INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY: APPLIED BUSINESS AND EDUCATION RESEARCH
2022, Vol. 3, No. 9, 1703 – 1716
http://dx.doi.org/10.11594/ijmaber.02.09.11

Research Article

Fibonacci Numbers as Hyperparameters for Image Dimension of a Convolutional Neural Network Image Prognosis Classification Model of COVID X-ray Images

Edwin M. Torralba*

Science Department Academy of Our Lady of Guam, 96910, USA Physics and Computer High School Teacher

Article history:
Submission September 2022
Revised September 2022
Accepted September 2022

**Corresponding author:*
E-mail:
edwin.torralba@aolg.edu.gu
edwinmtorralba@gmail.com

ABSTRACT

In recent years, convolutional neural networks (CNNs) have achieved amazing success in a variety of image categorization tasks. However, the architecture of CNNs has a significant impact on their performance. The designs of the most cutting-edge CNNs are frequently hand-crafted by experts in both CNNs and the topics under investigation. As a result, it's tough for users who don't have a lot of experience with CNNs to come up with the best CNN architecture for their individual image categorization challenges. This work investigates the application of the Fibonacci numbers to efficiently solve picture classification challenges by utilizing the hyperparameter of image dimension of COVID and non-COVID x-ray images. The suggested algorithm's greatest strength is the development of a CNN model that can be utilized for COVID viral prognosis using x-ray images to supplement existing COVID pandemic testing techniques. The proposed approach is tested using the metrics of training time, accuracy, precision, recall, and F1-score on commonly used benchmark image classification datasets. According to the experimental data, the CNN model with an image dimension of 55 x 55 surpasses the other CNN models in terms of training time, accuracy, recall, and F1-score. Several issues were raised about how to choose the best CNN models for prognostic picture categorization.

***Keywords**: convolutional neural network, COVID x-ray image classification model, image dimension hyperparameters, Fibonacci numbers,machine learning*

Background

Big data in the medical and healthcare industry possesses enormous potential and has the ability to substantially benefit hospitals and science in general. Indeed, healthcare data analytics has the ability to cut treatment costs, forecast epidemic breakouts, prevent preventable diseases, and generally improve people's

How to cite:
Torralba, E. M. (2022). Fibonacci Numbers as Hyperparameters for Image Dimension of a Convolutional Neural Network Image Prognosis Classification Model of COVID X-ray Images. *International Journal of Multidisciplinary: Applied Business and Education Research*. *3*(9), 1703 – 1716. doi: 10.11594/ijmaber.03.09.11

quality of life. The average human lifetime is growing globally, posing significant hurdles for current therapy delivery technologies. Health experts can gather a lot of data and utilize it for their decision-making.

In a clinical setting, a huge amount of data is generated in medical checkups (Seki et al., 2021), medical laboratory tests (Maissen-haelter et al., 2018), and medical transcription activities (Shafique et al., 2019). Data can be used to create customized care plans, thereby increasing patient satisfaction with medical treatment and care quality (Islam et al., 2018). Data is frequently separated into two categories: structured data and unstructured data. As the name implies, structured data is information that can be recorded and shown in a consistent and ordered fashion. Validation of this sort of data against anticipated plausible parameters is straightforward, as is analysis and interpretation. This category of healthcare data includes data that has been coded using a standardized coding system such as ICD-CM, LOINC, or SNOMED. Additionally, structured data may include numerical values such as height, weight, blood pressure, and pulse rate, as well as categorical or ordinal values such as blood type or disease diagnosis stages. As compared with structured data, unstructured data is frequently in the form of free text, voice narratives, and image data, which most analytics software is incapable of collecting and analyzing numerically to draw relevant insights. Analyzing and interpreting unstructured data is far more challenging than structured data. Free texts are not as easily classified as structured numerical data points. For instance, a blood pressure measurement is denoted by a little number. Clinical information, on the other hand, such as patient symptoms during the doctor's visit, is frequently documented as unstructured text. A doctor's note that contains medical symptoms and prescriptions may need to be read by someone because of the possible spelling mistakes, and abbreviations.

The COVID Pandemic

During the third week of April 2020, the World Health Organization received over 2.6 million confirmed cases across 213 countries and regions (World Health Organization, 2020). WHO classified the COVID-19 epidemic as a public health emergency of international concern on January 30, 2020, succeeding Ebola in 2019, Zika in 2016, Polio in 2014, and H1N1 in 2009 (Shrivastava, 2020). The virus's rapid global growth and increasing death toll have sparked major concerns about the virus's global proliferation. Due to the rapid increase of confirmed cases, the World Health Organization (WHO) declared the global COVID-19 outbreak a pandemic on March 11, 2020 (World Health Organization, 2020). COVID-19 can be transmitted from person to person and animal to animal, and infection can arise as a result of exposure to symptomatic or asymptomatic people. The pandemic has disrupted the operations of healthcare institutions, especially in attending to the medical needs of patients (Bernacki et al., 2021). Problems involve reduced in-person medical consultations (Baum et al., 2021), postponement in diagnosis and treatment (De Luca et al., 2022), cancellation of dental appointments (Long & Corsar, 2020), increased online consultations (Iyengar et al., 2020), and modifications to medications or treatments (Kaye et al., 2020), but they go beyond these new normal medical practices. Health policymakers and medical professionals were also faced with an unusual and ethical dilemma in the early stages of the COVID pandemic (Menon & Padhy, 2020). They didn't know how to treat COVID patients without a known cure or treatment. Early detection through rapid testing, case tracking, and isolation were the best strategies implemented during the early phases of the pandemic to reduce transmission and severe COVID cases (Hashmi & Asif, 2020; Kang et al., 2020).

One of the challenges that countries have to face as they open their economies even after the development of COVID vaccines is the cost of testing. With numerous government institutions, business establishments, and offices demanding a negative result prior to admission, obtaining a COVID test has only become a significant obstacle. The COVID-19 test is available in two distinct forms. The more sensitive molecular PCR tests, which must be analyzed in a laboratory, will find COVID sooner. Because of the rapid spread of Omicron, a person may have to wait several days for a PCR test, as the laboratories that analyze the specimens are also overburdened (Del Rio et al., 2022). Antigen tests, which provide results in minutes, can be performed at home. However, as the omicron variant ravages the whole world, countries are facing shortages of COVID antigen testing kits (Mellor, 2022), making it more challenging to identify and isolate patients who are infected with the virus. However, the COVID pandemic generated an immense amount of unstructured data that can be used by healthcare policymakers to manage the pandemic. One of the potential unstructured data sources that can be utilized to manage the pandemic is the x-ray images of COVID positive patients. Early detection of the virus is critical for proper COVID treatment. Artificial intelligence can be tapped to develop models that can be used for early detection of the virus. Additionally, artificial intelligence models can increase the productivity of medical professionals by accurately delineating infections in X-ray and CT images, allowing for subsequent measurement (Shi et al., 2020). In fact, one doesn't have to be a medical professional or a radiologist to analyze and evaluate x-ray image datasets. A person can develop a machine learning model that predicts whether an x-ray image is COVID positive or not as long as she/he has the appropriate datasets and programming background (Tizhoosh & Fratesi, 2021).

Machine Learning and Computer Vision Algorithms

Machine learning is a branch of artificial intelligence that focuses on statistical models and analytical algorithms (Johnson et al., 2018). It is used by systems to carry out tasks without specific guidelines, relying instead on hidden patterns, sequences, and inference. Computers perform machine learning tasks with little input from software developers. It makes decisions based on data and lets it be used in new ways in a wide range of industries, especially in the discipline of healthcare. On the other hand, computer vision is the practice of employing computers to comprehend digital images and movies. It aims to automate operations that are currently being performed by human vision. This includes techniques for capturing, analyzing, manipulating, and evaluating digital images, as well as data extraction from the real world in order to generate information (Colyer et al., 2018). Additionally, it encompasses subdomains such as visual recognition, camera tracking, and motion detection, making it applicable to fields such as medicine, transportation, and object modeling. Simply defined, computer vision utilizes a device equipped with a camera to capture images or movies and then perform analysis. Computer vision and machine learning are two applications that have grown progressively inextricably linked. Machine learning has aided in the advancement of computer vision in terms of prognosis, diagnosis, localization, lateralization, and segmentation in medical image analysis (Yuan et al., 2022). Computer vision, in turn, has expanded the coverage of machine learning in signal and image processing for clinical practice (den Eynde et al., 2022). Several machine learning algorithms have already been utilized to perform image classification in the healthcare industry, such as support vector machines (Jimoh et al., 2022), k-means (Khan et al., 2021), decision trees (Ghiasi & Zendehboudi, 2021), and artificial neural networks (Kumar & Kumar, 2021). Through this study, a convolutional neural network algorithm will be made and used to make an image prognosis classification model for the COVID-19 virus that can be used to find the virus in x-ray images.

Convolutional Neural Network

Yann LeCun is widely regarded as the inventor of convolutional neural networks. In 1988, he developed the first convolutional neural network, dubbed "LeNet." LeNet was used to perform tasks such as detecting postcodes and figures (LeCun, 1988). A Convolutional Neural Network (ConvNet/CNN) is a type of deep learning algorithm that can accept an image as an input, allocate significance (trainable weights, functions, and biases) to different regions in the image, and distinguish between them and can be used for image classification (Han et al., 2018) and human activity tracking (Basavaiah & Patil, 2020). Convolutional networks require fewer extensive preprocessing operations than other classification techniques (Arora & Kansal, 2019). The structure of a ConvNet is akin to the connectivity network of neurons in the biological brain of a person and was inspired by the visual cortex's arrangement. Individual neurons react to changes exclusively within a specific portion of the visual field referred to as the receptive field. A cluster of such fields will overlap to completely fill the visual field (Strisciuglio & Petkov, 2021). However, one of the challenges of implementing the CNN algorithm is the determination of the appropriate hyperparameters of the model (Li et al., 2021). Different datasets necessitate distinct sets of hyperparameters for accurate prediction. The enormous number of hyperparameters makes it hard for machine learning developers to pick which to utilize (Zhang et al., 2019). There seems to be no definitive answer for how many layers are optimal in CNN, how many neurons are optimal per layer, or which optimizer is optimal for all datasets. Hyperparameter tuning is necessary to determine the optimal set of hyperparameters from which to generate the model from a given dataset. The goal of hyperparameter tuning is to optimize the accuracy of the model as well as the utilization of the computing power of the device (Rajagopal et al., 2020).

Fibonacci Number for Image Dimension

Ideally, the higher the image dimension, the more accurate the CNN model will produce. However, it will compromise the computational power of the device with respect to the central processing unit, graphical processing unit, and random-access memory. Thus, the model should be tuned up to find the least number of image dimensions without compromising the accuracy of the CNN model. In this study, the image dimension as an input hyperparameter will be tuned up using the selected numbers in the Fibonacci sequence. The Fibonacci sequence is a collection of numbers, each of which is the sum of the two preceding numbers. The series begins with 0 and 1 and continues indefinitely: 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, and so on. This mathematical equation can be used to describe the Fibonacci sequence: $X_{n+2} = X_{n+1} + X_n$. Another critical number in mathematics is concealed within the Fibonacci series, which is the golden ratio. It is a number, like pi, which has an indefinite fractional progression with no patterns. It begins with phi = 1.61803. It is calculated as the proportion of an A x B rectangle where A/B equals (A + B)/A. Numerous artists view this as the ideal proportion for an art based on mathematical principles (Stone, 2020; Pedersen, 2020; Scheps, 2020). When a number in the Fibonacci sequence is divided by the preceding number in the series (for example, 5/3), the fraction becomes increasingly close to the golden ratio as the Fibonacci sequence increases in size.

Research Objectives

The objective of this study is to determine the best CNN model with respect to accuracy and the computing capacity of the device. This study used the numbers 21, 34, 55, 89, and 144 of the Fibonacci sequence as the hyperparameters of the image dimension to identify the best CNN model. Thus, the following image dimensions will be used in this study: 21 by 21, 34 by 34, 55 by 55, 89 by 89, and 144 by 144.

Methods

This study utilized the combination of positive and negative COVID x-ray images to develop a CNN model (El-Shafai & Abd El-Samie, 2020). All of the x-ray images went through the 7-step CNN process. The first process is the creation of training sets and validation sets of x-ray images (Lalwani et al., 2022).

Dividing the Datasets into Training and Test Sets

The positive and negative Covid x-ray image datasets were divided into training and validation sets using a ratio of 80:20. The training dataset consists of 7,487 x-ray images while the validation dataset consists of 1,870 COVID x-ray images. The training set is a collection of image data which is used to train the CNN model and teach it how to recognize hidden features and patterns in the x-ray image. The same training data is supplied to the neural network repeatedly in each epoch, and the model proceeds to learn the attributes of the data. An epoch is a word used in machine learning that refers to the number of passes the algorithm has made across the full training dataset. Batches are commonly used to group data sets. Another term for an epoch is "iteration," referring to the process of running one batch of datasets through the algorithm. The training set should include a diverse collection of input data so that the model may be trained in all settings and predict any future data sample. The validation set consists of a set of data that is different from the training set, and it is used to test the performance and accuracy of the CNN model during training. This validation assessment contains information that we may use to fine-tune the model's hyperparameters and configurations. It acts like a critic, telling the CNN model whether the training is progressing on the right path. After every epoch, the model learns on the training set and is assessed on the validation set. The major goal of dividing the dataset into one validation set is to avoid overfitting (Schmitt et al., 2018), which happens when a model gets extremely good at categorizing samples in the training set but is unable to produce precise classifications on data it has never seen before.

Building the Convolutional Layers of Neural Network

The second step is to build the convolutional layers for the neural network (Chauhan et al., 2018). Multiple hidden layers in CNN aid the extraction of information from an image. Convolutional layers perform a convolution on the input before forwarding the output to the next layer. In image classification, every image is projected as a pixel value matrix. The pixels in a convolution's receptive area are all converted to a fixed value of either 1 or 0. If you apply a convolution to an image, for example, you will reduce the size of the image while also bringing all of the information in the area together into a pixel value. The convolutional layer's final output is a vector. We can employ several types of convolutions depending on the nature of the problem we're trying to solve and the features we want to learn. This layer is where we manipulate the image dimension of our input data using the following dimensions: 21 by 21, 34 by 34, 55 by 55, 89 by 89, and 144 by 144.

Building the Rectifier Linear Units

The third step is to build the rectifier linear units' layer (ReLU). The method in the rectifier linear layer removes all negative values from the segmented image and replaces them with zero. When the network input exceeds a particular threshold, the rectifier linear function is activated. As a result, when some input is less than zero, the output is also zero. When the input exceeds a particular threshold, however, the dependent variable and the input have a linear relationship. This means it can speed up the process of training a convolutional neural network (Lin & Shen, 2018). The goal of using the rectifier function is to make the images look non-linear (Nayak et al., 2020). The reason for this is that images and pictures are non-linear by nature. An image has a number of non-linear elements, such as transformations between pixels, adjustments in the borders, and saturation of colors. The rectifier tends to remove the linearity even further, compensating for any linearity that may be imposed on an image during the convolution process.

Building the Pooling Layer

The fourth step is to build a pooling layer. Pooling layers are employed in CNN to consolidate the information acquired by that same convolutional layer feature map (Arif et al., 2019). Through compression or summarizing the characteristics in the feature map, it fundamentally aids in the minimization of overfitting by the CNN model's training time (Apostolopoulos et al., 2020). Pooling layers frequently use the largest or mean values of inputs to down sample the data. The convolutional layers utilize feature maps to highlight the data's features, and feature maps are particularly sensitive to their coordinates in the image data. For example, if an object in the image shifts by chance, the convolutional layers will have to recreate the feature map. Since the older CNN model preserved the previous model, the algorithm might provide miscalculated predictions for classifying images. Downsampling of the map is required to tackle this problem, and the pooling layers in the CNN model are used to synthesize the features from the feature map. Using polling algorithms like average pooling and max pooling, the CNN algorithm can find out how many attributes are on a feature map on average and how many are the most common.

Flattening the Image and Connecting the Layers of CNN Model

The fifth step is the flattening of the images and connecting all the layers of the CNN model. Flattening is the process of turning data into a one-dimensional array for use in the next layer (More et al., 2021). To construct a single continuous feature vector, the CNN architecture must flatten the output of the convolutional layers. It entails translating the pooled feature image map obtained during the pooling process into a one-dimensional vector. It's also linked to the complete classification algorithm, which is referred to as a "fully connected layer." Attaching the artificial neural network to the current convolutional neural network is the task of the connection layer process. The artificial neural network's layers are superseded by a fully connected layer, which is a form of hidden layer. The neurons in a fully linked layer are all interconnected to the neurons in the subsequent layer. In a convolutional neural network, the fully connected layer's job is to discover certain features in an image and develop a model that will be used to classify images. The neurons in the convolutional layers are associated with a unique feature that may be represented in a picture. The score that the neuron sends to the next layer represents the probability of the attribute's presence in the image.

Compiling and Training the CNN Model using the Training Dataset

The sixth step is to compile and train the convolutional network model. Compilation is a phase in the convolutional neural network that converts the previously defined sequence of layers into an efficient and fast series of matrix operations and transformations. Compilation can be thought of as a pre-computation phase that allows the machine to train the model using the algorithm and the image datasets as the inputs. A convolutional neural network is trained in two phases: the forward phase, in which the input is sent entirely through the network, and the backward phase, in which the input is transmitted entirely through the network (Traore et al., 2018). Gradients were back propagated (backprop) while weight parameters were updated in the second phase. Throughout the forward sequence, each layer will store any data required for the backward phase, like image inputs and approximate features. This implies that any backward phase must be succeeded by a forward phase. Every layer collects a gradient and will also transfer back a gradient during the backward phase. It will acquire the loss gradient in relation to its outputs and send the loss gradient in relation to its inputs. The training time of CNN models was processed using the Levene Test of Homogeneity to assess whether the data is homogenous or not. This study utilized the Analysis of Variance (ANOVA) and Post-Hoc Test through SPSS to compare the mean of the training time for 50 epochs of each CNN model.

Implement the Trained CNN Model using the Test Dataset

The final step is to make a prediction using a test dataset consisting of 130 COVID positive x-ray images and 50 COVID negative images. The trained model will be tested and implemented on the test set. A confusion matrix was constructed to determine the type 1 and type 2 errors, accuracy, precision, recall, and F1-score of each model. The type 1 error refers to the predicted positive COVID x-ray images of the CNN model when the actual COVID x-ray images are negative. The type 2 error refers to the predicted negative COVID x-ray images when the actual COVID x-ray images are positive. The accuracy of a CNN image prognosis classifier is basically the frequency with which it produces the correct prediction. Accuracy is measured by the ratio of the number of successful predictions to the total number of forecasts. Precision is the ratio of the overall quantity of successfully categorized COVID positive x-ray images divided by the total number of anticipated COVID positive x-ray images. The proportion of the overall number of correctly identified positive COVID x-ray images to all positive COVID x-ray images is known as the "recall." The F1 score is the symmetrical mean of precision and recall, and it ranges from 0 to 1. The accuracy, precision, recall, and F1-score of the CNN prognosis image classification model can be interpreted as ideal metrics if the value is within the range of 0.80 to 1.00. The following are the equations of accuracy, precision, recall, and F1-score used to measure the performance of the models.

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \quad 1$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad 2$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad 3$$

$$F1 - score = 2 \text{ x } [\frac{(\ Recall \text{ x } Precision)}{(Recall + Precision}] \quad 4$$

This study utilized the Python 3.8 programming language, Jupyter notebooks, tensorflow, keras, pandas, and numpy libraries to develop the convolutional neural network models. All of the convolutional neural networks have 2 hidden layers and 233 neuron nodes for each hidden layer. The optimizer that was used in the model is adaptive moment estimation or Adam. Adam is a first-order gradient-based stochastic objective function optimization algorithm based on adaptive estimations of lower-order moments. Adam is a lot of work, but it doesn't use a lot of memory, isn't affected by gradient diagonal rescaling, and is good for problems with a lot of data or hyperparameters (Kingma & Ba, 2017).

Results and Discussion

This section discusses the simulation results achieved using the convolutional neural network classification model for COVID and non-COVID images, which were trained using specified Fibonacci numbers as hyperparameters of image dimension. Table 1 shows the total and the average training time for each model.

Table 1. Training Time of CNN Models

CNN Models by image dimensions	Number of X-ray images in Training Sets	Number of X-ray images in Validation Sets	Total Number of X-ray images	Total Training Time in seconds (50 epochs)	Average Training Time per epoch in seconds
21 by 21	7,487	1,870	9,357	5146	102.92
34 by 34	7,487	1,870	9,357	4910	98.20
55 by 55	7,487	1,870	9,357	5221	104.42
89 by 89	7,487	1,870	9,357	6604	132.08
144 by 144	7,487	1,870	9,357	8383	167.66

Based on the results of the Levene Test of Homogeneity, the dataset of training time is heterogeneous, as shown in Table 3. Thus, Dunnett T3, as the post hoc test, was adapted along with ANOVA to describe the difference between the means of the training time of CNN models. Table 3 shows the results of the analysis of variance, while Table 4 shows the post-hoc test.

Table 1. Test of Homogeneity of Variances

LEVENE STATISTIC	DF1	DF2	SIG.
26.37	4	245	.000

Table 2. Analysis of Variance for Training Time

	SUM OF SQUARES	DF	MEAN SQUARE	F	SIG.
BETWEEN GROUPS	171076.46	4	42769.11	271.37	0.00
WITHIN GROUPS	38612.76	245	157.60		
TOTAL	209689.22	249			

Table 3. Dunnett T3 Post-Hoc Test

(I) CNN MODEL	(J) CNN MODELS	MEAN DIFFERENCE (I-J)	STD. ERROR	SIG.
21 BY 21	34 by 34	4.720*	.822	.000
	55 by 55	-1.500	1.639	.987
	89 by 89	-29.160*	2.488	.000
	144 by 144	-64.740*	2.623	.000
34 BY 34	21 by 21	-4.720*	.822	.000
	55 by 55	-6.220*	1.707	.005
	89 by 89	-33.880*	2.533	.000
	144 by 144	-69.460*	2.667	.000
55 BY 55	21 by 21	1.500	1.639	.987
	34 by 34	6.220*	1.707	.005
	89 by 89	-27.660*	2.903	.000
	144 by 144	-63.240*	3.020	.000
89 BY 89	21 by 21	29.160*	2.488	.000
	34 by 34	33.880*	2.533	.000
	55 by 55	27.660*	2.903	.000
	144 by 144	-35.580*	3.553	.000
144 BY 144	21 by 21	64.740*	2.623	.000
	34 by 34	69.460*	2.667	.000
	55 by 55	63.240*	3.020	.000
	89 by 89	35.580*	3.553	.000

***.** The mean difference is significant at the 0.01 level.

The results of the analysis of variance and the post-hoc test show that there is a significant difference between the training times of the 5 CNN models. There is no significant difference in the training time between the CNN models of 21 by 21 and 55 by 55 image dimensions. On the other hand, results suggest that the higher the image dimensions, the more time it needs to train the CNN model. Table 5 shows the confusion matrix of the CNN models, which indicates the number of actual values versus predicted values.

Table 4: Confusion Matrix

CNN Models by image dimensions	True Positive	True Negative	False Negative	False Positive	N
21 by 21	65	49	65	1	180
34 by 34	96	42	34	8	180

CNN Models by image dimensions	True Positive	True Negative	False Negative	False Positive	N
55 by 55	121	31	9	19	180
89 by 89	103	25	27	25	180
144 by 144	116	20	14	30	180

Based on the confusion matrix, the CNN model with an image dimension of 55 by 55 has the highest value of accuracy at 0.85, with 9 x-ray images falling into the category of "false positive". In terms of precision, the CNN model with an image dimension of 21 by 21 got the highest predicted positive COVID x-ray images at 0.98, with only 1 x-ray image wrongly predicted as positive. For the recall, the CNN model with an image dimension of 55 by 55 captured the least number of "false negatives" with a value of 0.93. Finally, the CNN model with an image dimension of 55 by 55 got the highest value of the F1-score at 0.89, which suggests that the same model has the best balance between precision and recall. The accuracy, precision, recall, and F1-score of the five CNN image classification models are summarized in Table 6.

Table 5: Metrics of CNN Models

CNN Models by image dimensions	Accuracy	Precision	Recall	F1-score
21 by 21	0.64	0.98	0.50	0.66
34 by 34	0.77	0.92	0.74	0.82
55 by 55	0.84	0.86	0.93	0.89
89 by 89	0.71	0.80	0.79	0.79
144 by 144	0.76	0.79	0.89	0.84

Discussion

Among the 5 CNN models, the CNN model with an image dimension of 55 by 55 has the highest accuracy. However, what if the 9 positive x-ray images that were wrongfully predicted as negative are actually x-ray images that are infected with a new variant of the COVID virus that is very contagious? The penalties for having an actual positive or false negative labeled incorrectly are extremely substantial in this outlined scenario. This predicament explains the concept of "accuracy paradox" in machine learning models. The accuracy of the CNN model with a 55 by 55 image dimension was high, implying the model's veracity. However, it was ineffective because it did not indicate how well the model predicted the desired class. Consider that the difficulty in using accuracy in the CNN models derives from the fact that the x-ray image testing dataset was not balanced with 130 observations of x-ray images of COVID positive and 50 without. This is also referred to as class disparity. The fact that most real-world datasets are imbalanced shows how important it is to be aware of the accuracy metric's flaws and be careful when using it to appraise classification models. Thus, policymakers and data scientists should look at the precision, recall, and F1-score metrics of the model before adapting it as a prognosis image classification model.

The confusion matrix results describe the following four scenarios in the context of the COVID prognosis image classification model with 55 by 55 image dimensions. The first is a "true positive scenario," in which the model accurately predicts COVID in a patient who already has COVID. The second scenario is a "false positive," in which the model predicts COVID in a patient who does not have COVID. The third is the "false negative scenario," in which the model predicts that a patient with COVID does not have COVID. The final scenario is the "true negative scenario," in which the model predicts that a patient without COVID will have no COVID. While the "true positive scenario" and the "true negative scenario" are desirable outcomes, the "false positive scenario" and "false negative scenario" are incorrect predictions with undesirable consequences. The "false positive scenario" implies that out of 50 patients who do not actually have COVID, the model indicates that 19 do. In reality, these 19 patients will almost certainly undergo costly and worthless treatments that will jeopardize their health. On the other hand, the "false negative scenario" implies that 9 of the 130 patients who actually have COVID were not detected by the model. As a result, these nine patients will not be found, spread the virus to their loved ones, and not get any treatment that could lead to the deterioration of their health. As one might expect, the consequences of the "false positive scenario" and "false negative scenario" are quite different, but nonetheless substantial. This is applicable not only to COVID x-ray image classification but also to a variety of other applications in the healthcare industry. While data scientists and machine learning engineers would prefer that all model results fall within the "true positive scenario" and "true negative scenario," no machine learning model is perfect in real life. Model predictions almost always contain "false positives" and "false negatives." The goal would then be to make sure there are as few false positive and negative results as possible. This can be done by using precision and recall.

In terms of precision, it is interesting to note that the CNN model with an image dimension of 21 by 21 that has the lowest accuracy among the 5 models got the highest precision rate at 0.98. The CNN model, with an image dimension of 55 by 55, only got a 0.86 precision rate, trailing the CNN model with an image dimension of 34 by 34 at 0.92. In terms of recall, the CNN model with an image dimension of 55 by 55 got the highest value at 0.93. As a policymaker or data scientist, you must decide when to use precision and when to use recall. Consider that precision and recall both quantify the classification performance of a model, but in distinct ways. Precision quantifies the magnitude of error introduced by false positives (FP), whereas recall quantifies the magnitude of error introduced by false negatives (FN). As a result, before deciding on a metric to use, one should weigh the relative importance of these two types of errors in the use-case of this study. Thus, the critical question is which type of error is more undesirable: is it a false positive or false negative? Several studies in healthcare would suggest that the most undesirable predicament is the false negative scenario for failing to diagnose a particular disease for proper treatment (Kader, et al., 2020; Joyce, et al., 2021; James, et al., 2021). Being untreated for COVID can endanger one's life and increase the probability of spread of the virus, which is probably worse than getting unwarranted treatment. It makes sense to choose a model with as few FNs as possible, as the effects of errors due to FNs are estimated to be large. That is, you should use recall, not precision. So, when should one employ precision? Many real-world datasets are unlabeled, which means policymakers have no idea which category each observation belongs to. One of the main advantages of utilizing machine learning to classify x-ray images is that it requires less human effort. As a result, it is preferred to use a CNN model that will predict as many true positives as possible. Precision should be prioritized over recall in such situations.

It's also feasible that the errors generated by FPs and FNs are equally bad for a particular use-case. As a result, both precision and recall should be maximized. Due to the sheer tradeoff between precision and recall, it is not possible to maximize both precision and recall simultaneously in practice. The value of precision decreases when the value of recall increases, and vice versa. F1-scores can be used to figure out which model is best based on a pair of precision and recall values for CNN models. Among the 5 CNN models, the CNN model with an image dimension of 55 by 55 has the highest F1-score at 0.89. In this case, if "false positives" and "false negatives" are considered unacceptable errors, the F1-score should be prioritized over accuracy, precision, and recall.

Conclusion

Based on the training time and the metrics of accuracy, recall, and F1-score, the CNN model with an image dimension of 55 by 55 is the best prognosis image classification model. It is important to note that in the healthcare field, undiagnosed prognosis is an important issue as it may lead to serious repercussions as COVID is a very contagious and infectious virus. Thus, recall and the F1-score are being prioritized over the other metrics. Furthermore, the number of Fibonacci patterns will further increase, and this might affect the performance of the model. This means that the image dimension of 55 by 55 can be used as a benchmark for more research to improve the metrics and reach the desired value of 0.90 or above for the accuracy, precision, recall, and F1-score. This study recommends exploring the configuration of other hyperparameters such as the number of hidden layers, the number of nodes, and the number of image dimensions. In addition, future research may explore the possibility of integrating other machine learning algorithms to enhance the performance metrics of the CNN models that were created in this study.

References

Apostolopoulos, I. D., Aznaouridis, S. I., & Tzani, M. A. (2020). Extracting Possibly Representative COVID-19 Biomarkers from X-ray Images with Deep Learning Approach and Image Data Related to Pulmonary Diseases. *Journal of Medical and Biological Engineering, 40*, 462-469. https://doi.org/10.1007/s40846-020-00529-4

Arif, S., Wang, J., Hassan, T., & Fei, Z. (2019). 3D-CNN-Based Fused Feature Maps with LSTM Applied to Action Recognition. *Future Internet, 11*(2), 42. https://doi.org/10.3390/fi11020042 .

Arora, M., & Kansal, V. (2019). Character level embedding with deep convolutional neural network for text normalization of unstructured data for Twitter sentiment analysis. *Social Network Analysis and Mining, 9*, https://doi.org/10.1007/s13278-019-0557-y.

Basavaiah, J., & Patil, C. M. (2020). Human activity detection and action recognition in videos using convolutional neural networks. *Journal of Information and Communication Technology, 19*(2), 157-183. https://doi.org/10.32890/jict2020.19.2.1

Baum, A., Kaboli, P. J., & Schwartz, M. D. (2021). Reduced In-Person and Increased Telehealth Outpatient Visits During the COVID-19 Pandemic. *Annals of Internal Medicine, 174*(1), 129-131. https://doi.org/10.7326/M20-3026

Bernacki, K., Keister, A., Sapiro, N., Joo, J., & Mattle, L. (2021). Impact of COVID-19 on patient and healthcare professional attitudes, beliefs, and behaviors toward the healthcare system and on the dynamics of the healthcare pathway. *BMC Health Services Research, 21*, https://doi.org/10.1186/s12913-021-07237-y

Chauhan, R., Ghanshala, K., & Joshi, R. (2018). Convolutional neural network (CNN) for image detection and recognition. In *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)* (pp. 278-282. https://doi.org/10.1109/ICSCCC.2018.8703316). IEEE

Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. I. (2018). A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System. *Sports medicine-open, 14*(1), 1-15. https://doi.org/10.1186/s40798-018-0139-y.

De Luca, P., Bisogno, A., Colacurcio, V., Marra, P., Cassandro, C., Camaioni, A., . . . Scarpa, A. (2022). Diagnosis and treatment delay of head and neck cancers during COVID-19 era in a tertiary care academic hospital: what should we expect? *European Archives of Oto-Rhino-Laryngology, 279*, 961-965. https://doi.org/10.1007/s00405-021-06834-1 .

Del Rio, C., Omer, S. B., & Malani, P. N. (2022). Winter of Omicron—the evolving COVID-19 pandemic. *JAMA, 327*(4), 319-320. doi:10.1001/jama.2021.24315.

den Eynde, J., Lachmann, M., Laugwitz, K.-L., Manlhiot, C., & Kutty, S. (2022). Successfully implemented artificial intelligence and machine learning applications in cardiology: State-of-the-art review. *Trends in Cardiovascular Medicine*, https://doi.org/10.1016/j.tcm.2022.01.010.

El-Shafai, W., & Abd El-Samie, F. (2020). Extensive COVID-19 X-Ray and CT Chest Images Dataset. *Mendeley Data*, http://dx.doi.org/10.17632/2fxz4px6d8.4.

Ghiasi, M., & Zendehboudi, S. (2021). Application of decision tree-based ensemble learning in the classification of breast cancer. *Computers in Biology and Medicine, 128*, https://doi.org/10.1016/j.compbiomed.2020.104089.

Han, D., Liu, Q., & Fan, W. (2018). A new image classification method using CNN transfer learning and web data augmentation. *Expert Systems with Applications, 95*, 43-56. https://doi.org/10.1016/j.eswa.2017.11.028.

Hashmi, H., & Asif, H. (2020). Early detection and assessment of covid-19. *Frontiers in medicine, 7*, 311. https://doi.org/10.3389/fmed.2020.00311.

Islam, M., Hasan, M., Wang, X., Germack, H. D., & Noor-E-Alam, M. (2018). A systematic review on healthcare analytics: application and theoretical perspective of data mining. *In Healthcare,* 6(2), 54. https://dx.doi.org/10.3390%2Fhealthcare6020054.

Iyengar, K., Upadhyaya, G. K., Vaishya, R., & Jain, V. (2020). COVID-19 and applications of smartphone technology in the current pandemic. *Diabetes & metabolic syndrome,* *14*(5), 733-737. https://dx.doi.org/10.1016%2Fj.dsx.2020.05.033.

James, L. J., Wong, G., Tong, A., Craig, J. C., Howard, K., & Howell, M. (2021). Patient preferences for cancer screening in chronic kidney disease: a best–worst scaling survey. *Nephrology Dialysis Transplantation*, https://doi.org/10.1093/ndt/gfab360.

Jimoh, R. G., Abisoye, O. A., & Uthman, M. B. (2022). Ensemble feed-forward neural network and support vector machine for prediction of multiclass malaria infection. *Journal of Information and Communication Technology, 21*(1), 117-148. https://doi.org/10.32890/jict2022.21.1.6.

Johnson, K. W., Torres Soto, J., Glicksberg, B. S., Shameer, K., Miotto, R., Ali, M., . . . Dudley, J. T. (2018). Artificial Intelligence in Cardiology. *Journal of the American College of Cardiology, 71*(23), 2668-2679. https://doi.org/10.1016/j.jacc.2018.03.521.

Joyce, D., Gracias, C., Murphy, F., Tubassam, M., Walsh, S., & O'Hanlon, S. (2021). Potentially undiagnosed cognitive impairment in patients with peripheral arterial disease: A systematic review of the literature. *Potentially undiagnosed cognitive impairment in patients with peripheral arterial disease: A systematic review of the literature*, https://doi.org/10.1016/j.surge.2021.06.007.

Kader, N., Clement, N. D., Patel, V. R., Caplan, N., Banaszkiewicz, P., & Kader, D. (2020). The theoretical mortality risk of an asymptomatic patient with a negative SARS-CoV-2 test developing COVID-19 following elective orthopaedic surgery. *The Bone and Joint Journal*, 1256-1260. https://doi.org/10.1302/0301-620X.102B9.BJJ-2020-1147.R1.

Kang, J., Jang, Y., Kim, J., Han, S.-H., Lee, K., Kim, M., & Eom, J. (2020). South Korea's responses to stop the COVID-19 pandemic. *American Journal of Infection Control,* *48*(9), 1080-1086. https://doi.org/10.1016/j.ajic.2020.06.003.

Kaye, L., Theye, B., Smeenk, I., Gondalia, R., Barrett, M. A., & Stempel, D. A. (2020). Changes in medication adherence among patients with asthma and COPD during the COVID-19 pandemic. *The journal of allergy and clinical immulogy. In practice, 8*(7), 2384-2385. https://dx.doi.org/10.1016%2Fj.jaip.2020.04.053.

Khan, A., Khan, S., Harouni, M., Abbasi, R., Iqbal, S., & Mehmood, Z. (2021). Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification. *Microsc Res Tech, 84*(7), 1389-1399. https://doi.org/10.1002/jemt.23694.

Kingma, D. P., & Ba, J. (2017, January). *Adam: A Method for Stochastic Optimization.* Retrieved April 26, 2022 from arXiv: https://arxiv.org/abs/1412.6980

Kumar, N., & Kumar, D. (2021). An improved grey wolf optimization-based learning of artificial neural network for medical data classification. *Journal of Information and Communication Technology, 20*(2), 213-248. https://doi.org/10.32890/jict2021.20.2.4.

Lalwani, P., Mishra, M., Chadha, a., & Sethi, P. (2022). Customer churn prediction system: a machine learning approach. *Computing, 104*, 271-294. https://doi.org/10.1007/s00607-021-00908-y.

LeCun, Y. (1988). A theoretical framework for back-propagation. In *Proceedings of the 1988 connectionist models summer school* (pp. 21-28).

Li, J.-Y., Zhan, Z.-H., Xu, J., Kwong, S., & Zhang, J. (2021). Surrogate-assisted hybrid-model estimation of distribution algorithm for mixed-variable hyperparameters optimization in convolutional neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, https://doi.org/10.1109/tnnls.2021.3106399.

Lin, G., & Shen, W. (2018). Research on convolutional neural network based on improved Relu piecewise activation function. *Procedia Computer Science, 131*, 977-984. https://doi.org/10.1016/j.procs.2018.04.239.

Long, L., & Corsar, K. (2020). The COVID-19 effect: number of patients presenting to The Mid Yorkshire Hospitals OMFS team with dental infections before and during The COVID-19 outbreak. *Br J Oral Maxillofac Surg, 58*(6), 713-714. Br J Oral Maxillofac Surg.

Maissenhaelter, B. E., Woolmore, A. L., & Schlag, P. M. (2018). Real-world evidence research based on big data. *Onkologe, 24*, 91-98. https://doi.org/10.1007/s00761-018-0358-3.

Mellor, S. (2022, January 10). *Covid inequality worsened by antigen test supply crunch.* Retrieved April 1, 2022 from Fortune: https://fortune.com/2022/01/10/omicron-testing-shortages-delays-covid-inequality/

Menon, V., & Padhy, S. (2020). Ethical dilemmas faced by health care workers during COVID-19 pandemic: Issues, implications and suggestions. *Asian journal of psychiatry, 51*, https://dx.doi.org/10.1016%2Fj.ajp.2020.102116.

More, S. S., Mange, M., Sankhe, M., & Sahu, S. (2021). Convolutional Neural Network based Brain Tumor Detection. In *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1532-1538. doi: 10.1109/ICICCS51141.2021.9432164).

Nayak, D., Das, D., Dash, R., Majhi, S., & Majhi, B. (2020). Deep extreme learning machine with leaky rectified linear unit for multiclass classification of pathological brain images. *Multimedia Tools and Applications, 79*, 15381–15396. https://doi.org/10.1007/s11042-019-7233-0.

Pedersen, E. (2020). The Beauty of Mathematical Order: a Study of the Role of Mathematics in Greek Philosophy and Modern Art Works of Piet Hein and Inger Christensen. *The Journal of Somaesthetics, 6*, https://doi.org/10.5278/ojs.jos.v6i1.3667.

Rajagopal, A., Joshi, G., Ramachandran, A., Subhalakshmi, R. T., Khari, M., Jha, S., . . . You, J. (2020). A deep learning model based on multi-objective particle swarm optimization for scene classification in unmanned aerial vehicles. *IEEE Access, 8*, 135383-135393. https://doi.org/10.1109/ACCESS.2020.3011502.

Scheps, R. (2020). Geometry and the Life of Forms. In *Mathematics in the Visual Arts* (pp. 29-52. https://doi.org/10.1002/9781119801801.ch2).

Schmitt, F., Banu, R., Yeom, I.-T., & Do, K. (2018). Development of artificial neural networks to predict membrane fouling in an anoxic-aerobic membrane bioreactor treating domestic wastewater. *Biochemical Engineering Journal, 133*, 47-58. https://doi.org/10.1016/j.bej.2018.02.001.

Seki, T., Takeuchi, M., & Kawakami, K. (2021). Eating and drinking habits and its association with obesity in Japanese healthy adults: retrospective longitudinal big data analysis using a health check-up database. *British Journal of Nutrition, 126*(10), 1585-1591. https://doi.org/10.1017/s0007114521000179.

Shafique, M., Khurshid, M., Rahman, H., Khanna, A., & Gupta, D. (2019). The role of big data predictive analytics and radio frequency identification in the pharmaceutical industry. *IEEE Access, 7*, 9013-9021. https://doi.org/10.1109/ACCESS.2018.2890551.

Shi, F., Wang, J., Shi, J., Wu, Z., Wang, Q., Tang, Z., . . . Shen, D. (2020). Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19. *IEEE reviews in biomedical engineering, 14*, 4-15. https://doi.org/10.1109/RBME.2020.2987975.

Shrivastava, S. R. (2020). 2019-nCoV Outbreak Declared as Public Health Emergency of International Concern: What Next? *International journal of preventive medicine, 11*, 65. https://dx.doi.org/10.4103%2Fijpvm.IJPVM_63_20.

Stone, A. (2020). Measurement as a tool for painting. *Journal of Mathematics and the Arts, 14*(1-2), 154-156. https://doi.org/10.1080/17513472.2020.1732786.

Strisciuglio, N., & Petkov, N. (2021). Brain-Inspired Algorithms for Processing of Visual Data. In *Brain-Inspired Computing. BrainComp 2019. Lecture Notes in Computer Science()* (Vol. 12339, pp. https://doi.org/10.1007/978-3-030-82427-3_8). Springer, Cham.

Tizhoosh, H. R., & Fratesi, J. (2021). COVID-19, AI enthusiasts, and toy datasets: radiology without radiologists. *European Radiology, 31*, 3553-3554. https://doi.org/10.1007/s00330-020-07453-w.

Traore, B., Kamsu-Foguem, B., & Tangara, F. (2018). Deep convolution neural network for image recognition. *Ecological Informatics, 48*, 257-268. https://doi.org/10.1016/j.ecoinf.2018.10.002.

World Health Organization. (2020). *Coronavirus disease (COVID-19) pandemic.* Retrieved April 1, 2022 from WHO | World Health Organization: https://www.who.int/emergencies/diseases/novel-coronavirus-2019

World Health Organization. (2020, March 11). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020.* Retrieved April 1, 2022 from WHO | World Health Organization: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

Yuan, J., Ran, X., Liu, K., Yao, C., Yao, Y., Wu, H., & Liu, Q. (2022). Machine learning applications on neuroimaging for diagnosis and prognosis of epilepsy: A review. *Journal of Neuroscience Methods, 368*, https://doi.org/10.1016/j.jneumeth.2021.109441.

Zhang, T., Gao, C., Ma, L., Lyu, M., & Kim, M. (2019). An empirical study of common challenges in developing deep learning applications. In *2019 IEEE 30th International Symposium on Software Reliability Engineering (ISSRE)* (pp. 104-115. https://doi.org/10.1109/ISSRE.2019.00020).