perform element-wise multiplication and summing, producing feature maps. The activation function employed is the Rectified Linear Unit (ReLU). It creates non-linearity by passing positive values through unaltered thresholding at zero for negative values. The number of parameters for these layers are 1792 and 36928 respectively. The equation (1) illustrate the mathematical formulation to calculate them where F, kH, kW and C represents no of filters, kernel height, kernel width and number of channels.

$$\textbf{\textit{Number of Params}} = \boldsymbol{F} \times [(\boldsymbol{kH} \times \boldsymbol{kW} \times \boldsymbol{C}) + \mathbf{1}] \quad (1)$$

Layer 3, or 2D-MaxPooling layer with a pool size of 2 × 2, is a down-sampling procedure used to minimize the spatial dimensions of the input with the goal of preserving the most significant information. It lowers the computational complexity and proceeds by choosing the maximum value within each designated pooling window. With the 'padding=='same'' parameter, it is assured that dimensions of the output and input are the same (zeros are inserted around the input if required). Layer 4 is the 2D-convolutional layer having 128 filters and 65536 neurons. The same filter size is used as in previous convolutional layers. Similarly, each layer's type, number of neurons, filter size or pool size and number of channels are described in the architectural diagram (Fig. 5). Layer 7 has a great importance as it is an encoded image output having dimensions of 224 × 224. Moreover, it's a 2D-convolutional layer with number of neurons 16384, number of filters 256, filter size of 3 × 3 and trainable parameters 295168. The sequence of layers in the decoder side is reversed as in encoder part, but 2D-Maxpooling is replaced with 2D-UpSampling layer.

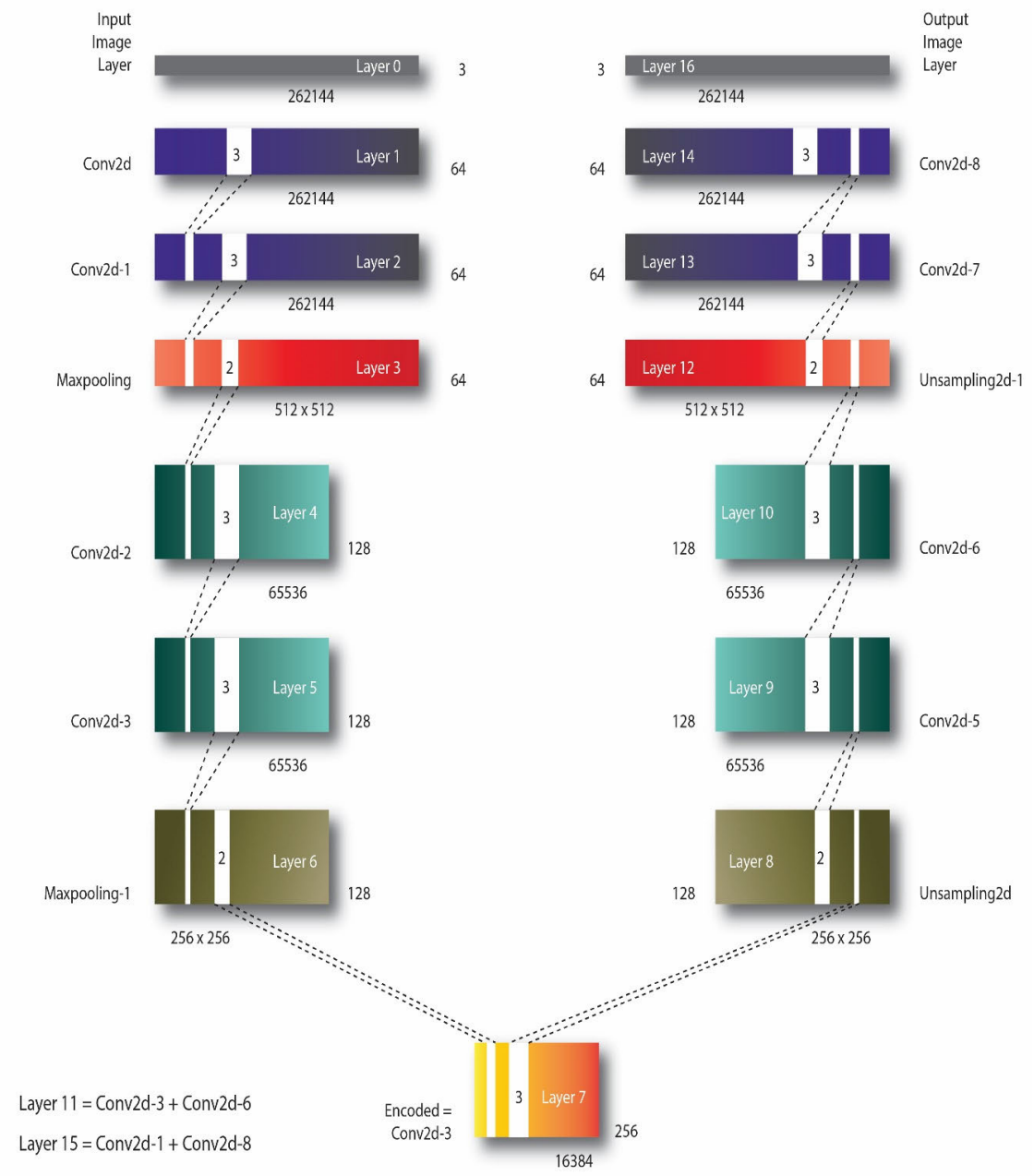

FIGURE 5. Architecture for image super resolution auto-encoder (ISRAE) [22].

The Learning Rate (LR) is an essential hyperparameter in the training of any neural network whose value has a great influence on the convergence and performance of the model. It is usually varied across several epochs to enable efficient optimization. At the start of training process, we see very high LR up to 6 epochs (Fig. 7), which enables the model to explore the parameter space and rapidly reduce the loss by making large modifications to its parameters. Through the examination of the learning rate curve, we gain insight into the rate at which a model is converging. A prominent initial reduction in loss, followed by a subsequent steadier decline, suggests a favorable convergence rate. In our case, we set it to 0.001. Other hyper parameters include batch size, epochs and target shape. These specifications are illustrated in Table 2.

B. IMAGE CLASSIFICATION

The reduced images from dual featured auto-encoder are used to classify them and the classification performance is evaluated on standard parameters such as accuracy, precision, recall and F1-measure [24]. Each of these serves for different purposes.

The ratio of accurately predicted occurrences to all of the dataset's instances is known as accuracy. When there are roughly equal numbers of cases in each class in the dataset,

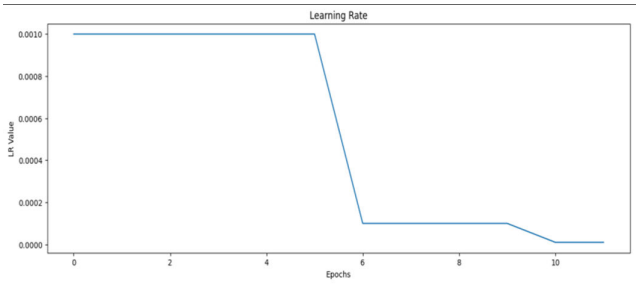

FIGURE 6. ISRAE learning rate over different no of epochs.

TABLE 2. Hyper-parameter specifications for ISRAE.

Hyper parameters	Values
Learning Rate	0.001
Epochs	200
Batch size	20
Target shape	224, 224
Down sample shape	80, 80

accuracy becomes valuable. In datasets that are unbalanced and exhibit a dominant class, it may be deceptive.

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} \quad (2)$$

Precision is another evaluation parameter that measures the positive predictions. It is the proportion of all positively anticipated instances to all accurately predicted instances.

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

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

$$F1 - Score = \frac{2 \times (precision \times recall)}{Precision + Recall} \quad (5)$$

Recall quantifies the model's accuracy in identifying positive examples. It is the proportion of all real positive instances to all correctly projected positive instances. Exceptional recall is desirable in medical diagnostics to guarantee that all true positive cases are identified whereas the harmonic mean of recall and precision is the F1 score. It combines recall and precision into one statistic. Image classification techniques are explained here in detail.

1) CNN-LENET5

Convolutional neural networks (CNNs) have emerged as a prevalent and efficacious technique for image classification in computer vision tasks, including image categorization. CNNs are well-suited for image classification tasks due to their ability to analyses and extract characteristics from images. Hyper-parameters of the Convolutional Neural Network (CNN) or any other deep learning algorithm, such as the learning rate, batch size, and regularization techniques (e.g., dropout) are required to be configured favorably and appropriately.

TABLE 3. CapsNet model summary.

Layer (type)	Output Shape	Param #
conv2d_3 (Conv2D)	(None, 26, 26, 128)	1280
reshape_3 (Reshape)	(None, 676, 128)	0
capsule_3 (Capsule)	(None, 28, 16)	57344
flatten_3 (Flatten)	(None, 448)	0
dense_6 (Dense)	(None, 128)	57472
dense_7 (Dense)	(None, 4)	516

Total params: 116612 (455.52 KB)
Trainable params: 116612 (455.52 KB)
Non-trainable params: 0 (0.00 Byte)

Convolutional layers serve as the fundamental components of Convolutional Neural Networks (CNNs). The application of a collection of learnable filters, often referred to as kernels, to the input image enables the process of local feature extraction. The filters undergo a sliding operation across the image, performing dot product computations, resulting in the generation of feature maps that accentuate certain characteristics of the image. After convolutional layers, pooling layers are commonly incorporated in order to decrease spatial dimensions while preserving significant characteristics. Pooling methods commonly used in neural networks include max pooling and average pooling which involves selecting the maximum and average value respectively within a given region. The values of the resulting feature map can be computed using the convolution formula shown below:

$$G\,[m, n] = (f * h)\,[m, n] = \sum_{j}\sum_{k} h\,[j, k]\, f[m-j, n-k] \quad (6)$$

Let f represent the input image, h represent the filter, and m and n represent the indices of rows and columns of the resulting matrix. The operator is a distinct form of matrix multiplication.

2) CAPSULE NETWORK

The Capsule Network, which was introduced by Geoffrey Hintonin 2012 [25] and modified in the later years [26], [27], [28], [29] has gained significant popularity as a convolutional neural network (CNN) model for the purpose of image classification. Neural clusters known as capsules are responsible for encoding information regarding the spatial connections between individual features in a picture in addition to the features themselves. Every capsule represents a particular aspect of an object along with its attributes including texture and position.

Dynamic routing is one of the fundamental ideas of capsule networks. Based on the concordance between their predictions and the actual output, this procedure enables capsules in one layer to communicate and convey information to pertinent capsules in the succeeding layer. In the data, this aids in preserving spatial links and hierarchical structures. By using a routing algorithm, capsule networks can adjust the

weights between capsules according to how well or poorly the predictions made by one capsule match the output of another capsule in a higher layer. The comprehension of spatial hierarchies is enhanced and the predictions are improved through this iterative approach. The network evaluates the agreement between capsules and their predictions to get a final classification after several iterations of dynamic routing. The summary of model is given in the Table 3.

3) LONG SHORT-TERM MEMORY (LSTM)

LSTM is an advanced form of a recurrent neural network (RNN) architecture [30]. The shortcomings of conventional RNNs in capturing long-term dependencies in sequential data are addressed by LSTMs.

In conventional RNNs, data is processed by a series of recurrent units [31], each of which processes data while keeping a hidden internal state. Standard RNNs, on the other hand, have the vanishing gradient problem, which makes it challenging for them to detect long-term dependencies. As a result, while processing sequential data, they have trouble remembering information from prior time steps. LSTMs [32] were developed to deal with this problem. Compared to conventional RNNs, they feature additional parts called gates and a more sophisticated structure. The LSTM may selectively retain or forget information thanks to these gates, which regulate the flow of information into and out of the memory cell.

LSTMs are generally made for processing sequential data i.e. text or time series data and rarely used for picture classification. Nevertheless, it is possible that the integration of their sequential information alongside image features [33] can offer further understanding or enhance performance, particularly in tasks that involve films [34], image sequences, or time-dependent visual data.

4) GATED RECURRENT UNIT (GRU)

GRU is a specific form of recurrent neural network (RNN) structure that is frequently employed for a range of tasks involving sequences, such as classification [35], [36], [37]. It is a modified version of the conventional recurrent neural network (RNN) that is especially adept at processing sequential data. Based on a single GRU, a small and efficient model for spectral-spatial classification on hyperspectral images have been proposed [38]. This method allows the spatial information to be fused as the GRU's initial hidden state, and the core GRU can learn spectral correlation within a full spectrum input. In this manner, a single GRU calculates and expands both spectral and spatial information simultaneously. The model's effectiveness can be evaluated on the test set by measures such as accuracy, precision, recall, F1-score, and confusion matrices.

5) TRIPLE LAYERED CONVOLUTIONAL ARCHITECHTURE (TLCA)

Transfer Learning using convolutional architectures (TLCA) proposed in our previous research [39] has been extremely efficient in image recognition tasks. TLCA has significantly improved the performance of large-scale image datasets. TLCA is an advanced version of Convolutional Neural Networks (CNNs) and excels in understanding local and global structural features in images.

TABLE 4. Pseudo code TLCA.

Algorithm3 TLCA
Parameters initialization
-DataFrame : *df*
-Kernel size of convolutional layer : k_c
-Kernel size of maxpooling layer : k_m
-Activation function : *AF*
-DownScaled size of Image : *input_downsize*
-Biased values for fully connected layers: [b_1, b_2, b_3, b_4]
-Weight values for fully connected layers: [w_1, w_2, w_3, w_4]
Image data processing loop
For each image in *df*
Open and resize: *input_downsize*
Convert to NumpyArray: *X-data*
Function TLCA(input) → output:
Conv1-output = Convolutional Layer (32, k_c, *AF*, *input_downsize*)
Conv2-output = Convolutional Layer (64, k_c, *AF*, *Conv1-output*)
Pooled1-output = MaxPooling Layer (64, k_m, *Conv2-output*)
Conv3-output = Convolutional Layer (64, k_c, *AF*, *Pooled1-output*)
Conv4-output = Convolutional Layer (128, k_c, *AF*, *Conv3-output*)
Pooled2-output = MaxPooling Layer (128, k_m, *Conv4-output*)
Conv5-output = Convolutional Layer (256, k_c, *AF*, *Pooled2-output*)
Conv6-output = Convolutional Layer (128, k_c, *AF*, *Conv3-output*)
Pooled3-output = MaxPooling Layer (128, k_m, *Conv4-output*)
flattened_output = Flatten(*Pooled3-output*)
fc1_output = FullyConnectedLayer(*flattened_output*, w_1, b_1)
fc2_output = FullyConnectedLayer(*fc1_output*, w_2, b_2)
fc3_output = FullyConnectedLayer(*fc2_output*, w_3, b_3)
output = FullyConnectedLayer(*fc3_output*, w_4, b_4)
return *output*

The complete design of the framework is shown in Figure 3 where the first layer consists of a convolutional layer with 32 output channels and a 3×3 kernel. The second segment of the first layer has a convolutional layer with 64 output channels and a similar kernel size. The third segment of the first layer incorporates a max-pooling layer with a 2×2 kernel. This configuration is repeated three times.

In the subsequent layers, a fully connected architecture is maintained, with neurons in the pattern of 73728-1024-512-64-4. Pseudo code for TLCA is given in Table 4. The network is assigned the task of acquiring extensive-scale or high-level features through learning. The network's architecture, which is composed of three layers, demonstrates remarkable proficiency in achieving image classification objectives. However, the large number of parameters in original image dataset that need to be acquired are potentially result in overfitting; so by using the compressed images via ISRAE, the task of classification is accompanied by improved accuracy with overfitting avoidance. It is encouraging to note

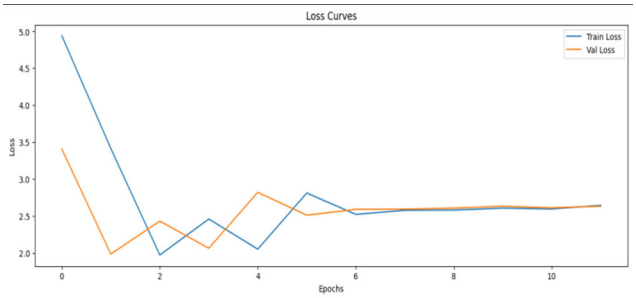

FIGURE 7. DFAE loss curve over no of epochs.

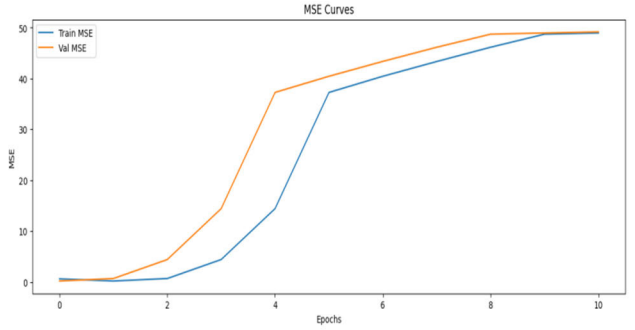

FIGURE 8. DFAE mean square error over no. of epochs.

TABLE 5. Evaluation of image classification techniques using reduced image dataset by DFAE.

Algorithm	Test Accuracy	Test Loss	Precision	Recall	F1-score
DFAE-CNN	50.00%	1.35	33.33%	50.00%	37.50%
DFAE-TLCA	**96.07%**	**0.73**	**94.00%**	**92.15%**	**91.96%**
DFAE-CapsNet	60.78%	0.76	78.31%	60.78%	53.86%
DFAE-LSTM	23.52%	1.38	05.53%	23.52%	08.96%
DFAE-GRU	25.49%	1.36	05.48%	25.52%	09.96%

that remarkable results can be obtained by TLCA utilizing a configuration of 20 epochs with a batch size of 32.

III. RESULTS

The process of histologically classifying human lung tissues involves analyzing tissue samples under a microscope in order to distinguish between various features. These features from high resolution images are reduced by auto-encoder in the first phase and then the reduced features are utilized to classify the histological images. Depending on the precise goals of the research, the results are evaluated as auto-encoder's performance and image classification performance.

A. PERFORMANCE EVALUATION OF AUTO-ENCODER

Auto-encoder is trained to develop a compact representation of input data that captures key properties from both the spectral and spatial domains. As it is helpful in all the tasks that require both spectral and spatial information, we found its application in image reconstruction, and feature extraction followed by image classification.

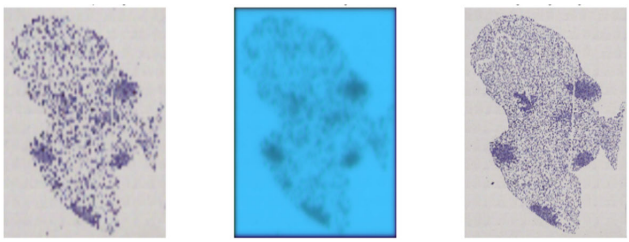

FIGURE 9. Three different illusions of an image of lung tissues (i) Downscaled by resize function, (ii) Downscaled by ISRAE and (iii) Original image.

The loss curve of an auto-encoder offers significant insights into the efficacy of its learning process in encoding and decoding data. Initially, there is a significant level of loss, which is expected to diminish progressively as the model acquires the ability to effectively recreate the input data (shown in the Fig. 7). The practice of monitoring the loss curve helps in the identification of over-fitting. With the persistent decrease in training loss while the validation loss exhibits an increase or remains constant, it indicates that the model is exhibiting over-fitting. However, in our case, convergence of training loss and validation loss shows no over-fitting. Mean Square Error Curve (MSE) also tells the same story as loss curve portrays (Fig. 8).

A reduced image by DFAE, its original image (input version) and the downscaled one (Fig. 9 (ii), (i) and (iii) respectively) has the dimensions of reduced super resolution image or the predicted image and downscaled image are the 224 × 224 and 80 × 80 respectively. To see how CNN behaves to both, we can observe the test accuracy of CNN and ISRAE-CNN in Table 5. In the first case the downscaled image is used to train, test and evaluate the model for classification and in later one the predicted Super Resolution image is used to train test and evaluate the CNN.

B. PERFORMANCE EVALUATION OF IMAGE CLASSIFICATION

The performance of DFAE-TLCA in Table 5 represents the output of TLCA when applied to reduced dimensional data from Dual Featured Auto-encoder. Its outstanding performance in terms of accuracy 96%, precision 94%, recall 92% and F1 measure 91% is leading DFAE-CNN, ISRAE-CapsNetwork, DFAE-LSTM and DFAE-GRU. Among all other methods applied, capsule network perform well on reduced data by DFAE and is comparable to best one i.e. DFAE-CNN. In this experiment, the lowest performance is from DFAE-GRU with test accuracy of 25.49%, precision of 05.48%, recall of 25.52% and F1-score of 9.96%.

DFAE-TLCA proved to be the most promising solution technique to our research objective 1. The graph of convergence in Fig. 10 for this technique shows how training and validation accuracy curves are overlapping each other, which provides insights towards our second objective. The reduced image data from DFAE means less number of features to learn. The DFAE-CNN (Fig. 10-i) contrary to DFAE-TLCA

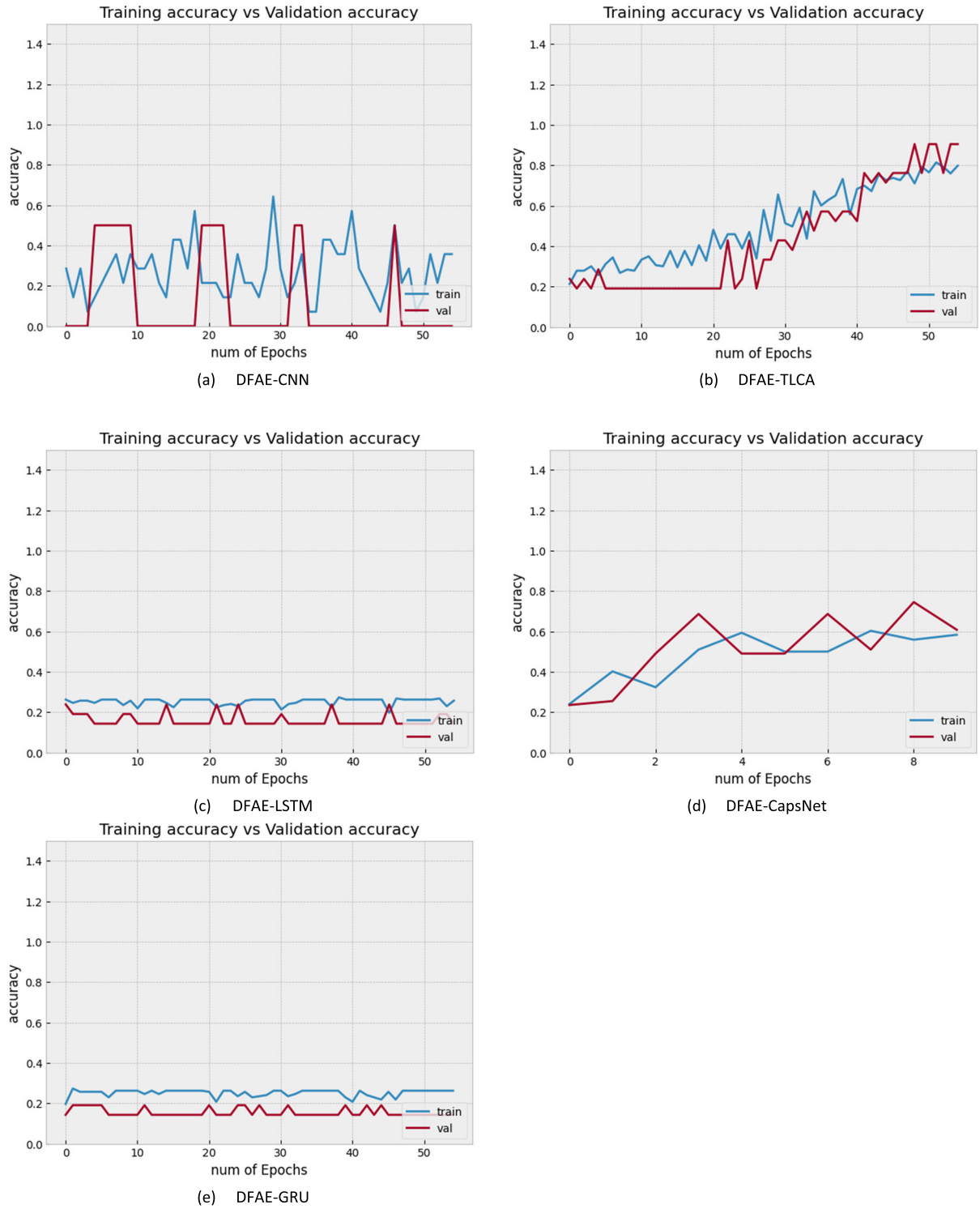

FIGURE 10. Graph of convergence for (i) DFAE-CNN, (ii) DFAE-TLCA, (iii) DFAE-LSTM, (iv) DFAE-CapsNet and (v) DFAE-GRU.

(Fig. 10-ii), is showing poor data fitting. The reason behind is a basic CNN lacks the depth or aptitude to identify intricate patterns in the data. Extra convolutional layers, filters, pooling, or thick layers can enhance the model's capacity to fit better and help it capture more complex features but upto a certain level and it will also add up to model's complexity.

IV. DISCUSSION

The performance of convolutional neural networks, advance neural networks TLCA, Capsule Networks, Long Short-Term Memory, and Gated Recurrent Units depends on their applicability, benefits, and possible drawbacks when using compressed input images from an auto-encoder.

Capsule networks were developed in order to address some shortcomings of conventional convolutional neural networks [40] in managing hierarchical relationships and spatial hierarchies inside images. CNNs are unable to identify an image's pose, texture, and deformations [41] which lead us to use TLCA. Unlike CNNs, which concentrate only on specific pixel patterns [42], capsule networks try to comprehend the spatial relationships between different elements of an object in an image. So capsule networks have the potential to perform better than simple CNN especially in certain situations like image categorization tasks. Translation invariance is a feature of CNNs [43], [44] that allows them to identify objects regardless of where they are in the image. This feature is crucial for a lot of image classification applications when an object's exact location could change. Conversely, Capsule Networks have trouble with translation invariance and need further architectural changes or routing techniques to deal with it. Moreover, hyperparameters and architectural design decisions have an impact on capsule networks. It can be challenging to find the ideal architecture and hyperparameter values, and the performance gains over more straightforward designs like CNNs are not always guaranteed.

Performance of image classification is enhanced undoubtedly by compressing first and then reconstructing images using an auto-encoder before feeding them into a triple layer convolutional neural network (TLCA). This method can be applied to jobs where the auto-encoder learns a more compact representation of the input data, such as picture enhancement or denoising [45], [46]. Using an auto-encoder to minimize the size of images means that some details are lost in the process of extracting the most significant features. The performance of the TLCA improves clearly (see Fig. 10) if these extracted features capture the data that is most pertinent to the current task. As compared to working directly with resized images (Fig. 10-i), the reduced image might have more prominent elements that the TLCA can better comprehend, improving accuracy or efficiency.

Furthermore, even CNNs perform better than LSTM and GRU models in image classification tasks, particularly when dealing with compressed pictures. LSTM models, are frequently employed in time series analysis and natural language processing applications [47], [48] because they are primarily built for processing sequential data. Although LSTM models can theoretically be used to process picture data by treating it like a series of pixels or patches, CNNs are a better option for this kind of application [49], which can be depicted from results (section III) also. GRUs share the same architectural designs as LSTM, despite some complexity differences. They are less effective at capturing the spatial linkages [50] and local patterns that are essential for comprehending images because of their sequential processing nature.

V. LIMITATIONS AND FUTURE DIRECTION

Researchers are always investigating novel architectures and strategies to increase performance in a variety of applications including image classification. There is quite a chance that sophisticated algorithms for classification may perform better on reduced features than conventional CNNs, like Transformer-based models [51]. They were introduced for tasks related to natural language processing, but have got consideration in the field of computer vision as well. Using self-attention mechanisms, the models like the Vision Transformer (ViT) [52] capture global dependencies in images. Another series of advance models: EfficientNet [53] aims to increase efficiency as well as accuracy. These models balance the size, depth, and width of the model using a technique called compound scaling [54], which produces extremely effective structures with enhanced performance. Regarding the constraints of the proposed dimensionality reduction technique, DFAE generates reconstructions that are either blurry or distorted as shown in Figure 9. In future, work can be done to improve reconstruction quality. Moreover, DFAEs demonstrate a notable susceptibility to noise present in the input data images.

VI. CONCLUSION

The goal of this research is to use auto-encoders to improve the classification accuracy of images from histological slides. With 96% accuracy and 94% precision in labeling these images, the suggested model, DFAE-TLCA, has shown impressive performance. The neural network architectures, the auto-encoders, which are specifically engineered for unsupervised learning, allowed the model to acquire significant representations of the input images. This learning process increased classification accuracy by lowering noise and extracting pertinent information.

The histopathological analysis through proposed model has a significant clinical importance. First of all, it is essential for the diagnosis of many diseases, including cancer. Pathologists can make more precise and dependable diagnoses with model's high accuracy and precision in identifying histology slide images. Histological slide examination and interpretation by hand can be laborious and prone to observer error. Pathologists can use the DFAE-CNN model as a decision assistance tool. When human specialists and machine learning algorithms work together, diagnostic precision can be increased and patient care can be improved overall. It can make easier to efficiently examine a huge number of histology slides, which will help with population-level research, better disease surveillance, the identification of risk factors, and the widespread adoption of preventative interventions can all result from this. One potential avenue for future exploration involves the investigation about the efficacy of proposed dimension reduction technique DFAE and classification model TLCA in exploring the 'Histopathologic Oral Cancer Detection' dataset images for the purpose of accurate

classification and analysis, addressing the challenges posed by high-dimensional data and complex image features.

REFERENCES

[1] M. K. K. Niazi, A. V. Parwani, and M. N. Gurcan, "Digital pathology and artificial intelligence," *Lancet Oncol.*, vol. 20, no. 5, pp. e253–e261, May 2019, doi: 10.1016/s1470-2045(19)30154-8.

[2] R. Nooreldeen and H. Bach, "Current and future development in lung cancer diagnosis," *Int. J. Mol. Sci.*, vol. 22, no. 16, p. 8661, Aug. 2021, doi: 10.3390/ijms22168661.

[3] T. E. Miller, A. M. Roche, and M. Mythen, "Fluid management and goal-directed therapy as an adjunct to enhanced recovery after surgery (ERAS)," *Can. J. Anesthesia*, vol. 62, no. 2, pp. 158–168, 2015, doi: 10.1007/s12630-014-0266-y.

[4] L. Hema, "Lungs lobes and fissures: A morphological study," *Int. J. Recent Trends Sci. Technol.*, vol. 11, no. 1, pp. 122–126, 2014.

[5] K. Al-Dulaimi, V. Chandran, K. Nguyen, J. Banks, and I. Tomeo-Reyes, "Benchmarking HEp-2 specimen cells classification using linear discriminant analysis on higher order spectra features of cell shape," *Pattern Recognit. Lett.*, vol. 125, pp. 534–541, Jul. 2019, doi: 10.1016/j.patrec.2019.06.020.

[6] A. E. Hozain, Y. Tipograf, M. R. Pinezich, K. M. Cunningham, R. Donocoff, D. Queen, K. Fung, C. C. Marboe, B. A. Guenthart, J. D. O'Neill, G. Vunjak-Novakovic, and M. Bacchetta, "Multiday maintenance of extracorporeal lungs using cross-circulation with conscious swine," *J. Thoracic Cardiovascular Surg.*, vol. 159, no. 4, pp. 1640–1653, 2020, doi: 10.1016/j.jtcvs.2019.09.121.

[7] W. Alakwaa, M. Nassef, and A. Badr, "Lung cancer detection and classification with 3D convolutional neural network (3D-CNN)," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 8, pp. 66–73, 2017, doi: 10.14569/ijacsa.2017.080853.

[8] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, "Medical image classification with convolutional neural network," in *Proc. 13th Int. Conf. Control Autom. Robot. Vis. (ICARCV)*, Dec. 2014, pp. 844–848, doi: 10.1109/ICARCV.2014.7064414.

[9] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 147–171, 2009, doi: 10.1109/RBME.2009.2034865.

[10] A. E. Minarno, K. M. Ghufron, T. S. Sabrila, L. Husniah, and F. D. S. Sumadi, "CNN based autoencoder application in breast cancer image retrieval," in *Proc. Int. Seminar Intell. Technol. Its Appl. (ISITIA)*, Jul. 2021, pp. 29–34, doi: 10.1109/ISITIA52817.2021.9502205.

[11] F. Khozeimeh, D. Sharifrazi, N. H. Izadi, J. H. Joloudari, A. Shoeibi, R. Alizadehsani, J. M. Gorriz, S. Hussain, Z. A. Sani, H. Moosaei, A. Khosravi, S. Nahavandi, and S. M. S. Islam, "Combining a convolutional neural network with autoencoders to predict the survival chance of COVID-19 patients," *Sci. Rep.*, vol. 11, no. 1, p. 15343, 2021, doi: 10.1038/s41598-021-93543-8.

[12] J. Borovec. (2021). *Histology (CIMA) Dataset*. Accessed: Apr. 22, 2024. [Online]. Available: https://www.kaggle.com/datasets/jirkaborovec/histology-cima-dataset

[13] C. Yan, G. Pang, X. Bai, C. Liu, X. Ning, L. Gu, and J. Zhou, "Beyond triplet loss: Person re-identification with fine-grained difference-aware pairwise loss," *IEEE Trans. Multimedia*, vol. 24, pp. 1665–1677, 2022, doi: 10.1109/TMM.2021.3069562.

[14] M. Ramamurthy, Y. H. Robinson, S. Vimal, and A. Suresh, "Auto encoder based dimensionality reduction and classification using convolutional neural networks for hyperspectral images," *Microprocessors Microsyst.*, vol. 79, Nov. 2020, Art. no. 103280, doi: 10.1016/j.micpro.2020.103280.

[15] G. Liu, L. Xie, and C.-H. Chen, "Unsupervised text feature learning via deep variational auto-encoder," *Inf. Technol. Control*, vol. 49, no. 3, pp. 421–437, Sep. 2020, doi: 10.5755/j01.itc.49.3.25918.

[16] Z. Tang, S. Wang, X. Chai, S. Cao, T. Ouyang, and Y. Li, "Auto-encoder-extreme learning machine model for boiler NOx emission concentration prediction," *Energy*, vol. 256, Oct. 2022, Art. no. 124552, doi: 10.1016/j.energy.2022.124552.

[17] F. M. Shiri, T. Perumal, N. Mustapha, R. Mohamed, M. Anuaruddin Bin Ahmadon, and S. Yamaguchi, "A survey on multi-resident activity recognition in smart environments," 2023, *arXiv:2304.12304*.

[18] C. R. García-Alonso, L. M. Pérez-Naranjo, and J. C. Fernández-Caballero, "Multiobjective evolutionary algorithms to identify highly autocorrelated areas: The case of spatial distribution in financially compromised farms," *Ann. Oper. Res.*, vol. 219, no. 1, pp. 187–202, Aug. 2014, doi: 10.1007/s10479-011-0841-3.

[19] T. A. Soomro, L. Zheng, A. J. Afifi, A. Ali, S. Soomro, M. Yin, and J. Gao, "Image segmentation for MR brain tumor detection using machine learning: A review," *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 70–90, 2023, doi: 10.1109/RBME.2022.3185292.

[20] D. E. Tobbala, A. S. Rashed, B. A. Tayeh, and T. I. Ahmed, "Performance and microstructure analysis of high-strength concrete incorporated with nanoparticles subjected to high temperatures and actual fires," *Arch. Civil Mech. Eng.*, vol. 22, no. 2, p. 85, 2022, doi: 10.1007/s43452-022-00397-6.

[21] M. Digra, R. Dhir, and N. Sharma, "Land use land cover classification of remote sensing images based on the deep learning approaches: A statistical analysis and review," *Arabian J. Geosci.*, vol. 15, no. 10, p. 1003, 2022, doi: 10.1007/s12517-022-10246-8.

[22] A. Aich. (2021). *Image-Super-Resolution-Autoencoders*. Accessed: Apr. 22, 2024. [Online]. Available: https://github.com/animikhaich/Image-Super-ResolutionAutoencoders/blob/main/PrototypeModel.ipynb

[23] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *Int. J. Data Mining Knowl. Manag. Process*, vol. 5, no. 2, pp. 1–11, 2015.

[24] M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: An overview," 2020, *arXiv:2008.05756*.

[25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[26] J. W. Johnson, *Adapting Mask-RCNN for Automatic Nucleus Segmentation*. New York, NY, USA: Springer, 2018.

[27] A. Deliège, A. Cioppa, and M. Van Droogenbroeck, "HitNet: A neural network with capsules embedded in a hit-or-miss layer, extended with hybrid data augmentation and ghost capsules," 2018, *arXiv:1806.06519*.

[28] C. Xiang, L. Zhang, Y. Tang, W. Zou, and C. Xu, "MS-CapsNet: A novel multi-scale capsule network," *IEEE Signal Process. Lett.*, vol. 25, no. 12, pp. 1850–1854, Dec. 2018, doi: 10.1109/LSP.2018.2873892.

[29] J. O'Neill, "Siamese capsule networks," 2018, *arXiv:1805.07242*.

[30] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Phys. D, Nonlinear Phenomena*, vol. 404, Mar. 2020, Art. no. 132306, doi: 10.1016/j.physd.2019.132306.

[31] D. Lee, M. Lim, H. Park, Y. Kang, J.-S. Park, G.-J. Jang, and J.-H. Kim, "Long short-term memory recurrent neural network-based acoustic model using connectionist temporal classification on a large-scale training corpus," *China Commun.*, vol. 14, no. 9, pp. 23–31, Sep. 2017, doi: 10.1109/CC.2017.8068761.

[32] R. C. Staudemeyer and E. Rothstein Morris, "Understanding LSTM—A tutorial into long short-term memory recurrent neural networks," 2019, *arXiv:1909.09586*.

[33] Y. Heryadi and H. L. H. S. Warnars, "Learning temporal representation of transaction amount for fraudulent transaction recognition using CNN, stacked LSTM, and CNN-LSTM," in *Proc. IEEE Int. Conf. Cybern. Comput. Intell. (CyberneticsCom)*, Nov. 2017, pp. 84–89, doi: 10.1109/CYBERNETICSCOM.2017.8311689.

[34] A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, and S. W. Baik, "Action recognition in video sequences using deep bi-directional LSTM with CNN features," *IEEE Access*, vol. 6, pp. 1155–1166, 2018, doi: 10.1109/ACCESS.2017.2778011.

[35] H. Luo, "Shorten spatial–spectral RNN with parallel-GRU for hyperspectral image classification," 2018, *arXiv:1810.12563*.

[36] E. Pan, Y. Ma, X. Dai, F. Fan, J. Huang, X. Mei, and J. Ma, "GRU with spatial prior for hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2019, pp. 967–970.

[37] S. Hao, W. Wang, and M. Salzmann, "Geometry-aware deep recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 2448–2460, Mar. 2021, doi: 10.1109/TGRS.2020.3005623.

[38] E. Pan, X. Mei, Q. Wang, Y. Ma, and J. Ma, "Spectral–spatial classification for hyperspectral image based on a single GRU," *Neurocomputing*, vol. 387, pp. 150–160, Apr. 2020, doi: 10.1016/j.neucom.2020.01.029.

[39] A. Ashraf, N. M. Nawi, and M. Aamir, "Adaptive feature selection and image classification using manifold learning techniques," *IEEE Access*, vol. 12, pp. 40279–40289, 2024.

[40] M. Kwabena Patrick, A. Felix Adekoya, A. Abra Mighty, and B. Y. Edward, "Capsule networks—A survey," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 1, pp. 1295–1310, 2022, doi: 10.1016/j.jksuci.2019.09.014.
[41] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, p. 30.
[42] H. Tang, B. Xiao, W. Li, and G. Wang, "Pixel convolutional neural network for multi-focus image fusion," *Inf. Sci.*, vols. 433–434, pp. 125–141, Apr. 2018, doi: 10.1016/j.ins.2017.12.043.
[43] E. Kauderer-Abrams, "Quantifying translation-invariance in convolutional neural networks," 2017, *arXiv:1801.01450*.
[44] C. Mouton, J. C. Myburgh, and M. H. Davel, "Stride and translation invariance in CNNs," in *Proc. Southern Afr. Conf. Artif. Intell. Res.* (Communications in Computer and Information Science), vol. 1342. New York, NY, USA: Springer, 2020, pp. 267–281, doi: 10.1007/978-3-030-66151-9_17.
[45] P. Zhou, J. Han, G. Cheng, and B. Zhang, "Learning compact and discriminative stacked autoencoder for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4823–4833, Jul. 2019, doi: 10.1109/TGRS.2019.2893180.
[46] P. Kumar Mallick, S. H. Ryu, S. K. Satapathy, S. Mishra, G. N. Nguyen, and P. Tiwari, "Brain MRI image classification for cancer detection using deep wavelet autoencoder-based deep neural network," *IEEE Access*, vol. 7, pp. 46278–46287, 2019, doi: 10.1109/ACCESS.2019.2902252.
[47] L. Yao and Y. Guan, "An improved LSTM structure for natural language processing," in *Proc. IEEE Int. Conf. Saf. Produce Informatization (IICSPI)*, Dec. 2018, pp. 565–569, doi: 10.1109/IICSPI.2018.8690387.
[48] D. Soutner and L. Müller, "Application of LSTM neural networks in language modelling," in *Lecture Notes in Computer Science (including Subseries Lecture Notes Artificial Intelligence Lecture Notes Bioinformatics)* (Lecture Notes in Artificial Intelligence), vol. 8082. New York, NY, USA: Springer, 2013, pp. 105–112, doi: 10.1007/978-3-642-40585-3_14.
[49] P. Dhruv and S. Naskar, "Image classification using convolutional neural network (CNN) and recurrent neural network (RNN): A review," in *Machine Learning and Information Processing* (Advances in Intelligent Systems and Computing), vol. 1101. New York, NY, USA: Springer, 2020, pp. 367–381, doi: 10.1007/978-981-15-1884-3_34.
[50] D. Linsley, J. Kim, V. Veerabadran, C. Windolf, and T. Serre, "Learning long-range spatial dependencies with horizontal gated recurrent units," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 152–164, doi: 10.32470/ccn.2018.1116-0.
[51] K. Han, A. Xiao, E. Wu, J. Guo, C. Xu, and Y. Wang, "Transformer in transformer," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 19, 2021, pp. 15908–15919.
[52] H. Yin, A. Vahdat, J. M. Alvarez, A. Mallya, J. Kautz, and P. Molchanov, "A-ViT: Adaptive tokens for efficient vision transformer," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 10799–10808, doi: 10.1109/CVPR52688.2022.01054.
[53] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks mingxing," *Can. J. Emergency Med.*, vol. 15, no. 3, p. 190, 2013.
[54] C. Lin, P. Yang, Q. Wang, Z. Qiu, W. Lv, and Z. Wang, "Efficient and accurate compound scaling for convolutional neural networks," *Neural Netw.*, vol. 167, pp. 787–797, Oct. 2023, doi: 10.1016/j.neunet.2023.08.053.

AMNA ASHRAF was born in Bahawalpur, Pakistan, in 1992. She received the bachelor's and master's degrees in computer engineering from the University of Engineering & Technology, Lahore, Pakistan, in 2014 and 2017, respectively. She is currently pursuing the Ph.D. degree with Universiti Tun Hussein Onn Malaysia (UTHM), under the kind supervision of Prof. Nazri Mohd Nawi and Muhammad Aamir. Her research interest includes artificial intelligence, with a particular emphasis on dimensionality reduction of image datasets. Additionally, she is exploring the application of AI techniques in processing hyper-spectral images obtained from satellites. She has emerged as a distinguished scholar in the fields of machine learning and AI. Her significant contributions to the academic community are evident through her extensive research and mentorship endeavors. With more than ten scholarly research articles published, her work is recognized. With her supervision of more than ten Master of Science students, she has demonstrated the ability to impart knowledge and guide the intellectual development of emerging researchers.

NAZRI MOHD NAWI received the bachelor's degree from Universiti Sains Malaysia (USM), the master's degree in computer science from University Teknologi Malaysia (UTM), and the Ph.D. degree in data mining from Swansea University, Wales, U.K. He is currently a Professor with the Department of Software Engineering, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), where he has been a Faculty Member, since 2001. His research interests include soft computing and data mining techniques, particularly in artificial neural networks, ranging from theory to design and implementation. In recent years, he has focused on better techniques for classification, analyzing, and hybridizing some new improvement on ANN using meta-heuristic techniques. He has successfully supervised few Ph.D. students and currently he is supervising eight Ph.D. students and published more than 100 papers in journals and conference proceedings. He has involved with many conferences and workshop program committees and serves as a reviewer for many outstanding journals and international conferences.

TARIQ SHAHZAD received the B.E. and M.S. degrees from COMSATS University Islamabad, Sahiwal, Pakistan, in 2006 and 2014, respectively, and the Ph.D. degree from the University of Johannesburg, South Africa, in 2021. He is currently a Research Fellow with the University of Johannesburg. He has published several research articles in journals and international conferences. His research interests include biomedical signals processing, machine learning, and computer break vision.

MUHAMMAD AAMIR received the master's degree in computer science from the City University of Science and Information Technology, Pakistan, and the Ph.D. degree in information technology from Universiti Tunn Hussien Onn Malaysia, in 2020. He has been a Research Data Scientist and Machine Learning Development Engineer with the University of Derby, U.K., since October 2020. A prior tenure of two years saw him serving as a Data Scientist with Xululabs LLC. His areas of expertise encompass data science, deep learning, and computer programming. A prolific contributor to academia, he has authored over 25 journal articles and conference papers, in addition to being the co-author of a book focused on data analysis.

MUHAMMAD ADNAN KHAN (Senior Member, IEEE) received the B.S. and M.S. degrees from International Islamic University, Islamabad, Pakistan, by obtaining Scholarship Award from Punjab Information & Technology Board, Government of Punjab, Pakistan, and the Ph.D. degree from ISRA University, Islamabad, by obtaining a Scholarship Award from the Higher Education Commission, Islamabad, in 2016. He is currently an Assistant Professor with the School of Software, Faculty of Artificial Intelligence and Software, Gachon University, Seongnam-si, Republic of Korea. Before joining Gachon University, he has worked in various academic and industrial roles in Pakistan. He has been teaching graduate and undergraduate students in computer science and engineering for the past 16 years. He is also guiding five Ph.D. scholars and six M.Phil. scholars. He has published more than 290 research articles with Cumulative JCR-IF more than 840 in reputed international journals and international conferences. His research interests include computational intelligence, machine learning, mud, image processing and medical diagnosis, and channel estimation in multi-carrier communication systems using soft computing.

KHMAIES OUAHADA (Senior Member, IEEE) was the Head of the Department, from December 2018 to December 2021. He is currently a Full Professor with the Department of Electrical and Electronic Engineering Science, Faculty of Engineering and the Built Environment. He lectures signal processing 4A after teaching two modules, telecommunications 3B, and signal processing 3B, which form a critical foundation for fourth-year exit level modules. In the past, he has taught other introductory and advanced modules and has done an excellent job organizing and lecturing. His lecture notes for electrotechnics 2A, electrotechnics 2B, and signals and systems 3A are still used by his successors. He was awarded the Vice-Chancellors Teaching and Learning Excellence Award, in 2016. This is a natural result of his hard work and passion as a Lecturer.