

## Article

# A Framework for Lung and Colon Cancer Diagnosis via Lightweight Deep Learning Models and Transformation Methods

Omneya Attallah <sup>1,\*</sup> , Muhammet Fatih Aslan <sup>2</sup>  and Kadir Sabanci <sup>2</sup> 

<sup>1</sup> Department of Electronics and Communications Engineering, College of Engineering and Technology, Arab Academy for Science, Technology and Maritime Transport, Alexandria 1029, Egypt

<sup>2</sup> Department of Electrical and Electronics Engineering, Karamanoglu Mehmetbey University, 70100 Karaman, Turkey

\* Correspondence: o.attallah@aast.edu

**Abstract:** Among the leading causes of mortality and morbidity in people are lung and colon cancers. They may develop concurrently in organs and negatively impact human life. If cancer is not diagnosed in its early stages, there is a great likelihood that it will spread to the two organs. The histopathological detection of such malignancies is one of the most crucial components of effective treatment. Although the process is lengthy and complex, deep learning (DL) techniques have made it feasible to complete it more quickly and accurately, enabling researchers to study a lot more patients in a short time period and for a lot less cost. Earlier studies relied on DL models that require great computational ability and resources. Most of them depended on individual DL models to extract features of high dimension or to perform diagnoses. However, in this study, a framework based on multiple lightweight DL models is proposed for the early detection of lung and colon cancers. The framework utilizes several transformation methods that perform feature reduction and provide a better representation of the data. In this context, histopathology scans are fed into the ShuffleNet, MobileNet, and SqueezeNet models. The number of deep features acquired from these models is subsequently reduced using principal component analysis (PCA) and fast Walsh–Hadamard transform (FHWHT) techniques. Following that, discrete wavelet transform (DWT) is used to fuse the FHWHT's reduced features obtained from the three DL models. Additionally, the three DL models' PCA features are concatenated. Finally, the diminished features as a result of PCA and FHWHT-DWT reduction and fusion processes are fed to four distinct machine learning algorithms, reaching the highest accuracy of 99.6%. The results obtained using the proposed framework based on lightweight DL models show that it can distinguish lung and colon cancer variants with a lower number of features and less computational complexity compared to existing methods. They also prove that utilizing transformation methods to reduce features can offer a superior interpretation of the data, thus improving the diagnosis procedure.

**Keywords:** CNN; deep learning; DWT; FHWHT; lung and colon cancer diagnosis; PCA



**Citation:** Attallah, O.; Aslan, M.F.; Sabanci, K. A Framework for Lung and Colon Cancer Diagnosis via Lightweight Deep Learning Models and Transformation Methods.

*Diagnostics* **2022**, *12*, 2926.

<https://doi.org/10.3390/diagnostics12122926>

Academic Editor: Dechang Chen

Received: 29 October 2022

Accepted: 22 November 2022

Published: 23 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cancer occurs as a result of the uncontrolled proliferation of abnormal cells in the body's organs or tissues. Cancer cells can occur in different organs or tissues of the body. According to the estimates made by the World Health Organization (WHO) for 2019, the primary or secondary cause of death before age 70 is cancer for 112 countries [1,2]. Furthermore, based on a report published by the International Agency for Research on Cancer (IARC) [3] in 2020, cancer is the primary or secondary cause of death in 134 countries. Cancer has more than 200 types [4]. Lung and colon cancers are anticipated to rank in the top three of the most prevalent cancer types in 2020, according to a statistical analysis done in America. The research also indicated that among all cancer diagnoses in America

in 2020, patients with lung and colon malignancies will experience the highest rates of death [5]. Lung and colon cancer rates are 11.4% and 18.0%, respectively, according to GLOBOCAN 2020 statistics [1]. Moreover, the WHO stated that around 4 million people globally developed colon or lung cancer in 2020. About 2.7 million people died from these cancers [6]. These statistical results prove that lung and colon cancer are common and deadly diseases worldwide.

Lung cancer might occur at the same time as colon cancer, as based on [7], about 17% of cases of these two cancers occur simultaneously. In addition, there is a considerable risk of cancer cells spreading between the two organs in the absence of an early diagnosis. Smoking is known to have a detrimental impact on the development of lung cancer, and it is believed that an unaware diet contributes to the development of colon cancer. In other words, the negative effect of lung cancer on the intestines can trigger colon cancer. Therefore, lung cancer can be seen as the second cancer disease in patients with colon cancer [7]. That is, a patient can get both lung and colon cancer at the same time. For this reason, it is vital to investigate both types of cancer in patients together and to diagnose them early [8].

Symptoms that can provide an early diagnosis of cancer are not direct markers of cancer. The most common symptoms such as fatigue, cough, muscle pain, etc., occur along with different types of diseases. The biggest tool that shows the presence of cancer is medical imaging devices [9]. Radiographic imaging methods such as mammography, histopathological imaging, computed tomography (CT), positron emission tomography (PET), magnetic resonance imaging (MRI), and ultrasound are frequently used for cancer detection [9,10]. Among them, histopathology images containing phenotypic information are indispensable for the diagnosis and evaluation of cancer diseases in clinics [10,11]. Manual analysis of such medical images by experts is a delicate and difficult task. It is therefore time consuming and requires a strong focus [10,12]. Moreover, the detection of cases is much more difficult in the case of early diagnosis because at the beginning of the disease, the symptoms are very vague and difficult to diagnose. When symptoms become evident, it is too late for early treatment [9]. Thanks to the advancements achieved in the field of artificial intelligence (AI), today, AI-based medical image analysis methods have assumed the role of a decision support mechanism due to both early diagnosis and support to doctors [13,14].

The automatic diagnosis methods are based on AI technology such as machine learning and deep learning. In this way, data analysis tasks based on expert knowledge have subsequently evolved into expert-independent and fully automatic diagnostic systems [15]. Several conventional machine-learning methods have been utilized to solve medical problems [16–22] and health-related applications [23,24]. However, these methods primarily require feature selection and feature extraction steps, and they suffer from the disadvantages of not using the appropriate feature extraction method and loss of information during feature extraction. On the other hand, deep learning (DL) has become popular in medical diagnostic applications due to both eliminating these disadvantages and strong discrimination ability [25]. Medical data are usually radiographic image data, and thus the convolutional neural network (CNN) is the well-known DL architecture commonly used to analyze medical images [26]. Recently, both pre-designed and pre-trained CNN models are frequently preferred due to their convenience and high performance. Examples of these models are AlexNet [27], SqueezeNet [28], ResNet [29], ShuffleNet [30], VGGNet [31], MobileNet [32], and GoogleNet [33]. These networks have been extensively used in the medical domain to analyze medical images and diagnose several types of diseases such as heart abnormalities [34], lung anomalies [35–39], brain disorders [40–42], breast cancer [43–45], genetic facial diseases [46], and eye diseases [47].

CNN models extract high-level features from raw data through their deep layers. In this way, complex and difficult data can be interpreted successfully thanks to CNNs. This success of CNN models is actually due to the complexity and depth of its architecture. As the complexity of the model increases, the number of parameters in the model also

increases [48]. Modern CNN models require millions of parameter updates during the training phase. However, the large number of parameters may negatively affect the generalization ability of the network and cause overfitting [49]. Reducing the number of features and utilizing lightweight DL models are solutions to prevent overfitting that might occur due to the complexity of the model [50]. In general, the number of features and parameters of pre-trained CNN models is high, which may cause overfitting and adversely affect classification and/or regression performance. Various feature reduction methods can be used to avoid overfitting; thus, unnecessary and/or redundant features are removed [51]. The most well-known size reduction algorithm is principal component analysis (PCA). Moreover, transformation methods such as fast Walsh–Hadamard transform (FHWT) and discrete wavelet transform (DWT) that analyze data and extract useful representations of the data can reduce their dimension. Thanks to feature reduction and transformation methods, features that do not contribute to network training are reduced, and solutions with less computational complexity are provided [52]. Although accuracy comes to the fore in AI-based medical diagnostic applications, the reliability and computational complexity of the system should also be considered. A network with less computation is preferable to a more complex network with similar accuracy.

This study presents a framework based on multiple lightweight DL models and transformation methods for the early detection of lung and colon cancer, which has a striking number of cases and deaths in both men and women worldwide. Experimental studies are performed on the five-class LC25000 dataset containing histopathological images for colon cancer and lung cancer. Unlike previous work, both accuracy and less computational complexity are adopted in principle. In this context, after preprocessing involving resizing and data augmentation, histological images are fed into ShuffleNet, MobileNet, and SqueezeNet models, which are lightweight DL architectures. PCA and FHWT methods are used to reduce the size and complexity of deep features obtained from these CNN models. Afterwards, reduced features of the FHWT obtained from the three DL models are fused using DWT. Additionally, PCA features attained using the three DL models are concatenated. As a result of PCA and FHWT reduction and fusion steps, the reduced features for each model are given as inputs to four different machine learning algorithms (LDA, QDA, SVM, and ensemble subspace discriminate (ESD)), both individually and combined. The classification accuracies obtained from machine learning algorithms show that the proposed methodology can distinguish lung and colon cancer variants with less computational complexity and high accuracy.

The key novelty and contributions of the introduced framework are:

- The proposed framework depends on lightweight deep learning models, including ShuffleNet, MobileNet, and SqueezeNet, in contrast to many of the earlier studies based on heavy deep learning models such as VGG.
- The majority of related studies relies on individual deep learning models for performing the diagnosis; nevertheless, the proposed framework utilizes three deep learning models of various structures.
- Most previous studies used spatial deep features extracted from the last pooling layer of the deep learning models directly to train machine learning classifiers, which may have huge dimensions and thus increase the complexity and duration of the classification process; however, the proposed framework employs two feature reduction approaches to diminish their dimension, including PCA and FHWT.
- The reduced sets of features generated after FHWT provides the spatial–frequency demonstration of the features, not only spatial information such as with previous methods.
- The proposed framework integrates the privileges of distinct architectures of three deep learning models by first merging the reduced features obtained using PCA for the three deep learning models.
- Furthermore, the three reduced sets of features produced after FHWT for ShuffleNet, MobileNet, and SqueezeNet are merged using DWT, which offers the time–frequency

representations of the features of lung and colon cancer, resulting in spatial–time–frequency representations, which usually improves diagnostic performance and further lowers the features dimension.

## 2. Literature Survey

The proposed method in this study is analyzed on the LC25000 lung and colon cancer histopathological image dataset published in 2020. Because lung and colon cancer cases are frequent, the dataset is new and contains a suitable number of images for deep learning, and many researchers have recently implemented deep learning-based applications on this dataset. Some of these are explained below.

Masud, Sikder, Nahid, Bairagi, and AlZain [9] presented a classification system for five types of lung and colon tissues by employing deep learning on histopathological images. First, pathological sample images were subjected to image sharpening. Then, the features were extracted from the images with a 2D-Fourier transform and 2D wavelet transform. These features were used to train a manually tuned CNN model. It was reported that the accuracy performance of this model was 96.33%. Alternatively, Kumar, Sharma, Singh, Madan, and Mehandia [8] compared two classification methods for lung and colon cancer. In the first presented method, texture-, color-, and shape-based features were extracted from histopathological images. These attributes were used for classification with various machine learning methods such as random forest, multilayer perceptron, support vector machine (SVM) with radial basis function, and gradient boosting. In the second method, transfer learning (TL) was used for feature extraction. For this purpose, seven pre-trained CNN models were used for feature extraction. The extracted deep features were applied to the same machine learning methods for classification. It was reported that the random forest algorithm had the best performance with attributes obtained by the TL method from DenseNet-121. The performance of this couple was 98.60% in terms of accuracy. Talukder et al. [53] presented a hybrid ensemble attribute-obtaining method for the identification of lung and colon cancer types. The deep feature extraction method and ensemble learning for image filtering were integrated. It was reported that the proposed hybrid model detected cancer possibility with a rate of 99.05%, whereas Mangal et al. [54] proposed a CNN-based diagnosis system. A shallow neural network architecture was employed for the classification of samples as squamous cell carcinomas, adenocarcinomas, and benign. For lung and colon samples, separate training was performed. It was reported that the success rates of the presented models were 97% and 96% for lung and colon, respectively. Similarly, Hatuwal and Thapa [55] presented a CNN-based histopathological image classification method for cancer diagnosis. A custom-shaped neural network was created and trained. It was reported that the training and validation accuracies were 96.11% and 97.20%, respectively. In contrast, Ali and Ali [56] proposed a multi-input capsule network model for the diagnosis of lung and colon cancer. CLB (convolutional layers block) and SCLB (separable convolutional layers block) convolutional layers were used in this model. By using the dual input technique, the model learned the features more efficiently. Their proposed model achieved 99.58% accuracy for lung and colon abnormalities based on histopathological images. On the other hand, Hasan et al. [57] used the deep CNN model for detecting and classifying colon cancer using colon images of the LC25000 dataset. In this colon data, there were 10,000 histopathological images of colon samples. The operations performed on the images split the data as 80% training, 10% validation, and 10% testing, data augmentation in the separated groups, normalization, and feature extraction with PCA. A ten-layered custom CNN model was constructed and trained. The images are classified as adenocarcinoma or benign. Success rates obtained from different pre-trained CNN models with transfer learning (TL) and classification operations and the results of the specified method were compared. Success rates obtained from different pre-trained CNN models with TL and classification operations and the results of the specified method were compared. In addition, Bukhari et al. [58] used three CNN models, ResNet18, ResNet30, and ResNet50, to classify digital images of colon tissue. The publicly available LC25000

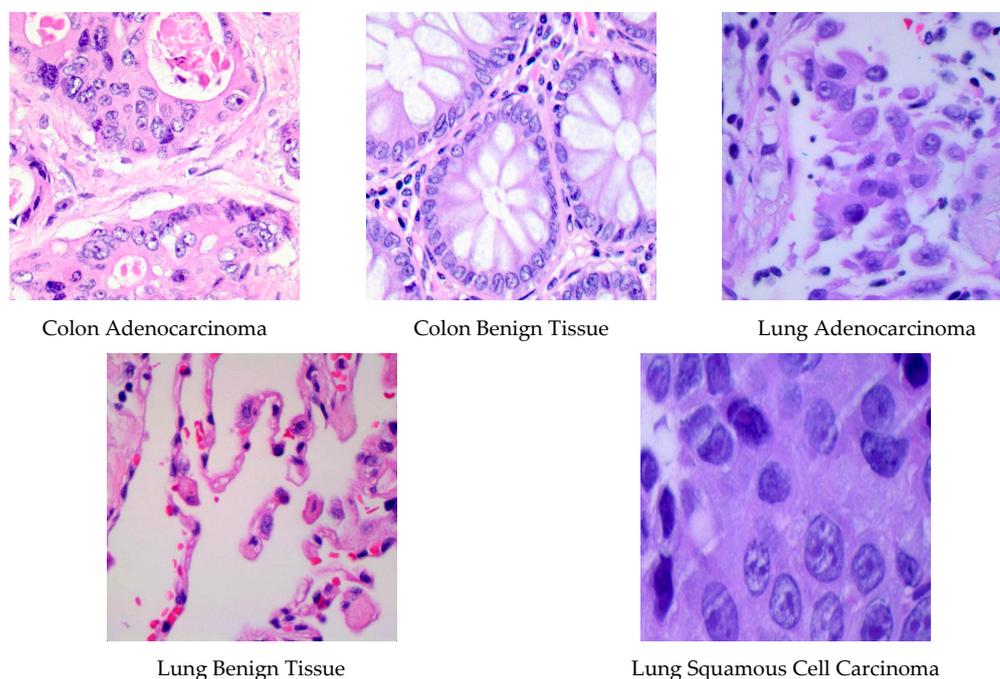
and (colorectal adenocarcinoma gland) CRAG datasets were used in this study. The colon images of the LC25000 dataset were used to train the models and for validation phases. The CRAG dataset was used for the training and testing phases. The highest classification accuracy of 93.13% was achieved with the Resnet50 model.

When previous studies are examined in general, deep features extracted with different CNN models are classified by either end-to-end DL or TL and machine learning techniques. Furthermore, these studies depended on DL models, which necessitate a high level of computational ability and resources. In addition, the majority of them relied on individual DL models to extract high-dimension features or perform diagnosis. However, our study, unlike previous studies, uses multiple lightweight CNN models and adopts feature transformation methods that perform reduction and provide a better representation of the data approaches for reliable and robust diagnosis.

### 3. Materials and Methods

#### 3.1. LC25000 Dataset

The LC25000 dataset, created by Borkowski et al. [59] in 2019, contains cancerous samples of lung and colon tissues. This dataset has been approved by the Health Insurance Portability and Accountability Act (HIPAA). It contains lung and colon tissue images. Lung images are classified into three classes: lung benign tissue, lung adenocarcinoma, and lung squamous cell carcinoma. Colon sample images consist of colon benign tissue and colon adenocarcinoma classes. Example for each class of lung and colon images are shown in Figure 1. In the dataset, which consists of five classes in total, there are 5000 samples in each class. As a result, this dataset contains a total of 15,000 images from three-class lung tissue and 10,000 images from two-class colon tissue.



**Figure 1.** Lung and colon image samples of the dataset.

#### 3.2. Feature Transformation Approaches

##### 3.2.1. Discrete Wavelet Transform

Wavelet transform (WT) is a powerful method for the time–frequency analysis of a signal. Although Fourier transform (FT) is also used for frequency analysis, time information is lost after FT. Therefore, it cannot be known which frequency occurs in which time period. Short-time FT (STFT), which is a modification of FT, also includes time resolution by applying FTs at a certain interval [60]. However, fixed-width windows cannot correctly analyze

non-stationary signals. WT performs a calculation with variable window sizes to provide more precise time–frequency resolution. In this way, the analysis of non-stationary signals, whose frequency changes over time, is performed in a powerful way [61]. A mother wavelet is determined for WT, and then the wavelet function is scaled and translated. After scaling and translation, the wavelet function is multiplied by the signal and summed over the time domain. These processes continue throughout the signal. In continuous wavelet transform (CWT), the transformation is performed by scaling and translating the wavelet function over the entire time interval. Because this requires excessive computation, the scaling and translation values are usually set as discrete values (discrete wavelet transform (DWT)). When DWT is applied to the image, operations are performed in two dimensions. As a result, four matrices are formed. Three of them (horizontal, vertical, diagonal) represent the detail coefficient, while one represents the approximate coefficient. The detail coefficients are the result of the high-pass filter, and the approximate coefficients are the result of the low-pass filter. Therefore, the edges are more prominent in the detail coefficients. Detail coefficients represent vertical, horizontal, and diagonal coefficients [37].

### 3.2.2. Principal Component Analysis

Features extracted from a dataset do not contain equal information for classification or regression. Some features are more related to output. Moreover, in artificial intelligence-based applications, high-dimensional features increase the memory requirement, computational cost, and the possibility of overfitting. For this reason, using feature selection or feature reduction steps before classification instead of using direct CNN models with many features can increase the efficiency of the network. To create a feature vector that better represents the output data, either useful features are selected, or different transformations are applied to existing features [62,63]. The most well-known method for size reduction is PCA.

PCA reduces the size of high-dimensional input features by calculating the correlation between them. However, during size reduction, it statistically analyzes the data and generates high variance features at the output. It uses orthogonal transformation techniques to achieve this. In general terms, the features obtained as a result of PCA are formed as a result of projection from multidimensional space to a lower dimensional space. As a result of PCA, new features are both less dimensional and more distinguishable. The features that result from PCA are called principal components. There is no correlation between the principal components, that is, the principal components are perpendicular to each other in the new space created by PCA [63,64].

### 3.2.3. Fast Walsh–Hadamard Transform

Walsh–Hadamard transform (WHT), which is one of the linear image transformations, is a square wave transform, which is frequently used in signal and image processing, orthogonal and non-sinusoidal. After conversion, a signal is decomposed into a series of functions. These functions are known as Walsh functions. Walsh functions only take two values, +1 and −1. Equation (1) shown below is used to apply the WHT to an array  $f(x)$  with  $n$  samples [65].

$$F(X) = \frac{1}{n} \times [H_n] \times f(x) \quad (1)$$

Here,  $F(X)$  represents the WHT applied to the  $f(x)$  signal.  $[H_n]$  is the  $n$ -dimensional Hadamard matrix. The Hadamard matrix is a square matrix of size  $n \times n$ , and for  $n = 2$ , it is as follows:

$$H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad (2)$$

WHT features simple and fast conversion and is therefore useful for real-time signal and image processing applications. It has the least computational cost among discrete orthogonal transformations [66,67]. With fast Walsh–Hadamard transform (FWHT), WHT is calculated more quickly and efficiently.

### 3.3. Proposed Framework

The proposed framework is made up of four stages, which are preprocessing of histopathological images, deep learning models training and feature extraction, feature reduction and incorporation, and finally classification. Firstly, the histopathological image sizes are altered, and then these images are augmented. Next, three pre-trained lightweight deep learning models including ShuffleNet, SqueezeNet, and MobileNet are constructed, and then deep features are extracted. Then, deep features extracted using each deep learning model are reduced using two feature reduction methods. Furthermore, for each reduction approach, the reduced deep features obtained from each deep learning model are incorporated. Finally, these incorporated features are used to train a number of machine learning classifiers. The four stages of the proposed framework are summarized in Figure 2.

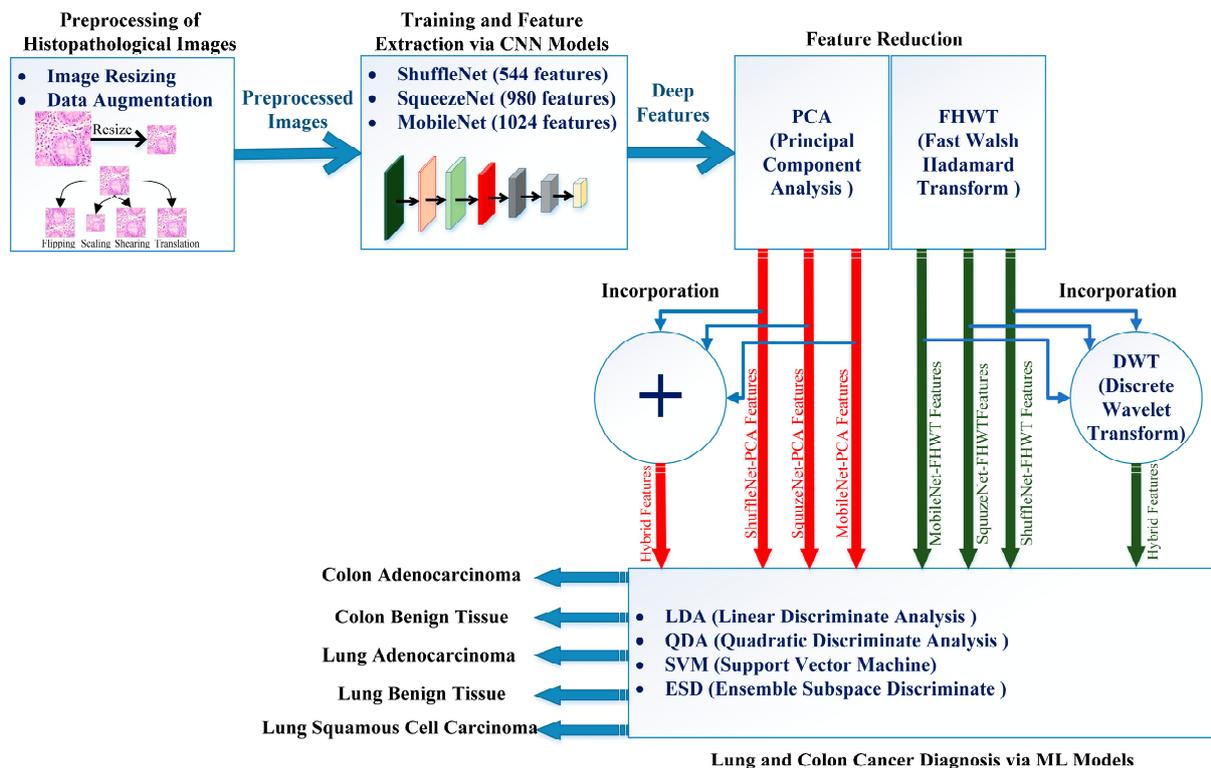


Figure 2. Stages of the proposed framework.

#### 3.3.1. Preprocessing of Histopathological Images

The input layer of each CNN accepts an input image of a specific size in order to start the training procedure. Thus, the histopathological images dimension of lung and colon cancer are initially altered to be similar to the input layer size of the three deep learning models. The input layer size for ShuffleNet, and MobileNet is  $224 \times 224 \times 3$ , while for SqueezeNet it is  $227 \times 227 \times 3$ . Next, in order to improve the training performance of the deep learning models and avoid overfitting, an augmentation procedure is required. Augmentation simply boosts the amount of training images available in a dataset, which allows the training models to learn more effectively [68]. Thus, several augmentation methods are utilized in this paper involving the following: scaling in x and y preferences in the range of [0.5, 2], flipping in both x and y orientations, translation in both x and y orientations with an angle range  $[-20, 20]$ , and shearing in both directions (x and y) within the range  $[-45, 45]$ .

#### 3.3.2. Deep Learning Models Training and Feature Extraction

As mentioned earlier, three pre-trained lightweight convolutional neural networks (CNNs) are used in this stage including ShuffleNet, SqueezeNet, and MobileNet. Pretrained

CNN means that the network has been previously trained with a large dataset. Transfer learning [69] allows pre-trained CNNs to realize representations from a large dataset of images, such as those in ImageNet, and then apply that knowledge to a related classification problem with smaller datasets, advancing the training process. Therefore, TL is used along with the three pre-trained CNNs. TL is used to modify the number of output layers to five to be equivalent to the number of categories in the dataset. In addition, other CNNs parameters are modified, which will be explained later. Then the three pre-trained CNNs are retrained with the lung–colon cancer dataset. Next, TL is further used to extract deep features from a particular layer of each CNN. In order to produce a more accurate image classification, the CNN first learns the simple patterns of the input image before moving on to learn the important elements of the input image in subsequent, deeper layers [35,36]. Thus, deeper layers are used for feature extraction. The layers used for features extraction are the layers called “node200” for ShuffleNet, “global\_average\_pooling2d\_1” for MobileNet, which are the last pooling layers of these CNNs, whereas for SqueezeNet, the layer “relu\_conv10” is utilized for feature extraction, which is a rectified linear unit layer immediately following the tenth convolutional layer. The length of each of the feature vectors extracted is 544, 980, and 1280 for Shuffle, SqueezeNet, and MobileNet, respectively.

### 3.3.3. Feature Reduction and Incorporation

The dimensions of the feature extracted in the previous steps are large, which could increase the complexity and the duration of the training of machine learning classifiers. Hence, in this stage, two feature reduction approaches are used and compared to lower features dimension and extract a significantly reduced representation of features. These methods include principal component analysis (PCA) and fast Walsh–Hadamard transform (FWHT). Each deep feature set extract from every CNN is analyzed utilizing PCA. Several numbers of principal components are used to study the effect of changing these numbers on diagnostic performance. Similarly, the deep feature sets of the CNNs are analyzed using FWHT, which presents the spectral representations of the input. Different reduced set extracts using FWHT are also employed to examine the influence of varying these numbers on the diagnostic performance. An incorporation step follows the two feature reduction procedures. For PCA, principal components obtained using the three CNNs are concatenated. On the other hand, the reduced feature sets of the three CNNs generated using FWHT are incorporated using DWT, which further diminishes their dimension.

### 3.3.4. Lung and Colon Cancer Diagnosis

Four machine learning classifiers are used in this stage to diagnose lung and colon cancer and identify the five classes of the lung and colon cancer dataset. These classifiers are LDA, QDA, and SVM with linear kernel function, and ensemble subspace discriminate (ESD). The number of learners and subspace for the ESD classifiers is 30 and 2, respectively, while its learning rate is 0.1. Five-fold-cross validation is employed to evaluate the performance of the proposed CAD.

## 4. Performance Measures and CNNs’ Parameters Setting

### 4.1. Performance Measures

Several evaluation criteria are used to assess the effectiveness of the proposed pipeline, including precision, accuracy, F1-score, sensitivity, Mathew correlation coefficient (*MCC*), and specificity. The ensuing formulae are used to calculate these metrics (3)–(8). The confusion matrix and receiving operating characteristic (*ROC*) are also computed.

$$Accuracy = \frac{TP + TN}{TN + FP + FN + TP} \quad (3)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (6)$$

$$F1 - Score = \frac{2 \times TP}{(2 \times TP) + FP + FN} \quad (7)$$

$$Specificity = \frac{TN}{TN + FP} \quad (8)$$

The true positive ( $TP$ ) represents the proportion of instances that are perfectly recognized as positive, the false negative ( $FN$ ) denotes the proportion of samples that have been mistakenly categorized as negative, the true negative ( $TN$ ) represents the proportion of instances that are correctly identified as negative, and the false positive ( $FP$ ) resembles the proportion of samples that are erroneously classified as positive.

#### 4.2. CNN's Parameters Setting

The learning rate, frequency of epochs, and minimum batch size were adjusted to 0.0001, 5, and 4, respectively, for retraining the CNNs for end-to-end classification. Furthermore, the validation frequency was set to 130 to calculate the training error per epoch. The stochastic gradient descent with a momentum technique was applied to train the three CNNs. All other hyperparameters remained the same. MATLAB R2020a was employed in implementing the proposed framework.

### 5. Results

#### 5.1. Results of PCA

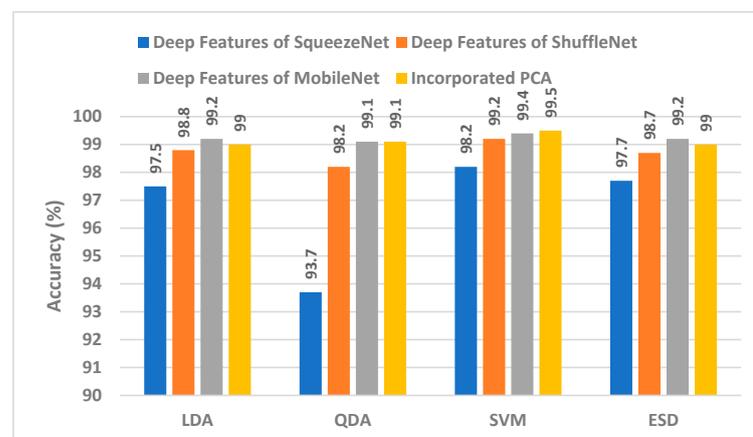
This section examines the performance of reduced features as a result of applying PCA to features extracted from CNN models. These reduced features were individually fed into four different machine learning methods. As a result of applying PCA to deep features, new features are formed according to the number of principal components determined. Therefore, the number of principal components has a direct impact on classification performance. According to the change in the number of principal components, the diagnostic performance of machine learning algorithms also changes. Table 1 shows the variation in the diagnostic performance of different machine learning algorithms as a result of processing the extracted features from SqueezeNet, ShuffleNet, and MobileNet CNN models with PCA with different principal component values. In general, all machine learning algorithms showed outstanding success for lung and colon cancer diagnosis. However, as can be seen from Table 1, the change in PCA values had a significant effect on the accuracy of machine learning methods, because as the PCA value changed, both the value of the features and the number of features changed. As the principal component value was increased from 5 to 30, the diagnostic accuracy of the ShuffleNet-LDA fabric increased by 4.9%. Accuracies for SqueezeNet-LDA and MobileNet-LDA structures increased by 4.2% and 1% with an increasing number of principal components. The accuracy changes for other models (ShuffleNet-QDA, SqueezeNet-SVM, MobileNet-ESD, etc.) are also clearly visible in Table 1.

When Table 1 is examined to compare the success of CNN model features, it is seen that machine learning algorithms classified MobileNet features more accurately. Figure 3 shows the accuracies obtained as a result of the classification of features extracted from all three CNN models by four different machine learning algorithms. The superior performance of MobileNet compared to others can also be seen in Figure 3. Figure 3 also includes the classification accuracies obtained from PCA reduction of the combined features of the three CNN models (incorporated PCA). PCA significantly reduced the combined features of the three CNN models. Figure 4 shows the features extracted from the CNN models and the number of new features obtained by PCA. Accordingly, hybrid features provided very little computational cost compared to CNN models. Note that in Table 1, the highest performance

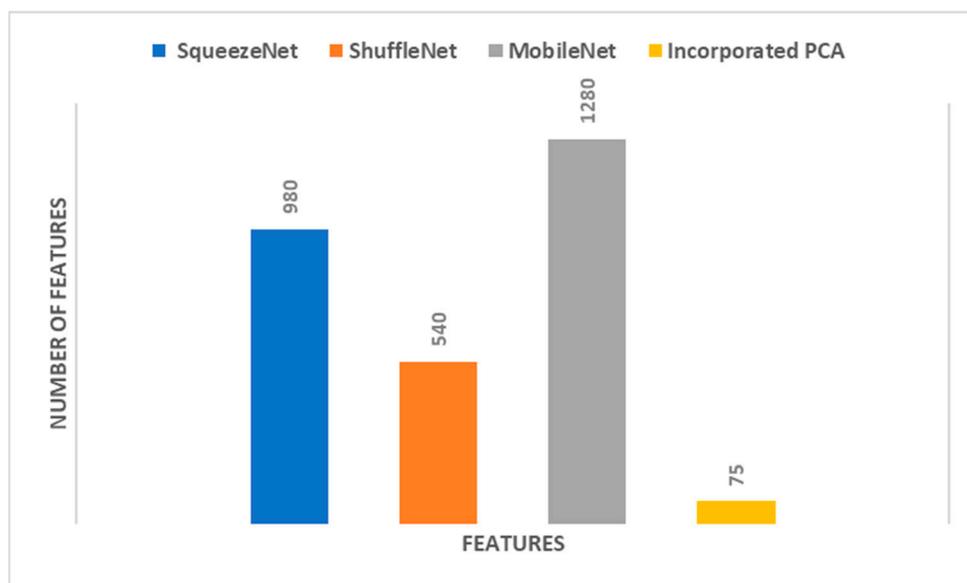
is attained with 15 PCA features for SqueezeNet, 30 PCA features for ShuffleNet, and 30 PCA features for MobileNet. Therefore, when concatenating these PCA features, the total number of incorporated PCA features will be 75, and that is the reason why in Figure 4, the number of PCA features is 75.

**Table 1.** Classification accuracies (%) for several numbers of principal components after the PCA process conducted for the deep features of the three CNNs.

SqueezeNet				
Features	LDA	QDA	Linear SVM	ESD
PCA = 5	93.4	96.6	97.2	93.0
PCA = 10	96.0	96.7	97.7	94.5
PCA = 15	97.5	96.7	98.1	96.8
PCA = 20	97.5	96.4	98.1	96.7
PCA = 25	97.5	96.1	98.1	96.9
PCA = 30	97.6	96.0	98.1	96.9
ShuffleNet				
PCA = 5	92.8	96.8	97.4	91.7
PCA = 10	95.4	97.7	98.2	93.4
PCA = 15	96.4	97.8	98.5	94.8
PCA = 20	97.1	97.8	98.7	95.5
PCA = 25	97.1	97.8	98.7	95.5
PCA = 30	97.7	97.9	98.8	96.6
MobileNet				
PCA = 5	97.4	98.5	98.7	97.5
PCA = 10	98.1	98.6	99.0	97.9
PCA = 15	97.9	98.7	99.1	97.8
PCA = 20	98.3	98.8	99.1	98.1
PCA = 25	98.3	98.9	99.2	98.1
PCA = 30	98.4	98.9	99.2	98.3

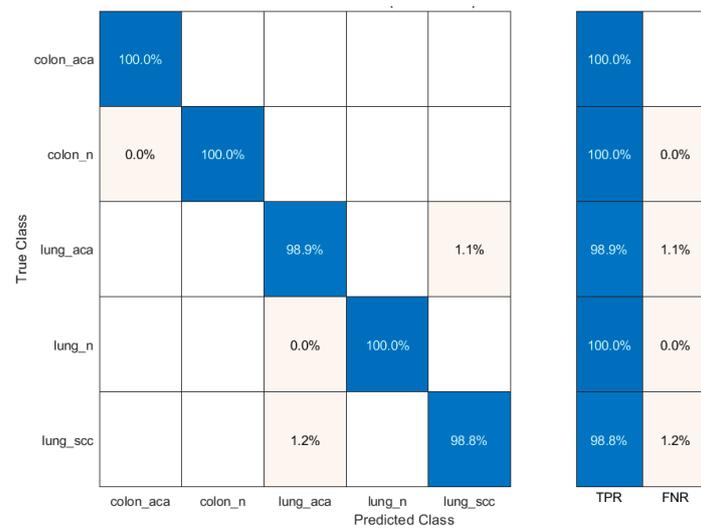


**Figure 3.** Classification accuracies (%) of deep features versus integrated PCA features of the three CNNs.

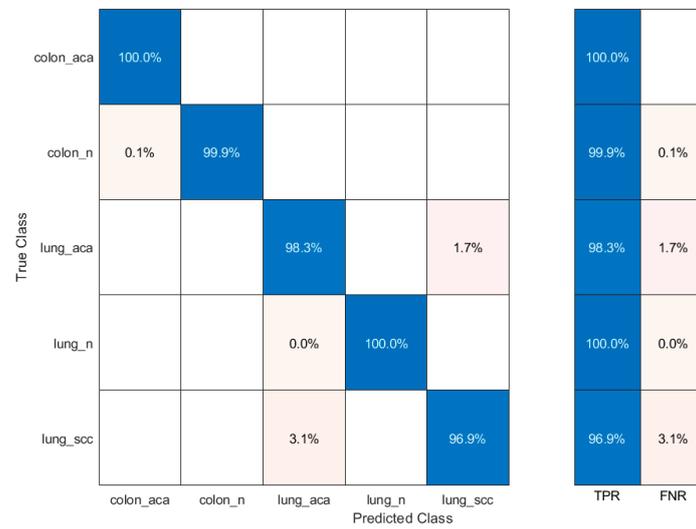


**Figure 4.** Number of features of each of the deep features extracted from the three CNNs compared to the incorporated reduced features of PCA used to train the four machine learning classifiers.

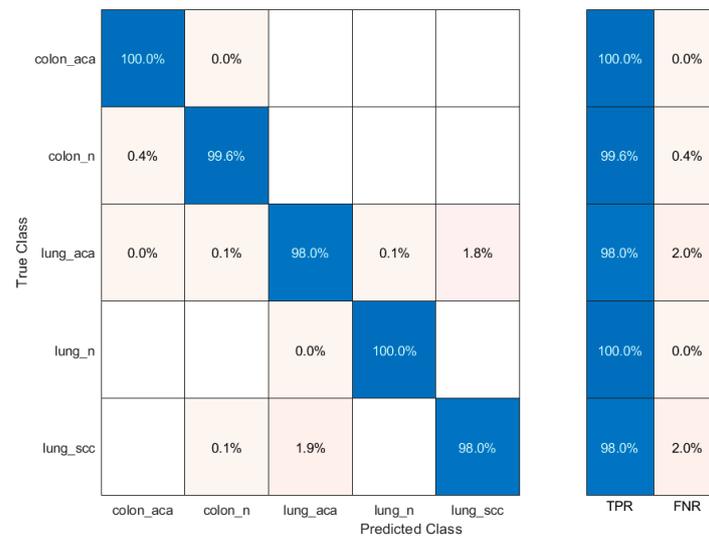
According to Figure 3, as a result of the classification of hybrid features with machine learning, the highest accuracy was obtained with SVM at 99.5%. Accuracies from other hybrid features were 99%, 99.1%, and 99% for LDA, QDA, and ESD, respectively. Confusion matrices showing the correct and incorrect classification percentages for each class are given in Figure 5. Considering the confusion matrices, all of the colon adenocarcinoma class types were correctly recognized by all machine learning algorithms. In addition, the lung benign tissue class contained 0.1% error in the result of QDA, while it was classified with 100% accuracy by other algorithms. According to the results of the four machine learning algorithms, lung squamous class samples were more difficult to recognize than other classes. Therefore, the error made for this class was higher than for other classes. Machine learning methods generally classify samples belonging to the lung squamous class as lung adenocarcinoma. As can be seen in Figure 5, ESD is the machine learning algorithm that made the most incorrect classifications in the images belonging to the lung squamous class. However, ESD correctly classified 96.7% of test samples belonging to the lung squamous class. Table 2 shows the different performance metrics obtained from the confusion matrices for each machine learning algorithm. When the sensitivity, specificity, precision, F1-score, and MCC metric values were analyzed, the SVM method with linear kernel function produced the most successful results. These results were 0.996, 0.999, 0.996, 0.996, and 0.994, respectively. The most erroneous results were obtained as a result of ESD. Unlike these metric values, receiver operating characteristic (ROC) curves, a popular method to evaluate the accuracy of medical diagnostic systems, were also plotted, as seen in Figure 6 for SVM, which was the most successful machine learning method. This curve was created by plotting the true positive rate against the false positive rate. For successful classification, the area under the curve (AUC) should be close to 1 [70]. As seen in Figure 6, the AUC value was calculated as 1 for all curves. According to these results, the proposed method based on CNN and PCA methods provides a low-cost, unbiased, and highly accurate diagnosis of lung and colon cancer.



SVM

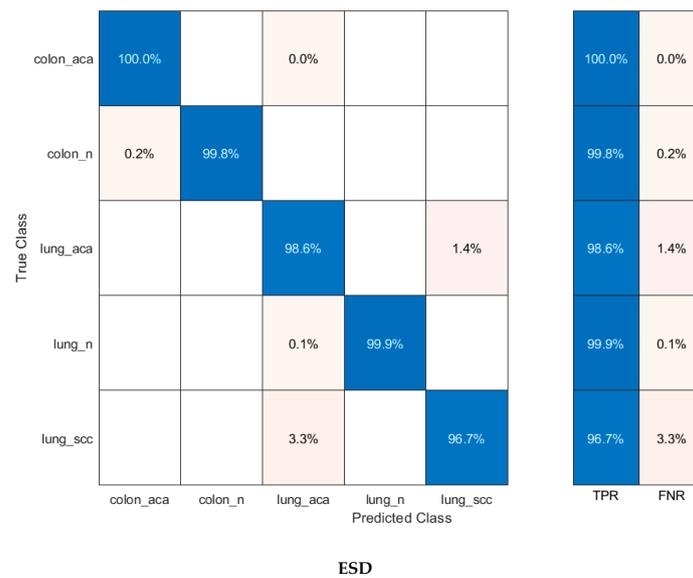


LDA



QDA

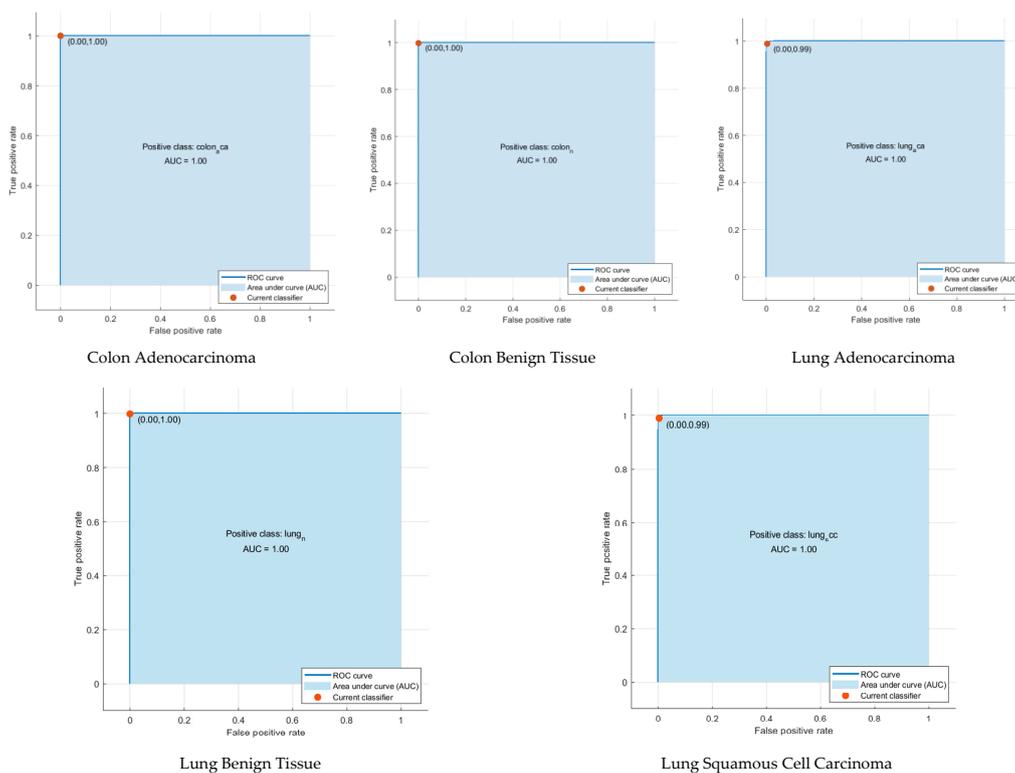
Figure 5. Cont.



**Figure 5.** Confusion matrices obtained as a result of classification of PCA-based incorporated CNN features with machine learning algorithms.

**Table 2.** Performance measures attained using the combined PCA coefficients of the three CNNs.

Model	Sensitivity	Specificity	Precision	F1-Score	MCC
LDA	0.991	0.998	0.991	0.991	0.988
QDA	0.991	0.998	0.991	0.991	0.989
Linear SVM	0.996	0.999	0.996	0.996	0.994
ESD	0.989	0.997	0.990	0.989	0.987



**Figure 6.** ROC curves of each class as a result of the CNN-PCA-SVM structure.

### 5.2. Results of FWHT

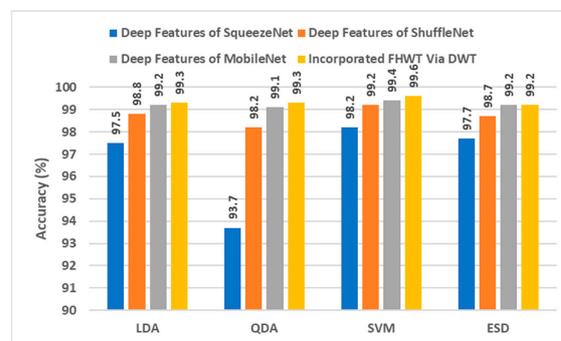
This section examines the performance of reduced features as a result of applying FHWT to features extracted from CNN models. FHWT, which is used for feature reduction, has a low computational cost and is preferred in real-time applications because it is a fast conversion tool. In this study, deep features extracted with CNN models were reduced by FHWT. Then these features were classified by machine learning methods. After feature reduction with FHWT, machine learning results differed based on the number of new features. To show the effect of feature reduction on diagnostic performance, different feature size values and corresponding machine learning algorithm accuracy values are shown in Table 3. When the feature sizes and accuracy values were examined, the increase in the number of features decreased the accuracy of the SqueezeNet model. For the ShuffleNet and MobileNet models, the higher the number of features, the higher the accuracy. Accordingly, as the number of features increased, the accuracy of the ShuffleNet-QDA and MobileNet-QDA structures increases by 4.1% and 2.8%, respectively. For SqueezeNet-QDA, increasing the number of features reduced the accuracy by 2.4%.

**Table 3.** Classification accuracies (%) for different feature sets after the process conducted for the deep features of the three CNNs.

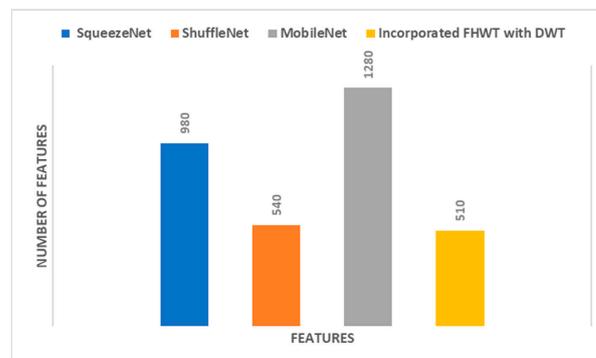
SqueezeNet				
Features	LDA	QDA	Linear SVM	ESD
50	97.7	95.3	98.0	97.6
40	97.8	96.5	98.1	97.6
30	97.8	96.9	98.1	97.6
20	97.8	97.5	98.2	97.6
10	97.8	97.7	98.2	97.6
ShuffleNet				
500	98.4	93.3	99.0	98.3
400	98.2	92.9	99.0	98.1
300	97.8	92.4	98.8	97.8
200	97.4	91.3	98.6	97.2
100	94.4	89.2	97.3	94.1
MobileNet				
500	98.7	97.1	99.2	98.7
400	98.5	97.0	99.2	98.5
300	98.4	96.8	99.1	98.3
200	97.8	96.2	98.8	97.8
100	96.8	94.3	98.1	96.7

When Table 3 is examined to compare the success of CNN models, MobileNet's superiority over other models is striking, as in the PCA analysis. This can be seen more clearly in Figure 7. Each bar represents a different CNN-machine learning structure so that the success of CNN features and machine learning algorithms can be compared more easily. Figure 7 compares the performance of the deep features of each CNN with the incorporated FHWT features of the three CNNs using DWT. All four machine learning methods classified MobileNet features more accurately. Accordingly, it is understood that MobileNet features represent cancer types more strongly. The most unsuccessful classification was carried out with SqueezeNet features. Figure 7 also includes classification accuracies of hybrid CNN features reduced by FHWT and incorporated with DWT. Accordingly, it is seen that hybrid features created using DWT were better classified by machine learning algorithms. The number of each CNN deep feature and the number of combined features created at the end of the incorporation are shown in Figure 8. This figure shows that the number of

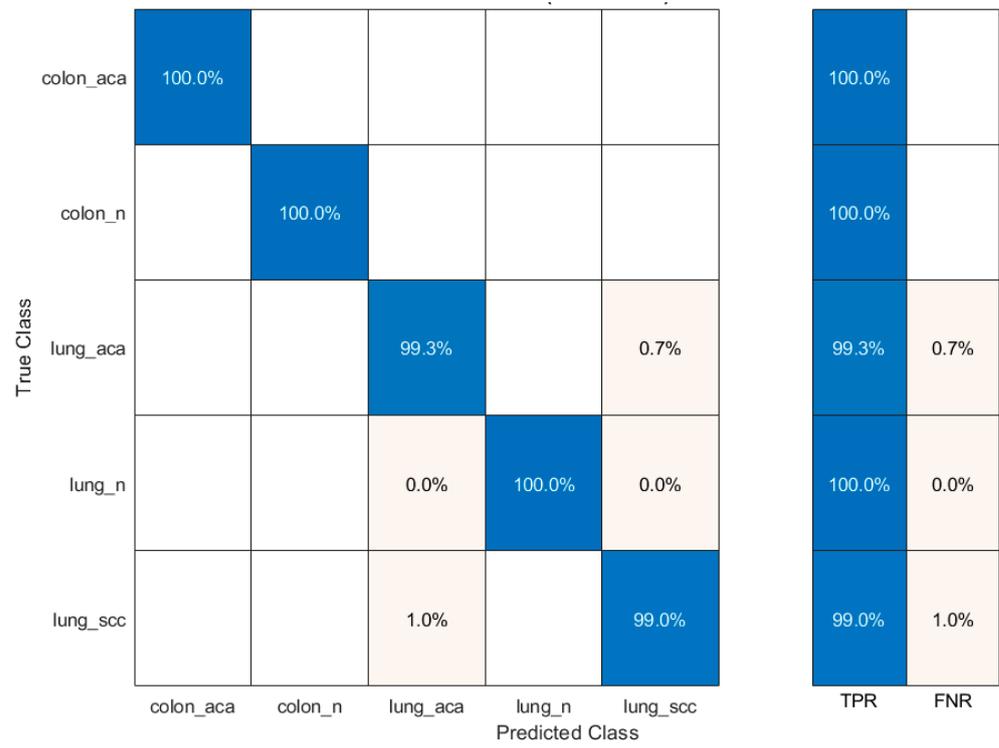
FHWT features combined with DWT was much less than other CNN models. Thanks to the feature reduction methods used, more successful results were obtained with fewer features. This also means less computational costs and more reliable results. Accordingly, although hybrid features include all CNN features, they contain fewer features than other models. Furthermore, Figure 7 shows that machine learning algorithms classified hybrid features more accurately than individual features. The most successful method of classifying hybrid features was SVM, and its accuracy value was 99.6%. The accuracies obtained from the hybrid features for LDA, QDA, and ESD were 99.3%, 99.3%, and 99.2%, respectively. The confusion matrices obtained during the classification of hybrid features are also shown in Figure 9. Based on the confusion matrices, colon adenocarcinoma was distinguished with 100% accuracy by all machine learning models. In addition, colon benign tissue and lung benign tissue classes were classified with 100% accuracy by SVM and LDA. Machine learning methods generally made the most inaccurate classification for the lung squamous and lung adenocarcinoma classes. In the misidentifications, some samples belonging to the lung squamous class were classified as lung adenocarcinoma. Similarly, some lung adenocarcinoma samples were assigned to the lung squamous class. The fact that the features in these two classes are similar to each other caused these misclassifications. ESD, which had the lowest average accuracy among machine learning algorithms, classified lung adenocarcinoma and lung squamous classes with 98.2% and 97.8% accuracy, respectively. Table 4 shows the different performance metrics calculated as a result of classification with machine learning algorithms. Similar to accuracy values, while the most successful metric values were provided by SVM, the method with the highest error rate was ESD. Sensitivity, specificity, precision, F1-score, and MCC metric values for SVM were 0.996, 0.999, 0.996, 0.996, and 0.995, respectively. In addition, as in the analysis with PCA, ROC curves were plotted for the most successful machine learning method, SVM, in this section. It is seen in Figure 10 that AUC values are 1 in these curves drawn for each class.



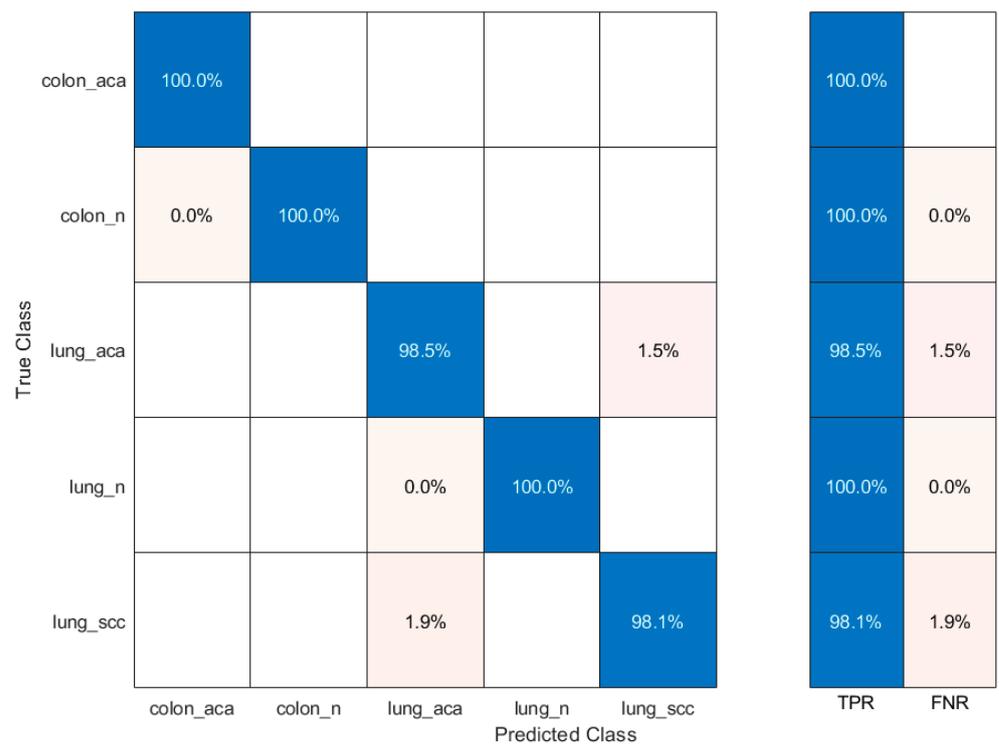
**Figure 7.** Classification accuracies (%) of deep features versus integrated FHWT features via DWT of the three CNNs.



**Figure 8.** The number of features of each of the deep features extracted from the three CNN compared to the incorporated reduced features of FHWT used to train the four machine learning classifiers.

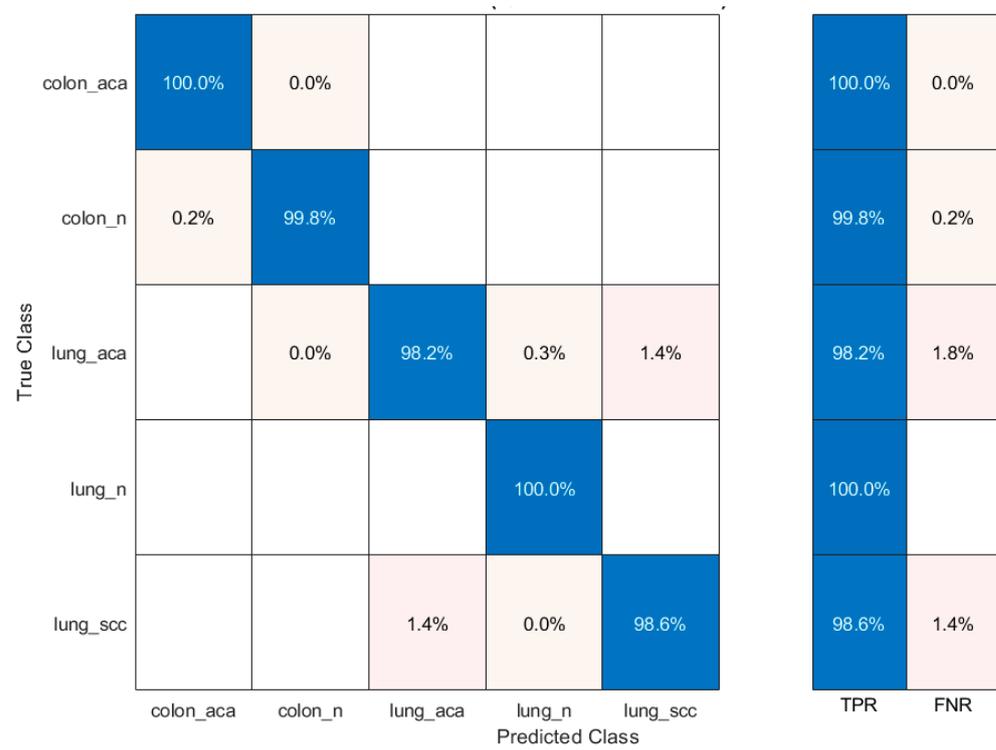


**SVM**

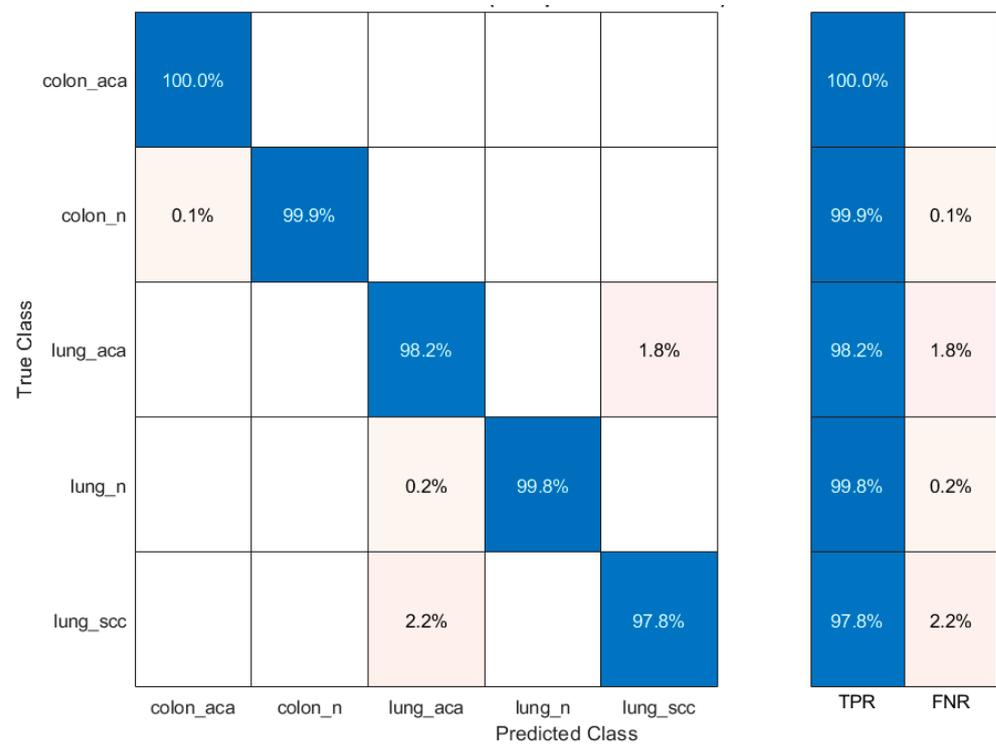


**LDA**

Figure 9. Cont.



**QDA**

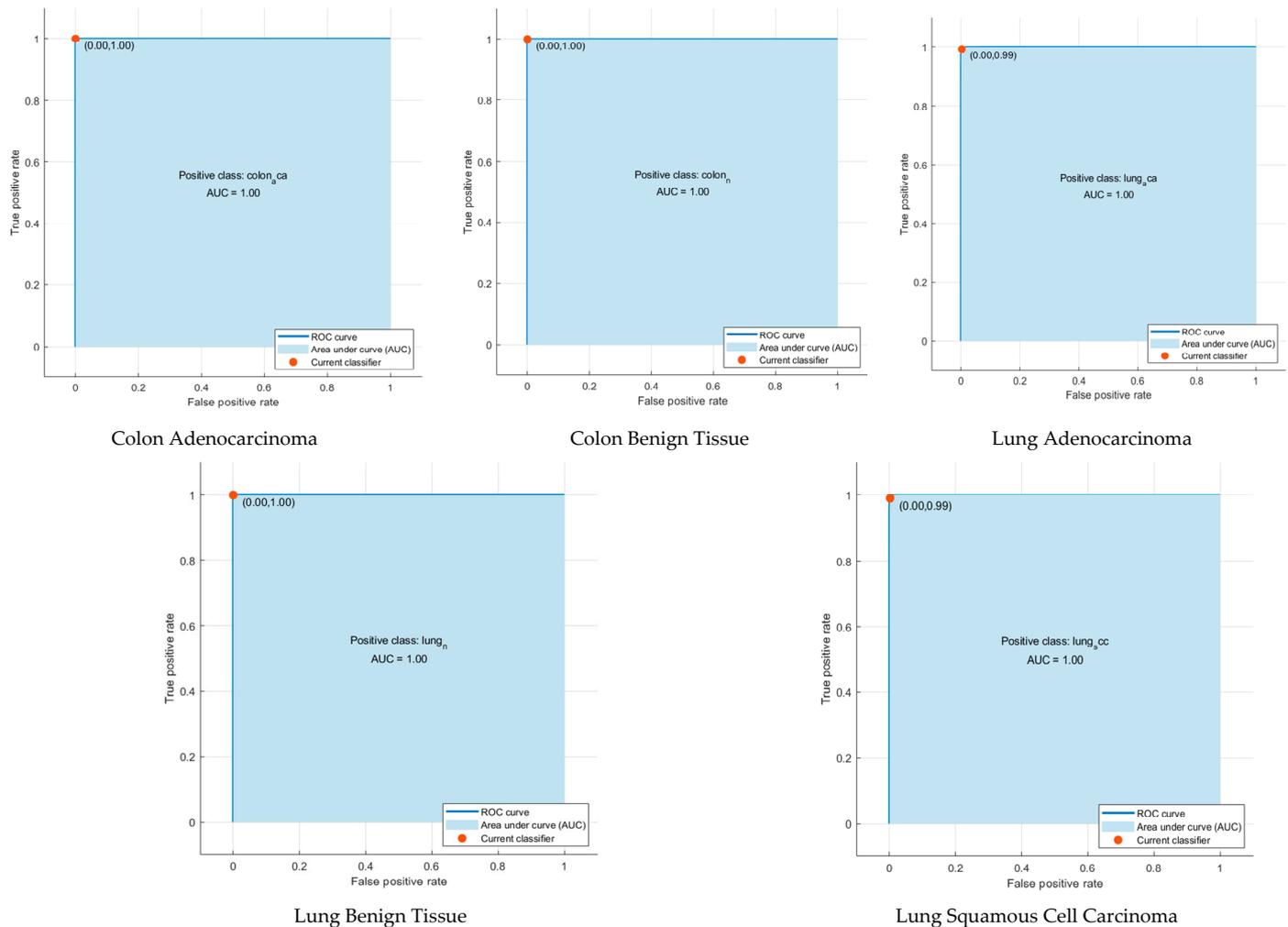


**ESD**

**Figure 9.** Confusion matrices obtained as a result of the classification of FHWT-based incorporated CNN features using DWT with machine learning algorithms.

**Table 4.** Performance measures attained using the combined FHWT coefficients of the three CNNs via the DWT method.

Model	Sensitivity	Specificity	Precision	F1-Score	MCC
LDA	0.993	0.998	0.993	0.993	0.991
QDA	0.993	0.998	0.993	0.993	0.991
Linear SVM	0.996	0.999	0.996	0.996	0.995
ESD	0.992	0.997	0.992	0.992	0.989



**Figure 10.** ROC curves of each class as a result of the CNN-FHWT-SVM structure.

### 5.3. Comparison with the Literature

A comparison was conducted in this section to compare the results of the proposed framework with methods available in the literature. Table 5 demonstrates the results of this comparison. The performance of the proposed framework was competitive with other models employed in the literature. This is because accuracies attained after PCA and FHWT fusion were 99.5% and 99.6%. These accuracies were higher than most of the methods in the literature based on using a single DL model having high computation capacity. Thus, the system can be used to assist pathologists in the early detection of lung and colon cancers.

**Table 5.** Comparative analysis with previous works.

Authors	Method	Accuracy (%)
Masud, Sikder, Nahid, Bairagi, and AlZain [9]	CNN + 2D Fourier transform and 2D wavelet transform	96.33
Hasan, Ali, Rahman, and Islam [57]	CNN + PCA	99.80
Kumar, Sharma, Singh, Madan, and Mehandia [8]	DenseNet121 + Random Forest	98.60
Talukder, Islam, Uddin, Akhter, Hasan, and Moni [53]	Deep feature extraction + Ensemble learning	99.05
Bukhari, Syed, Bokhari, Hussain, Armaghan, and Shah [58]	ResNet50	93.13
Mangal, Chaurasia, and Khajanchi [54]	CNN	97.00
Hatuwal and Thapa [55]	CNN	97.20
Ali and Ali [56]	Capsule network model	99.58
Proposed method	PCA + CNN + SVM	99.5
	FHWT + CNN + SVM	99.6

## 6. Discussion

In this study, CNN models with lower computational costs compared to previous studies are used to extract features from lung and colon cancer data. Experimental studies aim to propose a methodology with low computational cost but high accuracy. However, it should be noted that more costly and deeper CNN models can provide more powerful features. Moreover, the hyperparameter settings that are effective in the success of CNN models are determined manually for this study. Hyperparameter optimization methods can be used for a more successful CNN model. Different CNN models represent the same data with different features. Therefore, three different feature vectors are generated from ShuffleNet, MobileNet, and SqueezeNet CNN models to provide feature diversity and compare the different models. The size of the feature vectors is reduced by PCA, FHWT, and DWT. Size reduction can result in the removal of useful features while reducing the computational cost. However, this study does not take this into account, at the expense of developing a less costly method. In addition, the number of newly created features during feature reduction with PCA and FHWT affects system performance. In this study, the feature size is determined manually. Instead, systems that optimize the number of features can be developed. The strength of the proposed method is combining reduced features extracted from individual CNN models to increase feature diversity. In this way, at the classification stage, hybrid features produce more stable results than individual CNN features. However, extracting features from different CNN models is a temporal disadvantage. Instead, designing a single comprehensive deep model can provide more efficient results.

The most important requirement in deep learning-based diagnostic studies is the generalization ability and reliability of the network. In order to increase the number of training images, data augmentation methods are employed in this study. Although data augmentation is useful for deep learning studies, the new images produced are artificial and are a modified version of the original image. Few original images and many augmented images negatively affect the reliability of the artificial intelligence system. For this reason, using a large number of raw data, which are compiled by combining different datasets, provides results closer to reality.

The limitations and shortcomings of the proposed method discussed above will be the focus of our future work. The results obtained in our study showed that the proposed method can be used in histopathological image analysis for lung and cancer detection. Table 5 shows the methods and classification accuracy of previous studies using the same LC25000 dataset. In general, deep learning-based methods have provided successful results

in previous studies. However, it is seen that the proposed method is comparable to previous studies in terms of accuracy.

## 7. Conclusions

Lung and colon cancer are the most common types of cancer worldwide. Early diagnosis is very effective in preventing deaths. In this study, a new methodology based on CNN and feature reduction methods is proposed for the early detection of lung and colon cancer cases. In the proposed methodology, the basic principle was to perform classification in an efficient and computationally low-cost manner. In this context, unlike previous studies, light CNN models were preferred. Instead of using a single model, features were extracted from the lung and colon images with the ShuffleNet, MobileNet, and SqueezeNet models to make comparisons, provide feature diversity, and benefit from hybrid features. PCA and FHWT and then DWT applied to these images reduced the computational cost, improved the variance between features, and offered a superior demonstration of the features used as input to machine learning classifiers. As a result of the application of individual methods to four different machine learning algorithms, high-accuracy diagnoses were achieved. In addition, with the classification of hybrid features obtained as a result of combining reduced features, more stable results were obtained compared to individual CNN models. The results from the proposed methodology showed that an application based on lightweight CNN models and feature transformation methods such as PCA, FHWT, and DWT provides superior accuracy, even though it contains fewer features and lower computational cost.

**Author Contributions:** Conceptualization, O.A. and M.F.A.; methodology, O.A.; software, O.A.; validation, O.A. and M.F.A.; formal analysis, O.A.; investigation, M.F.A. and K.S.; resources, K.S.; data curation, O.A.; writing—original draft preparation, O.A., M.F.A. and K.S.; writing—review and editing, O.A., M.F.A. and K.S.; visualization, O.A. and M.F.A.; supervision, O.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data utilized in this work can be found at <https://www.kaggle.com/datasets/andrewmvd/lung-and-colon-cancer-histopathological-images> (accessed on 5 September 2022).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
2. World Health Organization (WHO). Global Health Estimates: Life Expectancy and Leading Causes of Death and Disability. Available online: <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death> (accessed on 17 October 2022).
3. International Agency for Research on Cancer. Available online: <https://www.iarc.who.int> (accessed on 15 October 2022).
4. Yadav, A.R.; Mohite, S.K. Cancer-A silent killer: An overview. *Asian J. Pharm. Res.* **2020**, *10*, 213–216. [[CrossRef](#)]
5. RL, S.; KD, M.; Jemal, A. Cancer statistics, 2020. *CA Cancer J Clin* **2020**, *70*, 7–30.
6. World Health Organization (WHO). Cancer. Available online: <https://www.who.int/news-room/fact-sheets/detail/cancer> (accessed on 17 October 2022).
7. Kurishima, K.; Miyazaki, K.; Watanabe, H.; Shiozawa, T.; Ishikawa, H.; Satoh, H.; Hizawa, N. Lung cancer patients with synchronous colon cancer. *Mol. Clin. Oncol.* **2018**, *8*, 137–140. [[CrossRef](#)] [[PubMed](#)]
8. Kumar, N.; Sharma, M.; Singh, V.P.; Madan, C.; Mehandia, S. An empirical study of handcrafted and dense feature extraction techniques for lung and colon cancer classification from histopathological images. *Biomed. Signal Process. Control* **2022**, *75*, 103596. [[CrossRef](#)]
9. Masud, M.; Sikder, N.; Nahid, A.-A.; Bairagi, A.K.; AlZain, M.A. A Machine Learning Approach to Diagnosing Lung and Colon Cancer Using a Deep Learning-Based Classification Framework. *Sensors* **2021**, *21*, 748. [[CrossRef](#)]

10. ud din, N.M.; Dar, R.A.; Rasool, M.; Assad, A. Breast cancer detection using deep learning: Datasets, methods, and challenges ahead. *Comput. Biol. Med.* **2022**, *149*, 106073. [[CrossRef](#)]
11. Wang, Y.; Coudray, N.; Zhao, Y.; Li, F.; Hu, C.; Zhang, Y.-Z.; Imoto, S.; Tsigirigos, A.; Webb, G.I.; Daly, R.J. HEAL: An automated deep learning framework for cancer histopathology image analysis. *Bioinformatics* **2021**, *37*, 4291–4295. [[CrossRef](#)]
12. Kumar, M.A.; Ram, P.G.; Suseendhar, P.; Sangeetha, R. An Efficient Cancer Detection Using Machine Learning Algorithm. *NVEO-Nat. Volatiles Essent. OILS J. | NVEO* **2021**, *8*, 6416–6425.
13. Aslan, M.F.; Sabanci, K.; Ropelewska, E. A CNN-Based Solution for Breast Cancer Detection With Blood Analysis Data: Numeric to Image. In Proceedings of the 2021 29th Signal Processing and Communications Applications Conference (SIU), Istanbul, Turkey, 9–11 June 2021; pp. 1–4.
14. Taspinar, Y.S.; Cinar, I.; Koklu, M. Classification by a stacking model using CNN features for COVID-19 infection diagnosis. *J. X-ray Sci. Technol.* **2022**, *30*, 73–88. [[CrossRef](#)]
15. Hamida, A.B.; Devanne, M.; Weber, J.; Truntzer, C.; Derangère, V.; Ghiringhelli, F.; Forestier, G.; Wemmert, C. Deep learning for colon cancer histopathological images analysis. *Comput. Biol. Med.* **2021**, *136*, 104730. [[CrossRef](#)] [[PubMed](#)]
16. Attallah, O. An effective mental stress state detection and evaluation system using minimum number of frontal brain electrodes. *Diagnostics* **2020**, *10*, 292. [[CrossRef](#)] [[PubMed](#)]
17. Karthikesalingam, A.; Attallah, O.; Ma, X.; Bahia, S.S.; Thompson, L.; Vidal-Diez, A.; Choke, E.C.; Bown, M.J.; Sayers, R.D.; Thompson, M.M. An artificial neural network stratifies the risks of Reintervention and mortality after endovascular aneurysm repair; a retrospective observational study. *PLoS ONE* **2015**, *10*, e0129024. [[CrossRef](#)] [[PubMed](#)]
18. Attallah, O.; Karthikesalingam, A.; Holt, P.J.; Thompson, M.M.; Sayers, R.; Bown, M.J.; Choke, E.C.; Ma, X. Using multiple classifiers for predicting the risk of endovascular aortic aneurysm repair re-intervention through hybrid feature selection. *Proc. Inst. Mech. Eng. Part H J. Eng. Med.* **2017**, *231*, 1048–1063. [[CrossRef](#)] [[PubMed](#)]
19. Attallah, O.; Karthikesalingam, A.; Holt, P.J.; Thompson, M.M.; Sayers, R.; Bown, M.J.; Choke, E.C.; Ma, X. Feature selection through validation and un-censoring of endovascular repair survival data for predicting the risk of re-intervention. *BMC Med. Inform. Decis. Mak.* **2017**, *17*, 115. [[CrossRef](#)]
20. Attallah, O.; Sharkas, M.A.; Gadelkarim, H. Fetal brain abnormality classification from MRI images of different gestational age. *Brain Sci.* **2019**, *9*, 231. [[CrossRef](#)]
21. Ozkan, I.A.; Koklu, M. Skin lesion classification using machine learning algorithms. *Int. J. Intell. Syst. Appl. Eng.* **2017**, *5*, 285–289. [[CrossRef](#)]
22. Aslan, M.F.; Ceylan, M.; Durdu, A. Segmentation of retinal blood vessel using gabor filter and extreme learning machines. In Proceedings of the 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, Turkey, 28–30 September 2018; pp. 1–5.
23. Ayman, A.; Attalah, O.; Shaban, H. An efficient human activity recognition framework based on wearable imu wrist sensors. In Proceedings of the 2019 IEEE International Conference on Imaging Systems and Techniques (IST), Abu Dhabi, United Arab Emirates, 9–10 December 2019; pp. 1–5.
24. Attallah, O.; Abougharbia, J.; Tamazin, M.; Nasser, A.A. A BCI system based on motor imagery for assisting people with motor deficiencies in the limbs. *Brain Sci.* **2020**, *10*, 864. [[CrossRef](#)]
25. Aslan, M.F.; Sabanci, K.; Durdu, A. A CNN-based novel solution for determining the survival status of heart failure patients with clinical record data: Numeric to image. *Biomed. Signal Process. Control* **2021**, *68*, 102716. [[CrossRef](#)]
26. Anwar, S.M.; Majid, M.; Qayyum, A.; Awais, M.; Alnowami, M.; Khan, M.K. Medical image analysis using convolutional neural networks: A review. *J. Med. Syst.* **2018**, *42*, 226. [[CrossRef](#)]
27. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Processing Syst.* **2012**, *25*, 1097–1105. [[CrossRef](#)]
28. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv* **2016**, arXiv:1602.07360.
29. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
30. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 6848–6856.
31. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
32. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
33. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
34. Attallah, O.; Ragab, D.A. Auto-MyIn: Automatic diagnosis of myocardial infarction via multiple GLCMs, CNNs, and SVMs. *Biomed. Signal Process. Control* **2023**, *80*, 104273. [[CrossRef](#)]
35. Attallah, O. ECG-BiCoNet: An ECG-based pipeline for COVID-19 diagnosis using Bi-Layers of deep features integration. *Comput. Biol. Med.* **2022**, *142*, 105210. [[CrossRef](#)]

36. Attallah, O. An Intelligent ECG-Based Tool for Diagnosing COVID-19 via Ensemble Deep Learning Techniques. *Biosensors* **2022**, *12*, 299. [\[CrossRef\]](#)
37. Attallah, O.; Samir, A. A wavelet-based deep learning pipeline for efficient COVID-19 diagnosis via CT slices. *Appl. Soft Comput.* **2022**, *128*, 109401. [\[CrossRef\]](#)
38. Attallah, O. A computer-aided diagnostic framework for coronavirus diagnosis using texture-based radiomics images. *Digit. Health* **2022**, *8*, 20552076221092543. [\[CrossRef\]](#)
39. Attallah, O. Deep Learning-Based CAD System for COVID-19 Diagnosis via Spectral-Temporal Images. In Proceedings of the 12th International Conference on Information Communication and Management, London, UK, 13–15 July 2022; pp. 25–33.
40. Attallah, O. CoMB-Deep: Composite Deep Learning-Based Pipeline for Classifying Childhood Medulloblastoma and Its Classes. *Front. Neuroinform.* **2021**, *15*, 663592. [\[CrossRef\]](#)
41. Attallah, O. MB-AI-His: Histopathological diagnosis of pediatric medulloblastoma and its subtypes via AI. *Diagnostics* **2021**, *11*, 359. [\[CrossRef\]](#)
42. Attallah, O.; Zaghlool, S. AI-Based Pipeline for Classifying Pediatric Medulloblastoma Using Histopathological and Textural Images. *Life* **2022**, *12*, 232. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Attallah, O.; Anwar, F.; Ghanem, N.M.; Ismail, M.A. Histo-CADx: Duo cascaded fusion stages for breast cancer diagnosis from histopathological images. *PeerJ Comput. Sci.* **2021**, *7*, e493. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Ghanem, N.M.; Attallah, O.; Anwar, F.; Ismail, M.A. AUTO-BREAST: A fully automated pipeline for breast cancer diagnosis using AI technology. In *Artificial Intelligence in Cancer Diagnosis and Prognosis, Volume 2: Breast and bladder cancer*; IOP Publishing: Bristol, UK, 2022.
45. Aslan, M.F.; Celik, Y.; Sabancı, K.; Durdu, A. Breast cancer diagnosis by different machine learning methods using blood analysis data. *Int. J. Intell. Syst. Appl. Eng.* **2018**, *6*, 289–293. [\[CrossRef\]](#)
46. Attallah, O. A deep learning-based diagnostic tool for identifying various diseases via facial images. *Digit. Health* **2022**, *8*, 20552076221124432. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Attallah, O. DIAROP: Automated Deep Learning-Based Diagnostic Tool for Retinopathy of Prematurity. *Diagnostics* **2021**, *11*, 2034. [\[CrossRef\]](#)
48. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaria, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [\[CrossRef\]](#)
49. Xu, Q.; Zhang, M.; Gu, Z.; Pan, G. Overfitting remedy by sparsifying regularization on fully-connected layers of CNNs. *Neurocomputing* **2019**, *328*, 69–74. [\[CrossRef\]](#)
50. Anowar, F.; Sadaoui, S.; Selim, B. Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE). *Comput. Sci. Rev.* **2021**, *40*, 100378. [\[CrossRef\]](#)
51. Naz, J.; Sharif, M.; Raza, M.; Shah, J.H.; Yasmin, M.; Kadry, S.; Vimal, S. Recognizing Gastrointestinal Malignancies on WCE and CCE Images by an Ensemble of Deep and Handcrafted Features with Entropy and PCA Based Features Optimization. *Neural Process. Lett.* **2021**. [\[CrossRef\]](#)
52. Garg, I.; Panda, P.; Roy, K. A Low Effort Approach to Structured CNN Design Using PCA. *IEEE Access* **2020**, *8*, 1347–1360. [\[CrossRef\]](#)
53. Talukder, M.A.; Islam, M.M.; Uddin, M.A.; Akhter, A.; Hasan, K.F.; Moni, M.A. Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning. *Expert Syst. Appl.* **2022**, *205*, 117695. [\[CrossRef\]](#)
54. Mangal, S.; Chaurasia, A.; Khajanchi, A. Convolution neural networks for diagnosing colon and lung cancer histopathological images. *arXiv* **2020**, arXiv:2009.03878.
55. Hatuwal, B.K.; Thapa, H.C. Lung cancer detection using convolutional neural network on histopathological images. *Int. J. Comput. Trends Technol* **2020**, *68*, 21–24. [\[CrossRef\]](#)
56. Ali, M.; Ali, R. Multi-input dual-stream capsule network for improved lung and colon cancer classification. *Diagnostics* **2021**, *11*, 1485. [\[CrossRef\]](#)
57. Hasan, I.; Ali, S.; Rahman, H.; Islam, K. Automated Detection and Characterization of Colon Cancer with Deep Convolutional Neural Networks. *J. Healthc. Eng.* **2022**, *2022*, 5269913. [\[CrossRef\]](#)
58. Bukhari, S.U.K.; Syed, A.; Bokhari, S.K.A.; Hussain, S.S.; Armaghan, S.U.; Shah, S.S.H. The histological diagnosis of colonic adenocarcinoma by applying partial self supervised learning. *MedRxiv* **2020**. [\[CrossRef\]](#)
59. Borkowski, A.A.; Bui, M.M.; Thomas, L.B.; Wilson, C.P.; DeLand, L.A.; Mastorides, S.M. Lung and colon cancer histopathological image dataset (lc25000). *arXiv* **2019**, arXiv:1912.12142.
60. Jha, D.; Kim, J.-L.; Choi, M.-R.; Kwon, G.-R. Pathological Brain Detection Using Weiner Filtering, 2D-Discrete Wavelet Transform, Probabilistic PCA, and Random Subspace Ensemble Classifier. *Comput. Intell. Neurosci.* **2017**, *2017*, 4205141. [\[CrossRef\]](#)
61. Sabancı, K.; Aslan, M.F.; Durdu, A. Bread and durum wheat classification using wavelet based image fusion. *J. Sci. Food Agric.* **2020**, *100*, 5577–5585. [\[CrossRef\]](#)
62. Nahiduzzaman, M.; Goni, M.O.F.; Anower, M.S.; Islam, M.R.; Ahsan, M.; Haider, J.; Gurusamy, S.; Hassan, R.; Islam, M.R. A Novel Method for Multivariant Pneumonia Classification Based on Hybrid CNN-PCA Based Feature Extraction Using Extreme Learning Machine with CXR Images. *IEEE Access* **2021**, *9*, 147512–147526. [\[CrossRef\]](#)

63. Azam, F.; Rizvi, A.; Khan, W.Z.; Aalsalem, M.Y.; Yu, H.; Zikria, Y.B. Aircraft Classification Based on PCA and Feature Fusion Techniques in Convolutional Neural Network. *IEEE Access* **2021**, *9*, 161683–161694. [[CrossRef](#)]
64. Samee, N.A.; Alhussan, A.A.; Ghoneim, V.F.; Atteia, G.; Alkanhel, R.; Al-antari, M.A.; Kadah, Y.M. A Hybrid Deep Transfer Learning of CNN-Based LR-PCA for Breast Lesion Diagnosis via Medical Breast Mammograms. *Sensors* **2022**, *22*, 4938. [[CrossRef](#)] [[PubMed](#)]
65. Prabha, K.; Sam, I.S. A novel blind color image watermarking based on Walsh Hadamard Transform. *Multimed. Tools Appl.* **2020**, *79*, 6845–6869. [[CrossRef](#)]
66. Meenakshi, K.; Rao, C.S.; Prasad, K.S. A robust watermarking scheme based Walsh-Hadamard transform and SVD using ZIG ZAG scanning. In Proceedings of the 2014 International Conference on Information Technology, Bhubaneswar, India, 22–24 December 2014; pp. 167–172.
67. Lakshmi, P.G.G.; Domic, S. Walsh–Hadamard Transform Kernel-Based Feature Vector for Shot Boundary Detection. *IEEE Trans. Image Process.* **2014**, *23*, 5187–5197. [[CrossRef](#)] [[PubMed](#)]
68. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 60. [[CrossRef](#)]
69. Guo, L.; Lei, Y.; Xing, S.; Yan, T.; Li, N. Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data. *IEEE Trans. Ind. Electron.* **2018**, *66*, 7316–7325. [[CrossRef](#)]
70. Aslan, M.F. A robust semantic lung segmentation study for CNN-based COVID-19 diagnosis. *Chemom. Intell. Lab. Syst.* **2022**, *231*, 104695. [[CrossRef](#)]