Volume 14, Issue 2, Page 240-246, 2025 https://doi.org/10.46810/tdfd.1661900 Research Article

BİNGÖL UNIVERSITY

Türk Doğa ve Fen Dergisi
Turkish Journal of Nature and Science
www.dergipark.gov.tr/tdfd

A Residual Neural Network with a Novel Orthogonal Regularization for Covid-19 Diagnosis using X-ray images

Kazım FIRILDAK[1*], Gaffari ÇELİK[2], Muhammed Fatih TALU[3]

[1] Fırat University, Department of Computer Technology, Elazığ, Türkiye
[2] Ağri Ibrahim Çeçen University, Department of Computer Technology, Ağrı, Türkiye
[3] İnönü University, Department of Computer Science, Malatya, Türkiye
Kazım FIRILDAK ORCID No: 0000-0002-1958-3627
Gaffari ÇELİK ORCID No: 0000-0001-5658-9529
Muhammed Fatih TALU ORCID No: 0000-0003-1166-8404

**Corresponding author: kfirildak@firat.edu.tr*

(Received: 20.03.2025, Accepted: 04.06.2025, Online Publication: 27.06.2025)

Keywords
Deep learning, Residual network, Orthogonal regularization, Covid-19

Abstract: Covid-19 is a viral infection that affects the respiratory tract and causes serious health problems on a global scale. Due to the high contagiousness of the disease, early detection and accurate classification are of great importance. In this study, a novel orthogonal regularization method is proposed to improve the detection accuracy of Covid-19 disease from X-ray images. The proposed regularization method, evaluated using ResNet110, improves the classification accuracy compared to traditional Orthogonal regularization approaches. In the experimental studies, the proposed method is compared with various regularization techniques and the highest classification success rate is achieved by increasing the test accuracy rate to 96.52%. In addition, it is observed that the proposed method optimizes the learning curve of the model, especially in the later stages of the training process, and increasing the test accuracy. In addition, compared to the existing orthogonal regularization methods for Covid-19 detection, the proposed approach improved the test classification performance by approximately 1% in accuracy, F1-score, sensitivity, recall and specificity metrics.

Yeni bir Orthogonal Düzgünleştirme Kullanan Artık Yapay Sinir Ağı ile X-Ray Görüntülerinden Covid-19 Tespiti

Anahtar Kelimeler
Derin öğrenme, Artık ağ, Ortagonal düzgünleştirme, Covid-19

Öz: Covid-19, solunum yollarını etkileyen ve küresel ölçekte ciddi sağlık sorunlarına neden olan viral bir enfeksiyondur. Bulaşıcılığı nedeniyle hastalığın erken teşhis ve doğru sınıflandırılması büyük önem taşımaktadır. Bu çalışmada, X-ışını görüntülerinden Covid-19 hastalığının tespit doğruluğunu artırmak için yeni bir ortogonal düzgünleştirme yöntemi önerilmiştir. ResNet110 ağına uygulanan yöntem, geleneksel ortogonal düzgünleştirme yaklaşımlarına kıyasla sınıflandırma doğruluğunu artırılmaktadır. Deneysel çalışmalarda, önerilen yöntem çeşitli düzgünleştirme teknikleriyle karşılaştırılmış ve test doğruluk oranını %96,52'ye çıkarılarak en yüksek sınıflandırma doğruluğu elde edilmiştir. Önerilen yöntemin özellikle eğitim sürecinin sonraki aşamalarında modelin öğrenme eğrisini optimize ettiği ve test doğruluğunu artırdığı da görülmüştür. Ayrıca, Covid-19 tespiti için mevcut ortogonal düzgünleştirme yöntemleriyle karşılaştırıldığında, önerilen yaklaşım test sınıflandırma performansını doğruluk, F1 puanı, duyarlılık, keskinlik ve özgüllük metriklerinde yaklaşık %1 oranında iyileşme sağlanmıştır.

1. INTRODUCTION

Covid-19 is a new mutated form of coronaviruses, a ribonucleic virus [1]. It, which has affected many countries of the world in recent years, spreads rapidly by human-to-human transmission [2]. According to the report published by the World Health Organization, 776.8 million confirmed cases and more than 7 million deaths have been reported [3]. Patients infected with this new coronavirus have symptoms such as fever, fatigue,

headache, shortness of breath [4]. The rapid spread and progression of the disease increases the importance of early diagnosis. In advanced stages, it can lead to serious complications up to pneumonia in the lungs.
Covid-19 is primarily diagnosed using Reverse Transcription Polymerase Chain Reaction (RT-PCR) test. However, the PCR test has an accuracy rate of approximately 70% in correctly identifying the disease [5]. Therefore, in addition to the RT-PCR test, X-ray and computed tomography (CT) imaging are commonly utilized by physicians [6]. CT is a widely used technique for disease detection; however, obtaining CT images is an expensive procedure. Given the rapid global spread of the disease, particularly in underdeveloped countries, the implementation of a more cost-effective diagnostic method is crucial. Although the sensitivity of detecting COVID-19 from chest X-ray images is below 70%, its affordability makes it a valuable tool in disease diagnosis when further developed and optimized [7,8].

Artificial intelligence applications are widely utilized in the medical field. Narin et al. implemented a hybrid deep learning approach on X-ray images to detect pneumonia caused by the Covid-19 [9]. In their study, COVID-19 pneumonia was identified with an accuracy of 98.0% using the InceptionV3 convolutional neural network, alongside ResNet50. Another study conducted by Wang et al. achieved a sensitivity of 91% in detecting COVID-19 using their proposed deep learning architecture [10]. Similarly, Horry et al. employed CT, X-ray, and ultrasound imaging for COVID-19 detection. They applied VGG19, a deep learning architecture, to these images, achieving a sensitivity of 86% for X-ray images, 100% for ultrasound images, and 84% for CT images [11]. Loey et al. utilized CT images to detect COVID-19. To enhance the performance of the deep learning architecture, the dataset containing a limited number of samples was augmented using a Convolutional Generative Adversarial Network (CGAN). Subsequently, the architectures were trained on both classical and CGAN-based augmented data using AlexNet, VGG16, VGG19, GoogleNet, and ResNet50 networks. Among these, the ResNet50 architecture achieved the highest accuracy of 82% with classical data augmentation [12].

In this study, a novel orthogonal regularization (OR) method is proposed. This method is subsequently applied to the ResNet110 network for the detection of COVID-19. The second section of the study outlines the materials and methods. The third section presents the experimental results and compares the performance of the proposed method with existing approaches. Finally, the fourth section discusses the obtained results and provides suggestions for future research.

2. MATERIAL AND METHOD

2.1. Covid-19 Dataset

The dataset consists of chest X-ray images of pneumonia caused by the COVID-19 virus, which has significantly impacted the world in recent years. It includes a total of three classes: normal, other pneumonia, and COVID-19 pneumonia [13,14]. The dataset contains 1,200 COVID-19, 1,341 normal, and 1,345 other pneumonia X-ray images. A total of 80% of the data was used for training, while the remaining 20% was allocated for testing. Figure 1 shows sample images of the Covid-19 dataset.

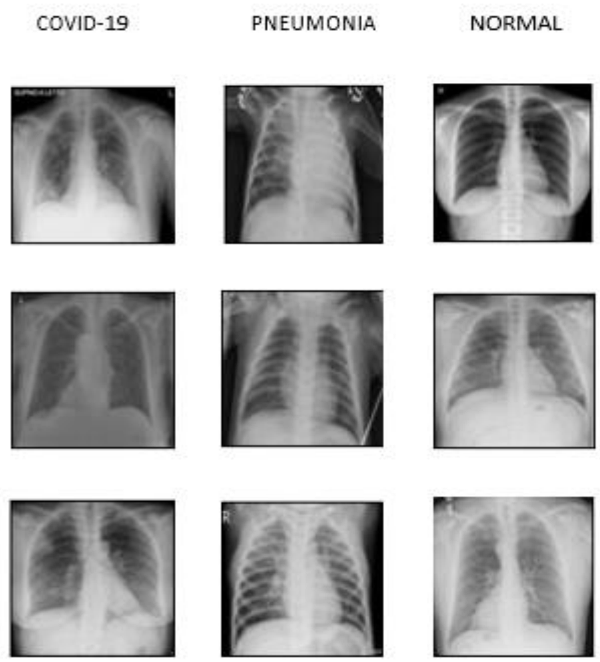

Figure 1. Covid-19 dataset X-ray images.

2.2. Residual Neural Network

Residual Neural Network (ResNet) is an architecture proposed to address the gradient vanishing problem [15]. Residual layers (residual blocks) enhance training efficiency by utilizing the output of a previous layer as an input to subsequent layers. The architecture of a residual layer is illustrated in Figure 2(a). The literature contains ResNet architectures with varying depths and widths. In this study, ResNet110 (Figure 2(b)) is used. This architecture incorporates bottleneck residual layers, and its depth is given by the formula p = 9n + 2, where n represents the total number of convolutions and p denotes the total depth. Thus, ResNet110 consists of 1.7 million parameters.

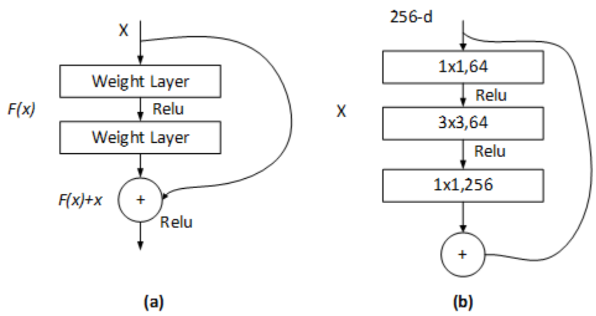

Figure 2. ResNet architectures. (a) Basic residual block, (b) Residual block used for ResNet110.

2.3. Orthogonal Regularization

Regularization is one of the key elements of deep learning, a subfield of machine learning. Regularization is usually applied as a penalty function added to the loss function [16]. The penalty function is defined as a method that aims to reduce the test error in the learning model but does not reduce the training error [17]. This definition is a restrictive statement for deep architectures. Because methods such as weight both training and test error [18]. As a result, all techniques used for better generalization and test accuracy of the neural network are called regularization.

One of the most significant challenges in artificial neural networks (ANN) is the vanishing or exploding gradients as the depth of the network increases. To address this issue, it has been proposed that the parameter matrix should approximate the Gram matrix [19]. While the stability of forward propagation is achieved through batch normalization, the uniform distribution of the error cannot be ensured in the backpropagation process [20]. An alternative approach to address this issue is the OR method [21–23]. OR methods in the literature have been proposed as an alternative to the classical weight decay regularization approach.

Orthogonality is defined as x, y being two vectors, $x, y \in R^n$ for $x \cdot y = 0 \rightarrow x \perp y$. Unit orthogonality (orthonormality) forces the vector norms to be equal to one: $\|x\| = 1 \quad and\ \|y\| = 1$. After this a priori information, it is necessary to mention how orthogonality is used in the ANN training process. In the feed-forward algorithm of ANN, kth layer output vector x is transformed $y = W^T x$ while moving to the $(k + 1)th$ layer input.

where W is called the linear transformation matrix (Equation 1). The condition that the norms of x and y are equal to each other during this transformation is called Norm-Preservation and is shown in Equations 1 and 2.

$$W = \begin{bmatrix} | & | & | \\ w_1 & \dots & w_n \\ | & | & | \end{bmatrix}_{mxn} \tag{1}$$

$$\|y\| = \sqrt{y^T y} = \sqrt{x^T W W^T x} = \sqrt{x^T x} = \|x\|,\ \ eğer\ WW^T = I \tag{2}$$

Orthogonal vectors are needed to preserve vector norm values. Therefore, the distance of the WW^T result to the unit matrix is obtained as the cost value. Bansal et al. proposed to classify four different categories of methods for regularization this cost value [21]. These are Soft Orthogonal (SO), Double Soft Orthogonal (DSO), Mutual Coherence (MC) and Spectral Restricted Isometry Property (SRIP) orthogonal regularization [20,24–27].

SO, the column vectors of the matrix W are required to be perpendicular to each other and have unit length. Accordingly, the regularization cost (R) is calculated by multiplying the distance of the result $W^T W \in R^{nxn}$ from the unit matrix $[I]_{nxn}$by a coefficient λ .The mathematical form of the SO method is given in Equation 3.

$$D(W) = \lambda \|W^T W - I\|_F^2 \tag{3}$$

The cost gradient is calculated as $4\lambda W(W^T W - I)$ and is used in the back propagation algorithm to update the parameters. W a matrix with rows m and columns n. The rank of the matrix is m if it is greater than or equal to m, n. This situation is called under complete matrix. In such cases, an orthogonality relation can be established. However, if it is greater than or equal to n, m, even if the rank of the matrix is m, this is called an over complete matrix. $W^T W \in R^{nxn}$ may not be identified from these matrices. To overcome these shortcomings, approaches have been developed that divide the weight matrix W into subspaces such as Stiefel manifold or Jakobi. These approaches reduce the columns of the over complete W matrix to lower dimensional subspaces, making the matrix easier to process and analyze [24].

DSO, the column vectors of the matrix W are required to be perpendicular to each other in two different vector spaces ($W^T W \in R^{nxn}$ ve $WW^T \in R^{mxm}$) and to be of unit length [21]. Accordingly, the cost function is defined as follows:

$$D(W) = (\|W^T W - I\|_F^2 + \|WW^T - I\|_F^2) \tag{4}$$

where W is weight matrix and has m rows and n columns. m is greater than n, the regularization loss is calculated according to the formula $\lambda\|W^T W - I\|_F^2$, n is greater than or equal to m, the formula $\lambda\|WW^T - I\|_F^2$ is used.

Another OR method is MC. W the MC value between the column vectors of the parameter matrix is calculated as shown in Equation 5 [26].

$$\mu_W = max_{i \neq j} = \frac{|\langle w_i, w_j \rangle|}{\|w_i\| + \|w_j\|} \tag{5}$$

For the MC method, w_i is the column vector ith of W. μ_W, is seen that in the range [0,1] and in the case of ortagonality, it approaches 0, and in other cases it approaches 1. The use of L_∞ is preferred because it is the vector element with the highest absolute value and plays the biggest role in increasing the consistency value [26]:

$$D(W) = \lambda \|W^t W - I\|_\infty \tag{6}$$

For MC L_∞ returns the largest value in the vector elements.

Regularization methods developed using spectral restricted isometry property (SRIP) give better results in statistical metrics and execution time than other methods[27,28].The regularization cost function used in this approach is as shown in Equation 7.

$$D(W) = \lambda \cdot \sigma(W^T W - I) \tag{7}$$

where λ is the regularization coefficient. $\sigma(W^T W - I)$ function returns the spectral norm of the $W^T W - I$ matrix and is calculated as shown in Equation 8.

$$
\begin{aligned}
u &\leftarrow (W^t W - I)v \\
v &\leftarrow (W^t W - I)u \\
\sigma(W^t W - I) &\leftarrow \frac{\|v\|}{\|u\|}
\end{aligned}
\tag{7}
$$

where v is a vector starting with random values in R^n space. Then the vector u and again v are computed iteratively. The spectral norm is obtained by the ratio of both vector norms.

2.4. Proposed Orthogonal Regularization Method

Orthogonal regularization approaches generally aim to approximate the weight matrix W as a Gram matrix. However, this approach weakens the regularization effect in overcomplete cases and negatively impacts the performance of the network. Furthermore, enforcing all weights to be orthogonal vectors hinders the model's convergence towards an optimal learning curve [27]. The proposed OR approach is based on enforcing column vectors to be binary while modulating orthogonality transitions.

In this context, Figure 3 presents the parameter images (filter/mask images) in the layers of three different CNN architectures (AlexNet [29], VGG16 [30], and ResNet50 [15]) trained on the ImageNet dataset. A careful analysis of the figure reveals that numerous binary images within the layers are nearly orthogonal to each other. This observation supports the hypothesis that the training process inherently seeks orthogonal pairs of binary images [31].

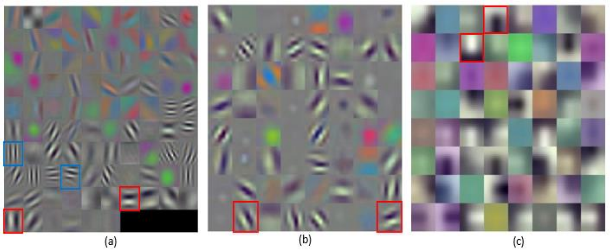

Figure 3. Hidden layer weight visualization (a) AlexNet,(b) ResNet50, (c) VGG16.

While classical regularization approaches force all parameter vectors in the layer to be orthogonal to each other by $W^T W - I$ operation, in the proposed approach, only binary vector pairs are forced to be orthogonal. The cost function of the proposed regularization approach is given below:

$$
D(W) = \lambda \sum ({w_i}^T w_{i+1} - 1), \quad i \in \{1, 3, 5 \dots\}
\tag{9}
$$

where λ is the regularization rate and w_i is the column vector i of W. The total loss function observed in the training/testing activities of the datasets is calculated as shown in Equation 10.

$$
H(W) = \alpha * K\,(W) + (1 - \alpha) * D(W)
\tag{10}
$$

K is the loss function and the cross entropy error value. D is the regularization loss. In the experiments D cost varies and the effect of regularization is analyzed. $H(W)$ is the total loss function. The total loss function of the proposed algorithm contains loop and condition expressions. This process forces the weights of the network to perform OR in binary layers, not in general. The main reason for this is to avoid the overcomplete situation. For these reasons, the derivative of the proposed method and the loss function are calculated by automatic derivative methods [32].

2.5. Performance Evaluation Metrics

In this study, the metrics Acc, Pre, Recall, F1-score and Spe and the confusion matrix are used to measure the experimental performance of the proposed models. In the confusion matrix, TP and TN values indicate the number of correctly classified samples, FP and FN indicate the number of incorrect predictions of the model. These metrics are given mathematically below [33,34].

$$Acc = (TP + TN)/(TP + TN + FP + FN) * 100 \tag{11}$$
$$Pre = TP/(TP + FP) \tag{12}$$
$$Recall = TP/(TP + FN) \tag{13}$$
$$F1 - score = (2 * Pre * Recall)/(Pre + Recall) \tag{14}$$
$$Spe = TN/(TN + FP) * 100 \tag{15}$$

where Acc, Pre, Recall, Spe and F1-score are derived from the confusion matrix. Acc is the ratio of the number of correctly predicted images for each class to the total number of images. Pre and Recall are the precision and sensitivity values of class detection, respectively. The higher these values are, the better the images belonging to the class are detected. F1-score is the harmonic mean of Pre and Recall.

The confusion matrix gives information about the actual and predicted classes in a classifier. The class performance of a model is evaluated using the matrix values in Figure 4 [33,35].

		Predicted Class	
		Positive Class	Negative Class
Actual Class	Positive Class	True Positive (TP)	False Negative (FN)
	Negative Class	False Positive (FP)	True Negative (TN)

Figure 4. Confusion matrix.

3. EXPERIMENTAL RESULTS

In this section, the effect of regularization on training and testing using the COVID-19 dataset with the ResNet110 architecture is analyzed. The training parameters utilized for ResNet110 are presented in Table 1. The Adam optimizer is employed, with the learning rate set to $10^{-2}$and the number of epochs set to 200. The obtained results are presented in Table 2.

Additionally, Figure 5 presents the accuracy curve of the proposed approach (ResNet110+OR) and the ResNet110 architecture during training.

Table 1. Training parameters of ResNet110

Parameters	ResNet110
Input layer	32x32x3
Intermediate layer	12 residual blocks and fully connected classifier
Output layer	10
Activation	ReLU
Optimization	Adam
Package size	128
Epoch	200
Learning rate	0.01

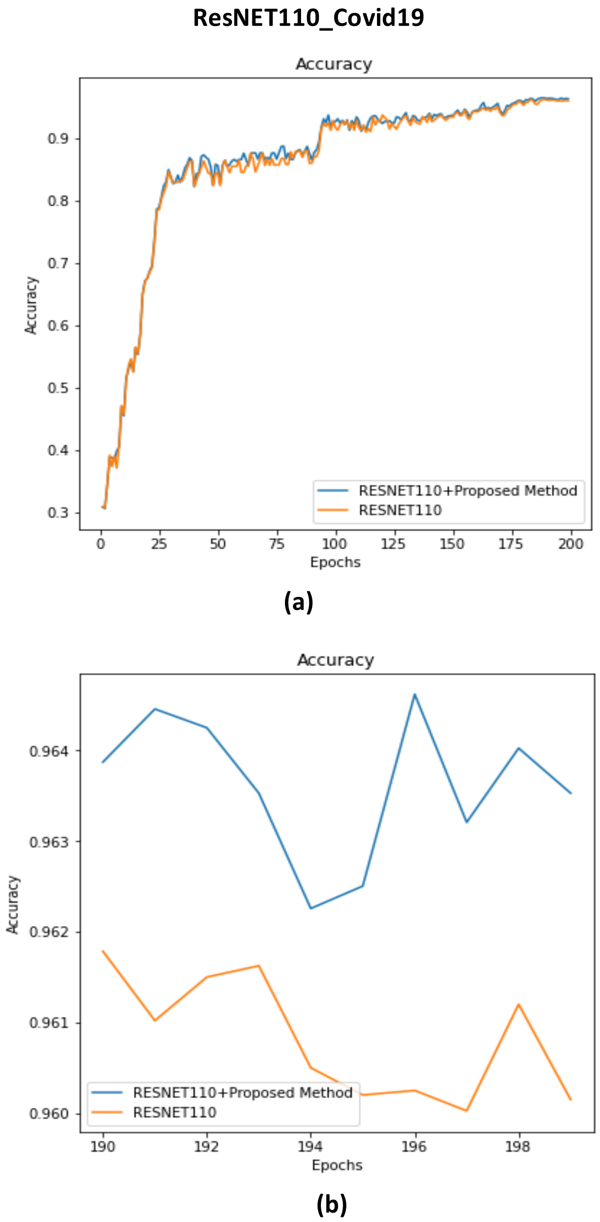

Figure 5. ResNet110 accuracy/epoch for Covid-19 dataset.

In Table 2, when the proposed OR method is compared with orthogonal regularization methods in the literature, the results indicate that the proposed method demonstrates an improvement in statistical metrics. While the accuracy of ResNet110 (a model without regularization) is reported as 96.01%, this increases to 96.54% with the proposed OR method. Precision (Pre) and Recall are measured at 96.52% and 96.51%, respectively, indicating that the proposed approach achieves better class discrimination compared to other methods. Furthermore, a specificity (Spe) of 98.27% was achieved with the proposed approach, demonstrating improved performance in reducing false positives.

Table 2. Test results obtained with ResNet110 architecture using different OR techniques

Method	Dog (%)	Pre (%)	Recall (%)	Spe (%)	F1 (%)
ResNet110	96.01	95.98	95.97	98.01	95.97
ResNet110+SO	96.24	96.25	96.22	98.13	96.24
ResNet110+DSO	96.14	96.12	96.09	98.07	96.10
ResNet110+MC	96.24	96.25	96.22	98.13	96.24
ResNet110+SRIP	96.41	96.38	96.36	98.20	96.37
Proposed OR	**96.54**	**96.52**	**96.51**	**98.27**	**96.51**

The proposed OR method provides a significant improvement in both classification accuracy and other key performance metrics. Specifically, the test accuracy improves by approximately 1%.

As shown in Figure 5, the effect of regularization is limited during the initial 25 iterations. However, after the 50th iteration, the effect of regularization becomes more pronounced in the model's test performance. These findings confirm that the proposed approach enhances classification accuracy by approximately 1%.

4. DISCUSSION AND CONCLUSION

In this study, the effectiveness of a new orthogonal regularization method developed for COVID-19 detection is investigated. The proposed OR method, which is added to the loss function of the ResNet110 architecture, aims to optimize the hidden layer weights as binary orthogonal vectors. Experimental results indicate that the regularization effect is limited during the first 25 iterations of training, but a significant performance gain is observed after the 50th iteration. In addition, the proposed method achieves 96.52% classification accuracy in detecting COVID-19 from X-ray images, demonstrating superior performance compared to orthogonal regularization approaches such as SO, DSO, MC, and SRIP. It is shown that the unified regularization approach improves learning efficiency by enhancing the classification performance of neural networks and can be utilized in detecting critical diseases such as COVID-19.

In future studies, the applicability of the method for detecting other diseases will be investigated by testing it on different deep learning architectures and larger datasets. Furthermore, by evaluating its performance in real-time applications within clinical environments, this study aims to contribute to the broader and more effective adoption of AI-driven medical imaging technologies for disease detection.

REFERENCES

[1] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med [Internet]. 2020 Feb 20;382(8):727–33. Available from: http://www.nejm.org/doi/10.1056/NEJMoa2001017

[2] Lewis D. COVID-19 rarely spreads through surfaces. So why are we still deep cleaning? Nature [Internet]. 2021 Feb 4;590(7844):26–8. Available from: https://www.nature.com/articles/d41586-021-00251-4.

[3] WHO. COVID-19 Epidemiological Update [Internet]. 2024. Available from: https://www.who.int/publications/m/item/covid-19-epidemiological-update---24-december-2024.

[4] Tuncer T, Ozyurt F, Dogan S, Subasi A. A novel Covid-19 and pneumonia classification method based on F-transform. Chemom Intell Lab Syst [Internet]. 2021 Mar;210:104256. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0169743921000241.

[5] Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. Radiology [Internet]. 2020 Aug;296(2):E115–7. Available from: http://pubs.rsna.org/doi/10.1148/radiol.2020200432

[6] Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for Typical Coronavirus Disease 2019 (COVID-19) Pneumonia: Relationship to Negative RT-PCR Testing. Radiology [Internet]. 2020 Aug;296(2):E41–5. Available from: http://pubs.rsna.org/doi/10.1148/radiol.2020200343

[7] Quan H, Xu X, Zheng T, Li Z, Zhao M, Cui X. DenseCapsNet: Detection of COVID-19 from X-ray images using a capsule neural network. Comput Biol Med [Internet]. 2021 Jun;133:104399. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0010482521001931

[8] Wong HYF, Lam HYS, Fong AHT, Leung ST, Chin TWY, Lo CSY, et al. Frequency and Distribution of Chest Radiographic Findings in Patients Positive for COVID-19. Radiology [Internet]. 2020 Aug;296(2):E72–8. Available from: http://pubs.rsna.org/doi/10.1148/radiol.2020201160

[9] Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. Pattern Anal Appl [Internet]. 2021 Aug 9;24(3):1207–20. Available from: https://link.springer.com/10.1007/s10044-021-00984-y

[10] Wang L, Wong A. COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images. 2020 Mar 22; Available from: http://arxiv.org/abs/2003.09871

[11] Horry MJ, Chakraborty S, Paul M, Ulhaq A, Pradhan B, Saha M, et al. COVID-19 Detection Through Transfer Learning Using Multimodal Imaging Data. IEEE Access [Internet]. 2020;8:149808–24. Available from: https://ieeexplore.ieee.org/document/9167243/

[12] Loey M, Manogaran G, Khalifa NEM. A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images. Neural Comput Appl [Internet]. 2020 Oct 26; Available from: https://link.springer.com/10.1007/s00521-020-05437-x

[13] Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub Z Bin, et al. Can AI Help in Screening Viral and COVID-19 Pneumonia? IEEE Access [Internet]. 2020;8:132665–76. Available from: https://ieeexplore.ieee.org/document/9144185/

[14] Rahman T, Khandakar A, Qiblawey Y, Tahir A, Kiranyaz S, Abul Kashem S Bin, et al. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. Comput Biol Med [Internet]. 2021 May;132:104319. Available from: https://linkinghub.elsevier.com/retrieve/pii/S001048252100113X

[15] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. IEEE; 2016. p. 770–8. Available from: http://ieeexplore.ieee.org/document/7780459/

[16] Bishop CM. Neural networks for pattern recognition. Oxford; 1995.

[17] Goodfellow IJ, Bengio Y, Courville A. Deep Learning. Cambridge, MA, USA: MIT Press; 2016.

[18] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM [Internet]. 2017 May 24;60(6):84–90. Available from: https://dl.acm.org/doi/10.1145/3065386

[19] Szu H, Scheff K. Gram-Schmidt Orthogonalization Neural Nets for. 2018.

[20] Zhang L, Li D, Guo Q. Deep Learning From Spatio-Temporal Data Using Orthogonal Regularizaion Residual CNN for Air Prediction. IEEE Access [Internet]. 2020;8:66037–47. Available from: https://ieeexplore.ieee.org/document/9056826/

[21] Bansal N, Chen X, Wang Z. Can We Gain More from Orthogonality Regularizations in Training Deep CNNs? 2018 Oct 22; Available from: http://arxiv.org/abs/1810.09102

[22] Xie D, Xiong J, Pu S. All You Need is Beyond a Good Init: Exploring Better Solution for Training Extremely Deep Convolutional Neural Networks with Orthonormality and Modulation. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. IEEE; 2017. p. 5075–84. Available from: http://ieeexplore.ieee.org/document/8100022/

[23] Rodríguez P, Gonzàlez J, Cucurull G, Gonfaus JM, Roca X. Regularizing CNNs with Locally Constrained Decorrelations. 2016 Nov 7; Available from: http://arxiv.org/abs/1611.01967

[24] Huang L, Liu X, Lang B, Yu A, Wang Y, Li B. Orthogonal weight normalization: Solution to

optimization over multiple dependent stiefel manifolds in deep neural networks. In: Proceedings of the AAAI Conference on Artificial Intelligence. 2018.
[25] Donoho DL. Compressed sensing. IEEE Trans Inf Theory [Internet]. 2006 Apr;52(4):1289–306. Available from: http://ieeexplore.ieee.org/document/1614066/
[26] Lu C, Li H, Lin Z. Optimized projections for compressed sensing via direct mutual coherence minimization. Signal Processing [Internet]. 2018 Oct;151:45–55. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0165168418301464
[27] Zhang Z, Ma W, Wu Y, Wang G. Self-Orthogonality Module: A Network Architecture Plug-in for Learning Orthogonal Filters. In: 2020 IEEE Winter Conference on Applications of Computer Vision (WACV) [Internet]. IEEE; 2020. p. 1044–8. Available from: https://ieeexplore.ieee.org/document/9093466/
[28] Zhang L, Li D, Guo Q. Deep Learning From Spatio-Temporal Data Using Orthogonal Regularizaion Residual CNN for Air Prediction. IEEE Access. 2020;8:66037–47.
[29] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;60(6):84–90.
[30] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations (ICLR 2015). Computational and Biological Learning Society; 2015. p. 1–14.
[31] Fırıldak K, Çelik G, Talu MF. Derin Ağlar İçin Yeni Bir Birimdik Düzgünleştirme Yaklaşımı. Adıyaman Üniversitesi Mühendislik Bilim Derg [Internet]. 2024 Apr 30;11(22):18–34. Available from: http://dergipark.org.tr/tr/doi/10.54365/adyumbd.1390894
[32] Griewank A, Walther A. Evaluating derivatives: principles and techniques of algorithmic differentiation. SIAM; 2008.
[33] Powers DMW. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. 2020 Oct 10; Available from: http://arxiv.org/abs/2010.16061
[34] Celik G. Detection of Covid-19 and other pneumonia cases from CT and X-ray chest images using deep learning based on feature reuse residual block and depthwise dilated convolutions neural network. Appl Soft Comput [Internet]. 2023 Jan;133:109906. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1568494622009553
[35] Dodia S, B. A, Mahesh PA. Recent advancements in deep learning based lung cancer detection: A systematic review. Eng Appl Artif Intell [Internet]. 2022 Nov;116:105490. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0952197622004808