IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 7 October 2023, accepted 21 November 2023, date of publication 30 November 2023,
date of current version 21 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3338228

RESEARCH ARTICLE

EfDenseNet: Automated Pulmonary Hypertension Detection Model Based on EfficientNetb0 and DenseNet201 Using CT Images

TARIK KIVRAK[1], JAGADISH NAYAK[2], (Senior Member, IEEE), MEHMET ALI GELEN[1], PRABAL DATTA BARUA[3], MEHMET BAYGIN[4], HILAL ERKEN PAMUKCU[5], SENGUL DOGAN[6], TURKER TUNCER[6], AND U. RAJENDRA ACHARYA[7,8], (Senior Member, IEEE)

[1]Department Cardiology, Fırat University Hospital, Fırat University, 23119 Elâzığ, Turkey
[2]Electrical and Electronics Engineering Department, Birla Institute of Technology and Science Pilani, Dubai Campus, Dubai, United Arab Emirates
[3]School of Business (Information System), University of Southern Queensland, Springfield, QLD 4350, Australia
[4]Department of Computer Engineering, Faculty of Engineering and Architecture, Erzurum Technical University, 25050 Erzurum, Turkey
[5]Department of Cardiology, Etlik Education and Training Hospital, 06170 Ankara, Turkey
[6]Department of Digital Forensics Engineering, Technology Faculty, Fırat University, 23119 Elâzığ, Turkey
[7]School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, QLD 4350, Australia
[8]Centre for Health Research, University of Southern Queensland, Springfield, QLD 4350, Australia

Corresponding author: Sengul Dogan (sdogan@firat.edu.tr)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Local Ethical Committee, Ethics Committee of Firat University (2022/04-06).

ABSTRACT Pulmonary hypertension (PH) is a chronic and progressive disease. We introduced a novel automated self-organized feature engineering architecture for PH detection, which was trained and refined using a new thoracic CT image dataset. This study's dataset includes 807 transverse contrast-enhanced CT images from 313 patients, categorized into four groups: Group 1 with 20 mmHg $\leq$ mean pulmonary artery pressure (mPAP) $<$ 25 mmHg; Group 2 with 25 mmHg $\leq$ mPAP $\leq$ 30 mmHg; Group 3 where mPAP $>$ 30 mmHg; and a control group with no PH. Our model consists of four primary stages: (i) generation of features based on combinations from nested patches, (ii) feature selection, (iii) classification and (iv) majority voting. CT images were segmented into nested patches, each being processed through pretrained EfficientNetB0 and DenseNet201 to derive four deep feature vectors, utilizing both the global average pooling and fully connected layers of these networks. These four extracted features underwent combinatorial operations, resulting in 15 feature vectors. Subsequently, these vectors were introduced to neighborhood component analysis, ReliefF, and Chi2 feature selectors. This process yielded 45 refined feature vectors with diminished data dimensions. These selected vectors were then processed through a support vector machine and k-nearest neighbors classifiers, producing 90 predictive vectors. By applying mode-based iterative majority voting to these vectors, an additional 88 voted prediction vectors were generated, leading to a total of 178 classifier-generated and voted prediction vectors. The optimal classification result was selected from these 178 vectors. With the use of 10-fold cross-validation, our model achieved a remarkable 97.27% overall accuracy for the 4-class classification on the study dataset. Owing to its reduced time complexity, this model is practical for CT-based PH screenings.

INDEX TERMS Artificial intelligence, computed tomography, mean pulmonary arterial pressure, pulmonary hypertension.

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

I. INTRODUCTION

A. BACKGROUND

Pulmonary hypertension (PH) is a group of diseases of various etiologies [1], [2] that cause high blood pressure in the pulmonary arteries [3], [4]. The chronically raised pulmonary pressure induces adverse remodeling in both cardiac and respiratory systems [5], and can manifest clinically as shortness of breath, fatigue, chest pain, and premature death [6], [7]. The disease is progressive; early and accurate diagnosis is important for prognostication and treatment decisions [8]: certain PH-specific drug therapy can ameliorate symptoms and retard disease progression [9]. The diagnosis of PH is established when invasively measured mean pulmonary artery pressure exceeds 20 mmHg [3], [4]. but the requisite right heart catheterization procedure carries risks and is expensive [6]. In suspect PH cases, noninvasive investigations like blood biomarkers, electrocardiography (ECG), chest X-ray, echocardiography, thoracic computed tomography (CT) and magnetic resonance imaging (MRI) may help corroborate the presence of morphological and/or hemodynamic changes in the heart chambers and lung vasculature consequent to PH-induced remodeling [10].

Artificial intelligence-enabled models are increasingly being used to facilitate disease diagnosis [11], especially in the classification of medical images, like CT, MRI, and X-ray [12]. While research interest in automated PH classification has burgeoned, the paucity of open-access PH datasets has limited the development of machine-learning methods. Aras et al. [13] proposed a ResNet-based deep learning approach for automatic PH classification using ECG signals. Trained on 5016 PH and 19454 non-PH ECG signals, their model attained modest 79%, 84%, and 89% sensitivity, specificity, and C-statistic for binary classification using a 70:10:20 hold-out validation strategy. Gudigar et al. [14] developed a classification model based on global weighted local binary pattern, fuzzy entropy and support vector machine (SVM) methods. On a small balanced dataset of 49 PH and 49 non-PH ultrasound images, they reported 91.77% classification accuracy. Ong et al. [15] compared decision rules versus machine learning approaches to PH screening. Using the electronic health records of 386 PH and 164 non-PH patients, they concluded that extant decision rules performed poorly.

The phenomenon of vasodilation in pulmonary hypertension results in the morphological enlargement of vessels with high pulmonary arterial pressure, which is discernible on CT imaging. Recognizing the potential of leveraging this morphological distinction, we endeavor to develop an artificial intelligence (AI)–based diagnostic system in this study. To achieve this objective, we introduce a novel nested-patch-based deep feature extraction model designed to automatically detect and classify the different types of vessel morphological changes associated with the disease.

In this study, we propose a novel self-organized deep feature engineering model using deep transfer learning. Our proposed model generates 178 outputs, and our architecture selects the output with the highest classification accuracy, making it a self-organized model. Furthermore, we evaluate the performance of the utilized models based on the findings of our own model.

B. MOTIVATION AND OUR ARCHITECTURES

We present an innovative automated model for PH diagnosis based on CT images, which has been lacking in the literature. In the domain of medical image classification, many deep learning models have been explored [16], [17] that relied on end-to-end training or transfer learning approaches. Here, we proposed a novel self-organized learning architecture that comprised (1) nested patch division; (2) two pretrained deep network architectures for multiple feature vector extraction; (3) combinatorial operation to generate an increased number of combinations of extracted feature vectors; (4) multiple feature selectors; (5) multiple classifiers; and (6) mode-based majority voting to automatically calculate the optimal model result. The model was developed on a new PH dataset comprising thoracic CT images divided into one control (no PH) and three PH classes of increasing severity. Our model attained an excellent 97.27% overall accuracy for 4-class classification.

C. NOVELTIES AND CONTRIBUTIONS

The primary aim of this study is to introduce an innovative self-organized machine learning model that utilizes a multiple output-based approach. To achieve this objective, our proposed system involves two convolution neural networks (CNNs), a combinator, three feature selectors, two classifiers, and an iterative information fusion model, generating multiple outputs. We also implemented a nested patch division technique to detect local abnormalities, enhancing the model's classification capability.

One notable feature of our model is its automatic selection of the best result. We evaluated the classification capability of our model for detecting pulmonary hypertension, and it demonstrated exceptional performance, achieving high classification accuracy.

We have listed innovations of the model below.

- To our knowledge, our model is the first nested patch-based CT image classification model.
- Our combination-based deep feature engineering framework is parametric and is amenable to adaptation for diverse datasets and classification task applications.

Our contributions are listed below.

- A novel self-organized learning and classification architecture was proposed. Our model attained high classification for detecting PH and grading its severity using noninvasive CT images. It can be implemented to facilitate high throughput screening of CT images to assist doctors with PH diagnosis.
- A new PH CT image dataset was collected and made publicly accessible. This will stimulate the development of PH diagnostic models.

- The proposed model demonstrated that new-generation PH detection assistants/applications can be developed since our model attained over 97% classification accuracy on real-world CT images.

II. DATASET

A new thoracic CT dataset comprising 807 transverse contrast-enhanced CT images of the pulmonary artery bifurcation (Figure 1) was curated from the imaging records of 313 patients who had attended the Cardiology Clinic, Firat University Hospital from 01/01/2016 to 31/12/2022. The retrospective analysis of these images and the relevant clinical records had been approved by the hospital ethics committee. The patients were stratified into one control group without known PH and three groups of patients with progressively more severe PH as quantified by the mean pulmonary artery pressure (mPAP) measured on invasive right heart catheterization [18]: Group 1, 20 mmHg $\leq$ mPAP $<$25 mmHg; Group 2, 25 mmHg $\leq$mPAP $\leq$30 mmHg; and Group 3, mPAP $>$30 mmHg. Patient characteristics and the number of CT images are summarized in Table 1. The CT dataset is available for download at https://www.kaggle.com/datasets/turkertuncer/ph-ct-v1 URL. Herein, the axial slice with the clearest relationship between the pulmonary artery and the aorta was selected.

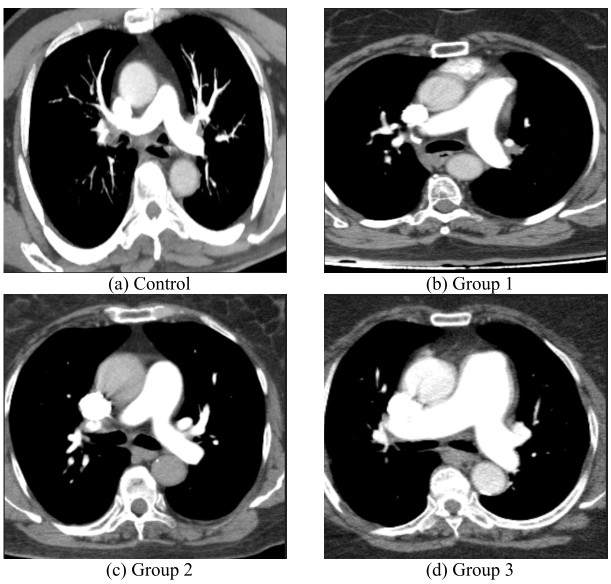


FIGURE 1. Example transverse thoracic (axial) CT images in the control (a) and pulmonary artery hypertension patient groups (b to d) in the study dataset. The bright inverted Y structure in the approximate center of each image is the contrast-enhanced lumen of the pulmonary artery, which is shown bifurcating into the right and left branch pulmonary arteries. In Group 3, the pulmonary artery dimensions appear to be comparatively larger (images are not depicted to scale).

III. THE EFDENSENET MODEL

The self-organized architecture comprised four main phases: (1) combination-based feature generation from nested patches; (2) feature selection; (3) classification; and (4) majority voting of results (Figure 2). First, CT images were divided into nested patches, each of which was fed to EfficientNetB0 [19] and DenseNet201 [20] pretrained convolutional neural networks (CNNs) to extract four deep feature vectors using the global average pooling and fully connected layers of both networks. The four extracted features next underwent a combinatorial operation, which produced 15 feature vectors. These were input to neighborhood component analysis (NCA) [21], ReliefF [22] and Chi2 [23] feature selection functions, which output 45 selected feature vectors, each with reduced data dimensionality. The selected feature vectors were fed to SVM [24] and k-nearest neighbors (kNN) classifiers [25], which calculated 90 prediction vectors. Applying mode-based iterative majority voting (IMV) [26] to the latter, an additional 88 voted prediction vectors were generated, yielding a total of 178 classifier-based plus voted prediction vectors. Finally, the best classification result was chosen from the 178 vectors. A detailed explanation of the individual phases is provided in the following sections.

TABLE 1. Patient characteristics and number of CT images stratified by group in the study dataset.

Group	Number of participants	Age (years)	Female: Male	Number of CT images
Control	114	42.7±6.1	68:46	210
Group 1	43	39.6±4.5	21:22	80
Group 2	65	41.3±5.3	39:26	130
Group 3	91	45.7±7.5	55:36	387
Total	313	42.3±5.9	183:130	807

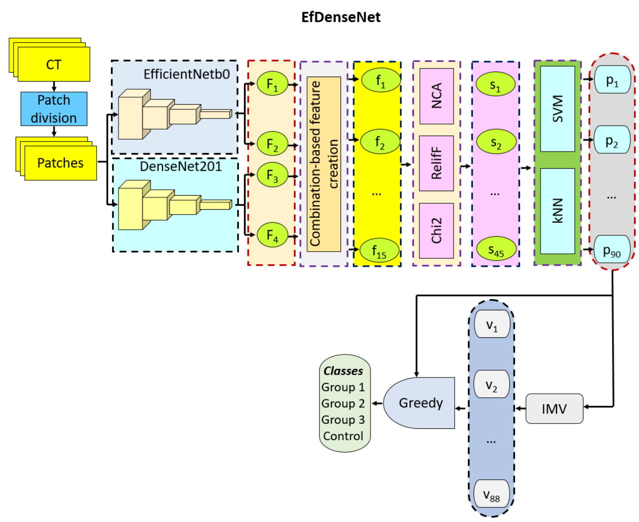


FIGURE 2. Block diagram of the EfDenseNet model. F: features extracted from deep learning architectures; f: Feature vectors generated using the combinator (combination-based feature vectors creation algorithm); s: selected meaningful features; p: classification results from classification algorithms; v: results from information fusion. See text for details.

A. FEATURE GENERATION

First, the CT images were divided into nested patches, which enabled local and global detailed feature extraction with fewer patches compared with fixed-length patches. Each nested patch was then fed to pretrained CNNs for deep feature extraction. CNNs are widely used in deep learning diverse visual processing tasks, including image classification, object detection, and segmentation. In our model, two CNN models—EfficientNetb0 and DenseNet201—that had been pretrained on ImageNet1k, dataset was deployed. EfficientNetB0 is the smallest in the EfficientNet series of CNN models that have been optimized in terms of the numbers of weights and parameters for enhanced accuracy and reduced computational cost. The pretrained model has been used successfully for various visual processing tasks [19]. Due to its small size and low computational requirements, EfficientNetB0 model is well suited for applications on mobile or low-power devices, as well as image processing tasks that require real-time or low-latency processing capabilities. DenseNet201 is a 201-layer CNN model based on the ResNet architecture, which contains dense connections. With dense connections, each layer is connected to all previous layers, which facilitates effective network learning using fewer parameters. Detailed information about the used CNNs is given below.

EfficientNetb0 [19]*:*

Input: 224 × 224

Convolutional Block: Convolutional layer + Batch Normalization + Swish activation.

MBConv Block (Mobile Inverted Residual Block): This block consists of depthwise convolution, expansion convolution, squeeze-and-excitation operation, and pointwise convolution. The number of layers and channels is determined based on the compound scaling coefficients.

Global Average Pooling: Reduces the spatial dimensions to 1 × 1. We have used this block to generate features.

Fully Connected: The output layer for classification. The used CNN was trained on ImageNet1k. Therefore, we have used this layer to get 1000 features.

DenseNet201 [20]*:*

Input: 224 × 224

Convolutional Layer: Initial convolutional layer + Batch Normalization + ReLU activation.

Dense Block: This block comprises multiple convolutional layers where each layer receives the concatenation of feature maps from all preceding layers in the block.

Transition Block: A bottleneck layer that reduces the number of channels and spatial dimensions using a convolutional layer and average pooling.

Global Average Pooling: Reduces the spatial dimensions to 1 × 1. We have used this block to generate features.

Fully Connected: The output layer for classification. By using this layer, we have generated 1000 features.

The global average pooling and fully connected layers of the above two deep learning networks were used in a self-organized manner to extract deep features, yielding four feature vectors from each input CT image (Table 2). The extracted feature vectors were input to a combinator to obtain 15 new feature vectors (Figure 3), a process that is defined by Equations (1) to (3) below.

$$f_k = Efb0\left(Im, L_k\right), \quad k \in \{1, 2\} \tag{1}$$

$$f_{k+2} = D201\left(Im, L_k\right) \tag{2}$$

$$f_{i+4} = C\left([f_1 f_2 f_3 f_4]\right), \quad i \in \{1, 2, \ldots, 11\} \tag{3}$$

where Im represents the image; f, feature vector; $Efb0(.,.)$, EfficientNetb0-based deep feature extraction function; $D201(.,.)$, DenseNet201 CNN-based deep feature generation function; L, layers (two layers were: L_1, fully connected layer; and L_2, global average pooling); and $C()$, combinator/combination function. By deploying Equations (1) - (3), 15 feature vectors have been extracted. The first four feature vectors have been generated fully connected (L_1) and global average pooling (L_2) layers of the pretrained EfficientNetb0 and DenseNet201 CNNs. The remainder 11 feature vectors have been generated using the combination.

TABLE 2. Details of the feature vectors extracted using EfficientNetb0 and DenseNet201.

Pretrained network	Layer	Extracted feature vector	Size
EfficientNetb0	Fully connected	f_1	1000
EfficientNetb0	Global average pooling	f_2	1280
DenseNet201	Fully connected	f_3	1000
DenseNet201	Global average pooling	f_4	1920

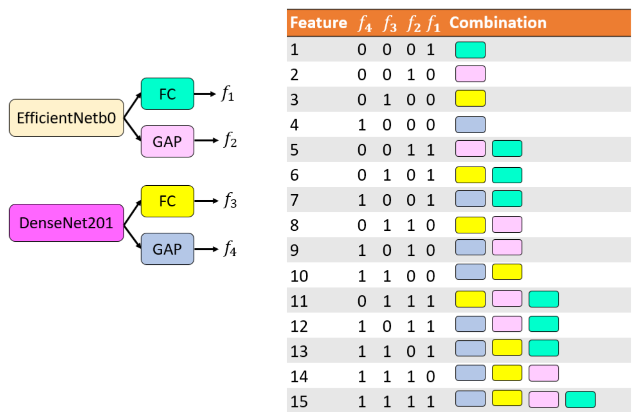


FIGURE 3. New feature vector generation by the combinator. Of 16 (=2^4) possible combinations, 15 new feature vectors were generated each containing at least one feature vector f_1 to f_4 extracted by the pretrained networks.

Nested patch division of input images, enabled the extraction of both local and global features (Figure 4). The steps are detailed below.

Step 1: Resize the input image to 256 × 256.

Step 2: Divide images into nested patches. Using expansion steps of size 32 × 32, four patches were obtained from a single image (Figure 4).

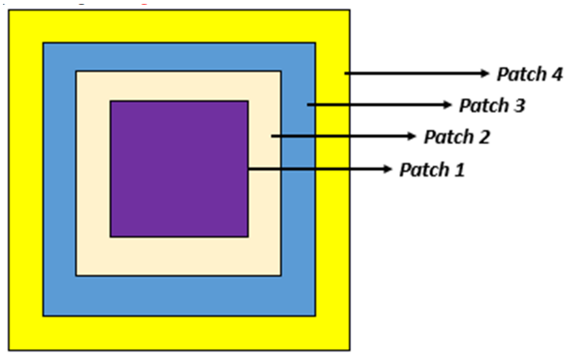


FIGURE 4. Nested patch division. Four nested patches of size 64 × 64, 128 × 128, 196 × 196, and 256 × 256 were created in this work.

Step 3: Apply deep feature extraction to each patch. This operation was performed layer-wise.

$$f_k = merge\begin{pmatrix} Efb0\left(Patch_1, L_k\right), \\ Efb0\left(Patch_2, L_k\right), \ldots, \\ Efb0\left(Patch_4, L_k\right) \end{pmatrix} \quad (4)$$

$$f_{k+2} = merge\begin{pmatrix} D201\left(Patch_1, L_k\right), \\ D201\left(Patch_2, L_k\right), \ldots, \\ D201\left(Patch_4, L_k\right) \end{pmatrix} \quad (5)$$

where ***merge***() represents the feature merge function, here, we resized the patch to a 224 × 224 sized image to use these patches as input for DensNet201 and EfficientNetb0.

Step 4: Input feature vectors to a combinator.

$$f_{i+4} = C\left(\left[f_1\, f_2\, f_3\, f_4\right]\right), \quad i \in \{5, 6, \ldots, 15\} \quad (6)$$

where ***C***(.) represents the combinator operation. The feature vectors extracted using the network layers were concatenated, which yielded 11 new feature vectors.

B. FEATURE SELECTION

To reduce data dimensionality, we used three established feature selection functions: neighborhood component analysis (NCA) [21], ReliefF [22], and Chi2 [23]. NCA is a distance-based algorithm that considers the neighborhood instances of each instance in the dataset and computes the distances between these instances. These distances are subsequently used to determine the weights of the features in the feature space, thereby facilitating optimal feature merging and selection of the best features that effectively separate the classes of instances in the dataset. NCA is able to handle high-dimensional datasets but does not account for nonlinear interactions between features in the dataset. ReliefF. ReliefF measures the importance of each feature in the dataset by running random samples and calculating the influence of neighboring samples in each sample. It creates a weight vector for each feature that is then used to determine the order of importance of the features. ReliefF can help reduce noise in the dataset, and owing to the random sampling approach, mitigate overfitting. ReliefF is effective for large and/or complex datasets and is comparatively less computationally intensive. The Chi-square (Chi2) test is a standard statistical test of the association between two categorical variables. It measures how much the observed data differs from the expected data and determines whether this difference is random. Where this difference is nonrandom, the null hypothesis is rejected and an association between the variables is accepted. To apply the test, some assumptions must be met: data are independent, expected values are sufficiently large, and variables are nominal or ordinal. The test is unreliable for small sample sizes but is highly effective for large sample sizes.

The steps of the multiple feature selection function-based feature selection in our model are detailed below.

Step 5: Apply each feature selector to each feature vector in turn to obtain a sorted feature index.

Step 6: Select the 250 most discriminative features from each feature vector based on the sorted indexes.

$$\begin{aligned} & ind^{t+15\times(h-1)} \\ & \quad = fs_h\left(f_t, y\right), \quad h \in \{1, 2, 3\},\ t \in \{1, 2, \ldots, 15\} \\ & s^{t+15\times(h-1)}\left(d, j\right) \\ & \quad = fs_h\left(d, ind^{t+15\times(h-1)}\left(j\right)\right), \\ & \quad\ d \in \{1, 2, \ldots, n\},\ j \in \{1, 2, \ldots, 250\} \end{aligned} \quad (7)$$

where fs represents the feature selector; f, feature vector; ind, feature index; s, selected feature vector; and n, number of CT images. In total, 45 (=3 × 15) selected feature vectors, each of length 250, were generated by applying NCA, ReliefF, and Chi2 to the 15 combinatorial extracted feature vectors.

C. CLASSIFICATION

Using a 10-fold cross-validation (CV) strategy, we input the selected feature vectors to established k-nearest neighbors (kNN) [25] and SVM [24] algorithms for classification. kNN is a distance-based function that considers k's closest data points and computes the average of their classes or features. Its performance is dependent on data point distribution within the dataset and the value of k value: the algorithm becomes more sensitive when k is small. In addition, the method of calculating distances between data points can impact performance. kNN is computationally expensive for larger datasets and is susceptible to the presence of outliers and noisy data. SVM can perform both classification and regression. It determines the best line of separation (margin) for data on a hyperplane based on support vectors, which are the data points closest to the margin. SVM algorithm is robust against outliers, which are farther from the support vectors, and effective for both linear and nonlinear data. For the latter, SVM uses kernel functions to classify the data in a linear hyperplane. In regression problems, SVM estimates the regression by fitting data points to a function on the hyperplane.

The steps of the multiple classifier-based classification are detailed below.

Step 7: Input each selected feature vector in turn to each classifier to calculate prediction vectors (outputs of the

TABLE 3. Transition table of EfDenseNet model.

Phase	Method	Parameters	Output
Feature generation	Image resizing	Width and height. We have used bilinear interpolation	256 × 256 images
	Patch division	32 × 32 expanding patches	4 nested patches 1st patch: 64 × 64 2nd patch: 128 × 128 3rd patch: 192 × 192 4th patch: 256 × 256
	Feature generation	Deep feature extraction with EfficientNetb0 and DenseNet201 We have used both CNNs with their default settings. These CNNs have been used to generate deep features. We have used fully connected and GAP layers of these networks. These CNNs were trained on ImageNet1k. Herein, EfficientNetb0 uses Swish and DenseNet201 uses ReLu activation function	4 feature vectors
	Combinator	Merging function	15 feature vectors with combinatorial fusion
Feature selection	NCA, ReliefF and Chi2	The 250 most discriminative features from each feature vector. We have used these feature extractors with default settings. The parameters of the feature selectors: NCA: Solver: Stochastic gradient descent, the number of iterations: half of the number of observations. ReliefF: Number of neighbors: 10, distance: Euclidean. *Chi2:* We have only used Chi2 statistical moment to select features.	45 (=3×15) selected feature vectors.
Classification	SVM and kNN	Test of selected feature vectors with each classification algorithm. *kNN* k:1, Distance: City block, Voting: None. *SVM* Kernel: Cubic, box constraint: 1, coding: one-vs-all *Validation* 10-fold CV	90 (=2×45) prediction vectors containing prediction labels.
Information fusion	Iterative majority voting	Test result vectors with iterative majority voting. *IMV* Sorting parameter: accuracy, voting function: mode, range of iteration: from 3 to 90 Number of voted vectors: 88=90-3+1 *Selection* The best accurate output	178 (=90+88) voted for prediction vectors containing prediction labels

TABLE 3. *(Continued.)* Transition table of EfDenseNet model.

classifiers) using a 10-fold CV. These classifiers are widely recognized and commonly used, proving to be the most effective classifiers for our specific problem.

$$p_g = kNN\left(s_g, y\right), \quad g \in \{1, 2, \ldots, 45\}$$
$$p_{g+45} = SVM\left(s_g, y\right) \tag{8}$$

where $kNN()$ represent kNN function; and $SVM()$, SVM function. Parameter settings for kNN were: k value,1; distance, L1-norm. For the SVM, the kernel type was cubic. In total, 90 (=45 × 2) prediction vectors were calculated.

D. INFORMATION FUSION

The 90 prediction vectors were merged and input to a mode-based iterative majority voting (IMV) operation [26] to calculate an additional 88 (=90-2) voted prediction vectors (the iteration range was 3 to 90; 3 being the minimum for mode-based majority voting). The 90 prediction vectors and 88 voted prediction vectors were then merged, and the overall best model result was chosen.

Step 8: Apply IMV (iteration range 3 to 90) to the 90 prediction vectors to calculate voted results.

Step 9: Calculate the accuracies of the combined 178 (=90+88) predictions and voted prediction and select the overall best result with the maximum accuracy.

IV. EXPERIMENTAL RESULTS

A. EXPERIMENTAL SETUP

The EfDenseNet model was implemented in the MATLAB platform on a personal computer with the following specifications: Intel i9-9900 CPU, 32 GB RAM, and Windows 11 operating system. The transition status of the model architecture is summarized in Table 3.

B. RESULTS

The self-organized EfDenseNet model yielded low rates of 4-class misclassification on the PH dataset (Figure 5) and excellent classification performance, attaining 97.27% overall 4-class classification accuracy (Table 5).

FIGURE 5. Confusion matrix for the EfDenseNet-based model. Herein, the white cell defines zero.

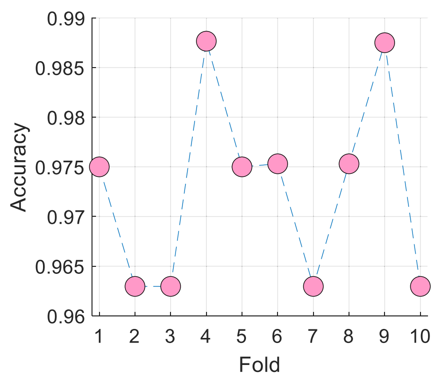


FIGURE 6. Fold-wise accuracies obtained for the proposed EfDenseNet.

TABLE 4. Class-wise and overall classification performance of EfDenseNet-based model on the PH dataset.

Class	Accuracy (%)	UAR (%)	UAP (%)	F1-score (%)
Control	-	99.05	98.11	98.58
Group 1	-	91.25	92.41	91.82
Group 2	-	94.62	96.85	95.72
Group 3	-	98.45	97.94	98.20
Overall	97.27	95.84	96.33	96.08

*UAR, unweighted average recall; UAP, unweighted average precision.

We employed a 10-fold cross-validation approach, and fold-specific classification accuracies of the proposed model are shown in Figure 6.

Figure 6 shows that the highest classification accuracy of 98.77% was achieved in the 4th fold and exceeded 96% in all folds.

C. EXPLAINABLE RESULTS

We applied gradient-weighted class activation mapping [27], [28] to generate heat maps showing areas of the input CT images that had the most impact on the classification results. The EfDenseNet focused mainly on the pulmonary artery bifurcation in the middle of the images to detect PH (see Figure 7), which aligns with the way radiologists visually examine the image to assess for morphological signs of PH. We developed EfDenseNet using MATLAB to obtain these results by merging EfficientNetb0 with DenseNet201. Our model has approximately 25.3 million total trainable parameters (~5.3 million from EfficientNetb0 and ~20 million from DenseNet201). The hyperparameters of this model are as follows:

Solver: Stochastic gradient descent with momentum
Mini-batch size: 32
Number of epochs: 30
Momentum: 0.9
Training-to-validation split ratio: 70:30

We have trained on this dataset, and the computed training and validation curves are displayed in Figure 7.

Our proposed model achieved a final validation accuracy of 94.12%. Using this model, we generated heatmaps, and the interpretable results obtained with EfDenseNet are presented in Figure 8.

Figure 8 indicates that the most significant features are located on the vessels. These features were extracted using our proposed deep feature engineering model. Considering the relatively small size of our dataset, owing to the uncommon occurrence of PH as a cardiac disorder, we opted for deep feature engineering over end-to-end deep learning to achieve superior classification performance. Additionally, training

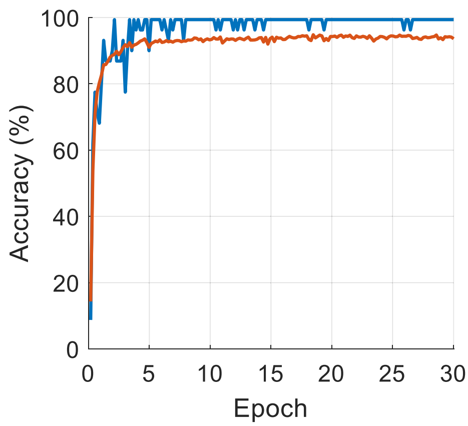


FIGURE 7. Training and validation curve of the EfDenseNet.

images were utilized to provide interpretable and explainable results. We have trained the proposed EfDenseNet model to show the effectiveness of our model (see Figure 7 and Figure 8).

D. TIME COMPLEXITY ANALYSIS

In this work, we employed a deep transfer learning technique, eliminating the need for training operations using CNNs. Pretrained EfficientNetb0 and DenseNet201 were utilized for feature extraction, resulting in a faster model compared to an end-to-end deep learning approach. The time complexity of the proposed model was computed and the calculations are presented below [29].

During the feature extraction phase, we employed two pretrained CNNs and a combinator to generate 15 feature vectors. As a result, the time complexity of this phase is given by $O(E + D + C)$, where E is the time complexity coefficient of EfficientNetb0, D denotes the time burden of pretrained DenseNet201, and C denotes the time complexity of the combinator, responsible for generating 11 additional feature vectors from the initial four.

To select the most relevant features, three feature selectors (NCA, ReliefF, and Chi2) were employed, resulting in a computational complexity of $O(mN + mR + mc)$, where m represents the number of feature vectors, and N, R, and c are the time complexity coefficients of NCA, ReliefF, and Chi2 feature selectors, respectively.

During the classification phase, two shallow classifiers, kNN and SVM, were employed. The computational complexity of this phase can be expressed as $O(sk+sS)$, where s is the number of selected feature vectors, k denotes the time burden of kNN, and S represents the time burden of SVM.

Furthermore, we introduced an information fusion phase, which resulted in a time complexity of $O(pR)$, where p and R represent the number of outputs and the range of the IMV, respectively.

Hence, the total computational complexity of the proposed model is given by $O(E + D + C + mN + mR + mc + sk + sS + pR)$. This indicates that the proposed model carries a linear time burden, as we employed deep transfer learning to develop a self-organized PH classification model in this work.

V. DISCUSSION

We have proposed a self-organized hybrid feature engineering model, EfDenseNet, that used pretrained EfficientNetb0 and DenseNet201 to extract multiple deep features from nested patches, which were then fed to downstream combinator, multiple feature selectors, multiple classifiers, and finally, iterative majority voting, to calculate the overall best result. We elected to use the combination of EfficientNetb0 and DenseNet201 based on preliminary testing, which demonstrated superior performance against other common deep networks (Figure 9).

We chose standard methods for the feature selection and classification phases without conducting preliminary tests. From the initial four EfDenseNet-extracted feature vectors, the combinator, three feature selectors and two classifiers increased the number of results to 15, 45, and 90, respectively. Of the resultant 90 classifier-based prediction vectors, the highest accuracy attained was 96.65% (Figure 10), which provided indirect support for the validity of our proposed EfDenseNet approach.

In the last phase of the model, we applied IMV to generate voted prediction vectors, which were combined with the prior classifier-based prediction vectors to calculate the optimal result. With this information fusion process, the overall best model accuracy was boosted to 97.27% (Figure 11).

A. COMPARISON OF PERFORMANCE OF VARIOUS MODEL ELEMENTS

We compared the relative performance of the various methods used for EfDenseNet feature extraction (Figure 10). Overall, DenseNet 201 slightly outperformed EfficientNetB0 (Figures 12a and 12b); and the fully connected layer slightly outperformed the global average pooling layer (Figure 12c).

Among the different feature selection functions, NCA produced the highest accuracy results (Figure 13).

Between the two classifiers, kNN yielded a better accuracy rate (Figure 14).

The analyses described above were performed using classifier-based results. With IMV, the performance was boosted. The overall best classification accuracy was attained at the second iteration, which was generated by voting on the four top single results, i.e., the 43rd, 10th, 40th, and 28th classifier-based prediction vectors, which had received input from the respective combinations of EfDenseNet-extracted feature vectors $f_1+f_2+f_3+f_4$, f_4, $f_1+f_3+f_4$, and f_3+f_4, respectively (see Table 3 for details of extracted feature vectors f_1 to f_4).

B. ABLATION STUDIES

To examine the relative contributions of various model components to the performance, we performed ablation studies according to the cases listed below.

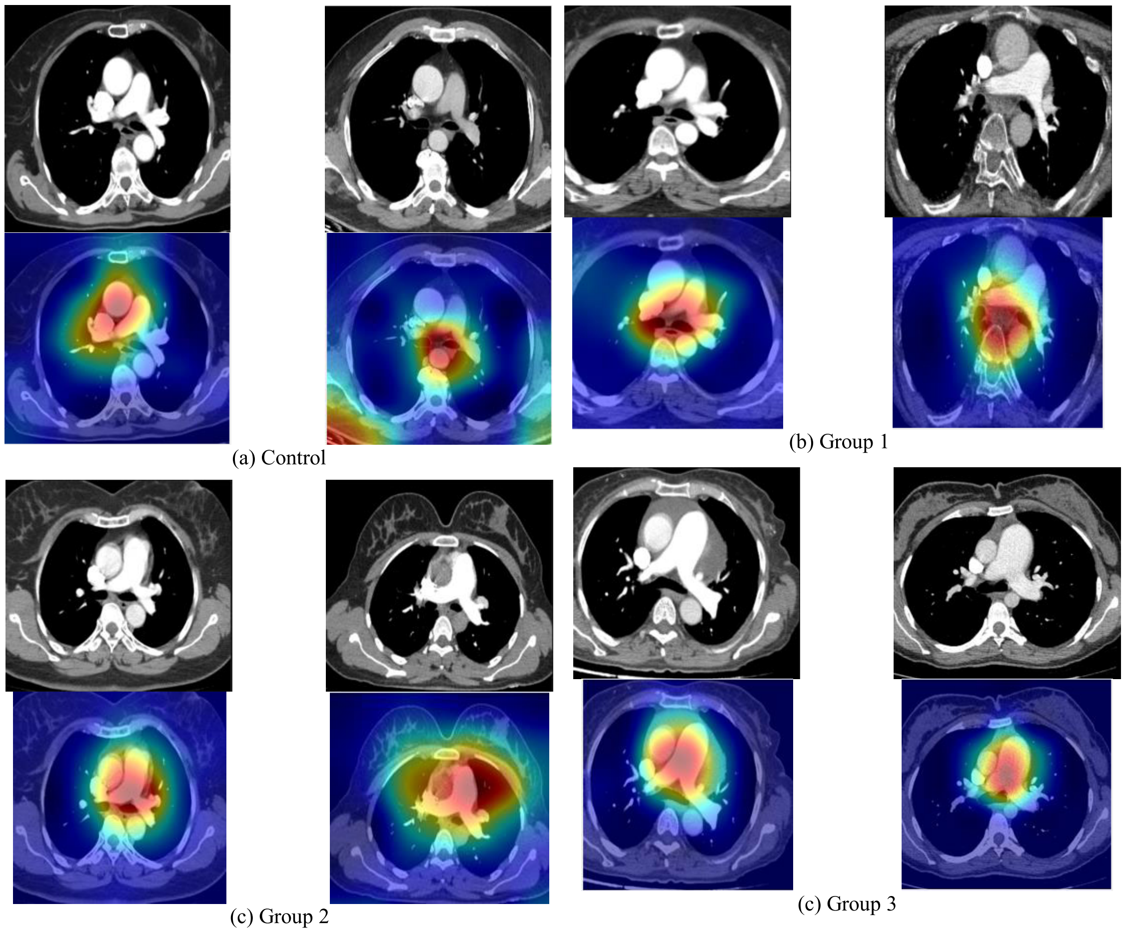


FIGURE 8. Examples of thoracic CT images with overlayed heat maps showing regions on the input images that contributed to the classification results (red color denotes the highest impact) in the various groups.

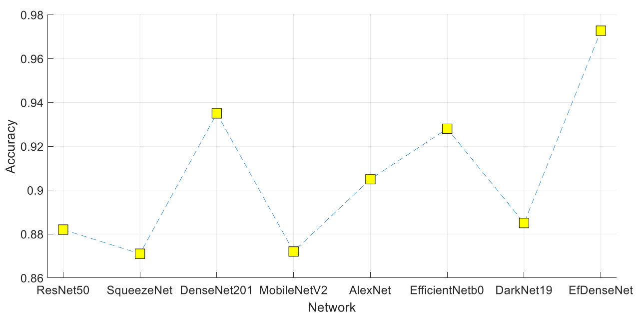


FIGURE 9. Comparison of classification results obtained using various pretrained networks. For fair comparison, all extracted feature vectors were input to NCA and kNN algorithms for feature selection and classification, respectively.

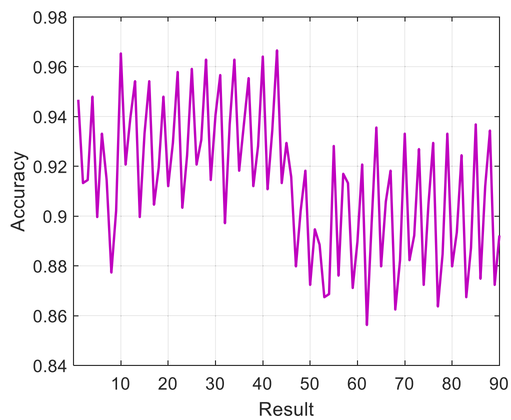


FIGURE 10. Accuracy results of all 90 kNN- or SVM-classifiers prediction vectors.

Case 1: EfficientNetb0 feature extractor + NCA selector + kNN classifier.

Case 2: DenseNet201 feature extractor + NCA selector + kNN classifier.

Case 3: Full model minus nested patch division.

Case 4: Full model.

The overall accuracies of Cases 1 to 4 were 92.32%, 93.80%, 96.65%, and 97.27%, respectively. These results suggested that DenseNet201 slightly outperformed EfficientNetb0 (Case 2 versus Case 1); and that inclusion of nest patch division conferred marginal improvement in accuracy

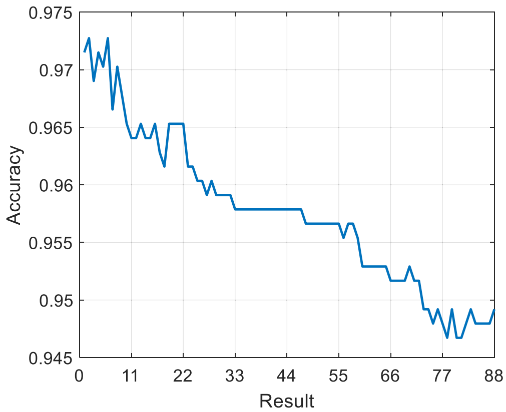


FIGURE 11. The calculated 88 voted results by deploying IMV algorithm.

(Case 4 versus Case 3). Moreover, the computed accuracies of these cases have been depicted in Figure 15.

As shown in Figure 15, we observed that DenseNet201 achieved better classification performance compared to EfficientNetb0. Furthermore, we conducted Case 3 to evaluate the classification accuracy of our architecture without using the nested patch division model. Interestingly, this case reported a classification accuracy of 96.65%, demonstrating the effectiveness of our proposed architecture in achieving high classification accuracy.

C. COMPARISON WITH THE LITERATURE

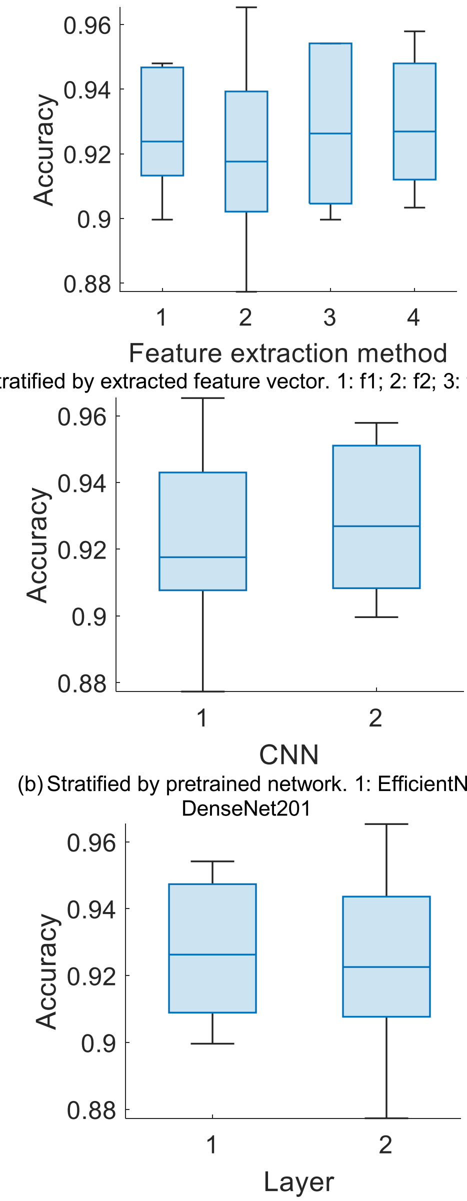
We performed a nonsystematic literature review of automated PH prediction models, which yielded studies based on various biomedical signals, serum and imaging biomarkers (Table 5). Notably, our study is the only one that used CT images and yielded the best accuracy for PH classification.

There can be noted from the existing literature that, our proposed model stands out as the first deep transfer learning approach applied to a PH dataset comprising four classes. Our model achieved an accuracy of 97.27% using our collected dataset. In comparison, other researchers focused on datasets with only two or three classes. However, we presented a more challenging problem than theirs, making our model's performance even more remarkable, surpassing the classification performance of other existing approaches.

Moreover, the p-values were computed for 45 generated feature vectors. The p-values for each feature vector were conducted using couples, as detailed in the literature [35], [36]. It is important to highlight that a p-value smaller than 0.05 signifies a statistically significant feature.

As a result, we determined the ratio of meaningful features by dividing the number of features with p-values less than 0.05 by the total number of features. This calculation was carried out using the computed p-values. The findings pertaining to the ratio of meaningful features, as established from the p-values, are presented in Figure 16.

According to the observations from Figure 14, the most distinguishable classes are Group 3 and Control. The classes that show similarity are the pairs Group 1 - Group 2, Group 2 - Group 3, and Group 1 - Control.



(a) Stratified by extracted feature vector. 1: f1; 2: f2; 3: f3; 4: f4.

(b) Stratified by pretrained network. 1: EfficientNetB0; 2: DenseNet201

(c) Stratified by network layer. 1: fully connected layer; 2: global average pooling layer

FIGURE 12. Classification accuracies of various elements of EfDenseNet deep feature extraction. See Table 3 for details of extracted feature vectors f1 to f4.

D. HIGHLIGHTS

Our research findings, highlights, and limitations are listed below.

Findings:

- The effectiveness of nested patch division was demonstrated in ablation studies.
- Among the elements used in the EfDenseNet feature extraction, DenseNet201 slightly outperformed Effi-

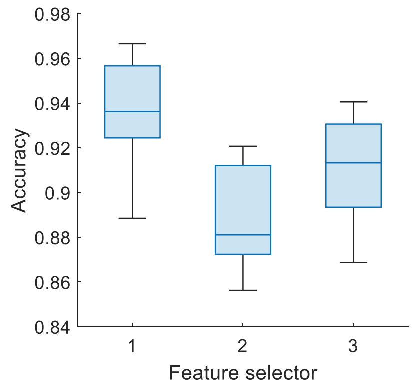


FIGURE 13. Classification accuracies stratified by feature selector. 1: NCA; 2: Chi2; 3: ReliefF.

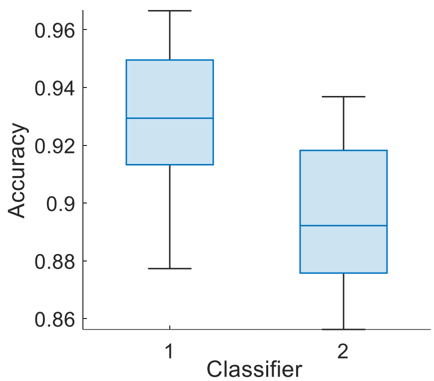


FIGURE 14. Classification accuracies stratified by feature selector. 1: kNN, 2: SVM.

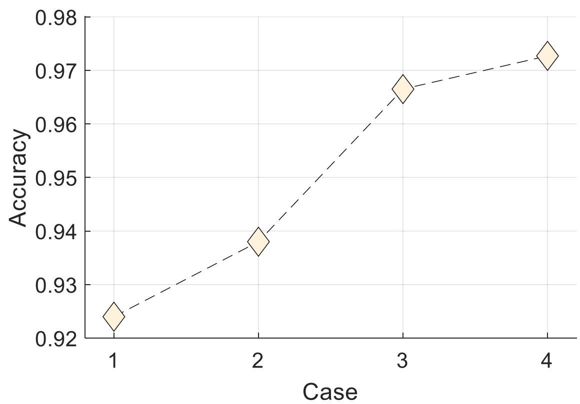


FIGURE 15. Ablation results obtained for various cases.

cientNetb0; and the fully connected layer slightly outperformed global average pooling later.

- Among feature selectors, NCA performed the best.
- Among classifiers kNN outperformed SVM.
- IMV-based voted results outperformed classifier-based results.

Highlights:

- A new PH dataset comprising 807 images in four classes (1 control, 3 PH of varying severity) was collected and made publicly accessible.

TABLE 5. Summary of published machine learning models for PH classification.

Work	Method	Subjects/ Classes	Data type	Split ratio	Results (%)
Swift et al. [30]	Multilinear principal component analysis	150 PH; 70 no PH	CMR	10-fold CV	Acc: 92.00 Sen: 96.00 Spe: 87.00
Ge et al. [31]	CNN, short-term Fourier transform, majority voting	161 normal; 161 CHD; 161 CHD-PH	Heart sounds	55:20 :25	Acc: 88.61
Bauer et al. [32]	Protein-based feature extraction, random forest	77 PH; 80 no PH	Serum protein biomarkers	10-fold CV	Acc: 81.10 Sen: 77.30 Spe: 86.50
Aras et al. [13]	CNN	5016 PH; 19454 no PH	ECG	70:10 :20	AUC: 89.00 Sen: 79.00 Spe: 84.00
Alabed et al. [33]	Image restoration, multilinear principal component analysis, SVM	PH (mortality prediction)	CMR	70:30	Acc: 83.00
Diller et al. [34]	Two CNNs combinations	450 PH; 308 dilated right ventricle; 67 healthy	Ultrasound with augmentation	67:33	Acc: 95.00
Our	EfDenseNet	114 control; 43 Group 1; 65 Group 2; 91 Group 3	Thoracic CT	10-fold CV	Acc: 97.27 UAR: 95.84 UAP: 96.33 F1: 96.08

**Acc: accuracy; Sen: sensitivity; Spe: specificity; AUC: area under curve; CHD: congenital heart disease; ECG: electrocardiogram; CMR: cardiac magnetic resonance.

- IMV in the final phase rendered our model self-organizing structure: from feature vectors generated by the EfDenseNet architecture, the best result was automatically calculated by mode-based IMV.
- We combined computationally efficient classical machine learning methods with pretrained CNNs, and attained high classification performance at low time complexity,
- The dynamic-sized nested patch structure enabled detailed local and global feature extraction.
- EfDenseNet model explainability was demonstrated.

Limitations:

- The study used a relatively smaller dataset, which may limit the generalizability of our results. We collected

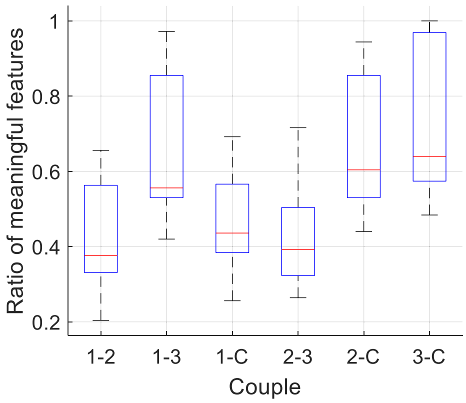


FIGURE 16. P-value analysis of the selected features. 1: Group 1; 2: Group 2; 3: Group 3; C: Control.

this dataset from a single medical center. Therefore, the collected dataset is not a large CT image dataset.

- We incorporated standard effective feature selectors and classifiers into our model without pre-testing. As our model is parametric, it is feasible to refine our model further by optimizing or substituting these elements.
- Our model achieved the highest classification performance among the compared models. However, it has approximately 25.3 million parameters, given that we proposed a hybrid CNN. Hence, our number of learnable parameters is larger than SqueezeNet (~1.2M), DenseNet201 (~20M), MobileNetV2 (~3.5M), EfficientNetb0 (~5.3M), and DarkNet19 (~20.8M). To address this issue, we employed a deep transfer learning technique.

VI. CONCLUSION

In this research, we introduced a novel deep feature engineering architecture, named EfDenseNet, which autonomously selects the best output based on classification accuracy. The model employs EfficientNetb0 and DenseNet201 to extract deep features. We compiled a new CT image dataset for PH detection, upon which the EfDenseNet was applied, achieving a classification accuracy of 97.27% across four classes. Furthermore, we analyzed the time complexity of our approach, determining that EfDenseNet operates with a linear time overhead. These findings demonstrate that our proposed model is a robust tool for PH classification.

VII. FUTURE WORKS

We aim to expand the PH dataset and refine the EfDenseNet architecture by integrating additional feature selectors and classifiers. The generalizability of our findings is constrained by the limited size of our dataset. Given that the CT procedure is noninvasive and widely accessible, we foresee the collection of larger datasets for independent validation of our model. In the future, we aspire to develop a universally applicable PH detection model utilizing more extensive and diverse CT image datasets. With the use of pre-trained comprehensive datasets, tools for PH detection can be devised. Furthermore, as highlighted, refining feature selection and classification strategies could enhance the model's performance. Ultimately, we plan to incorporate cutting-edge methodologies with explainable artificial intelligence and uncertainty quantification to use the model in a noisy environment [37].

DECLARATIONS

FUNDING

This research received no external funding.

INSTITUTIONAL REVIEW BOARD STATEMENT

The study was approved by the local ethical committee, Ethics Committee of Firat University (2022/04-06).

INFORMED CONSENT STATEMENT

Informed consent was obtained from all subjects involved in the study.

DATA AVAILABILITY STATEMENT

The CT dataset is available for download at https://www.kaggle.com/datasets/turkertuncer/ph-ct-v1.

ACKNOWLEDGMENT

We gratefully acknowledge the Ethics Committee, Firat University data transcription.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] M. M. Hoeper, V. V. McLaughlin, A. M. Al Dalaan, T. Satoh, and N. Galiè, "Treatment of pulmonary hypertension," *Lancet Respiratory Med.*, vol. 4, no. 4, pp. 323–336, 2016.

[2] V. V. McLaughlin, "Treatment goals of pulmonary hypertension," *J. Amer. College Cardiol.*, vol. 62, no. 25S, pp. D73–D81, 2013.

[3] M. M. Hoeper and M. Humbert, *The New Haemodynamic Definition of Pulmonary Hypertension: Evidence Prevails, Finally*, vol. 53. Lausanne, Switzerland: European Respiratory Society, 2019.

[4] S. A. Mandras, H. S. Mehta, and A. Vaidya, *Pulmonary Hypertension: A Brief Guide for Clinicians*, vol. 95, 9th ed. Amsterdam, The Netherlands: Elsevier, pp. 1978–1988.

[5] M.-R. Ghigna and P. Dorfmüller, "Pulmonary vascular disease and pulmonary hypertension," *Diagnostic Histopathol.*, vol. 25, no. 8, pp. 304–312, 2019.

[6] F. A. Klok, F. Couturaud, M. Delcroix, and M. Humbert, "Diagnosis of chronic thromboembolic pulmonary hypertension after acute pulmonary embolism," *Eur. Respiratory J.*, vol. 55, no. 6, Jun. 2020, Art. no. 2000189.

[7] A. Frost, "Diagnosis of pulmonary hypertension," *Eur. Respiratory J.*, vol. 53, no. 1, 2019, Art. no. 1801904.

[8] N. P. Nickel, K. Yuan, P. Dorfmuller, S. Provencher, Y.-C. Lai, S. Bonnet, E. D. Austin, C. D. Koch, A. Morris, F. Perros, D. Montani, R. T. Zamanian, and V. A. de Jesus Perez, "Beyond the lungs: Systemic manifestations of pulmonary arterial hypertension," *Amer. J. Respiratory Crit. Care Med.*, vol. 201, no. 2, pp. 148–157, Jan. 2020.

[9] H. Kankaanranta, T. Harju, M. Kilpeläinen, W. Mazur, J. T. Lehto, M. Katajisto, T. Peisa, T. Meinander, and L. Lehtimäki, ''Diagnosis and pharmacotherapy of stable chronic obstructive pulmonary disease: The Finnish guidelines,'' *Basic Clin. Pharmacol. Toxicol.*, vol. 116, no. 4, pp. 291–307, Apr. 2015.

[10] A. Miri, I. Kalla, and F. Seedat, ''Occult pulmonary arterial hypertension in patients with previous pulmonary tuberculosis,'' *Afr. J. Thoracic Crit. Care Med.*, vol. 26, no. 4, p. 133, Dec. 2020.

[11] L.-Q. Zhou, ''Artificial intelligence in medical imaging of the liver,'' *World J. Gastroenterol.*, vol. 25, no. 6, p. 672, 2019.

[12] S. Chakraborty and K. Mali, ''An overview of biomedical image analysis from the deep learning perspective,'' in *Research Anthology on Improving Medical Imaging Techniques for Analysis and Intervention*. PA, USA, 2023, pp. 43–59, doi: 10.4018/978-1-6684-7544-7.ch003.

[13] M. A. Aras, S. Abreau, H. Mills, L. Radhakrishnan, L. Klein, N. Mantri, B. Rubin, J. Barrios, C. Chehoud, E. Kogan, X. Gitton, A. Nnewihe, D. Quinn, C. Bridges, A. J. Butte, J. E. Olgin, and G. H. Tison, ''Electrocardiogram detection of pulmonary hypertension using deep learning,'' *J. Cardiac Failure*, vol. 29, no. 7, pp. 1017–1028, Jul. 2023.

[14] A. Gudigar, U. Raghavendra, T. Devasia, K. Nayak, S. M. Danish, G. Kamath, J. Samanth, U. M. Pai, V. Nayak, R. S. Tan, E. J. Ciaccio, and U. R. Acharya, ''Global weighted LBP based entropy features for the assessment of pulmonary hypertension,'' *Pattern Recognit. Lett.*, vol. 125, pp. 35–41, Jul. 2019.

[15] M. Ong, J. G. Klann, K. J. Lin, B. A. Maron, S. N. Murphy, M. D. Natter, and K. D. Mandl, ''Claims-based algorithms for identifying patients with pulmonary hypertension: A comparison of decision rules and machine-learning approaches,'' *J. Amer. Heart Assoc.*, vol. 9, no. 19, Oct. 2020, Art. no. e016648.

[16] L. Cai, J. Gao, and D. Zhao, ''A review of the application of deep learning in medical image classification and segmentation,'' *Ann. Transl. Med.*, vol. 8, no. 11, p. 713, Jun. 2020.

[17] W. Wang, ''Medical image classification using deep learning,'' in *Deep Learning in Healthcare*. Cham, Switzerland: Springer, 2020, pp. 33–51, doi: 10.1007/978-3-030-32606-7_3.

[18] Z. Higgs, D. Macafee, B. Braithwaite, and C. Maxwell-Armstrong, ''The seldinger technique: 50 years on,'' *Lancet*, vol. 366, no. 9494, pp. 1407–1409, Oct. 2005.

[19] M. Tan and Q. Le, ''EfficientNet: Rethinking model scaling for convolutional neural networks,'' in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 6105–6114.

[20] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, ''Densely connected convolutional networks,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.

[21] G. Jacob, R. Sam, H. Geoff, and S. Ruslan, ''Neighbourhood components analysis,'' in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 17, 2004, pp. 513–520.

[22] M. Robnik-Šikonja and I. Kononenko, ''Theoretical and empirical analysis of ReliefF and RReliefF,'' *Mach. Learn.*, vol. 53, nos. 1–2, pp. 23–69, Oct. 2003.

[23] H. Liu and R. Setiono, ''Chi2: Feature selection and discretization of numeric attributes,'' in *Proc. 7th IEEE Int. Conf. Tools With Artif. Intell.*, Nov. 1995, pp. 388–391.

[24] W. S. Noble, ''What is a support vector machine,'' *Nature Biotechnol.*, vol. 24, no. 12, pp. 1565–1567, Dec. 2006.

[25] L. E. Peterson, ''K-nearest neighbor,'' *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.

[26] A. Dogan, M. Akay, P. D. Barua, M. Baygin, S. Dogan, T. Tuncer, A. H. Dogru, and U. R. Acharya, ''PrimePatNet87: Prime pattern and tunable q-factor wavelet transform techniques for automated accurate EEG emotion recognition,'' *Comput. Biol. Med.*, vol. 138, Nov. 2021, Art. no. 104867.

[27] S. Yin, L. Wang, M. Shafiq, L. Teng, A. A. Laghari, and M. F. Khan, ''G2Grad-CAMRL: An object detection and interpretation model based on gradient-weighted class activation mapping and reinforcement learning in remote sensing images,'' *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 3583–3598, 2023.

[28] H. W. Loh, C. P. Ooi, S. Seoni, P. D. Barua, F. Molinari, and U. R. Acharya, ''Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022),'' *Comput. Methods Programs Biomed.*, vol. 226, Nov. 2022, Art. no. 107161.

[29] I. Chivers, J. Sleightholme, I. Chivers, and J. Sleightholme, ''An introduction to algorithms and the big O notation,'' in *Introduction to Programming With Fortran*, vol. 77. Cham, Switzerland: Springer, 2015, pp. 359–364, doi: 10.1007/978-3-319-17701-4_23.

[30] A. J. Swift, H. Lu, J. Uthoff, P. Garg, M. Cogliano, J. Taylor, P. Metherall, S. Zhou, C. S. Johns, S. Alabed, R. A. Condliffe, A. Lawrie, J. M. Wild, and D. G. Kiely, ''A machine learning cardiac magnetic resonance approach to extract disease features and automate pulmonary arterial hypertension diagnosis,'' *Eur. Heart J. Cardiovascular Imag.*, vol. 22, no. 2, pp. 236–245, Jan. 2021.

[31] B. Ge, H. Yang, P. Ma, T. Guo, J. Pan, and W. Wang, ''Detection of pulmonary arterial hypertension associated with congenital heart disease based on time–frequency domain and deep learning features,'' *Biomed. Signal Process. Control*, vol. 81, Mar. 2023, Art. no. 104451.

[32] Y. Bauer, S. de Bernard, P. Hickey, K. Ballard, J. Cruz, P. Cornelisse, H. Chadha-Boreham, O. Distler, D. Rosenberg, M. Doelberg, S. Roux, O. Nayler, and A. Lawrie, ''Identifying early pulmonary arterial hypertension biomarkers in systemic sclerosis: Machine learning on proteomics from the DETECT cohort,'' *Eur. Respiratory J.*, vol. 57, no. 6, Jun. 2021, Art. no. 2002591.

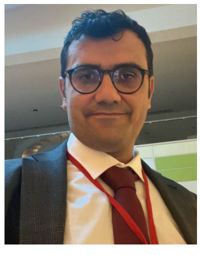
[33] S. Alabed, J. Uthoff, S. Zhou, P. Garg, K. Dwivedi, F. Alandejani, R. Gosling, L. Schobs, M. Brook, Y. Shahin, D. Capener, C. S. Johns, J. M. Wild, A. M. K. Rothman, R. J. van der Geest, R. Condliffe, D. G. Kiely, H. Lu, and A. J. Swift, ''Machine learning cardiac-MRI features predict mortality in newly diagnosed pulmonary arterial hypertension,'' *Eur. Heart J. Digit. Health*, vol. 3, no. 2, pp. 265–275, Jul. 2022.

[34] G.-P. Diller, M. L. Benesch Vidal, A. Kempny, K. Kubota, W. Li, K. Dimopoulos, A. Arvanitaki, A. E. Lammers, S. J. Wort, H. Baumgartner, S. Orwat, and M. A. Gatzoulis, ''A framework of deep learning networks provides expert-level accuracy for the detection and prognostication of pulmonary arterial hypertension,'' *Eur. Heart J. Cardiovascular Imag.*, vol. 23, no. 11, pp. 1447–1456, Oct. 2022.

[35] S. Athey, D. Eckles, and G. W. Imbens, ''Exact p-values for network interference,'' *J. Amer. Stat. Assoc.*, vol. 113, no. 521, pp. 230–240, 2018.

[36] S. Hosseini, B. Turhan, and M. Mäntylä, ''A benchmark study on the effectiveness of search-based data selection and feature selection for cross project defect prediction,'' *Inf. Softw. Technol.*, vol. 95, pp. 296–312, Mar. 2018.

[37] S. Seoni, V. Jahmunah, M. Salvi, P. D. Barua, F. Molinari, and U. R. Acharya, ''Application of uncertainty quantification to artificial intelligence in healthcare: A review of last decade (2013–2023),'' *Comput. Biol. Med.*, vol. 165, Oct. 2023, Art. no. 107441.



TARIK KIVRAK received the master's degree in medical faculty from Ondokuz Mayıs University, Samsun, Turkey, in 2006, and the dual Doctor (Medical Specialist) degree in cardiology from Marmara University, İstanbul, Turkey, in 2007 and 2013, respectively. He is currently an Associate Professor with the Cardiology, Medical Faculty, Firat University. His research interests include feature heart failure, pulmonary hypertension, cardiac MRI, and artificial intelligence. He has been working actively on developing algorithms in machine learning applied to visual and voice surveillance for cardiology data.

JAGADISH NAYAK (Senior Member, IEEE) received the bachelor's degree in electronics and communication engineering from Karnataka University, the Master of Technology degree in digital electronics and communication from the Manipal Academy of Higher Education, Manipal, and the Ph.D. degree from the National Institute of Technology Karnataka (NITK) Surathkal. His Ph.D. dissertation was titled, "Automated Detection of Eye Abnormalities and Patient Data Handling." Earlier to the present position, he was an Associate Professor with the Manipal Academy of Higher Education and a Faculty Member with NITK Surathkal. He is also with a renowned company Bradma of India Ltd., Bengaluru, as a Customer Support Engineer. He is currently an Associate Professor with the Birla Institute of Technology and Science Pilani, Dubai Campus. He has guided one Ph.D. student. He has a total of 27 years of experience both in industry, teaching, and research. He also handled some projects in the field of medical image processing. Recently, he handled a couple of projects in the Internet of Things. He has published around 18 research articles in reputed international journals, three book chapters, 20 papers in international conferences, and ten papers in national conferences. His research interests include signal processing and its application, machine learning methods for medical signals and images, microelectronics, VLSI design, embedded systems design, and the Internet of Things. He is a member of several professional bodies. He is also a Reviewer for many international journals, specifically the *Journal of Medical Systems*.

MEHMET ALI GELEN received the Graduate degree from the Faculty of Medicine, Fırat University, in 2018. He started his cardiology specialist training with Fırat University, in 2018, where he is currently a Research Assistant with the Faculty of Medicine, Department of Cardiology. His research interests include heart failure, pulmonary hypertension, cardiac MRI, and artificial intelligence.

PRABAL DATTA BARUA received the Ph.D. degree in information system from the University of Southern Queensland. He is currently an Adjunct Professor with the University of Southern Queensland and an Honorary Industry Fellow with the University of Technology Sydney. He is also an Academic and Accredited Research Supervisor with the University of Southern Queensland with 12 years of teaching experience. He received research support from the Queensland Government Innovation Connections under the Entrepreneurs Program to research "Cancer recurrence using innovative machine learning approaches." He is an Industry Leader in ICT entrepreneurship in Australia and sitting as an ICT Advisory Panel Member of many organizations. He has published several articles in the Q1 journals. His research interests include AI technology development in health, education, agriculture, and environmental science.

MEHMET BAYGIN received the B.S., M.S., and Ph.D. degrees in computer engineering from Fırat University, Elâzığ, Turkey, in 2010, 2013, and 2018, respectively. He is currently an Assistant Professor Doctor with the Department of Computer Engineering, College of Engineering, Erzurum Technical University, Turkey. His research interests include machine learning, computer vision, signal processing, blockchain, and photovoltaic systems.

HILAL ERKEN PAMUKCU received the Graduate degree from the Faculty of Medicine, Ankara University, in 2008. She became a Cardiology Specialist with the Faculty of Medicine, Ankara University, in 2013. She is currently with the Department of Cardiology, Etlik Training and Research Hospital. Her research interests include heart failure, pulmonary hypertension, cardiac MRI, and artificial intelligence.

SENGUL DOGAN received the master's degree in bioengineering and the Ph.D. degree in electrical and electronics engineering from the Fırat University, Elâzığ, Turkey, in 2007 and 2011, respectively. She is currently an Associate Professor with the Digital Forensics Engineering, Technology Faculty, Fırat University. Her research interests include computer forensics, mobile forensics, image processing, and signal processing. She has been working actively on developing algorithms in machine learning for biomedical data.

TURKER TUNCER received the master's degree in electronics and computer sciences and the Ph.D. degree in software engineering from Fırat University, Elâzığ, Turkey, in 2011 and 2016, respectively. He is currently an Associate Professor with the Digital Forensics Engineering, Technology Faculty, Fırat University. His research interests include feature engineering, image processing, signal processing, information security, and pattern recognition. He has been working actively on developing algorithms in machine learning applied to visual surveillance and biomedical data.

U. RAJENDRA ACHARYA (Senior Member, IEEE) received the Ph.D., D.Eng., and D.Sc. degrees. He is currently a Professor with the School of Mathematics, Physics and Computing, University of Southern Queensland, Australia, a Distinguished Professor with the International Research Organization for Advanced Science and Technology, Kumamoto University, Japan, an Adjunct Professor with the University of Malaya, Malaysia, and an Adjunct Professor with Asia University, Taiwan. His research interests include biomedical imaging and signal processing, data mining, and visualization, and the applications of biophysics for better healthcare design and delivery. His funded research has accrued cumulative grants exceeding six million Singapore dollars. He has authored over 800 publications, including 750 in refereed international journals, 42 in international conference proceedings, and 17 books. He has received over 75,000 citations on Google Scholar, with an H-index of 137. According to the Essential Science Indicators by Thomson, he consistently ranked among the top 1% of Highly Cited Researchers in computer science for the last eight years (2016–2023). He also sits on the editorial boards of multiple journals and he has served as a guest editor on several AI-related issues. For more information visit the link: https://scholar.google.com.sg/citations?user=8FjY99sAAAAJ&hl=en.