IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 10 February 2024, accepted 28 March 2024, date of publication 9 April 2024, date of current version 1 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3386634

RESEARCH ARTICLE

DFPT-CNN: A Dual Feature Extraction and Pretrained CNN Synergy for Minimal Computational Overhead and Enhanced Accuracy in Multi-Class Medical Image Classification

DINAH ANN VARUGHESE AND SRIADIBHATLA SRIDEVI, (Senior Member, IEEE)
Vellore Institute of Technology, Vellore 632014, India

Corresponding author: Sriadibhatla Sridevi (sridevi@vit.ac.in)

This work was supported by the Vellore Institute of Technology.

ABSTRACT In the advanced computer vision era, Convolutional Neural Network (CNN) plays a pivotal role in image processing, as they excel at automatically extracting important patterns, and structures, for accurate analysis across diverse domains. However, achieving higher accuracy often leads to intensifying computational and timing demands. To address the challenge, this research introduces a novel dual feature extraction methodology. This approach is implemented using two distinct feature extraction modules, employed at different stages of the model: 1) Edge Gradient-Dimensionality Reduction (EGDR) module which encapsulates the extraction of pixel edge gradient features from the raw input frame, leading to a dimensionality reduction by a factor of 0.5; 2) Subtle Local Feature Extraction (SLFE) pooling algorithm module, prioritizes the extraction of local and subtle features over maximum or average feature content. The combination of these two stages proves particularly effective in enhancing classification accuracy while minimizing computational overhead and training duration. Subsequently, comprehensive training, validation, and testing were conducted on a selected multi-class chest computed tomography medical image dataset using various state-of-the-art CNN architectures such as VGG-16, InceptionV3, ResNet50 to identify the most suitable model for further experimentation with the proposed method. The proposed CNN-SLFE framework with EGDR module achieved a significant reduction of 17.94% in computational time compared to non-EGDR module, and concurrently enhanced the classification accuracy with an improvement factor of 1.17 compared to existing CNN frameworks with EGDR module.

INDEX TERMS Classification accuracy, computation acceleration, convolutional neural network, feature extraction, medical images, pooling algorithm.

I. INTRODUCTION

Artificial Intelligence (AI) and Machine Learning (ML), are dedicated fields to the creation of intelligent systems capable of learning and acting independently. Deep Learning (DL) being a subset of ML, employs artificial neural networks to learn from data. Convolutional Neural Network (CNN), a type of DL algorithm, has proven effective for image processing [1], [2], [3] in particular. The ability of CNN to capture spatial information from images makes it an ideal tool for the analysis of medical images [4], [5], [6], including tasks such as segmentation of tissues and anatomical structures, classification of disorders, detection and classification of lesions and tumours [7], [8], [9], prediction of survival rates, diagnosis of various types of cancer [10], [11]. These images often contain subtle details that can be challenging for medical experts to identify. This wide range of applications have the potential to revolutionize

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

healthcare by enhancing the accuracy and efficiency of diagnostic processes [12], [13], [14].

CNNs are primarily designed to work with two-dimensional image data, but they can also be adapted for one-dimensional and three-dimensional data [2]. The fundamental component of CNN is convolutional layer. A convolution involves applying a filter to the input, resulting in an activation. When the same filter is repeatedly applied to the input, it produces a feature map. This map indicates the locations and strength of a detected feature in an input, such as an image, results in highly specific feature extraction. After the Convolution layer, Rectified Leaky Unit (ReLU) commonly used activation function will be applied to each of the feature maps, which introduces a non-linearity function and allows the network to learn complex functions rather than linear functions. ReLU creates sparsity in the network by setting all negative activations to zero as shown in Fig.1.

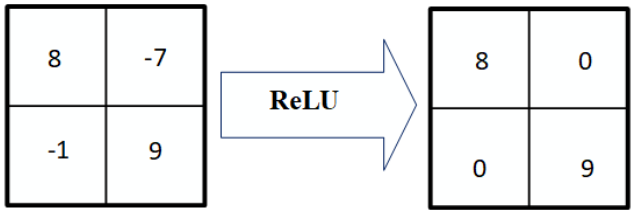

FIGURE 1. ReLU activation function.

In addition to convolutional layers, CNN architecture contains pooling layers and fully connected (FC) layers. The pooling layer down-samples the feature map dimensionality for reducing the computational complexity. Thus, by reducing the resolution of the feature maps, pooling helps to prevent the over-fitting issue. Then, the FC layer makes the final prediction. The network learns to optimize these filters through back-propagation and gradient descent. This process allows CNN to effectively analyse complex image data and make accurate predictions. Max and average (AVG) pooling are the two widely used techniques in image classification. Max pooling focuses solely on the highest value in each pooling area, which may lead to significant data being overlooked. On the other hand, AVG pooling can blur sensitive image details by averaging all pixel values in the pooling area, making it challenging for CNN to learn classification features [15].

Prior to feeding of input images into CNN layers, data preparation is an essential step. Preprocessing helps to minimize noise and superfluous data, and mitigate over-fitting which is a common issue in deep learning. Some of the widely used image preprocessing techniques include: Normalisation, which normalises the image pixel values to a range of [0, 1] or [−1, 1]; Resizing which scales the image to a specified resolution and reduce the computational cost of CNN training; Cropping which removes unwanted parts of the image, allowing the CNN to concentrate on the most significant portion of the image. Another preprocessing technique called Data Augmentation, creates new images by flipping, rotating, and cropping the original images and thereby increases the size of the dataset [16], [17].

Traditional methods of image preprocessing and pooling have certain limitations. For instance, they may not fully capture the complexity and variability inherent in image data. Noise reduction techniques might inadvertently remove important details, while pooling methods could lead to loss of spatial information. To address these challenges, this work proposes a novel dual feature extraction method. This method is designed to be implemented at two distinct stages of the model - one at the preprocessing stage and the other at the pooling layer of the CNN framework. In the initial stage, an Edge Gradient Dimensionality Reduction (EGDR) is employed using a two-fold approach which results in a more refined and accurate representation of the underlying structures and patterns and reduced noise sensitivity. The resulting image representation is characterized by a reduced pixel count and a halved frame size. The latter stage, used an innovative Subtle Local Feature Extraction (SLFE) pooling algorithm, to capture more generalized features across different images. As the CNN plays well in image classification tasks, and to validate the efficacy of the two proposed approach, a multi-class chest computed tomography (CT) medical image classification dataset of lung cancer is considered. The primary contributions of this paper are as follows:

- The EGDR module streamlines refined data representation, which is not only memory efficient but also facilitates faster processing and accurate analysis. This makes it well-suited for applications with limited computational resources or real-time requirements.
- In the scenarios where images may have varying lighting conditions or contrast levels, the SLFE module can normalize these variations. This allows the CNN to concentrate on the underlying features more effectively, enhancing the depth of feature extraction and improving the overall performance of the model.
- The proposed dual feature extraction method effectively handles images with varying qualities and characteristics, making it a versatile tool for image classification tasks. This adaptability can extend the applicability of the method to diverse datasets and domains.
- By integrating the EGDR and SLFE modules into the CNN framework, the method significantly improves the classification accuracy. It effectively captures both subtle and local features in the images, leading to a more comprehensive understanding of the image content. This results in improved performance in distinguishing between different classes in the dataset, thereby enhancing the overall classification performance.

TABLE 1. Comparison of medical image classification using various CNN TL approaches.

Reference	Method/Model	Inference(%)
[18]	CNN-VGG16 Xception	Accuracy-87 Accuracy-82
[19]	TL-CNN	Accuracy-90.7
[21]	TL-CNN	Accuracy-84.15
[22]	CNN-Boosting Architecture for FP reduction	Accuracy-86.7, Sensitivity-74.4
[23]	Dense CNN+ADABOOST classifier	Accuracy-90.85
[24]	CNN+CapsNet classifier	Accuracy-94, Sensitivity- 94.5, Specificity- 95
[25]	Hybrid CNN+SVM classifier	Accuracy-97
[26]	Hybrid Deep CNN and MAX-GAP pooling layer	Accuracy-81.87
[9]	Data Augmentation+CNN	Accuracy-92.8
[27]	Streamlined Sequential CNN	Accuracy-89.89
[28]	CNN-Hybrid parallelization+ NNLU activation function	Accuracy-98
[29]	CNN+MAX Pooling+ Batch Normalization	Accuracy-96.5%
[30]	Data pre-processing, ROI feature extraction to CNN	Accuracy-79.3
[31]	ROI feature extraction+ feature processing+ MLP	Accuracy-88.55% Sensitivity-86.81% Specificity-86.95%
[32]	Texture feature extraction, Haralick features on VGG16	Accuracy-93
[33]	Transformer module+ CNN model+Feature fusion branch	Accuracy-96.35

The rest of this article is structured as follows: Section II reviews related works, Section III describes the proposed methodologies, Section IV presents the implementation and results, and Section V concludes with final observations.

II. REVIEW OF RELATED WORKS

The advent of DL techniques, particularly through TL approach, has shown its effectiveness in several research works [18], [19], [20].CNNs, in particular, have brought about a revolution in medical field with their ability to provide fast interpretation with high accuracy, and became an invaluable tool in the medical imaging. This synergy between deep learning and CNNs has reshaped the landscape of medical diagnostics and imaging technologies. A comprehensive review on this synergy was conducted, with key findings systematically documented and tabulated on Table 1, underscoring the potential of these techniques in enhancing the accuracy and efficiency of medical image analysis.

The work in [34], demonstrated the effectiveness of TL CNN frameworks on chest CT images for early detection of lung cancer, especially in reducing the model False Positive rate (FPR) to 1.97 per scan, which signifies instances of incorrect classification of negative classes as positive. The study in [21] employed three types of neural networks, namely CNN, Deep Neural Network (DNN), and Stacked Auto-Encoders (SAE), and evaluated their performance on the Lung Image Database Consortium, LIDC-IDRI database, which contains annotated images of benign and malignant nodules. The CNN network achieved the highest accuracy of 84.15%, sensitivity of 83.96%, and specificity of 84.32% among the three networks. Furthermore, a novel automated R-CNN framework for pulmonary nodule detection and a CNN based boosting architecture for the reduction of FPR was proposed in [22], achieved sensitivity of 86.7% on the LUNA16 dataset. The work in [23] focused on the early detection of lung cancer using CT images. It utilized a densely connected CNN along with an adaptive boosting algorithm (ADABOOST) at the classification layer. The study aimed to classify lung images as normal or malignant, and achieved an accuracy of 90.85%.

Another study [24] presented LCD-CapsNet, a novel DL framework that combined a CNN and a Capsule Neural Network (CapsNet) for lung cancer detection and classification. This model was trained and tested on the LIDC dataset, consisting of 4335 images. LCD-CapsNet outperformed CapsNet, achieving an average precision of 95%, recall of 94.5%, F1-Score 94.5%, specificity 99.07%, and accuracy of 94% for benign and malignant data. The study [25] introduced a TL method that classified lung CT images into normal, benign, and malignant categories using a hybrid model of AlexNet, VGG, and GoogleNet, along with a multi-class SVM classifier. The model achieved an impressive accuracy of 97%, whereas the work in [26] used a hybrid deep CNN consisting of 22 convolutional layers with the inclusion of MAX and GAP pooling layers achieving a maximum accuracy of 81.87% for non-augmented dataset. Another study [9] presented a CNN-based medical image classification method which included a Max pooling layer for disease diagnosis and leveraged TL and data augmentation to enhance the CNN model's performance on a small CXR dataset. This resulted in an impressive accuracy of 99.5% on the training dataset and 92.8% on the test dataset. However, the data augmentation process, which generates additional training samples by applying various transformations to the original dataset, requires extra computational resources for processing during training and extends the duration of the training epochs, thereby prolonging the overall training time of the model.

In [27], a streamlined sequential CNN architecture was proposed and compared with two pre-trained CNN models (VGG-16 and InceptionV3) for diagnosing diseases from chest X-ray images, achieving a peak accuracy of 89.89%. In the work conducted by [28], a hybrid parallelization strategy, incorporating both model and data parallelism, was applied to CNN to speed up its operation, with a normalized non-linearity activation function (NNLU) and replacing

FC layers with Global Average Pooling (GAP) layers. These changes significantly improved accuracy and increased computation speed by 3.62 times with a batch size of (512,128), achieving a validation accuracy of approximately 98%. The authors in [29] used 1000 chest CT scans, trained CNN model with a batch size of 32, a 3 × 3 max pooling size, 0.25 drop out, batch normalization, and early stopping to achieve a binary classification accuracy of 96.5%.

Authors in [31] proposed a method to identify and categorize lung nodules in CT scans. It involves three image processing steps followed by the extraction of features from regions of interest (ROI). These features are then analyzed using various machine learning algorithms. The study found that a multi-layer perceptron classifier achieved the best results, with an accuracy of 88.55%, a sensitivity of 86.81%, and a specificity of 86.95%. The study in [30] focussed on extracting ROI features from the raw images using a computer vision model. A modified Inception v3 model was trained with these features, achieving a binary classification accuracy of 79.3%. Furthermore, [32] proposed a texture feature extraction method using Haralick features to identify abnormal features from images. These features were then used to train certain pre-trained models, such as VGG-16, ResNet50, and Inception V3, for the detection of COVID-19 from chest X-ray and CT scan images. Among the evaluated TL models, VGG-16 achieved the highest accuracy of 93%, sensitivity of 90%, and specificity of 91%, demonstrating the significant application of CNN in lung nodule detection and classification. A classification framework integrating transformer module to extract the spectral information, CNN model for the spatial feature extraction and a feature fusion branch in parallel, named as Fusion Transformer (FUST), was proposed in [33] for image classification tasks, achieved an overall accuracy of 96.35%.

In the context of medical image classification using various TL CNN approaches, the introduction of fast convolution algorithms has become an emerging trend on recent research works. The authors in [35] and [36] adopted certain fast convolution algorithms on various layers of CNN architecture to enhance the speed on hardware platforms, thereby enabling real-time computation. The overall CNN performance has improved to 2479.6 Giga Operations/second compared to prior works discussed in [35]. The study in [37], carried out one dimensional CNN for anomaly detection of bearing faults and the model weights were trained using one bit with the knowledge distillation and binarization algorithm, achieved an accuracy of 98.5%. Another study [38] applied a data quantization technique named as Equal distance Quantization (ENQ) to a VGG16 model, achieving 86.25% accuracy with lower computational cost. Furthermore, the research in [39], adopted light weight CNN with strong lego filters for the feature extraction, helps to reduce the memory as well as computational cost of the model, achieved an accuracy of 91.40%. This advancement plays a crucial role in optimizing the efficiency of CNNs, making them more practical and effective for tasks such as medical image analysis. This development is a significant stride towards achieving efficient accuracy coupled with reduced computation time in the realm of medical image analysis. Thus, the synergy between TL, CNNs, and fast convolution algorithms is paving the way for emerging advancements in medical diagnostics and imaging technologies.

III. METHODOLOGY

In this work, the proposed Dual Feature Extraction Pretrained CNN (DFPT-CNN) model is aimed at optimizing feature extraction and refining classification processes in two stages. Firstly, an EGDR approach utilizes the Winograd convolution algorithm, a fast convolution method to reduce the dimensionality of raw frames by 0.5 times, and Sobel edge detection filter to extract the relevant features, highlighting the significant edges and gradients within the images; in the next stage, these refined features are applied to a modified CNN model with the proposed SLFE module prioritizing the extraction of subtle and local features. This dual feature extraction approach, achieved a balance between computational efficiency and classification accuracy, ultimately enhancing the efficacy of image classification tasks, especially in domains requiring real-time processing and high accuracy. The DFPT-CNN model is illustrated in Fig.2.

The proposed model undergoes a training process for several epochs. Once the training phase is completed, the model is then applied in an inference procedure using new datasets. This inference stage involves the estimation of a variety of performance metrics to evaluate the model's effectiveness. The metrics used for this assessment include sensitivity, specificity, precision, recall, F1 score, and accuracy level. The details of these metrics are provided in Table 2, where TP, TN, FP, FN indicates True Positive, True Negative, False Positive and False Negative respectively. These measures provide a comprehensive understanding of the model's performance across different aspects.

A. EDGE GRADIENT DIMENSIONALITY REDUCTION (EGDR) MODULE

The visual representation of the EGDR module is shown in Fig.3. In this module, the raw image frame is divided into several 4 × 4 non-overlapping tiles and each 4 × 4 tile along with 3 × 3 kernel is transformed to Winograd domain using precomputed set of matrices, facilitating efficient matrix multiplications. After this, the intermediate output is transformed back to original domain to obtain the final convolution output. Finally, all the intermediate outputs are appended to form a new framework with reduced dimensions. In this work, Sobel edge detector is chosen as 3 × 3 kernel. It is a widely employed tool in image processing for edge detection tasks [40]. It consists of two 3 × 3 convolution kernels, one focussing on detecting vertical edges and the other on horizontal edges. These kernels are designed in such a way that they highlight the regions in an image where there is a significant change in pixel intensity, indicative of an edge or boundary. Thus, the integration of Winograd

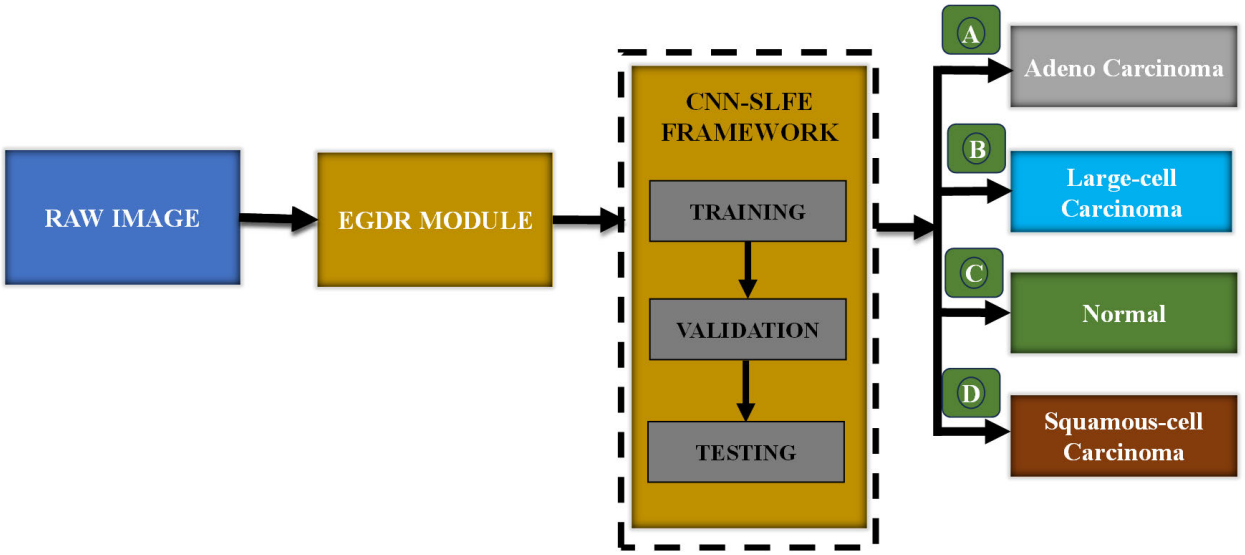

FIGURE 2. Proposed DFPT-CNN model.

TABLE 2. Performance metrics definition and its equations.

Performance Metric	Definition	Equation
Sensitivity	Number of positive cases that are correctly predicted by the model	(TP)/(TP+FN)
Specificity	Number of negative cases that are correctly predicted by the model	(TN)/(TN+FP)
Precision	Number of positive cases that are true	(TP)/(TP+FP)
Recall	Number of positive cases that are predicted by the model	(TP)/(TP+FN)
F1Score	Harmonic mean of Precision and Recall	(2*Precision*Recall)/(Precision+Recall)
Accuracy	Number of all cases that are correctly identified by the model	(TP+TN)/(TP+TN+FP+FN)

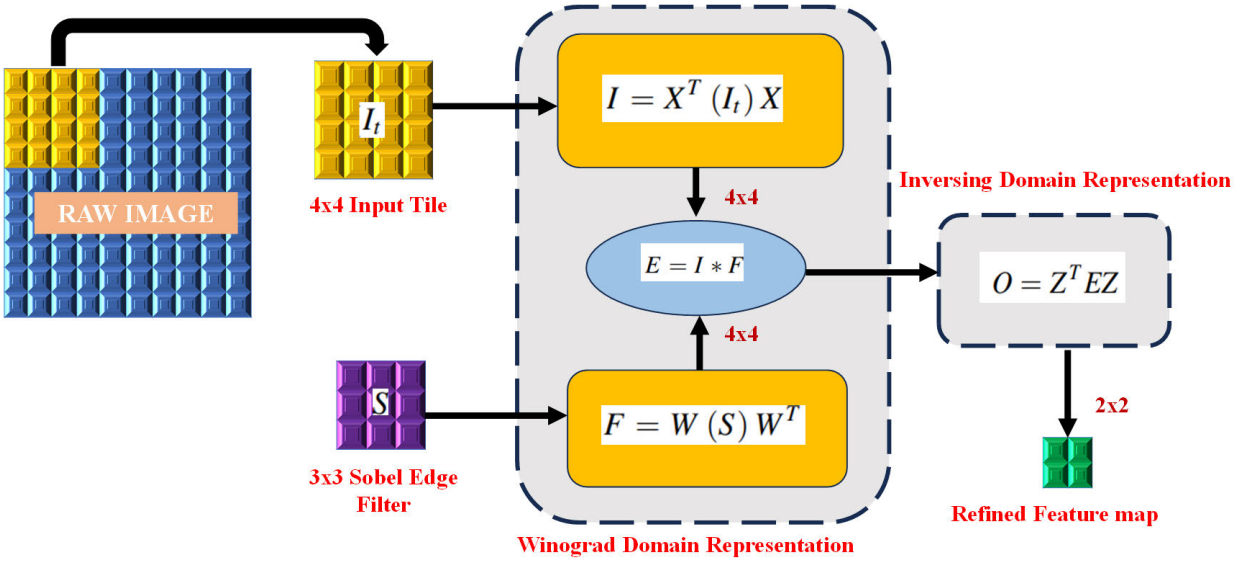

FIGURE 3. EGDR module.

feature extractor algorithm in conjunction with the Sobel edge detector in the EGDR module, constitutes a powerful synergy in image processing.

As the module works on small, non-overlapping tiles of the input data, allows for efficient processing of smaller data, reducing memory requirements and improving cache utilization. The final output, O can be computed as (1)

$$O = Z^T EZ \tag{1}$$

TABLE 3. Precomputed values of X,W and Z.

X				W			Z	
1	0	0	0	1	0	0	1	0
0	1	-1	1	0.5	0.5	0.5	1	0
-1	1	1	0	0.5	-0.5	0.5	1	-1
0	0	0	-1	0	0	1	-1	0

where, E is calculated by taking the dot product between I and F as shown in (2).

$$E = I * F \tag{2}$$

$$I = X^T (I_t) X, F = W (S) W^T \tag{3}$$

I and F represent the Winograd domain representation of an input tile of size 4×4, I_t and Sobel filter 3×3, S respectively, as (3). The X, W, and Z are precomputed values and they are given in Table 3 [41], [42].

Hence, the overall equation for a small tile WinSobel-FEA, $O_{2\times2}$ can be written as (4):

$$O_{2\times2} = \{Z^T_{2\times4} \left[\left(X^T (I_t) X \right)_{4\times4} . \left(W (S) W^T \right)_{4\times4} \right] Z_{4\times2}\} \tag{4}$$

B. PROPOSED SUBTLE LOCAL FEATURE EXTRACTION (SLFE) POOLING ALGORITHM

In the realm of feature extraction in CNN, pooling plays a pivotal role in reducing the spatial dimensions while retaining critical information.

The proposed SLFE module, shown in algorithm 1, aims to enhance the performance and robustness of image classification models. The SLFE algorithm for image processing is a novel global-to-local anomaly detector that uses a Gaussian mixture model to learn a probabilistic model of normal images. This model can then be used to detect anomalies in new images by identifying pixels that have a low probability of occurrence under the model. The proposed pooling methodology begins with the computation of the global average of the convolved feature map (GLA), capturing the overall intensity distribution. The parameter N indicates the convolved feature map dimensionality, i and j are used to indicate the row and column elements respectively. $x_{i,j}$ points to each element in the feature map. Subsequently, a defined pooling window is applied to group pixels for pooling operations and the parameter W represents the group count. For each window, a critical decision is made: the maximum pixel intensity value within the window, M, is compared with the previously computed GLA. If the M is lower than the GLA, then the pooled value for that window is set to GLA, signifying that the local features do not deviate significantly from the overall information. On the other hand, if the maximum value exceeds the GLA, a more intricate approach is employed. Here, the absolute deviation of each pixel intensity from the GLA within the window is calculated as Pooled Value, emphasizing local variations from the established GLA.

Algorithm 1 Proposed SLFE

Read: Convolved feature map, *NXN*
Step 1: Calculate $GLA = (1/(N * N)) * \sum \sum x_{i,j}$
Step 2: Non-overlapped input tile creation of size 2X2
Step 3: Assign $W \leftarrow (N * N)/4$
Step 4: Initialisation: $k = 1$
Step 5:
if $M \geq GLA$ **then**
  $PooledValue \leftarrow \sum_{i=0}^{1} \sum_{j=0}^{1} |x_{i,j} - GLA|$
else
  $PooledValue \leftarrow GLA$
Step 6: $k = k + 1$,
Step 7: Repeat *Steps 5 to 6*, till $k = W$

SLFE pooling functionality incorporates a combination of global average pooling and a conditional pooling strategy based on pixel intensity comparisons. It is aimed to achieve a balance between preserving the global information and capturing the local deviations from the average. By implementing this, we aim to extract important features from the convolved feature map, considering both global and local characteristics. This approach will be useful in scenarios where capturing local deviations from the average are essential for accurate representation of features in the data.

Moreover, SLFE estimation can also contribute to reducing over-fitting. Over-fitting occurs when a model learns to perform very well on the training data but fails to generalize to the unseen data. By focusing on generalized features rather than specific pixel values, GLAD can help create models that are more robust and perform better on new, unseen images. Thus, their ability to extract meaningful and generalized features from images are enhanced, thereby improving their performance and robustness.

IV. IMPLEMENTATION AND RESULTS

To evaluate the effectiveness of the proposed model, a dataset of medical images was utilized. Specifically, images of lung cancer epithelial cells from chest CT, which were sourced from the publicly available database. The implementation of this methodology involves four set of experiments:

- In the initial set of experiments, a multi-class medical image classification dataset related to lung cancer is carefully chosen and divided for training, validation, and testing. It is then fed into three different pretrained CNN frameworks such as VGG16, InceptionV3, and ResNet50, all of which are designed for classification tasks. The model that uses fewer parameters and exhibits notable accuracy is subsequently employed for further experimentation. Since the inference is conducted on an unseen dataset without additional feature processing, the threshold for testing accuracy is set to exceed 75% of the training accuracy. This ensures that the model generalizes well to unseen data and maintains a satisfactory level of performance.

TABLE 4. Results of pretrained CNN models on custom dataset.

Model	VGG-16	ResNet50	InceptionV3
Number of layers	16	50	48
Number of Parameters	14,815,044	22 23,989,124	22,327,076
Training Accuracy	99.45	75.53	99.86
Validation Accuracy	89.62	81.97	87.43
Inference Accuracy	78.95	73.68	86.84
Time Elapsed (Sec)	433	448	575

- In the second phase, the SLFE pooling module is incorporated into the chosen CNN model to validate the efficacy with other existing pooling methodologies. During this phase, it is imperative to ensure that the testing accuracy, sensitivity, and specificity thresholds are set at over 90% to achieve precise detection of image labels.
- The third set deals with the application of the EGDR module to the raw input frames. This involves the combined potency of Winograd convolution algorithm and Sobel edge detection filter to extract the refined features from the raw image frames for further training, validation and testing of the proposed model.
- Finally, it involves a performance analysis of the proposed CNN-SLFE framework with and without the EGDR module, on the selected dataset. The analysis is conducted in terms of computation time and various performance metrics such as accuracy, sensitivity, specificity, precision, recall and F1 score. The integration of the EGDR module aims to improve computation speed through frame size reduction while ensuring that the model maintains a minimum threshold level of 85% across all evaluation metrics. This is crucial for maintaining lower false positive rate and false negative rate values.

A. DATASET

For the proposed work, a dataset composed of 921 chest CT images were utilized. This dataset is categorized into four distinct classes: three are confirmed cancer cases - adenocarcinoma (class 0), large cell carcinoma (class 1), and squamous cell carcinoma (class 3), and the fourth class consists of normal images (class 2) without any signs of the disease as shown in Fig.4. Out of the total 921 images, a majority (738 images) were allocated for training purposes. These training images were distributed across all four classes. The remaining 183 images were set aside for the testing process. To ensure a balanced and effective model training, we followed a specific distribution strategy for the images in each class. The dataset was structured in such a way that 80 percentage of the images from each class were used for training the model, while the remaining 20 percentage were used for testing its performance. After the training phase, a set of 76 images from the training data was used for validation process. This approach allowed the researcher to train the model on a diverse set of data and provided a robust set for evaluating its accuracy and effectiveness.

B. TRAINING AND INFERENCE PROCEDURES AND ITS RESULTS

1) SELECTION OF APPROPRIATE PRETRAINED MODEL

Given the limited size of the medical dataset, pretrained models were leveraged to achieve higher accuracy with the dataset, shown in Fig.5. Three distinct pretrained CNN models: VGG-16, ResNet50, and InceptionV3 were trained on cancerous image inputs to identify the most suitable model for classifying the given image classes. This training phase was conducted over several epochs to ensure that the model learnt and adapted to the nuances of the data. Python was utilized as the programming language of choice and a Generic Processing Unit (GPU) was relied on for backend computation. This combination allowed the researchers to efficiently handle the computational demands of training.

Upon completion of the training phase, the trained parameters were saved onto a .h5 file. This file serves as a record of the model's learned parameters and can be used for future inference without needing to retrain the model. The next phase involved model prediction, where new dataset comprising four classes were used. The .h5 file containing the trained parameters was used to classify these new images. During this phase, the accuracy of each model is determined and the total number of parameters used by each was tallied.

As per the results from Table 4, it is evident that the InceptionV3 model stands out due to its high inference accuracy. However, the count of parameters and time elapsed are significantly higher when compared to the VGG-16 model. To address this issue, when the VGG-16 model is trained for 100 epochs, the results generated were promising, with training, validation, and inference accuracy of 99.86%, 91.80%, and 88.15% respectively. This experiment showed that by increasing the number of training epochs, model's accuracy can be enhanced while keeping the number of parameters constant. Encouraged by these results, the researchers decided to use the VGG-16 model for further studies.

2) CNN MODELING USING SLFE POOLING MODULE AND OTHER EXISTING POOLING METHODOLOGIES

In this second phase of work, the selected VGG model is considered for the training, validation and testing to find the effectiveness of the proposed SLFE pooling module in a CNN framework. The placement of SLFE module is depicted in the Fig.6.

The model underwent training over 40 epochs with Adam Optimizer and was tested on 183 images using a GPU Tesla T4 taking about 1 second or approximately 5 milliseconds per image. As the model advanced through its training epochs, its performance was carefully monitored. The resulting accuracy curve depicted in Fig.7 provides a clear picture of how each pooling method's effectiveness evolved over time. During the initial training, all methods showed low accuracy. By epoch 5, all methods improved, with AVG and SLFE pooling showing slightly higher accuracy. By epoch

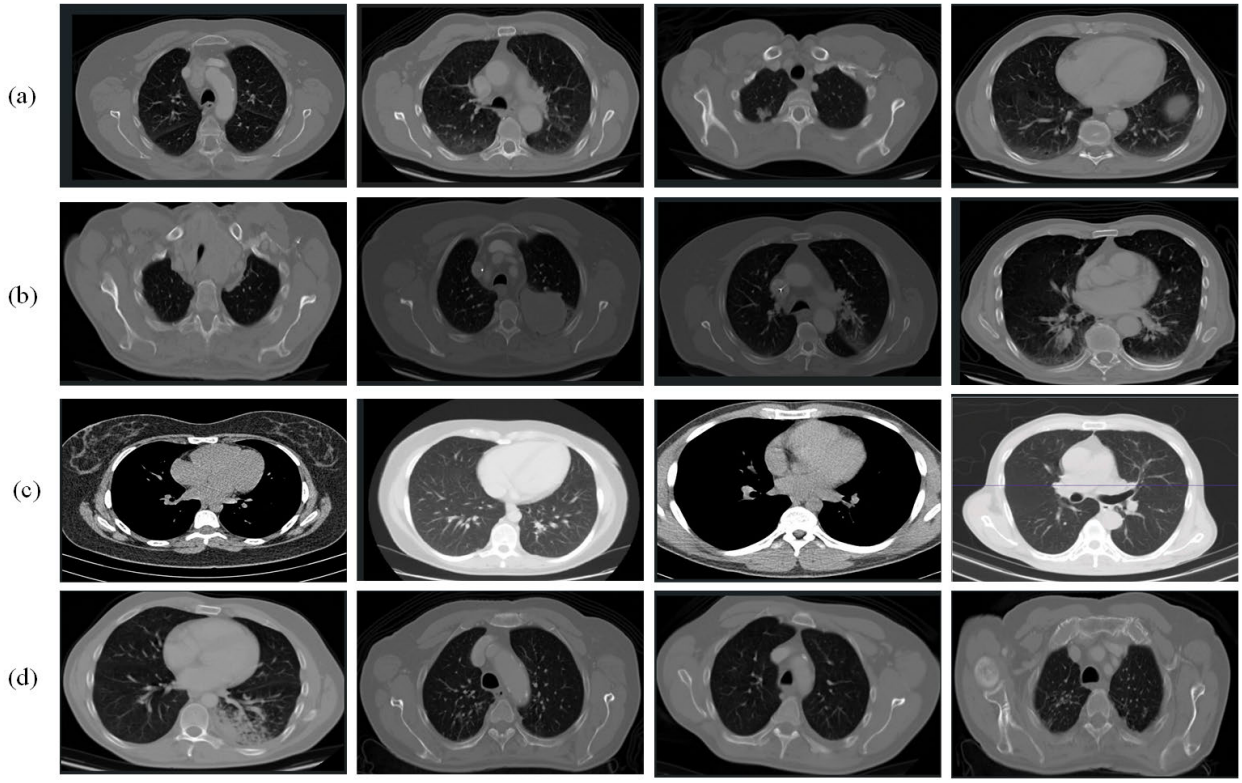

FIGURE 4. Multi-class chest CT lung cancer image dataset: (a)adenocarcinoma (b)large cell carcinoma (c)normal (d)squamous cell carcinoma.

TABLE 5. Evaluation metrics of raw images with different pooling approaches.

		Different Pooling Approaches			
Image Nature	Evaluation Metrics	MAX	AVG	AAD	SLFE*
Raw Images	Computation Time (Sec)	343	447	512	379*
	Accuracy	93.44	93.44	95.08	96.72*
	Sensitivity	91.74	91.74	95.00	95.64*
	Specificity	97.66	97.75	98.30	98.30*
	Precision	94.95	94.27	95.22	97.10*
	Recall	91.73	91.73	94.99	95.64*
	F1 Score	92.75	92.49	94.88	96.29*

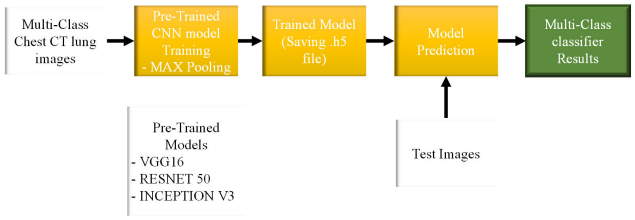

FIGURE 5. Workflow using pretrained models.

15, MAX and AVG pooling reached same accuracy level, while SLFE pooling remained competitive. At epoch 30, SLFE led in accuracy, demonstrating its strength in learning from the training data. Finally, by epoch 40, AAD and SLFE maintained peak performance while all other methods converged towards high accuracy values. Among AAD and SLFE, SLFE demonstrated superior accuracy levels. Therefore, the proposed SLFE pooling has an exceptional ability to adapt and improve its classification accuracy over time. Encouraged by these positive results, the more extensive evaluation of the model's performance using a wide range of performance metrics was conducted. The obtained performance metrics values are highlighted in Table 5.

The performance metrics of several works along with the proposed SLFE pooling method, are tabulated in Table 6. It is clearly evident that the proposed pooling approach outperforms the performance metrics of [26], [29], [30], and [15] in all aspects. This indicates that the SLFE pooling method can significantly influence the performance of the model. But, the proposed method exhibits a minor trade-off in sensitivity when compared to [32]. Despite this, the SLFE pooling method contributes a good balance of accuracy, sensitivity, specificity, and precision values; enabling the model to generalize better across various unseen datasets. It's often challenging to optimize all performance metrics simultaneously, and improving one metric may lead to a

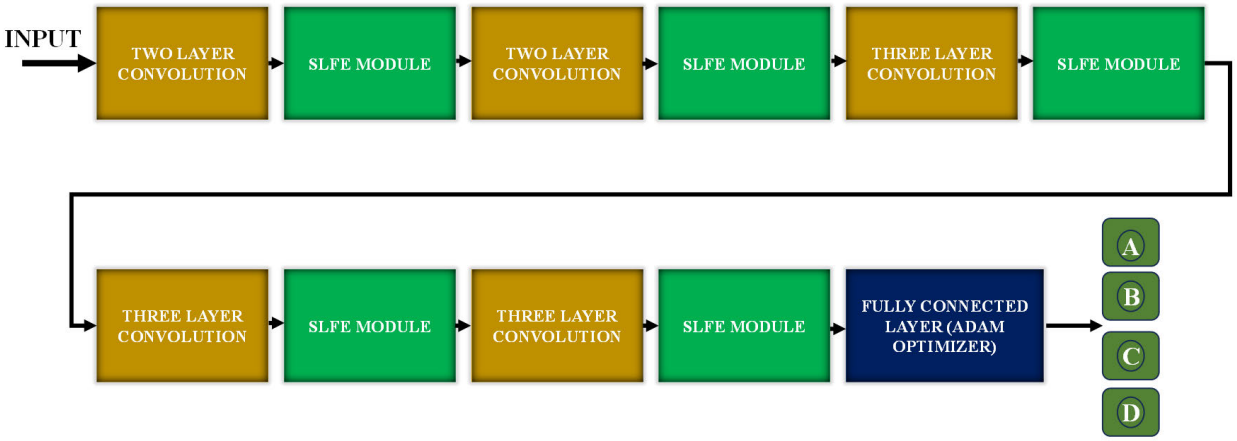

FIGURE 6. CNN-SLFE framework.

TABLE 6. Comparative classification accuracy analysis of SLFE pooling with existing pooling approaches for medical images.

	Ref [26]	Ref [30]	Ref [32]	Ref [29]	Ref [15]	Proposed SLFE*
Method	MAX, GAP	MAX	MAX	MAX	AAD	SLFE*
Accuracy	81.87	79.3	93.8	96.5	87.95	96.72*
Sensitivity	83.01	83	99.73	57.36	86.92	95.64*
Specificity	82.0	67	90	45.27	86.41	98.30*
Precision	91.6	55	93	51	85.05	97.10*

decrease in another. However, the overall performance of a model is usually considered more important than individual metrics. Hence, the proposed balanced model is suggested as a promising approach for future studies.

3) FEATURE EXTRACTION USING EGDR MODULE: CREATION OF SYNTHESIZED IMAGE

EGDR module is designed to extract relevant features from the input image, which are then used for model training. The effectiveness of a feature extractor can vary depending on the specific characteristics of the dataset and the complexity of the features being extracted [43], [44]. The visual representation of the synthesized image of each class is shown in Fig.8. Step-by-step breakdown of the process:

- Image Preprocessing: The raw image, which is typically in a three-dimensional format, is first converted to a grayscale representation. This grayscale image is then resized to match the target CNN model size of 224×224 pixels.
- Tile Division: The resized image is divided into several 4×4 tiles. The division is done in such a way that no two tiles overlap with each other.
- Feature Extraction: Each tile undergoes feature extraction using the Winograd algorithm combined with the vertical Sobel edge detector value. This process generates a 2×2 output tile for each input tile.
- Image Synthesis: The output tiles are appended to create a synthesized image of size 112×112 pixels as shown in Fig. 8. This size reduction is achieved because each 4×4 input tile is transformed into a 2×2 output tile.

4) CNN MODELING USING EGDR MODULE AND DIFFERENT POOLING APPROACHES

In the final phase of the work, the model proposed in Fig.2 undergoes an evaluation process to quantitatively determine the extent to which the proposed methodology meets the defined objective. The synthesized image from the EGDR module is fed into the CNN-SLFE framework training and validation process in which several epochs are carried out and later used for model prediction. Thus, the performance of the synthesized image on the proposed framework testing process is calculated.

Also, a comprehensive comparative analysis was conducted for four different VGG CNN models with MAX, AVG and AAD along with the proposed SLFE pooling. This approach facilitated a detailed understanding of how each method influences feature aggregation within the model. Upon completion of the training, the model was then utilized in an inference procedure with unseen datasets.

From Table 7, it is clearly proven that the proposed SLFE pooling method, when combined with the feature extractor, EGDR module has demonstrated excellent performance metrics in chest CT analysis. Achieving an accuracy of 89.62%, sensitivity of 89.42%, specificity of 96.13%, precision of 92.66%, recall of 89.42% and an F1 score of 90.35%, the proposed methodology outperforms conventional pooling techniques by a significant margin. The proposed DFPT-CNN model led to commendable 17.94% reduction in computation time, while still maintaining a better accuracy. The FPR and FNR for each class has been computed across various EGDR-CNN frameworks. As observed from Table8 and

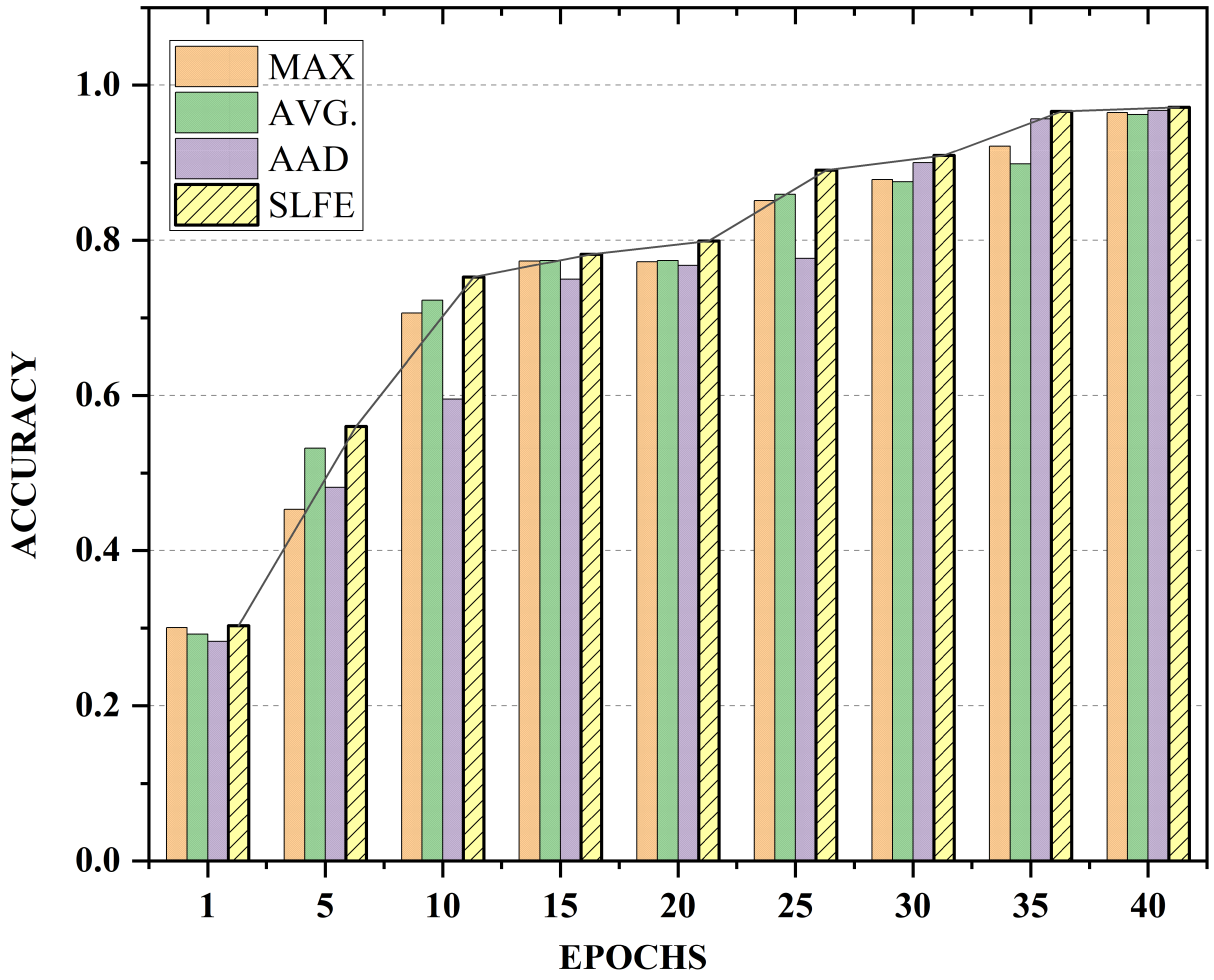

FIGURE 7. Accuracy representation of different pooling methods.

TABLE 7. Evaluation metrics of proposed DFPT-CNN model with different pooling approaches.

		Different Pooling Approaches			
Image Nature	Evaluation Metrics	MAX	AVG	AAD	SLFE*
Proposed EGDR module Feature Extracted Images	Computation Time (Sec)	310	330	424	311*
	Accuracy	76.50	81.97	83.61	89.62*
	Sensitivity	74.99	83.36	84.92	89.42*
	Specificity	91.09	93.90	94.50	96.13*
	Precision	87.51	85.83	86.69	92.66*
	Recall	74.98	83.36	84.92	89.42*
	F1 Score	76.23	83.06	82.83	90.35*

TABLE 8. Comparison of False Positive Rate (FPR) of proposed DFPT-CNN model with other CNN frameworks.

	Different Pooling Approaches			
Class Label	EGDR-CNN-MAX	EGDR-CNN-AVG	EGDR-CNN-AAD	Proposed Model*
AdenoCarcinoma	0.3416	0.0166	0.1083	0.1333*
Large cell Carcinoma	0	0.0065	0.1118	0.0065*
Normal	0	0.0419	0	0*
Squamous cell Carcinoma	0.0149	0.1791	0	0.0149*
Average	0.0891	0.0610	0.0550	0.0386*

Table9, the proposed DFPT-CNN model demonstrates a lower FPR and FNR, suggesting a reduced likelihood of incorrectly classifying the negative classes as positive and positive classes as negative respectively. Thus, the adaptability of the SLFE pooling method, based on pixel intensity comparisons with the global average value of the

TABLE 9. Comparison of False Negative Rate (FNR) of proposed DFPT-CNN model with other CNN frameworks.

	Different Pooling Approaches			
Class Label	EGDR-CNN-MAX	EGDR-CNN-AVG	EGDR-CNN-AAD	Proposed Model*
AdenoCarcinoma	0	0.3333	0.0476	0.0158*
Large cell Carcinoma	0.3225	0.2258	0	0.0967*
Normal	0.025	0.025	0.025	0.025*
Squamous cell Carcinoma	0.6530	0.0816	0.5306	0.2857*
Average	0.2501	0.1664	0.1508	0.1058*

TABLE 10. Comparison of SLFE and EGDR modules with different pretrained models.

Feature Extraction Modules	Different Pretrained Models	Accuracy(%)	Precision(%)	Recall(%)	F1Score(%)
Raw Images with predefined pool	VGG16	93.44	94.95	91.73	92.75
	VGG19	94.53	95.65	93.64	94.41
	ResNet50	68.85	79.51	69.73	69.18
	InceptionV3	90.16	91.96	88.85	90.00
	EfficientNet	26.77	6.73	25	10.60
Raw Images with SLFE	VGG16	96.72	97.10	95.64	96.29
	VGG19	95.62	95.23	95.84	95.43
	ResNet50	81.42	82.45	80.78	80.80
	InceptionV3	92.89	94.44	92.47	93.19
	EfficientNet	34.42	8.60	25	12.80
EGDR with Predefined pool	VGG16	76.50	87.51	74.98	76.23
	VGG19	68.85	75.52	68.26	63.51
	ResNet50	46.99	42.94	55.54	42.10
	InceptionV3	76.50	84.89	73.77	76.21
	EfficientNet	16.93	4.23	25	7.24
Proposed EGDR with SLFE	VGG16	89.6*	92.66	89.42	90.35
	VGG19	86.88	86.90	88.04	86.98
	ResNet50	56.28	36.01	50	40.29
	InceptionV3	79.78	85.19	77.95	79.99
	EfficientNet	26.77	6.69	25	10.56

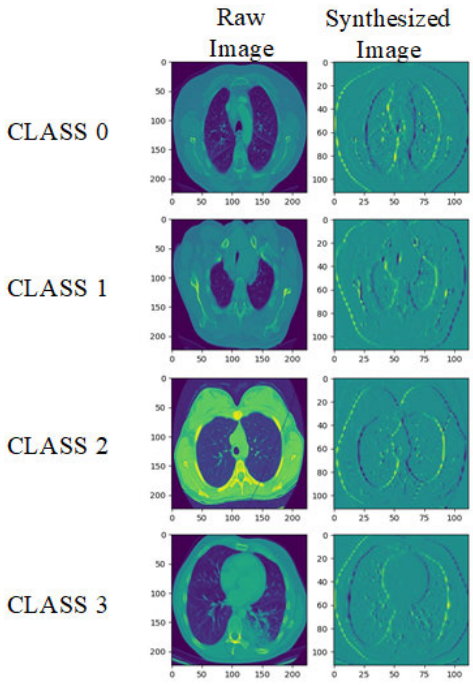

FIGURE 8. EGDR module: Feature extraction representation.

feature maps, has shown a remarkable ability to detect both global trends and subtle local variations in chest CT data. Moreover, integrating the Winograd convolution algorithm with the Sobel vertical edge detector has markedly enhanced the feature extraction process. By focusing on vertical edges, this approach effectively emphasizes important anatomical structures, which are vital in medical imaging. This combination of innovative techniques not only optimizes computational resources but also refines the detection of key patterns within CT images. These results highlight the importance of this work and suggest potential advancements in chest CT analysis.

In addition to the above, a comprehensive comparison of dual feature extraction approach on five different pretrained models was conducted. Each of these models was trained and tested independently under four different feature extraction modules and the results are tabulated in Table 10, providing a clear overview of the accuracy, precision, recall, and F1 score values for each model. The analysis revealed that the model incorporating the SLFE pooling method consistently outperformed those using predefined pooling methods. This superior performance can be attributed to the SLFE unique approach of comparing pixel intensities with the global average value of the feature maps, allowing it to effectively capture both global trends and subtle local variations in the chest CT data. Among the evaluated models, VGG16 demonstrated superior performance compared to the other models, namely VGG19, ResNet50, InceptionV3, and EfficientNet. This suggests that VGG16, when combined with the SLFE pooling method and EGDR module, provides a highly effective solution for chest CT data analysis. This

finding underscores the importance of choosing the right model and configuration for specific tasks and datasets.

Despite the impressive performance demonstrated by the proposed SLFE pooling method combined with the feature extractor EGDR module in chest CT analysis, there are certain limitations to acknowledge. While achieving commendable accuracy, sensitivity, specificity, precision, recall, and F1 score, it's essential to consider that without the EGDR module, all performance metrics were initially above 90%. With the introduction of frame size reduction, these metrics dropped below the 90% threshold, though still maintained above 85%. This suggests a trade-off between computational speed and performance metrics, where the model sacrifices some accuracy for faster computation. Nevertheless, the adaptability of the proposed methodology across various datasets offer promising avenues for further exploration and improvement in prediction analysis.

V. CONCLUSION

To ensure optimal classification accuracy while minimizing computational overhead and timing demands, this research implemented a unique dual feature extraction approach. Four experiments were conducted on a multi-class lung cancer classification dataset to validate the efficacy of the proposed methodology. Firstly, three pretrained CNN architectures such as VGG16, ResNet50, and InceptionV3 were evaluated to identify the optimal model in terms of parameters and classification accuracy and the results showed that VGG16 was the most suitable choice. In the second experiment, the model underwent training, validating, and testing with four different pooling methods-MAX, AVG, AAD, and the proposed SLFE to assess the ability of the proposed pooling algorithm for feature extraction. The classification accuracy values obtained from the four pooling methods confirmed that the proposed pooling approach yielded a noteworthy increase in accuracy, outperforming the MAX and AAD pooling methods by 3.5% and 1.72%, respectively. Subsequently, efforts focused on reducing computation overhead and training time by adopting the Edge Gradient-Dimensionality Reduction (EGDR) approach during preprocessing. In the final experiment, the features extracted from the EGDR approach was fed to the proposed CNN model with SLFE module for training, validation, and testing. The testing process was conducted using an unseen dataset. The comparative performance analysis of the proposed model with other EGDR-CNN frameworks, demonstrated noteworthy results. The proposed model achieved the classification accuracy of approximately 1.172 times greater to that of the CNN-MAX framework. The computational time of the proposed CNN-SLFE framework is reduced by 17.94% when compared to other CNN frameworks. This underscores the effectiveness of the proposed methodology in achieving higher classification accuracy with reduced computational demands. This research presents a promising framework for classification tasks especially in the medical domain, offering substantial contributions to the field's advancement.

REFERENCES

[1] M. M. Hasan, M. M. Hossain, M. M. Rahman, A. Azad, S. A. Alyami, and M. A. Moni, "FP-CNN: Fuzzy pooling-based convolutional neural network for lung ultrasound image classification with explainable AI," *Comput. Biol. Med.*, vol. 165, Oct. 2023, Art. no. 107407.

[2] S. H. Khan, J. Iqbal, S. A. Hassnain, M. Owais, S. M. Mostafa, M. Hadjouni, and A. Mahmoud, "COVID-19 detection and analysis from lung CT images using novel channel boosted CNNs," *Expert Syst. Appl.*, vol. 229, Nov. 2023, Art. no. 120477.

[3] J. Zhang, Z. Shangguan, W. Gong, and Y. Cheng, "A novel denoising method for low-dose CT images based on transformer and CNN," *Comput. Biol. Med.*, vol. 163, Sep. 2023, Art. no. 107162.

[4] G. Ali, A. Dastgir, M. W. Iqbal, M. Anwar, and M. Faheem, "A hybrid convolutional neural network model for automatic diabetic retinopathy classification from fundus images," *IEEE J. Translational Eng. Health Med.*, vol. 11, no. 1, pp. 341–350, Jun. 2023.

[5] M. Obayya, M. A. Arasi, N. Alruwais, R. Alsini, A. Mohamed, and I. Yaseen, "Biomedical image analysis for colon and lung cancer detection using tuna swarm algorithm with deep learning model," *IEEE Access*, vol. 11, pp. 94705–94712, 2023.

[6] S. Solanki, U. P. Singh, S. S. Chouhan, and S. Jain, "Brain tumor detection and classification using intelligence techniques: An overview," *IEEE Access*, vol. 11, pp. 12870–12886, 2023.

[7] M. A. Ottom, H. A. Rahman, and I. D. Dinov, "Znet: Deep learning approach for 2D MRI brain tumor segmentation," *IEEE J. Translational Eng. Health Med.*, vol. 10, no. 1, pp. 1–8, May 2022.

[8] M. Subramanian, J. Cho, V. E. Sathishkumar, and O. S. Naren, "Multiple types of cancer classification using CT/MRI images based on learning without forgetting powered deep learning models," *IEEE Access*, vol. 11, pp. 10336–10354, 2023.

[9] A. Heidenreich, "Modern approach of diagnosis and management of acute flank pain: Review of all imaging modalities," *Eur. Urology*, vol. 41, no. 4, pp. 351–362, Apr. 2002.

[10] J. Guo, W. Cao, B. Nie, and Q. Qin, "Unsupervised learning composite network to reduce training cost of deep learning model for colorectal cancer diagnosis," *IEEE J. Translational Eng. Health Med.*, vol. 11, no. 1, pp. 54–59, Jun. 2023.

[11] D. R. Sarvamangala and R. V. Kulkarni, "Convolutional neural networks in medical image understanding: A survey," *Evol. Intell.*, vol. 15, no. 1, pp. 1–22, Mar. 2022.

[12] M. Momeny, A. A. Neshat, A. Gholizadeh, A. Jafarnezhad, E. Rahmanzadeh, M. Marhamati, B. Moradi, A. Ghafoorifar, and Y.-D. Zhang, "Greedy autoaugment for classification of mycobacterium tuberculosis image via generalized deep CNN using mixed pooling based on minimum square rough entropy," *Comput. Biol. Med.*, vol. 141, Feb. 2022, Art. no. 105175.

[13] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, "Medical image analysis using convolutional neural networks: A review," *J. Med. Syst.*, vol. 42, no. 11, pp. 1–13, Nov. 2018.

[14] M. M. Rahman, B. C. Desai, and P. Bhattacharya, "Medical image retrieval with probabilistic multi-class support vector machine classifiers and adaptive similarity fusion," *Computerized Med. Imag. Graph.*, vol. 32, no. 2, pp. 95–108, Mar. 2008.

[15] K. Khalil, O. Eldash, A. Kumar, and M. Bayoumi, "Designing novel AAD pooling in hardware for a convolutional neural network accelerator," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 30, no. 3, pp. 303–314, Mar. 2022.

[16] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," 2017, *arXiv:1712.04621*.

[17] W. Cheng, I.-C. Lin, and Y.-Y. Shih, "An efficient implementation of convolutional neural network with CLIP-Q quantization on FPGA," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 10, pp. 4093–4102, Oct. 2022.

[18] E. Ayan and H. M. Ünver, "Diagnosis of pneumonia from chest X-ray images using deep learning," in *Proc. Sci. Meeting Elect.-Electron. Biomed. Eng. Comput. Sci. (EBBT)*, Apr. 2019, pp. 1–5.

[19] D. S. Kermany et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, Feb. 2018.

[20] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," in *Classification in BioApps*. Cham, Switzerland: Springer, 2018, pp. 323–350.

[21] Q. Song, L. Zhao, X. Luo, and X. Dou, "Using deep learning for classification of lung nodules on computed tomography images," *J. Healthcare Eng.*, vol. 2017, pp. 1–7, Jun. 2017.
[22] H. Xie, D. Yang, N. Sun, Z. Chen, and Y. Zhang, "Automated pulmonary nodule detection in CT images using deep convolutional neural networks," *Pattern Recognit.*, vol. 85, pp. 109–119, Jan. 2019.
[23] N. Kalaivani, N. Manimaran, S. Sophia, and D. Devi, "Deep learning based lung cancer detection and classification," in *Proc. IOP Conf. Series, Mater. Sci. Eng.*, vol. 994, 2020, p. 12026.
[24] A. R. Bushara, R. S. V. Kumar, and S. S. Kumar, "LCD-capsule network for the detection and classification of lung cancer on computed tomography images," *Multimedia Tools Appl.*, vol. 82, no. 24, pp. 37573–37592, Oct. 2023.
[25] S. Nigudgi and C. Bhyri, "Lung cancer CT image classification using hybrid-SVM transfer learning approach," *Soft Comput.*, vol. 27, no. 14, pp. 9845–9859, Jul. 2023.
[26] N. Faruqui, M. A. Yousuf, M. Whaiduzzaman, A. K. M. Azad, A. Barros, and M. A. Moni, "LungNet: A hybrid deep-CNN model for lung cancer diagnosis using CT and wearable sensor-based medical IoT data," *Comput. Biol. Med.*, vol. 139, Dec. 2021, Art. no. 104961.
[27] M. Hariri and E. Avşar, "COVID-19 and pneumonia diagnosis from chest X-ray images using convolutional neural networks," *Netw. Model. Anal. Health Informat. Bioinf.*, vol. 12, no. 1, p. 17, Mar. 2023.
[28] G. Habib and S. Qureshi, "GAPCNN with HyPar: Global average pooling convolutional neural network with novel NNLU activation function and HYBRID parallelism," *Frontiers Comput. Neurosci.*, vol. 16, Nov. 2022, Art. no. 1004988.
[29] R. Tandon, S. Agrawal, R. Raghuwanshi, N. P. S. Rathore, L. Prasad, and V. Jain, "Automatic lung carcinoma identification and classification in CT images using CNN deep learning model," in *Augmented Intelligence in Healthcare: A Pragmatic and Integrated Analysis*. Cham, Switzerland: Springer, 2022, pp. 143–166.
[30] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, and B. Xu, "A deep learning algorithm using CT images to screen for corona virus disease (COVID-19)," *Eur. Radiol.*, vol. 31, pp. 6096–6104, Feb. 2021.
[31] G. A. P. Singh and P. K. Gupta, "Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans," *Neural Comput. Appl.*, vol. 31, no. 10, pp. 6863–6877, Oct. 2019.
[32] V. Perumal, V. Narayanan, and S. J. S. Rajasekar, "Detection of COVID-19 using CXR and CT images using transfer learning and Haralick features," *Int. J. Speech Technol.*, vol. 51, no. 1, pp. 341–358, Jan. 2021.
[33] W. Zeng, W. Li, M. Zhang, H. Wang, M. Lv, Y. Yang, and R. Tao, "Microscopic hyperspectral image classification based on fusion transformer with parallel CNN," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 6, pp. 1–12, Jun. 2023.
[34] A. Masood, P. Yang, B. Sheng, H. Li, P. Li, J. Qin, V. Lanfranchi, J. Kim, and D. D. Feng, "Cloud-based automated clinical decision support system for detection and diagnosis of lung cancer in chest CT," *IEEE J. Translational Eng. Health Med.*, vol. 8, no. 1, pp. 1–13, Oct. 2020.
[35] Y. Liang, L. Lu, Q. Xiao, and S. Yan, "Evaluating fast algorithms for convolutional neural networks on FPGAs," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 39, no. 4, pp. 857–870, Apr. 2020.
[36] S. Kala, B. R. Jose, J. Mathew, and S. Nalesh, "High-performance CNN accelerator on FPGA using unified winograd-GEMM architecture," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 27, no. 12, pp. 2816–2828, Dec. 2019.
[37] Z.-S. Syu and C.-H. Lee, "One-dimensional binary convolutional neural network accelerator design for bearing fault diagnosis," *IEEE Sensors J.*, vol. 24, no. 3, pp. 3649–3658, Feb. 2024.
[38] J. Wang, J. Lin, and Z. Wang, "Efficient hardware architectures for deep convolutional neural network," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 6, pp. 1941–1953, Jun. 2018.
[39] Y. Tang, Q. Teng, L. Zhang, F. Min, and J. He, "Layer-wise training convolutional neural networks with smaller filters for human activity recognition using wearable sensors," *IEEE Sensors J.*, vol. 21, no. 1, pp. 581–592, Jan. 2021.
[40] C. N. Karthick and P. Nirmala, "Smart edge detection technique in X-ray images for improving PSNR using Sobel edge detection algorithm with Gaussian filter in comparison with Laplacian algorithm," *CARDIOMETRY*, no. 25, pp. 1751–1757, Feb. 2023.
[41] M. A. Sp, K. S. V. Pradyumna, C. Nithesh, D. A. Varughese, and S. Sridevi, "Evaluating Winograd algorithm for convolution neural network using verilog," in *Proc. IEEE Int. Symp. Smart Electron. Syst.*, Dec. 2022, pp. 582–585.
[42] D. Varughese and S. Sridevi, "Optimum resource utilization for the implementation of FPGA-based fast convolutional algorithms for CNN modelling," in *Proc. IEEE Women Technol. Conf. (WINTECHCON)*, Sep. 2023, pp. 1–6.
[43] L. Peng, C. Yang, Y. Chen, and W. Liu, "Predicting CircRNA-disease associations via feature convolution learning with heterogeneous graph attention network," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 6, pp. 1–11, Jun. 2023.
[44] Z. Shi, Z. Liao, and H. Tabata, "Enhancing performance of convolutional neural network-based epileptic electroencephalogram diagnosis by asymmetric stochastic resonance," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 9, pp. 1–12, Sep. 2023.

DINAH ANN VARUGHESE received the Bachelor of Technology degree in electronics and communication engineering from Mahatma Gandhi University, Kerala, India, in 2009, and the Master of Technology degree in VLSI and embedded system from CUSAT University, Kerala, in 2011. She is currently pursuing the Ph.D. degree with the Department of Micro and Nanoelectronics, Vellore Institute of Technology, Vellore, Tamil Nadu, India. Her research interests include hardware acceleration of convolutional neural networks, medical image processing, machine learning, and deep learning.

SRIADIBHATLA SRIDEVI (Senior Member, IEEE) received the B.Tech. degree in electronics and communication engineering and the M.Tech. degree in communication engineering from Jawaharlal Nehru Technological University, Hyderabad, India, and the Ph.D. degree in VLSI design from Andhra University, Visakhapatnam, India. Since August 2013, she has been with the School of Electronics Engineering, Vellore Institute of Technology (VIT), Vellore, India, where she is currently a Professor with the Department of Micro and Nano Electronics. She has 25 years of teaching experience. Four scholars have completed Ph.D. studies successfully under her guidance, and also guiding three Ph.D. scholars. She has published 55 papers in international journals and conferences. Her current research interests include low power memory design, low power digital design, and VLSI signal processing. She is a Life Member of the Instrument Society of India (ISOI).