Retinopathy grading with deep learning and wavelet hyper-analytic activations

Raja Chandrasekaran1 · Balaji Loganathan2

Accepted: 12 March 2022 / Published online: 25 April 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
Recent developments reveal the prominence of Diabetic Retinopathy (DR) grading. In the past few decades, Wavelet-based DR classification has shown successful impacts and the Deep Learning models, like Convolutional Neural Networks (CNN’s), have evolved in offering the highest prediction accuracy. In this work, the features of the input image are enhanced with the integration of Multi-Resolution Analysis (MRA) and a CNN framework without costing more convolution filters. The bottleneck with conventional activation functions, used in CNN’s, is the nullification of the feature maps that are negative in value. In this work, a novel Hyper-analytic Wavelet (HW) phase activation function is formulated with unique characteristics for the wavelet sub-bands. Instead of dismissal, the function transforms these negative coefficients that correspond to significant edge feature maps. The hyper-analytic wavelet phase forms the imaginary part of the complex activation. And the hyper-parameter of the activation function is selected such that the corresponding magnitude spectrum produces monotonic and effective activations. The performance of 3 CNN models (1 custom, shallow CNN, ResNet with Soft attention, Alex Net for DR) with spatial–Wavelet quilts is better. With the spatial–Wavelet quilts, the Alex Net for DR has an improvement with an 11% of accuracy level (from 87 to 98%). The highest accuracy level of 98% and the highest Sensitivity of 99% are attained through Modified Alex Net for DR. The proposal also illustrates the visualization of the negative edge preservation with assumed image patches. From this study, the researcher infers that models with spatial–Wavelet quilts, with the hyper-analytic activations, have better generalization ability. And the visualization of heat maps provides evidence of better learning of the feature maps from the wavelet sub-bands.

Keywords Diabetic retinopathy · CNN · Wavelets · Spatial–wavelet inputs · Complex activations

1 Introduction

One primary problem with Diabetes is the increase in the number of patients globally. In 2014, the World Health Organization (WHO) stated that the number of people with diabetes would increase to 300 million by the year 2025 [1]. Heretofore WHO has stated that in the year 2020, about 422 million people worldwide will have diabetes, among them the majority of them will be living in low-and middle-income countries. The numbers have increased drastically, compared to the expectations. A challenging problem that arises in this domain is the development of Diabetic Retinopathy (DR), which constitutes the common cause of blindness [2]. One of the problems is that diabetes damages blood vessels in the retina [3]. And DR is caused by the change of blood vessels [4]. As the retinal anatomy transposes, new attributes of diseases such as micro-aneurysms (MA), exudates, and hemorrhages appear in the retina [3]. It is a medical condition where the retina gets damaged as the fluid leaks from blood vessels into the retina [3]. DR may lead to blindness [5] which is usually because of vascular leakage or ischemia [6]. Unfortunately, the patients with diabetes are in need to undergo DR screening annually as almost 50% of the diabetic patients may develop DR [7]. In general, patients with either type 1 or type 2 diabetes are at endangerment. The cause of this endangerment reveals that the patients with type 1 or type...
2 may face developing neurovascular complications, which can guide diabetic retinopathy [3].

Previous researches suggest that there are five stages of DR. This research focuses on grading these five stages.

**Class 1** corresponds to the normal eye as shown in Fig. 1 (sub-image 1).

**Class 2** corresponds to Mild DR as shown in Fig. 1 (sub-image 2). In the stage of Mild DR, the fundus bulges like a ball. They are called Micro-Aneurysms (MA) in blood vessels. The widened blood vessels resemble balloons [8] and may drip into the eye [3]. To conclude, the fundus has at least one MA, dot, bleeding in all four fundus quadrants [9].

**Class 3** corresponds to Moderate DR as shown in the Fundus Image in Fig. 1 (sub-image 3). During this stage, the blood vessels protract vision of the eyes. This can lead to diabetic macular edema (DME) [3].

**Class 4** corresponds to severe DR as shown in the Fundus Image in Fig. 1 (sub-image 4). More blood vessels in the eye become lumped. Consequently, the retina signals to receive new blood vessels [3].

**Class 5** corresponds to Proliferative DR as shown in the Fundus Image in Fig. 1 (sub-image 5). In the retina, new fragile blood vessels proliferate and they may begin to bleed. As a result, scar tissues may get new shapes, which leads to retinal detachment [3]. The hypoxia leads to the release of vaso-proliferative factors which stimulate new blood vessel formation to provide better oxygenation in the retina [9].

A common strategy on the manual diagnosis of DR is time-consuming and exigencies the availability of Ophthalmologists round the clock, whereas the automation of DR diagnosis enables mass-screenings, which shall be the only fortune to reduce the potential blindness in the future.

For DR diagnosis both spatial and frequency domain features are significant [10], we apply both spatial and wavelet domain inputs to the CNN. The work [11] employs conventional wavelet transform along with other transforms and [12] proposes analytic Wavelet transform for DR grading. Gayathri et al., in [10] extracted Haralick and Complex Wavelet transform (CWT)-based features for DR grading. Considering the advantages of Hyper-analytic wavelet Transform (HWT) over the CWT [13], we propose HW activation functions for the wavelet domain inputs. Several works have used wavelet sub-bands in detecting MA [14], Hemorrhages, Lesions with the objective of DR grading [15].

In automation, the Wavelet-based DR classification plays an important role. There are some growing appeals for the recently evolved Deep Learning models like Convolutional Neural Networks (CNN’s) as they offer the highest prediction accuracy. With the evolution of Deep Learning, the DR grading has attained the highest metrics with CNN’s. In this study, the researcher includes Wavelet sub-bands along with the original images as inputs to the CNN.

Though the Wavelet domain information is a significant representation of the features related to DR, the spatial representation becomes indispensable for CNN. Hence, this study highlights a trendy approach of stacking the Wavelet sub-bands along with the original images. In [16], the original image was stacked with the transformed image for achieving the level of accuracy greater than 99% with the chest X-ray images to diagnose Covid-19 infection. This is the motivation for our work to stack the Wavelet sub-bands along with the original image.

As there is a stack of spatial domain images, wavelet sub-bands (inclusive of LF scaling functions and HF wavelet functions) a need arises to formulate activation function for each “domain” and “sub-bands.” In the frequency domain, both positive and negative coefficients are significant [17]. Varshney et al., in [18], emphasize the negative spectral coefficients and the importance to save them from dismissal by activations. But the bottleneck with the conventional spatial domain activation functions is that they nullify the negative coefficients. For the HF wavelet sub-bands both positive and negative coefficients are significant. Our research is one of the very few attempts addressing negative wavelet coefficients.

For the Detailed coefficient sub-bands (HF wavelet sub-bands), a Hyper-analytic Wavelet (HW) phase Activation function is selectively applied. The phase spectra of HW coefficients (imaginary part of the activation function) preserve the edges (that occur as negative coefficients). We also showcase the dismissal of negative co-efficient by conventional activation functions and the preservation of significant feature maps (that occur as negative coefficients) by the

---

Fig. 1 Diabetic Retinopathy stages (From Left to Right—Class 1 to Class 5)
formulated HW phase activations, with appropriate illustrations.

Further, as the convolution filters need to act separately upon the spatial and Wavelet domain inputs, the authors have incorporated an explicit boundary handling condition.

2 Literature survey

In automation of DR diagnosis, the texture features of the image in the spatial domain are not sufficient for DR classification as they incorporate hidden frequency information. The Wavelet transform uses both spatial and frequency information [10]. In [14], the authors mention that the compounding of the sub-bands of Wavelet transformed images with the lesion template to detect MA achieves better prediction. Pradeep Kumar et al., [12] proposed a flexible analytic Wavelet transform to diagnose DR, with wavelet sub-band grouping methods. For the IDRID dataset, the authors achieved an average accuracy of 97.5 percent for DR classification. DME is an extreme complication from DR. In [11] Rajendra Acharya et al., propose a sequence of transforms—Radon Transform, DWT, DCT for the diagnosis and achieve an accuracy of 97.01 percent with the MESSIDOR dataset. The simplest Haar wavelet is used on top hat transformed image. Approximate sub-bands are taken DCT to extract significant spectral features. The work establishes the significance of wavelet transforms in DR grading. Gayathri et al., in [10] extracted Haralick and Complex Wavelet transform (CWT)-based features for DR grading. The Anisotropic Dual tree CWT extracts directional features from images in the KAGGLE repository. The work strongly emphasizes that the texture features in the spatial domain are insufficient as the diagnostic feature maps exhibit significant spectral/wavelet domain features. In the proposed work we mine both the spatial and wavelet domain features for each image. And considering the advantages of Hyper-analytic wavelet transform (HWT) over the CWT, we propose HW activation functions for the wavelet domain inputs.

Recently Deep Learning models, like CNN’s have evolved offering drastic improvement in prediction accuracy. CNN’s exhibit better diagnosis in clinical images [19], like microcalcification detection in mammograms [20], Glaucoma diagnosis [21], DR diagnosis [22]. In the work [23], conventional AlexNet is employed to grade DR. The CNN has 5 convolution layers, 3 dense layers for feature extraction and classification, respectively. Neovascularization is diagnosed to grade DR, which is the growth of new blood vessels with Proliferative DR [24]. The process is done with green channel extraction, Gaussian filtering, morphological erosion and dilation to segment the blood vessels. And the classification is executed with conventional VGG-16 architecture. The model is executed with 2200 images from the KAGGLE dataset.

The authors in [25] use conventional VGG-16 architecture with a spatial pyramid pooling that pools the feature maps to adapt to the size received by the dense layer of VGG 16. Further, it uses a Network-in-network (NiN) architecture for additional nonlinearity. With all these modifications, the work attains a precision rate of 67 percent with 35, 216 fundus images from the Kaggle repository. Wang Hu Chen et al., in [26] mention the problem of over-fitting in the Deep learning architectures even on large datasets. The work attempts to integrate the learnings of multi-scale shallow CNN’s (each with convolution, pooling and dense layers). A novel learning integration that enhances strong base layers is adopted. With all these modifications, the work attains an accuracy of 92 percent with 35, 126 fundus images from the Kaggle repository. The proposed work on the other hand incorporates MRA in the CNN framework with appropriate activation functions for the wavelet domain inputs. With such modifications, the proposed work attains better accuracy on the same dataset.

There is a wide choice of research works, which attempt to feed the spatial domain inputs. Very few research works attempt to feed Wavelet domain features as a stand-alone to the classifiers. Several face-detection algorithms [27] recommend montage of images stacked together. Combining the spatial and spectral features may dramatically improve to testify the performance of classification [28]. A large number of existing studies reveal that several methods seek to combine spatial and spectral information to generate spatial–spectral features [29].

In this work, the primary motivation for stacking the input images with its Wavelet sub-bands originates from [16], as the authors claim that stacking of the original image with the transformed image contributes to feature efficiency. For the wavelet sub-bands both positive and negative coefficients are significant. The authors take necessary concerns to avoid the nullification of significant negative coefficients of the wavelet domain by formulating an appropriate activation function.

The activation function in a CNN transforms the data from a linear to nonlinear space which is very significant in both machine learning and deep learning. In Machine learning, saturated activation functions like Sigmoidal, tanh functions are used. The new unsaturated variant, the ReLU has been prominent in the CNN’s as per its best outcomes [18]. The voluminous literature applying the same activation function ReLU invariably to all types of data (right from text-based data mining to computer vision) and for all applications indicates a research gap. The role of signal processing researchers is yet to be intensified in the data science applications to formulate the functions rather than a choice-based. In recent years, the formulation of activation functions specific to the data or the application is emerging.
In this context, Sheng Qian et al. in [30] have implemented mixed activation functions, gated activation functions, hierarchical activation functions, adaptive activation functions. The gated combination provides specific combinations of different functions; in adaptive type, the slope of the activation function is allowed to be learned data specific [30].

The characteristics and behavior of an activation function are significant, for instance, parabolic tangent and sigmoid functions are preferred in long short-term memory (LSTM) and other sequential neural networks [18] as per the nature of the application. The activation functions with learnable slope parameters in [18] and [30] are of first-order. Polynomial functions with higher order are also researched and they exhibit curvature [31]. When the coefficients of these higher-order polynomials are higher i.e., near to unity the weights grow extreme [32]. In this work, the Hyper-analytic activation function is not a polynomial rather it involves a complex activation with the Hyper-analytic wavelet phase as the imaginary term.

The Fourier domain activation function can introduce non-linearity into Fourier domain CNN’s [33]. The authors in [17] discuss the importance to construct CNN activation functions in the frequency domain, with a strong emphasis to activate both positive and negative coefficients. Adaptive functional learning is proposed by Varshney et al., in [18], with special emphasis on the negative values. The negative part of PFLU [34] produces sparse outputs for the negative inputs. Our research is one of the very few attempts addressing negative wavelet coefficients and the preservation of significant negative coefficient.

We showcase the dismissal of negative co-efficient by conventional activation functions and the preservation of significant feature maps (that occur as negative coefficients) by the formulated HW phase activations, with appropriate illustrations.

Thus, in this work, we have experimented a step ahead with activation functions in the wavelet domain with unique characteristics for the LF and HF sub-bands. The phase term which forms the imaginary part of a complex activation function preserves the significant edge feature maps that are negative.

The Hyper-analytic Wavelet (HW) phase Activation function is made to selectively act on the Detailed (HF) sub-bands as most of the energy encompassed by the detailed coefficients sub-bands is “Edges.” The main advantage is that the negative co-efficient terms (with a high magnitude which correspond to edges in an image) are transformed instead of dismissal, unlike the conventional Relu function, the frequency domain activation function proposed in [17] and the Wavelet domain activation function proposed in this work for Approximate (LF) sub-bands.

Hyper-analytic Wavelet fosters the directional selectivity by the linear combination of Hilbert transform of different Wavelet sub-bands [35]. In [13], the researchers claim that the HWT optimizes the directional selectivity. Equivalent directional selectivity is possible to accomplish by adding/subtracting the different HWT pairs [35]. In [36], the researchers combine the magnitude and phase spectra for defining texture, discriminating HF edges in an image. Hence, in the activation function proposed for the Detailed Wavelet sub-bands, the Hyper-analytic wavelet phase term (the imaginary part of the complex activation) preserves the edges that occur as negative coefficients. This activation function evolves as non-pointwise suiting the wavelet domain.

Anyhow, the issues that need further investigation are experimentation with different mathematical transforms as inputs to CNN, identifying appropriate activation functions for the same. The organization of the manuscript is as follows. In Sect. 3, the Wavelet sub-bands as input to the CNN models, the spatially alternative activation function for different domains, boundary handling conditions (to enable the convolution filters to act separately for the spatial and Wavelet domain inputs) are explained. In Sect. 4, the results of the CNN’s with spatial—Wavelet domain inputs, the outcomes of the Hyper-analytic Wavelet (HW) phase Activation function are illustrated. In Sect. 5 the discussion on the evaluation metrics, generalization ability of the CNN’s with spatial—Wavelet domain inputs is done and the work is concluded in Sect. 6.

3 Methods

The contribution of the classical feature extraction methods by the Wavelet transforms, cutting edge CNN methods are widely researched and published. The CNN’s offer extremely high metrics for classification. The clinical and AI community seek further attainments by bridging the gaps. Section 3 discusses the database, the stacking of the spatial and Wavelet domain inputs, the CNN models implemented, mathematical formulation and implementation of the hybrid activation functions.

The images are extracted from the Kaggle repository [37]. In the proposed work, the researchers have experimented with 35,126 images with 80% of the images are used for training. The categories are Class 0—No DR with 25,810 images, Class 1—Mild DR with 2443 images, Class 2—Moderate DR with 5292 images, Class 3—Severe DR with 873 images, Class 4—Proliferative DR with 708 images.

The images and the wavelet sub-bands are stacked in the fashion of \([m, n][4 \times \frac{m}{2}, \frac{n}{2}]\) where the right side entities of size \([m/2, n/2]\) correspond the Approximate coefficient, (A) or LF sub-bands; Detailed coefficient or HF sub-bands (Horizontal (H), Vertical (V) and Diagonal (D)), respectively. These stacks are subjected to a convolution layer, Relu layer,
drop out followed by dense layer and softmax activation function. The CNN filters act upon the Wavelet sub-bands also.

Here, the detailed spectral information is helpful to map the distribution of materials of interest and the fine spatial resolution enables the analysis of small spatial structures [38].

3.1 The CNN model

The Convolution operation is executed in the Convolution Layer over the input images with 30 filters. Figure 2 shows that the CNN models (a custom CNN, a ResNet with attention, Alex Net for DR) with multi-resolution inputs, the different layers and hyper-analytic wavelet phase activation function is shown. In the Custom CNN, shown in the first sub-block of Fig. 2, there are 30 filters of size $3 \times 3$ and single striding is done. The purpose of Rectified Linear Unit (RELU) is to transform the features in a linear space to a non-linear space. The product size does not change with respect to the feed-in size. It can be viewed as a threshold function in which the product is zero, if the input is negative.

Under certain calculations, this study requires a spatially alternating activation function that is alternative over the spatial coordinates and acts as a spatial domain activation function for the content in the first half of the quilt and as a Wavelet domain activation function for the other half. As shown by the "spatially alternating activation function" block in Fig. 3, the Wavelet domain random offset Relu ($R_{\text{Psi}}$) is applied to the Approximate Wavelet sub-bands and a Hyper-analytic Wavelet (HW) phase activation function ($R_{\text{Hyper \Psi}}$) is applied to the Detailed Wavelet sub-bands to preserve the edges. The loss function applied is Categorical Cross-Entropy. And Stochastic Gradient Descent optimizer is applied to train the layers with a learning rate of 0.001.

As overfitting is a major drawback of any deep learning network, in the first drop-out layer, 30% of the neurons are dropped out to avoid model-overfitting. The dimension of the Feature Maps has to be reduced for minimizing the computations. In Max-pooling, an even-sized window is moved over the image with stride 2. And at each position, the maximum value is retained, thereby retaining only significant maps. The CNN model in Fig. 2 possesses a drop of 30% of the feature maps. Another Drop-out layer follows, before the Dense layer. The Dense Layer or the Fully Connected layer is the terminal layer that resembles the back-propagation stage of an Artificial Neural Network (ANN). The output is an M-dimensional vector where "M" is the number of classes.

Based on the theories implied by the literature [33] and [17], we extend the idea of frequency domain to wavelet

---

Fig. 2 The CNN models with Multi-Resolution Inputs
domain activation function. Hence in this work, the authors process the Wavelet Domain Random offset Relu and the hyper-analytic wavelet phase activation function as the real and imaginary parts separately. And a hyper-parameter is introduced in the imaginary term to sustain the activation function with the properties of an activation function.

3.1.1 The spatially-alternating activation function

The inputs possess both spatial domain and Wavelet domain representation, with spatial domain content which occupies the first half of the tile. The author has used the popular “Relu” as the spatial domain activation function. And for the sub-bands, a Wavelet domain random offset Relu, is selectively applied on the Approximate (LF) sub-bands and a Hyper-analytic Wavelet (HW) phase activation function is selectively applied on the Detailed (HF) sub-bands. The formulation of activation functions for different “domains” and “sub-bands” is shown in Fig. 3.

Mathematically, the spatially alternating-Activation function is expressed in Eq. (1). Let (x, y) be the pixel location in an image of size (m, n).

The proposed Activation function is expressed as,

\[
\begin{align*}
R_f(x, y) &= r(x, y), \quad \text{if } x \leq \frac{m}{2} \text{ or } y \leq \frac{n}{2} \\
R_\psi A_{k(i,j)}^l &= r(x, y), \quad \text{if } x \leq \frac{m}{2} \text{ or } y \leq \frac{3n}{4} \\
R_{\text{Hyper-}\psi}(D_{k(i,j)}) &= \text{otherwise}
\end{align*}
\]

(1)

In Eq. (1) and Fig. 3, \(r(x, y)\) represents the popular Relu function used for the Spatial domain. Here, \(R_\psi\) represents the Wavelet Domain random offset Relu, which acts as the unsaturated activation function [39] to extract the approximate sub-band Wavelet characteristics by selecting only the positive coefficients. This strategy normalizes the negative coefficient terms reducing the computational complexity in the frequency domain. This is suitable only for the Approximate sub-band and hence it is represented as \(R_\psi(A_{k(i,j)}^l)\) in Fig. 3. \(R_{\text{Hyper-}\psi}\) represents the Hyper-analytic Wavelet phase Random Offset Relu. It preserves the edges that occur as negative values in the detailed sub-bands.

Wavelet domain random offset relu The activation function selects only the positive Fourier coefficient terms alike the frequency domain random offset Relu proposed in [17]. Inspired from the same, the author proposes a Wavelet domain random offset Relu to accomplish MRA in the CNN framework. For the Approximate Coefficient sub-band (LF sub-band), it shall act as the unsaturated activation function to extract the nonlinear Wavelet-domain characteristics by selecting only the positive Wavelet coefficient terms. Let \(R_\psi\left(A_{k(i,j)}^l\right)\) be the output of the Wavelet domain Relu, where "K" is the resolution that is dealt with in the particular stage of the multi-resolution analysis and “l” represents the level of Wavelet transformation. The notations “i,” “j” represent the feature map associated with the i, jth spatial coordinate. And let \((m \times n)\) represent the size of the input feature map.

Before defining the Wavelet domain random offset Relu, the Wavelet domain representation of the Image understudy is given by \(A_k^l(i,j)\) in Eq. (2).

The Wavelet transform of I (x, y) is represented as \(A_k^l(i,j)\)

\[
A^l_k(i,j) = \sum_{i=1}^{m} \sum_{j=1}^{n} h(x)h(y)I(x,y)^{l-1} \frac{1}{2^{(i+2l\cdot y+2j)}}
\]

(2)

where \(h(.)\) represents the filter kernel.

Then, the output of the Wavelet-domain Relu is represented as given in Eq. (3) in which the absolute value of the Wavelet coefficients is added,

\[
R_\psi\left(A_{k(i,j)}^l\right) = \frac{1}{2}\left[A_{k(i,j)}^l + |A_{k(i,j)}^l|\right]
\]

(3)

where \(l. l\) is the absolute value of the wavelet sub-band. This Wavelet-domain Relu is shown as the II output in Fig. 3.

Hyper-analytic wavelet phase activation function In this course of research work, the authors propose an HW phase activation function for the detailed coefficients Wavelet sub-bands.

The Frequency domain Relu in [17] and Wavelet Domain random offset Relu proposed in this work normalize the negative coefficient terms of the output neuron’s characteristic parameters but the negative co-efficient terms correlate with the edge feature maps. The horizontal (H), Vertical (V), Diagonal (D) edges contribute to most of the energy contained in corresponding Wavelet sub-bands. Hence, for the Detailed Wavelet sub-bands, the HW phase activation function is proposed, in which the absolute value of the phase spectra of HW coefficients forms the imaginary part of the complex activation function. Thus, the author good fits to preserve the edges that occur as negative values in the detailed Coefficient sub-bands.

Hyper-analytic wavelet transform and formulation of the activation function Through this study, the authors clearly present their experiment as the directional selectivity is enhanced by the 2D HWT by the linear combination of Hilbert transform of different Wavelet sub-bands [35]. In [13, 36], the researchers claim that magnitude and phase spectra are combined to define texture, discriminative HF edges in an image.
The researchers in [40] state that the implementation of the HWT is to apply the real Wavelet transform of the image f(x,y) (I term in Eq. (4)) and then to apply the Hilbert transform to assess the analytical signal associated with the data (II, III, IV term in Eq. (4)). The output is complex coefficients \( \Psi_{\text{hyper}}(x,y) \) as shown in Eq. (4) and Fig. 3.

\[
\Psi_{\text{hyper}}(x, y) = \Psi(x, y) + iH_x[\Psi(x, y)] + jH_y[\Psi(x, y)] + kH_{xy}[\Psi(x, y)]
\]  

(4)

- In Eq. 4, \( \Psi(\cdot) \) is the real term that stands for the 2D-DWT.
- \( H_x, H_y, H_{xy} \) correspond to the Hilbert-Transform applied to the Wavelet sub-bands along the rows, columns, diagonal direction, respectively.
- Let these imaginary terms \( H_x[\Psi(x, y)] \) be shortly represented as hyper-\( \Psi_y \) and \( H_y[\Psi(x, y)] \) be shortly represented as hyper-\( \Psi \), as they preserve the phase spectra along with the H, V, D direction, respectively, when combined with the real term.
- It is worthy to note that each term of Eq. (4) has 4 sub-bands A, H, V, D.

As the researcher has executed in the research work [13], the phase spectra along with the 3 directions are obtained by the phase difference between the terms as given in Eqs. (5), (6), (7).

\[
\psi(Hyper - \Psi_H) = \Phi(\Psi_H(x, y), hyper - \Psi_{x-H})
\]

(5)

\[
\psi(Hyper - \Psi_V) = \Phi(\Psi_V(x, y), hyper - \Psi_{y-V})
\]

(6)

\[
\psi(Hyper - \Psi_D) = \Phi(\Psi_D(x, y), hyper - \Psi_{xy-D})
\]

(7)

- In Eq. (5), (6), (7), \( \Phi(\cdot) \) represents the phase-spectra;
- \( \Psi_H(\cdot), \Psi_V(\cdot), \Psi_{xy}(\cdot) \) represent the H, V, D sub-bands of the first/real term of Eq. (4), which is the DWT, not taken Hilbert Transform.
- Hyper-\( \Psi_{x-H} \) correspond to the H sub-band of II term of Eq. (4); Hyper-\( \Psi_{y-V} \) correspond to the V sub-band of III term of Eq. (4); Hyper-\( \Psi_{xy-D} \) correspond to the D sub-band of IV term of Eq. (4).
- The output terms \( \Phi(Hyper-\Psi) \) with sub-scripts H, V, D preserve the phase spectra/edge features along the Horizontal, Vertical and Diagonal directions, respectively.
In the HW phase activation function (LHS terms in Eqs. (8), (9), (10)) proposed to be applied to the Detailed Wavelet sub-bands, the absolute value of the phase spectra of HW coefficients is added in the activation function as shown in Eqs. (8), (9), (10). The parameters \( a_1, a_2, a_3 \) offer a choice to trade-off the significance for the phase spectra in H, V, D directions, respectively. And the Hyper-analytic wavelet phase term is represented in the imaginary part of the rectangular form of the Hyper-analytic wavelet activation function.

\[
R (\Psi_H; \emptyset(Hyper - \Psi_H)) = \frac{1}{2} [\Psi_H(x, y) + |\Psi_H(x, y)|] + j \alpha_H |\emptyset(Hyper - \Psi_H)| \\
(8)
\]

\[
R (\Psi_V; \emptyset(Hyper - \Psi_V)) = \frac{1}{2} [\Psi_V(x, y) + |\Psi_V(x, y)|] + j \alpha_V |\emptyset(Hyper - \Psi_V)| \\
(9)
\]

\[
R (\Psi_D; \emptyset(Hyper - \Psi_D)) = \frac{1}{2} [\Psi_D(x, y) + |\Psi_D(x, y)|] + j \alpha_D |\emptyset(Hyper - \Psi_D)| \\
(10)
\]

In Eqs. (8), (9), (10), the first two terms are similar to the Wavelet Domain Random Offset Relu proposed in Sect. 3.1.1.1 and are mentioned as \( R_{\Psi}(H^1_k(i,j)) \) in Fig. 3, which acts as an unsaturated activation function that nullifies negative coefficients. The imaginary term, the absolute value of the phase spectra of HW coefficients offers significant values for the edge feature maps. And then the magnitude spectrum is estimated which forms the absolute activation function. Particularly for the negative coefficient inputs, the first two terms of the activation function nullify to Zero, only the phase spectral value is preserved for the edges.

Properties of the proposed activation function In order to establish the general properties like differentiability, monotonicity of activation functions, the functions in Eqs. (8), (9), (10) are generalized as given in Eq. (11).

\[
f(x, \emptyset(Hyper - x)) = \frac{1}{2} [x + |x|] + j \alpha |\emptyset(Hyper - x)| \\
(11)
\]

where “\( x \)” is considered to be the input.

\[
i.e., \quad f(x, \emptyset(Hyper - x)) = x + j \alpha |\emptyset(Hyper - x)| for positive x \\
= j \alpha |\emptyset(Hyper - x)| for negative x \\
(12)
\]

Differentiating the generalized activation function in Eq. (12)

\[
f'(x, \emptyset(Hyper - x)) = x' + \alpha |\emptyset'(Hyper - x)| \\
\quad for positive x \\
\approx |\emptyset'(Hyper - x)| for negative x \\
(13)
\]

From Eq. (13), it is inferred that the proposed activation function is differentiable for both positive and negative inputs. In Eq. (13), \( x' \) represents the spatial domain differentiation of gray levels which is nonzero. And the term \( \emptyset' \) in Eq. (13) represents the phase derivative.

Regarding the phase derivative, Pei Dang et al., in [41] establish that the phase derivatives are nonzero for complex-valued signals. And Tao Qian et al., in [42] prove that the phase derivative of the inner functions is always nonzero. As it can be inferred from Eqs. (5), (6), (7) that the inner functions are the spatial coordinates, the phase derivatives with respect to these inner functions in Eq. (13) are nonzero.

As this is a non-pointwise activation function, addressing monotonicity is not straightforward. Considering the generalized representation of the function for positive and negative inputs as given in Eq. (12), it is required to analyze the condition as mentioned in Eq. (14) and then to tune the hyperparameter.

\[
|f(x, \emptyset(Hyper - \Psi))| = |x + \alpha |\emptyset(Hyper - \Psi)|| \approx x \\
\quad if \quad \alpha \ll \ll \ll x, \ for \ positive \\
= |\alpha |\emptyset(Hyper - \Psi)| for negative x \\
(14)
\]

As per Eq. (14), monotonicity can be preserved for positive inputs and the phase term can be made less impactful for negative inputs by choosing the value of the hyper-parameter “\( \alpha \)” much lesser than the input “\( x \)”.

Hence, the new activation function proposed in Eqs. (8), (9), (10) are fine-tuned considering the hyper-parameter \( \alpha \) as input \( (x) \) divided “\( K \)” with \( K > 1 \). If the \( K \) value is selected too high, then the activation function reduces to the value of the input, i.e., there is no activation. Tuning of “\( K \)” is done by trading off between effective activation and monotonicity. Thus, the proposed activation function is modified as given in Eq. (15).

\[
R (\Psi; \emptyset(Hyper - \Psi)) = \frac{1}{2} [\Psi(x, y) + |\Psi(x, y)|] + \frac{\psi}{K} |\emptyset(Hyper - \Psi)| \\
(15)
\]
3.2 Boundary handling conditions with the convolution layer

As the convolution filters need to act separately upon the spatial and Wavelet domain inputs, the author has incorporated an explicit boundary handling condition. The author incorporates a switching function that produces nil output when the convolution operation encompasses a region with mixed inputs. The switching function required for this work is simpler as the author deals with boundaries which fixed overlapping conditions (among the image-Wavelet stacks) as shown in equ 16, where (m, 2n) represent the size of the quilt with spatial–wavelet domain stacks.

Hence, the author employs a switching function \( s(m,n) \) given by Eq. (16),

\[
s(m, n) = \begin{cases} 
0 & \text{for } \left\{ \begin{array}{l} \forall m, n \\
\forall m, \forall n \geq n \\
\forall m, 3n/2 \\
\forall m, 3n/2 < n \\
\end{array} \right. \\
1 & \text{else} 
\end{cases}
\]

(16)

And then the convolution operation is modified as given by Eq. (17),

\[
g(x, y) = \sum_m \sum_n s(m, n)w(m, n)f(x + m, y + n) \tag{17}
\]

In Eq. (17), \( g(.) \) represents the convolution layer output; whereas \( w(m, n) \) represents the convolution filter and \( f(x + m, y + n) \) represents the area of the image \( f(x, y) \) encompassed by the convolution filter. The terms \( f(.) \), \( w(.) \) and \( g(.) \) are common for convolution in the spatial domain and the function \( s(m, n) \) is the switching function that produces nil output when the convolution operation encompasses a region with mixed inputs.

4 Results

Metrics of the CNN’s implemented with Spatial Wavelet inputs and novel activations: With the Spatial Wavelet quilts and the novel activation function, better results are achieved compared with the conventional methods. The fact is established with the training and testing metrics.

Figure 4 shows the training accuracy of CNN 1, 2, 3 with Spatial & stacked Wavelet inputs.

Ablation Study—Table 1 shows the ablation study for the test inputs, attained with the conventional model (Spatial domain inputs) versus the proposed model (with Spatial Wavelet inputs and hybrid activations). The CNN 1 is CNN in Fig. 2 with no attention maps or deep layers; CNN 2 is ResNet with Attention Layer in [44]; CNN 3 is Modified AlexNet for DR detection in [23].

- **During the training phase**, the CNN’s with Spatial Wavelet quilts have an improvement in accuracy of 1% (95% to 96%), 14% (83% to 98%), 3% (96% to 99%) for the CNN models 1 to 3, respectively, compared to the accuracy of the same models with Spatial domain inputs. It can be inferred with normal line versus dotted lines from Fig. 4.
- In the CNN model 2 (Res-Net with Attention Layer), the highest improvement of 14% is attained with Spatial Wavelet inputs with hybrid activation functions.
- During the testing phase, the performance of the CNN’s with Spatial Wavelet quilts and novel activations are better compared to the conventional CNN’s.
- The CNN 2 (Res-Net with Attention Layer) with Spatial & Wavelet quilts, with hyper-analytic wavelet phase activation function, have an improvement in accuracy of 14% (79% to 93%). This CNN proposed in [44] consists of, three residual blocks for training and a soft attention layer to select the most significant features, two fully connected layers. The soft-attention layer selects the image regions that are important to categorize. The integration of Wavelet domain inputs with the soft-attention layer (with hyper-analytic wavelet phase activation function) exhibits the better accuracy of 93% in the testing phase.
- The highest accuracy of 98% and the highest Sensitivity of 99% are attained with CNN 3 (Modified Alex-Net for DR). Moreover, the sensitivity of CNN 3 with stacked wavelet inputs, with hyper-analytic wavelet phase activation function, has an improvement by 12% compared to the same models with normal inputs. The integration of Wavelet domain inputs with the Alex-Net exhibits the best typical values of accuracy, sensitivity and highest improvement in sensitivity.
- The Sensitivity influences True Positive and AuROC values with direct proportionality, which is very significant for diagnosing the process of a disease. In the Area under Receiver Operating Characteristics Curve (AuROC), True Positive is plotted versus False Positive. It can be inferred that the RoC graph of the CNN’s 3, as shown in Fig. 5.
- The typical best value of the area under the curve is 0.87 for CNN 3 with stacked Wavelet inputs and novel activations.

As discussed in Table 1, the metrics for all the 3 CNN’s in the testing phase with Spatial & Wavelet are better compared to the same CNN’s with spatial domain inputs alone.
5 Discussions

5.1 Discussion on the performance metrics of conventional and proposed activation function for wavelet domain:

The authors have discussed the importance of a dedicated activation function for the wavelet domain inputs in a CNN in Sects. 1, 2 and 3. In Table 2, the metrics of the test inputs for the CNN’s with 1. Spatial domain inputs, 2. Spatial & Wavelet inputs with conventional and proposed activations are compared. As a primary observation, the metrics of the proposed hyper-analytic wavelet phase activation function are the highest with spatial and wavelet stacks as the input. The performance of the CNN’s with spatial domain input alone is lesser compared to that of the CNN’s with spatial and wavelet stacks. And ultimately the CNN’s with Relu as activation function with wavelet stacks perform less than the spatial domain inputs also. This provides a strong inference for the unsuitability of Relu as an activation function for the wavelet domain. The FD Relu proposed in [17] is better than Relu for wavelet domain inputs but the proposed hyper-analytic wavelet phase activation function provides the best metrics.

5.2 Discussion on the model generalization ability for the CNN’s with spatial & wavelet domain inputs

The author of this study attempts to compare testing accuracy with the training accuracy in order to infer the generalization ability of the CNN models with stacked quilts. The summary obtained by comparing the performance of all the 3 models, with the training (values from Fig. 4) and testing phase (values from Table 1) is given in Table 2.

For all the 3 CNN models with spatial inputs (LHS of Table 2) and also stacked inputs (RHS of Table 2), the testing accuracies are less compared to training accuracies. This is the expected outcome as the model performance reduces with “unseen” inputs. Generally, the extent of reduction in testing accuracy reveals the model over-fitting. For the CNN models 1 to 3, with spatial inputs (LHS of Table 2), the degradation in testing phase accuracy is 14%, 4%, 9%, respectively, i.e., the test phase accuracy is on an average 10% less than the training phase accuracy, which is a clear indication of model over-fitting, i.e., the model learns the limited set of training inputs. On the other hand, for the 3 CNN models with Spatial & Wavelet quilts (RHS of Table 2), the average test phase accuracy is 3% less than the training phase accuracy. Henceforth, there is an indication of model generalization.

5.3 Illustration of the hyper-analytic wavelet phase activation function

It is observed that the metrics of the CNN’s with Spatial & Wavelet domain inputs are higher. To enable simultaneous processing of spatial and Wavelet domain inputs, the author proposes a HW phase Activation function, which is the primary contribution of the work. For the Detailed Wavelet sub-bands, the absolute value of the phase spectra of HW coefficients forms the imaginary part of the complex activation function. This is the actual gap filled by the proposed activation function that, unlike Relu, FD Relu [17] which nullifies the significant negative feature maps, the phase term in the proposed activation function preserves the edges that occur as negative values.

And most of the energy encompassed in Detailed Wavelet sub-bands corresponds to edges feature maps. Here, the author illustrates the avoidance of negative edge nullification with assumed image patches and figures.

In the illustration with assumed image patches, leaky Relu is also considered for comparison.

The Leaky Relu is defined by Eq. (18),

\[
    f(x) = \begin{cases} \ x \text{ for } x \geq 0 \\ K \times x \text{ for } x < 0 \end{cases}
\]

(18)

where “K” is a scalar constant.
Table 1: Ablation Study—Metrics of the test inputs with the CNN’s (with 1. Spatial domain inputs, 2. Spatial & Wavelet inputs with conventional and proposed activations)

| Reference | CNN model | CNN as explained in reference and implemented | No of images | Spatial domain inputs | Spatial and wavelet domain inputs |
|-----------|-----------|-----------------------------------------------|--------------|-----------------------|-----------------------------------|
| A conventional CNN with no attention maps or deep layers | CNN1 – Custom CNN in Fig. 2 | 2 convolution layers, Max-pool and drop-out | 35,126 images of 5 classes from the Kaggle dataset 80%-training 20%-testing | 0.81 | 0.78 | 0.85 | 0.89 |
| [44] CNN 2-ResNet with Attention Layer | 6 learned layers, 4 Convolution layer, 2 Fully Connected Layer, Soft-max layer | | | 0.79 | 0.71 | 0.73 | 0.93 |
| [23] CNN 3-AlexNet for DR | 2 Convolution layers, Pool, Convolution layer, Pool, Convolution layer, Pool, Flatten, FC1, FC2, Output layer | | | 0.87 | 0.75 | 0.87 | 0.98 |

Fig. 5: ROC of the CNN’s 3 (with 1. regular inputs (shown in LHS), 2. Spatial & Wavelet inputs with hyper-analytic activations (shown in RHS))
Table 2 Illustration of model
Over-fitting for CNN’s with spatial inputs and model generalization for CNN’s with stacked inputs

|                      | CNN’s with spatial inputs | CNN’s with stacked Wavelet inputs & Hyper-analytic wavelet phase activations |
|----------------------|---------------------------|--------------------------------------------------------------------------------|
| Training accuracy (%)| Training accuracy (%)     | Testing accuracy (%)                                                               |
|                      | Testing accuracy (%)      |                                                                                   |
| CNN 1                | 95                        | 81                                                                                |
| CNN 2                | 83                        | 79                                                                                |
| CNN 3                | 96                        | 87                                                                                |
|                      | 96                        | 95                                                                                |
|                      | 98                        | 93                                                                                |
|                      | 99                        | 98                                                                                |

Fig. 6 An assumed image Patch and its detailed (H, V, D) Wavelet sub-bands

As it can be inferred from Eq. (13), the Leaky Relu also transforms negative values, it is used to compare and illustrate the significance of the proposed Hyper-analytic Wavelet phase Activation Function.

The top segment of Fig. 6 shows an assumed image patch with assumed H, V, D edges shown in green and “under-lined.” The second to last segment of Fig. 6 shows the H, V, D Wavelet sub-band outputs. It is possible to trace the edges in the Wavelet sub-bands also (shown in green and “under-lined,” with a change from positive to negative coefficient values).

The top-left sub-figure of Fig. 7 shows the Horizontal Wavelet sub-band, in which the negative values that correspond to the edges/significant feature maps are shown in Green and “under-lined”; the negative values that correspond to the “no-edges”/less-significant feature maps are shown in Red within brackets.

- As the Wavelet Domain random offset Relu applied to H sub-bands (II sub-figure in Fig. 7), both the significant and less-significant, negative values are nullified. This is suitable only for approximate sub-band as it is not specific to edge feature maps.
- As the Leaky-Relu is applied to H sub-band (III sub-figure in Fig. 7), the less-significant negative values are also transformed leading to insignificant feature maps.
- It can be inferred by mapping the values within brackets in the I sub-figure of Fig. 7 to the values within brackets in the III sub-figure of Fig. 7.
- As the Hyper-analytic Wavelet Activation function is applied to the H sub-band (IV sub-figure in Fig. 7), the less-significant, negative values are nullified, whereas the significant, negative values that correspond to edges alone are transformed.
- This is offered by the imaginary part of the complex activation function (phase spectra of the HW coefficients), which is significant in value only for the edge feature maps. Thus, the Hyper-analytic Wavelet Activation function can be inferred as a trade-off between the Wavelet Domain random offset Relu proposed for the approximate co-efficient sub-band and the leaky-Relu.
- The same effects can be observed with the application of the activation functions (Wavelet Domain random offset Relu, Leaky Rely and Hyper-analytic Wavelet phase Activation Function) to the vertical and diagonal sub-bands. The vertical sub-band and the output of the activation functions to the vertical sub-band are shown in Fig. 8. The discussion about the activation functions in Fig. 7 holds good for the vertical sub-bands in Fig. 8.

5.4 Qualitative evaluation of the proposed CNN multi-resolution inputs and novel activations

With the evaluation of the quantitative improvement in the metrics in Sect. 4 and establishing the appropriate formulation of the Hyper-analytic Wavelet Activation function in Sect. 5.3, with the exhibition of the feature maps of different activations for assumed edges. Further, the authors showcase
the “learning” of CNN as heat maps by tapping their weights as a qualitative study. DL models are “black boxes” by construction, because the features used for prediction are learned internally and not engineered beforehand. To gain insight into the inner workings of the DL models, the researchers in [43], constructed attribution maps created by means of guided back-propagation.

In [44], the authors designed a deep neural network that combines a ResNet and a soft-attention architecture. The different color patches showcase the network has learned different patches with different importance. The recent works disclose that some of the approaches assist the user to infer the discriminatory regions within an image, as they are vital for the CNN in order to conclude to a certain class [44].

The specific comments on the heat maps presented in the research works are as follows. In [44], as the objective is to detect the refractive error (myopia, hyperopia), the representative examples of heat maps related to the refractive error are shown and discussed. The observation in every image is that the macula is a prominent feature highlighted. In addition, diffuse signals such as retinal vessels and cracks in retinal pigment are also highlighted. Leastwise, there was not an obvious difference in the heat maps for different severities of refractive error.

In [43], the objective is to diagnose Macular thickening to assess DME, the heat maps have signals located inside the macula or close by, focusing on hemorrhages, exudates, and vessel contours. Hemorrhages and MAs are the most commonly detected and associated together with a single label [45].

In this research, the heat maps, nothing but the values of the weights before the final decision layer [46] of the CNN corresponding to images in different levels of severity are shown in Fig. 9. It is mentioned in [45], that the “Red” color corresponds to the regions, which are significant for hypothesizing a class and the “Blue” color corresponds to the regions, which are least-significant for hypothesizing a class.

The ablation study is presented as, in Fig. 9, heat map of the original image (sub-image 2), a Heat map of the H, V, D sub-bands (sub-images 3,4,5) are shown. In all the images of Fig. 9, the Micro-aneurysms, which are “small red patches” in the original image are encircled “thin” and “Exudates” which are “larger and brighter patches” are encircled “thick.”

As we can infer that the heat maps of the original image show “red patches” as evidence of learning the “exudates”
Fig. 9 Class 4 Image (Image 1), Heat map of the original Image (Image 2), Heat map of the H, V, D sub-bands (Image 3,4,5) (Best viewed in color)—Micro-aneurysms are encircled “thin” and “Exudates” are encircled “thick”

(circled thickly at the bottom), anyhow the heat maps of the H, V sub-bands show better discriminatory H and V “red patches” (circled thickly at the bottom). These red patches are induced in the heat maps as evidence of better learning of the feature maps from the H, V wavelet sub-band inputs compared to the original image input.

The micro-aneurysms are better learned from the H, V wavelet sub-band inputs compared to the original image as the input. This is evident from the induced “red feature maps” in the “thin” circles of H, V sub-bands’ heat maps (sub-images 3,4). Compared to this, the induction is almost-nil in the heat map of the original image for these “micro-aneurysms.” And some Exudate features are learned only from the diagonal sub-band as input.

6 Conclusion

The present study confirms the integration of multi-resolution analysis in a CNN framework in Retinal Image feature extraction and classification. The texture features of the image in the spatial domain are not sufficient for DR classification as it contains hidden frequency information. Combining the spatial and Wavelet features improvises the performance of classification, with the hyper-analytic activation function that is formulated to privilege the wavelet inputs with appropriate nonlinear mapping. The novel HW phase activation function avoids nullification of edge features as negative coefficients in the detailed coefficient Wavelet sub-bands, and only the significant, negative feature maps that correspond to edges are transformed. During the testing phase, with the stacked inputs, the AlexNet for DR, has an improvement in accuracy of 11% (87% to 98%). The highest accuracy of 98% and the highest Sensitivity of 99% are attained with Modified Alex Net for DR with spatial wavelet quilts. The heat maps provide evidence of better learning of the feature maps from the wavelet sub-band (inputs) compared to the original image. In fact, some exudate features are learned only from the diagonal sub-bands, which is inferred from the induced heat map. And the test phase accuracy of the models with stacked inputs manages to be as good as the training phase. Hence, there is an indication that the models with spatial wavelet quilts have better generalization ability, avoiding over-fitting, which needs to be investigated analytically.

The author of this work witnesses the following future directions to be explored.

- The hyper-parameter of the novel activation can be optimized for further improvement in the metrics, constrained to the properties of conventional activation functions.
- The stacking approach discloses new avenues for down-sampled, transformed inputs being fed to a CNN, along with the original input, without costing new convolution filters. The model can be extended and experimented with different mathematical transforms, hybrid mathematical transforms like application of DWT to Radon transform sinograms as proposed in [47], conventional and Optimized Wavelets suggested in research works like [48] and [49]. And deriving appropriate activation functions for different transforms shall be a challenging research problem.

Acknowledgements The author with a deep sense of gratitude would thank the supervisor for his guidance and constant support rendered during this research.

Funding The authors received no specific funding for this study.
Declarations

Conflict of interest I am enclosing herewith the electronic version of my research paper entitled “Retinopathy Evaluation Using Deep Learning with Wavelet Hyper-Analytical Function” for favor of publishing. I hereby declared my interest toward publishing in “The Visual Computer” journal.

The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Agrawal, A., Bhatnagar, C., Jalal, A.S.: A survey on automated microaneurysm detection in diabetic retinopathy retinal images. Inf. Syst. Comput. Netw. IEEE. 2, 24–29 (2013)
2. Vashist, P., Singh, S., Gupta, N., Saxena, R.: Role of early screening for diabetic retinopathy in patients with diabetes mellitus: an overview. Indian Community. Med. Publ. Indian. Assoc. Prev. Soc. Med. 36, 247–254 (2011)
3. Diaz, V.: Diabetic retinopathy is blood vessel damage in the retina that happens as a result of diabetes. Medical News Today. https://www.medicalnewstoday.com/articles/183417 (2020)
4. Peterman, S.: The Four Stages of Diabetic Retinopathy. Gris. Worr. Hom. https://www.griswoldhomecare.com (2015).
5. Mateen, M., Jen, H., Hassan, M., Naqshbandi, N., Sun, S., Hayat, S.: Automatic detection of diabetic retinopathy: a review on datasets, methods and evaluation metrics. IEEE Access 8, 48784–48811 (2020)
6. Bloomgarden, Z.T.: Screening for and managing diabetic retinopathy: current approaches. Am. J. Health. Syst. Pharm. 64, 8–14 (2007)
7. Aiello, L.P., Gardner, T.W., King, G.L.: Diabetic retinopathy. Diab Car. 21, 143–156 (1998)
8. Koetting, C.: The four stages of diabetic retinopathy. Modrn. Optomte (2019). https://modernoptom.com (2019).
9. Viswanath, K., Mcgavin, D.D.: Diabetic retinopathy: clinical findings and management. Community Eye. Health 16, 21–24 (2003)
10. Gayathri, S., Krishna, A.K., Gopi, V.P., Palanisamy, P.: Automated binary and multiclass classification of diabetic retinopathy using haralick and multiresolution features. IEEE Access 8, 57497–57504 (2020)
11. Rajendran Acharya, U., Mookiah, M.R.K., et al.: Automated diabetic macular edema (DME) grading system using DWT, DCT Features and maculopathy index. Comput. Biol. Med. 84, 59–68 (2017)
12. P.K. Chaudhary, R. B. Pachori.: Automatic diagnosis of different grades of diabetic retinopathy and diabetic macular Edema Using 2D-FBSE-FAWT. IEEE Trans. Instrum. Meas. (2022)
13. Raja, C., Gangatharan, N.: Incorporating phase information for efficient glaucoma diagnoses through hyper-analytic wavelet transform. Soft Comput. Probl. Solving 2, 329–343 (2015)
14. Muthu, R.K.M., Rajendran, A., Chua, K.C., Choo, M.L.: Computer aided diagnosis of diabetic retinopathy: a review. Comput. Biol. Med. 43, 2136–2155 (2013)
15. Quellec, G., Lamard, M., Josselin, P.M., Cazuguel, G., Cochen, B., Roux, C.: Optimal wavelet transform for the detection of microaneurysms in retina photographs. IEEE. Trans. Med. Image 27, 1230–1241 (2008)
16. Mesut, T., Burhan, E., Zafer, C.: COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches. Comput. Biol. Med. 121, 103805 (2020)
17. Lin, J., Ma, L., Cui, J.: A frequency-domain convolutional neural network architecture based on the frequency-domain randomized offset rectified linear unit and frequency-domain chunk max pooling method. IEEE. Access 8, 98126–981155 (2020)
18. Varshney, M., Singh, P.: Optimizing nonlinear activation function for convolutional neural networks. Signal Image Vidio Process. 7, 1–8 (2021)
19. Gondal, W.M., Kohler, J.M., Grzeszick, R., Fink, G.A., Hirsch, M.: Weakly-supervised localization of diabetic retinopathy lesions in retinal fundus images. IEEE. Image. Proc. 2069–2073 (2017)
20. Xi, P., Guan, H., Shu, C.: An integrated approach for medical abnormality detection using deep patch convolutional neural networks. Vis. Comput. 36, 1869–1882 (2020)
21. Chen, X., Xu, Y., Wong, D.W.K., Wong, T.Y., Liu, J.: Glaucoma detection based on deep convolutional neural network. IEEE Eng. Med. Biol. Soc. 6, 715–718 (2015)
22. Sun, Y.: The neural network of one-dimensional convolution-an example of the diagnosis of diabetic retinopathy. IEEE Access 7, 69657–69666 (2019)
23. Shanti, T., Sabenjian, R.S.: Modified Alexnet architecture for classification of retinopathy images. Comput. Electron. Eng. 76, 56–64 (2019)
24. Saranya, P., Prabakaran, S., et al.: Blood vessel segmentation in retinal fundus images for proliferative diabetic retinopathy screening using deep learning. Vis. Comput. (2021)
25. Khan, Z., et al.: Diabetic retinopathy detection using VGG-NIN a deep learning architecture. IEEE Access 9, 61408–61416 (2021)
26. Chen, W., Yang, B., Li, J., et al.: An approach to detecting diabetic retinopathy based on integrated shallow convolutional neural networks. IEEE Access 8, 178552–178562 (2020)
27. Akoglu, T.: Can an algorithm recognize montage portraits as human faces? IEEE. Signal. Proce. Mag. 28, 160–158 (2010)
28. Shu, L., McIsaac, K., Osinski, G.R.: Hyperspectral image classification with stacking spectral patches and convolutional neural networks. IEEE Trans. Geosci. Remote Sens. 56, 5975–5984 (2018)
29. Ghamisi, P., Dalla, M.M., Benediktsson, J.A.: A survey on spectral–spatial classification techniques based on attribute profiles. IEEE Trans. Geosci. Remote Sens. 53, 2335–2353 (2014)
30. Qian, S., Liu, H., Liu, C., Wu, S., San Wong, H.: Adaptive activation functions in convolutional neural networks. Neurocomputing 272, 204–212 (2018)
31. Obia, S., Gong, X., Aloufi, A., Hu, P., Takabi, D.: Effective activation functions for homomorphic evaluation of deep neural networks. IEEE Access 8, 153098–153112 (2020)
32. Yu, Y., Adu, K., Tashi, N., Anokye, P., Wang, X., Ayidzoe, M.A.: Rmaf: Relu-memristor-like activation function for deep learning. IEEE. Access 8, 72727–72741 (2020)
33. Lin, J., Ma, L., Yao, Y.: A frontier domain training framework for convolutional neural networks based on the fourier domain pyramid pooling method and fourier domain exponential linear unit. IEEE. Access 7, 116612–116631 (2019)
34. Zhu, M., Min, W., Wang, Q., Zou, S., Chen, X.: PFLU and FPFLU: Two novel non-monotonic activation functions in convolutional neural networks. Neurocomputing 429, 110–117 (2021)
35. Adam, L., Naforinta, C., Boucher, J.M., Isar, A.: A Bayesian approach of hyperanalytic wavelet transform based denoising. Int. Signal Proc. 3, 1–6 (2007)
36. Çelik, T., Tjahjadi, T.: Multiscale texture classification and retrieval based on magnitude and phase features of complex wavelet sub-bands. Comput. Electron. Eng. 37, 729–743 (2011)
37. Rosebrock A.: Diabetic Retinopathy. Kaggle. https://www.kaggle.com (2015)
38. Fauvel, M., Tarabalka, Y., Benediktsson, J.A., Chanussot, J., Tilton, J.C.: Advances in spectral–spatial classification of hyperspectral images. Proc. IEEE. 101, 652–675 (2012)
39. Wang, X., Qin, Y., Wang, Y., Xiang, S., Chen, H.: ReLtanh: an activation function with vanishing gradient resistance for SAE-based DNNs and its application to rotating machinery fault diagnosis. Neurocomputing 363, 88–98 (2019)

40. Nafornita, C., Firoiu, I., Boucherb, J.M., Isar, A.: A new watermarking method based on the use of the hyperanalytic Wavelet transform. In: Proc. SPIE. 33, 7000 (2008).

41. Dang, P., Qian, T.: Analytic phase derivatives, all-pass filters and signals of minimum phase. IEEE Trans. Signal Proc. 59, 4708–4718 (2011)

42. Qian, T.: Boundary derivatives of the phases of inner and outer functions and applications. Math. Methods Appl. Sci. 32, 253–263 (2009)

43. Arcadu, F., Benmansour, F., Maunz, A., Willis, J., Haskova, Z., Prunotto, M.: Deep learning algorithm predicts diabetic retinopathy progression in individual patients. NPJ Digital Med. 2, 1–9 (2019)

44. Varadarajan, A.V., Poplin, R., Blumer, K., Angermueller, C., Led- sam, J., Chopra, R., Webster, D.R.: Deep learning for predicting refractive error from retinal fundus images. Int. Ophthal Vis. Sci. 59, 2861–2868 (2018)

45. Grinsven, M.J.V., Ginneken, B.V., Hoyn, C.B., Theelen, T., Sanchez, C.I.: Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images. IEEE Trans. Med. Image 35, 1273–1284 (2016)

46. Kwasigroch, A., Jarzembinski, B., Grochowski, M.: Deep CNN based decision support system for detection and assessing the stage of diabetic retinopathy. Int. Interdisc. Wor. IEEE. 111–116 (2018)

47. Acharya, U.R., Mookiah, M.R.K., Koh, J.E., Tan, J.H., Bhandary, S.V., Rao, A.K., Laude, A.: Automated diabetic macular edema (DME) grading system using DWT, DCT features and maculopathy index. Comput. Biol. Med. 84, 59–68 (2017)

48. Raja, C., Gangatharan, N.: A hybrid swarm algorithm for optimizing glaucoma diagnosis. Comput. Biol. Med. 63, 196–207 (2015)

49. Raja, C., Gangatharan, N.: Optimal hyper analytic wavelet transform for glaucoma detection in fundal retinal images. J. Electron. Eng. Technol. 10, 1899–1909 (2015)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.