

# Grayscale image statistics of COVID-19 patient CT scans characterize lung condition with machine and deep learning

Sara Ghashghaei<sup>1</sup>  | David A. Wood<sup>2</sup>  | Erfan Sadatshojaei<sup>3</sup>  | Mansooreh Jalilpoor<sup>1</sup> 

<sup>1</sup>Medical School, Shiraz University of Medical Sciences, Shiraz, Iran

<sup>2</sup>Department of Research, DWA Energy Limited, Lincoln, LN5 9JP, UK

<sup>3</sup>Department of Chemical Engineering, Shiraz University, Shiraz, 71345, Iran

## Correspondence

David A. Wood, Department of Research, DWA Energy Limited, Lincoln, LN5 9JP, UK.  
Email: [dw@dwasolutions.com](mailto:dw@dwasolutions.com)

Edited by Yi Cui

## Funding information

None

## Abstract

**Background:** Grayscale image attributes of computed tomography (CT) of pulmonary scans contain valuable information relating to patients with respiratory ailments. These attributes are used to evaluate the severity of lung conditions of patients confirmed to be with and without COVID-19.

**Method:** Five hundred thirteen CT images relating to 57 patients (49 with COVID-19; 8 free of COVID-19) were collected at Namazi Medical Centre (Shiraz, Iran) in 2020 and 2021. Five visual scores (VS: 0, 1, 2, 3, or 4) are clinically assigned to these images with the score increasing with the severity of COVID-19-related lung conditions. Eleven deep learning and machine learning techniques (DL/ML) are used to distinguish the VS class based on 12 grayscale image attributes.

**Results:** The convolutional neural network achieves 96.49% VS accuracy (18 errors from 513 images) successfully distinguishing VS Classes 0 and 1, outperforming clinicians' visual inspections. An algorithmic score (AS), involving just five grayscale image attributes, is developed independently of clinicians' assessments (99.81% AS accuracy; 1 error from 513 images).

**Conclusion:** Grayscale CT image attributes can be successfully used to distinguish the severity of COVID-19 lung damage. The AS technique developed provides a suitable basis for an automated system using ML/DL methods and 12 image attributes.

## KEYWORDS

computed tomography analysis, confusion-matrix analysis, COVID-19 lung feature recognition, grayscale image attributes, visual versus algorithmic classification

## Highlights

- Grayscale image statistics of CT scans can effectively classify lung abnormalities
- Graphical trends of grayscale statistics distinguish visual assessments COVID-19 classes
- Machine/deep learning algorithms predict severity from image grayscale attributes
- Algorithmic class systems can be established using just five grayscale attributes
- Confusion matrices provide detailed insight to algorithm prediction capabilities

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Chronic Diseases and Translational Medicine* published by John Wiley & Sons Ltd on behalf of Chinese Medical Association.

## 1 | INTRODUCTION

The COVID-19 virus continues to have devastating impacts worldwide involving millions of lives lost, repeated nationwide lockdowns, and substantial damage to the global economy. Methods that can quickly and accurately classify the severity of patient lung damage caused by COVID-19 offer the potential to assist in establishing and rapidly implementing appropriate treatment plans. Most pulmonary examinations now include computed tomography (CT) scanning as they can be executed rapidly and painlessly. Analysis of CT is crucial in establishing the degree to which lung functionality is impacted by various diseases, including cancers, various forms of pneumonia and, since early 2020, COVID-19.<sup>1,2</sup> Radiological image analysis, both X-ray and CT, help to diagnose COVID-19 and provide more details of patient conditions combined with virus' nucleic acid by real-time reverse transcription-polymerase chain reaction (rRT-PCR) tests.<sup>3</sup> Moreover, CT scan analysis can achieve up to 98% sensitivity in diagnosing COVID-19.<sup>4</sup> Lung CT-scan-slice analysis also provide useful insight by identifying and/or confirming the nature of pulmonary impacts caused by COVID-19.<sup>5</sup> However, it takes time for radiological/medical assessments of CT scans to be conducted, making automated classification techniques desirable.<sup>6</sup>

Thoracic CT images taken from COVID-19 patients are typically used to support the diagnosis of the condition, establishing prognoses and treatments for patients.<sup>7</sup> They are also used to monitor symptom progression and patient responses during treatment.<sup>8</sup> The CT images tend not to be routinely used to quantify the severity of lung impacts caused by COVID-19, although some studies have focused on that objective.<sup>9</sup> Attempts to achieve this aided by deep- and/or machine-learning (DL/ML) have to overcome a number of challenges.<sup>10,11</sup>

Pulmonary-CT-scan images display a range of distinguishing features in some COVID-19 sufferers, that vary according to the conditions severity. Granularly opaque image areas, chaotic arrangements of lineations, agglomerated masses, opaque alveoli, inverse halo and/or atoll features, thick-polygonal forms with variable and indistinct linear opacities are some of the common features observed.<sup>12</sup> Lung features observed in radiological images are likely to be affected by some pre-existing abnormalities in some patients acting to confuse COVID-19 CT-scan interpretations.<sup>13</sup> Nevertheless, linking characteristic features observed in CT scan images has the potential to be exploited for systematically classifying the severity of lung conditions in COVID-19 sufferers.<sup>14</sup> In particular, DL/ML techniques can be exploited to identify the common lung features associated with COVID-19 with some degree of accuracy.<sup>15-17</sup> Some DL methods have applied convolutional neural networks (CNN) combined with

automated feature-extraction routines applied directly to the CT images to achieve this.<sup>18,19</sup> Such approaches require effective image-segmentation processing.<sup>16,20,21</sup>

Customizing effective and rapid CT-scan-image analysis offers the potential to accelerate COVID-19 patient diagnosis, and treatment plan implementation and improve patients' long-term prognoses. CNN coupled with effective feature-extraction algorithms offers one route for such analysis.<sup>10</sup> However, such approaches often ignore the underlying relationships between the various grayscale-image attributes associated with the distinctive CT image features typically observed in COVID-19 patients.

Many DL studies using X-ray and CT-scan images are focused on the early diagnosis of COVID-19.<sup>22,23</sup> This involves a binary assessment assigning a negative or positive result.<sup>24,25</sup> The novelty of this study is that it considers in detail the grayscale image attributes of lung CT images from patients with and without COVID-19. Moreover, it links the combinations of those grayscale statistical attributes to the severity of the pulmonary effects impacting those patients with COVID-19.

The research objective of this study was to automate the interpretation of thoracic CT-scan images using grayscale image statistical attributes to provide a reliable classification of the severity of lung conditions afflicting COVID-19 sufferers. It does this applying CNN and several ML algorithms and comparing their performance. The analysis initially focuses on predicting a clinicians' visual assessments of the CT scan images. It does this in terms of a binary distinction between patients with and without COVID-19 and classifying the degree of severity of the lung impacts. The study additionally develops an effective algorithmic classification of the severity of lung impacts of COVID-19 based on CT-scan image grayscale statistics. It assesses how that novel classification system could be adapted to provide an effective and rapid automated expert system for classifying the severity of lung impacts from pulmonary CT images. DL and ML models are applied to predict that algorithmic classification. The study evaluates the prediction accuracy of a combination of the proposed algorithmic classification of lung conditions with DL/ML prediction models. It also compares the results of the visual and algorithmic classification analysis of the CT images.

## 2 | METHOD

### 2.1 | Overview of proposed CT-image analysis

Grayscale statistics extracted from CT pulmonary-parenchyma images can be exploited to provide valuable insight into the lung's alveolar tissue employed in

respiratory processes, and this study focuses on that information. CT-image slices are rapidly analyzed to yield values for a range of grayscale statistical attributes. Interpretations considering combinations of these statistics with graphic representations identify distributions related to the severity of lung impacts related to COVID-19 that, to some extent, overlap with each other. However, DL/ML methods can be applied to combinations of these overlapping, grayscale-attribute distributions to accurately the severity of lung abnormalities of each CT-image slice assessed.

## 2.2 | Ethical approval

The CT analyses of human patients used in this study were performed in accordance with the relevant guidelines and regulations of Namazi Medical Centre and the Medical School of Shiraz University of Medical Sciences (Shiraz, Iran). In particular, the consent for analysis was obtained from each patient and their anonymity was guaranteed. Consequently, the confidentiality of the biographical information pertaining to each patient has been maintained at all times. All CT analytical protocols applied were approved by the Medical School of Shiraz University of Medical Sciences.

## 2.3 | CT-scan slice procurement

The suites of lung CT images were obtained from a representative set of patients. Most of them were shown by laboratory testing to be afflicted with COVID-19. However, images from several patients with negative COVID-19 tests were included in the suite of images compiled. These patients were all under treatment at the Namazi Medical Centre, Shiraz, Iran. The Philips Ingenuity CT-scanning equipment (Philips) located at that medical center provided suites of CT-scan-slice images (0.625 mm thickness) for all patients assessed. Such multislice-CT-scan equipment is now widely used to rapidly deliver high-resolution images.<sup>26</sup>

CT images considered are all noncontrasted, which are typical of those widely used to assess various lung conditions in addition to those caused by coronaviruses. The CT-scan-slice images utilized are derived from 49 COVID-positive patients displaying a substantial range of visible features known to be commonly associated with COVID-19. Additionally, multiple CT-scan-slice images associated with eight COVID-negative patients are also assessed for comparative purposes. The COVID-negative patients consisted of four women with ages ranging from 18 to 68 years, together with four men with ages varying from 24 to 71 years. These images were compared with those from 23 female COVID-positive patients with ages ranging from 22 to 74

years, together with 26, COVID-positive, male patients with ages ranging from 32 to 86 years. These patients' images were selected to provide well-balanced distributions of patients' genders and ages.

## 2.4 | CT-scan-slice image assessment

### 2.4.1 | Classifications involving visual assessment by clinicians

The suites of CT-slice images from each patient were assessed by a clinician (visual inspection). That assessment resulted in each image being assigned a number from 0 to 4 (five classes in total) to distinguish the degree of severity of visible, abnormal lung features they exhibited. Those five classes are defined as follows: *Class 0*, COVID-19-negative test lacking visual traces of abnormal lung features; *Classes 1–4*, apply to all patients with COVID-19-positive tests; *Class 1*, none, or trace, indications visible of abnormal lung features; *Class 2*, minimal but visible indications of abnormal lung features; *Class 3*, sporadic but pervasive visible signs of abnormal lung features; *Class 4*, extensive visible signs of severe abnormal lung features.

This visual-scoring (VS) classification scheme is considered as the initial objective function for predictive DL/ML algorithms assessed using grayscale-image attribute statistics.

Example CT-scan-slice images with associated rectangular extract images for VS Classes 0 and 1 (minimal or no visual sign of lung abnormalities) are displayed in Supporting Information: Figure S1. Rectangular-extract segments from the images displayed in Supporting Information: Figure S1 are representative of those used for grayscale—statistical analysis are shown in Supporting Information: Figure S2.

### 2.4.2 | Rectangular-extract-image segment analysis

A total of 392 rectilinear image extracts were compiled from the 49 COVID-19-positive individuals (approximately eight extracts/person). Additionally, 121 rectilinear image extracts were compiled from the eight COVID-19-negative individuals (approximately 15 extracts/person).

The dimensions of the 513 rectilinear images range from 2000 to 80,000 pixels, with average dimensions of approximately 25,000 pixels. This range occurs because the CT-image extracts are positioned to sample just the parenchyma portion of each lung. The rectilinear extracts are also positioned to avoid the pleura, diaphragm, or mediastinum portions of the lung. Consequently, areal extent in each CT image extract varies.

Every pixel in a monochromatic CT image possesses one grayscale value varying from 0 (black) to 255 (white). Values on that scale from 1 to 254 represent progressively lighter shades of gray. The pixel values are expressed in an 8-bit-integer digital format. This simple grayscale-image intensity scale means that the collective sets of pixels making up an image form grayscale distributions. Those distributions can be assessed with statistical confidence as thousands of pixels are involved in each image extract. OpenCV functions coded in Python are employed in this study to conduct a comprehensive grayscale statistical analysis of each image extract.<sup>27</sup> These are (1) the number of pixels present in an image; (2) average pixel grayscale value (a number between 0 and 255 on the grayscale); (3) the number of pixels in an image possessing its average pixel grayscale value; (4) percent of pixels in an image possessing its average pixel grayscale value; (5) variance of all pixel grayscale values in an image (a number between 0 and 255 on the grayscale); (6) ratio of pixel variance to pixel average pixel in an image; (7) standard deviation of all pixel grayscale values in an image (a number between 0 and 255 on the grayscale); (8) standard error of the mean of all pixel grayscale values in an image; (9) minimum grayscale value of all pixels in an image; (10) tenth percentile (P10) of the grayscale values of all pixels in an image; (11) fiftieth percentile (P50) of the grayscale values of all pixels in an image; (12) ninetieth percentile (P90) of the grayscale values of all pixels in an image; (13) maximum grayscale value of all pixels in an image.

Clearly, the number of pixels in an image extract that possesses the average pixel grayscale value (attribute#3), and the standard error of the mean (attribute#8), are statistics that are dependent on the image extract size. Attribute#8 is calculated as the

grayscale standard deviation divided by the square root of the number of pixels in the image extract. Apart from the two mentioned, all other grayscale attribute statistics listed are not dependent of image size.

The standard error of the mean values are an indicator of the degree of uncertainty associated with the average (mean) grayscale value of each image. The standard error of the mean values is all <0.7 for the images analyzed. Compared to the grayscale range of 0–255, a standard error of the mean of <0.7 confirms that there is a very limited uncertainty range associated with the average (mean) grayscale value of all the image extracts.

### 2.4.3 | Deep/machine learning (DL/ML) models configured to exploit grayscale statistics

Twelve of the grayscale attribute statistics determined for each image extract are used as input variables for the DL/ML analysis to predict the five VS Classes (0–4) as a dependent variable. The number of pixels in each image is not used as an input variable as it is dependent on image size. Each DL/ML classifier is configured and executed in Python. The objective function minimized by the DL/ML algorithms is the mean squared error (MSE) for between the actual (clinician determined) VS class ( $VS_{act}$ ) and the predicted VS class ( $VS_{pred}$ ) taking into account the full suite of 500 and 13 image extracts.

The 10 ML methods and 1 DL method employed are described in Table 1.

These algorithms are well known and extensively applied for many industrial and academic purposes, including image analysis and classification. Published

**TABLE 1** Description of deep learning/machine learning (DL/ML) algorithms applied to evaluating CT-scan-extract images based on grayscale attribute input variables.<sup>28–45</sup>

Algorithm name	Abbreviation	Description
Adaptive Boosting “Adaboost”	ADA	Boosted decision tree <sup>28,29</sup>
Convolutional Neural Network	CNN	DL neural network <sup>18,19</sup>
Decision Tree	DT	A single tree configuration <sup>30–32</sup>
Extreme Learning Machine	ELM	Single-layer feed-forward neural network <sup>33–35</sup>
Gaussian Process Classification	GPC	Gaussian probability applying Laplace approximation <sup>36,37</sup>
K-nearest Neighbor	KNN	Data matching algorithm <sup>15,38</sup>
Multilayer perceptron	MLP	Classic neural network <sup>39</sup>
Naïve Bayes Classifier	NBC	Simple Bayes’ theorem classifier <sup>15,40</sup>
Random Forest	RF	Tree-based ensemble method using bagging <sup>41,42</sup>
Quadratic Discriminant Analysis	QDA	Nonlinear classifier <sup>43</sup>
Support Vector Classifier	SVC	Non-probabilistic classifier <sup>44,45</sup>

examples of their specific applications to classify image datasets are cited in Table 1. Distinct mathematical, statistical, and logical concepts are associated with these algorithms. Most of these algorithms derive and build their predictions upon hidden regression relationships, driven by correlations between the input variables and dependent variables. The K-nearest Neighbor (KNN) algorithm is an exception in that it uses data-record matching to establish its predictions rather than correlations.

Each of the DL/ML algorithms requires its hyperparameters to be optimized. These algorithm control values were established using trial and error and grid search techniques (SciKit-Learn).<sup>46</sup> The configurations of each algorithms, expressed in terms of their optimized hyperparameters, are listed in Supporting Information: Table S1.

Trial-and-error and K-fold cross-validation analysis was used to determine the optimum allocation of data records between the training subset and testing subset evaluated by each of the DL/ML methods. That analysis indicated that a random 80% allocation of data records to the training subset with the remaining 20% of data records assigned to the testing subset worked effectively with the CT-scan-extract image data set.

#### 2.4.4 | Statistical metrics used to assess the classification accuracy of DL/ML models

The following statistical measures of classification accuracy are used in this study to determine the VS classification performance achieved by each DL/ML method applied.

*Mean square error (MSE)*

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n ((X_i) - (Y_i))^2, \quad (1)$$

where  $X_i$  is the known value and  $Y_i$  is the predicted value of the  $i$ th data record, and  $n$  is the number of data records in the data set.

*Root mean squared error (RMSE)*

$$\text{RMSE} = \sqrt{\text{MSE}}. \quad (2)$$

*Mean absolute error (MAE)*

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |X_i - Y_i|. \quad (3)$$

*Percent deviation ( $PD_i$ )*

$$\text{PD}_i = \frac{X_i - Y_i}{X_i} \times 100. \quad (4)$$

*Average percent deviation (APD)*

$$\text{APD} = \frac{\sum_{i=1}^n \text{PD}_i}{n}. \quad (5)$$

*Absolute average percent deviation (AAPD)*

$$\text{AAPD} = \frac{\sum_{i=1}^n |\text{PD}_i|}{n}. \quad (6)$$

*Standard deviation (SD)*

$$\text{SD} = \sqrt{\frac{\sum_{i=1}^n (D_i - D_i \text{mean})^2}{n - 1}}, \quad (7)$$

where  $D_i$  is  $(X_i - Y_i)$  for the  $i$ th data record.

$D_i$  mean is the mean of the  $D_i$  values of all the data records in a data set.

$$D_i \text{ mean} = \frac{1}{n} \sum_{i=1}^n (X_i - Y_i). \quad (8)$$

*Correlation coefficient (R)*

$$R = \frac{\sum_{i=1}^n (X_i - X_{\text{mean}}) (Y_i - Y_{\text{mean}})}{\sqrt{\sum_{i=1}^n (X_i - X_{\text{mean}})^2 \sum_{i=1}^n (Y_i - Y_{\text{mean}})^2}}. \quad (9)$$

$R$  between variables  $X_i$  and  $Y_i$  is expressed on a scale between  $-1$  and  $+1$ .

*Coefficient of determination*  $= R^2$ , which is expressed on scale between  $0$  and  $1$ .

#### 2.4.5 | Workflow of the methodology applied to assess CT images

Figure 1 summarizes the workflow applied in this study to assess and classify thoracic CT images in terms of their COVID-19 impacts.

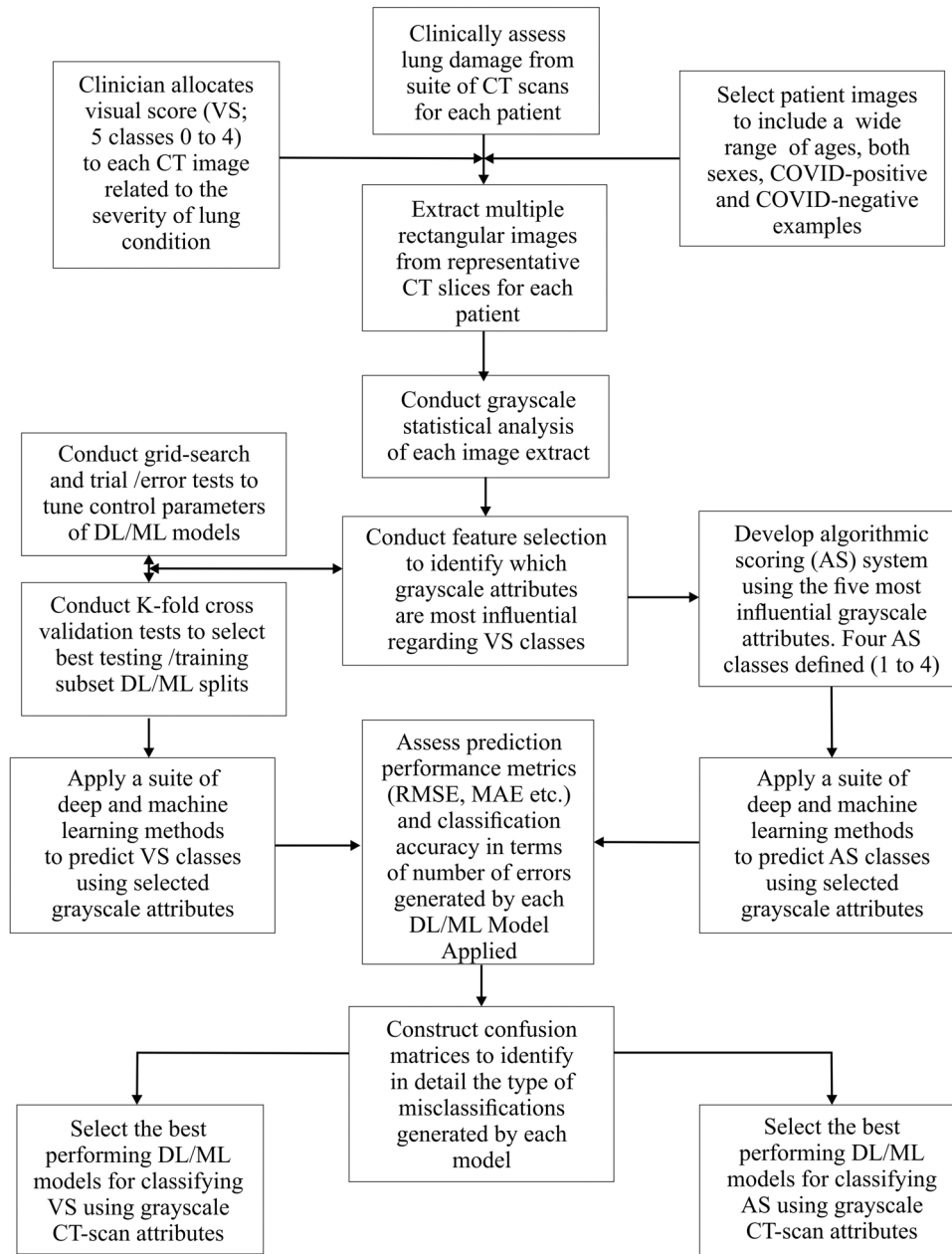
## 3 | RESULTS

### 3.1 | Grayscale image statistics

The grayscale statistical-attribute distributions of the CT-extract images evaluated are described in Supporting Information: Table S2. Figure 2 displays and compares box and whisker plots for selected CT-extract-image grayscale attribute distributions relating to the 121 COVID-19-negative individuals and the 392 COVID-19-positive individuals.

Pixel P50, pixel average, and pixel P90 grayscale statistics display some the most obvious distinctions between distributions for COVID-19-negative and COVID-19-positive individuals (Figure 2). It is these more obvious differences, plus the more subtle differences associated with other grayscale attribute statistics that the DL/ML models use in their VS class predictions.

## Workflow Applied to Predict Severity of COVID-19-Related Lung Conditions from Grayscale Statistical Analysis of CT Scan Images

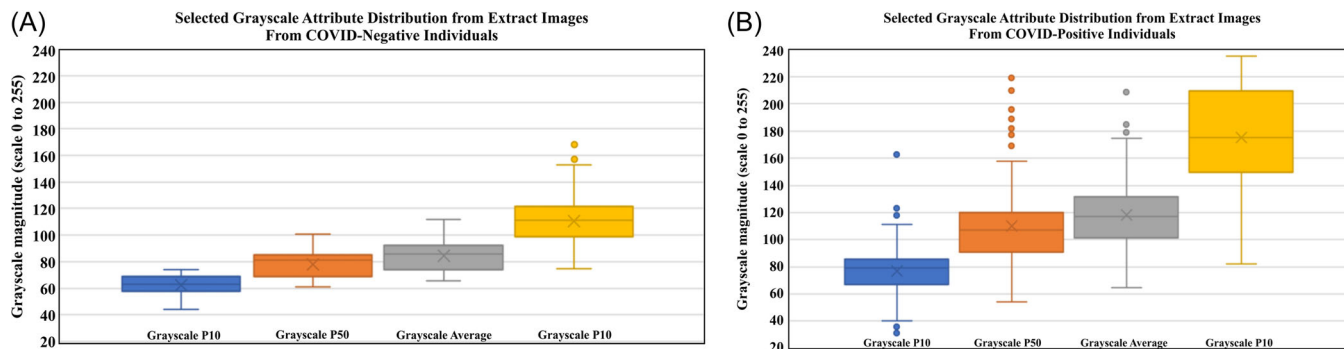


**FIGURE 1** Workflow diagram describing the procedures applied in the multiclass lung-condition-severity prediction analysis of computed tomograph images using grayscale statistics.

Some of the grayscale statistics relating to the CT-extract images also display strong correlations with their VS classification (Table 2). The Pearson's correlation coefficient ( $R$ ) values are highly positive for VS versus grayscale, P50, P90, average, variance, and standard deviation.  $R$  values for VS versus P10, standard error, and variance/average ratio are moderately positive. On the other hand, the  $R$  value for VS versus the percentage of pixels at the average grayscale value is highly negative. The dispersion of the grayscale-attribute distributions, and their correlations with VS, are

indicative of systematic variations occurring in these attributes that are spread across the VS classes throughout the CT-extract-image data set.

Spearman's rank correlation coefficient ( $\rho$ ) is also calculated for VS versus the grayscale statistical attributes and displayed in Table 2. A feature of the  $R$  statistic is that an assumption is made that the two-variable distributions compared both conform to normal distributions, that is, they display parametric behavior. The nonparametric  $\rho$  statistic avoids that assumption as it is calculated based on the ranked



**FIGURE 2** Box-and-whisker plots of selected grayscale attributes: (A) distributions for 121 COVID-19-negative computed tomography (CT) images; (B) distributions for 392 COVID-19-positive CT images. The boxes bracket the second and third quartiles of the distributions, the crosses are the averages, the lines within the boxes are the medians, the ends of the whiskers are minimum and maximum values, and the dots are considered as outlying values. P10, 10th percentile; P50, 15th percentile; P90, 19th percentile.

**TABLE 2** Correlation coefficients for 513 CT-scan-extract images grayscale statistical attributes versus VS.

Image grayscale statistical attributes	Attribute identifier	Pearson correlation coefficient ( <i>R</i> ) with VS	Spearman correlation coefficient ( <i>p</i> ) with VS
Number of Pixels in Extract	1	-0.0882	-0.1191
Grayscale Average	2	0.7834	0.8059
Number of Pixels at Grayscale Average	3	-0.5140	-0.5552
Percentage of Pixels at Grayscale Average	4	-0.7456	-0.7978
Grayscale Variance	5	0.6749	0.7184
Grayscale Variance/Average	6	0.3740	0.4117
Grayscale Standard Deviation	7	0.6911	0.7184
Grayscale Standard Error	8	0.4148	0.4539
Grayscale Minimum	9	0.1837	0.2373
Grayscale P10	10	0.5511	0.5640
Grayscale P50	11	0.7242	0.7534
Grayscale P90	12	0.8667	0.8776
Grayscale Maximum	13	0.3162	0.2236

Note: The Spearman correlation coefficient uses the same formula as the Pearson correlation coefficient (Equation 9) but uses rank positions in the distributions rather than absolute values.

Abbreviations: CT, computed tomography; VS, visual scores.

position of each data record in the respective distributions instead of their absolute values. For VS versus the grayscale attributes the *p* and *R* values are in relatively close agreement (Supporting Information: Figure S3). This implies that the grayscale attribute distributions are relatively symmetrical and not highly skewed.

### 3.2 | CT-extract image grayscale attribute dispersions across VS classes

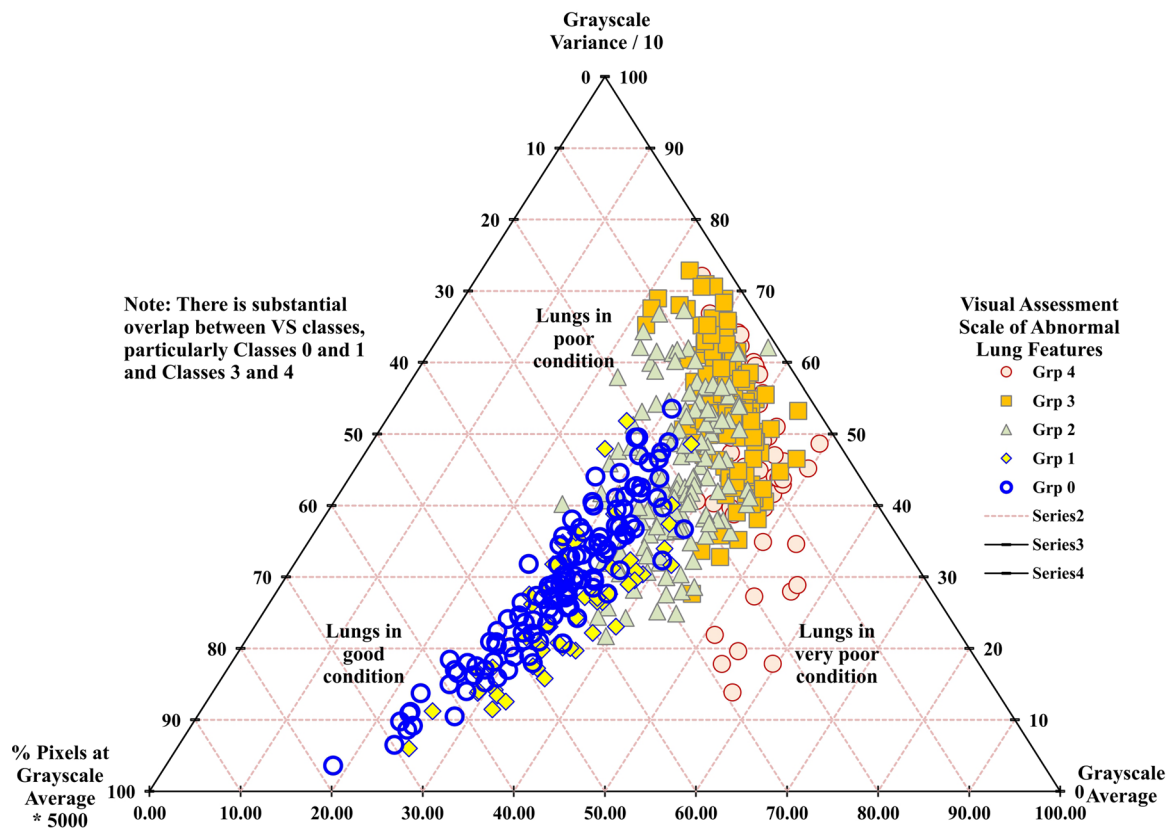
Figure 3 displays a ternary plot for three grayscale attributes: average; variance; and, pixel% at the average

value. The scaling factors applied to the grayscale attributes serve to centralize the spread of data points on this ternary plot. In such plots, the actual values of the three variables are normalized to sum to a constant value (the value of 100 is used for Figure 3).

These normalized values cannot then vary in the diagram independently of each other. Thus, if the normalized values of two of the variables plotted are known, the value of the third can be determined. This configuration creates only two degrees of freedom in the normalized system, allowing the three variables to be displayed in the two-dimensional ternary plots (Figure 3).

The three grayscale attributes (average; variance; pixel % at the average) perform relatively good discrimination

## Lung CT-scan-extract Image Selected Grayscale Attribute Relationships With Visual Classification Displayed



**FIGURE 3** Ternary plot of three grayscale statistical attributes displaying high correlation coefficients ( $R$  and  $p$ ) with visual scores (VS) class (Table 2). VS Class 0 = COVID-negative; VS Classes 1–4 refer to increased lung abnormalities in COVID-positive individuals.

between the VS classes (Figure 3). There is an almost continuous progression from VS = 0 (lower-left in ternary plot) to VS = 4 (middle-right in ternary plot). VS Classes 0 and 3 appear in distinct regions on the ternary plot. Nevertheless, overlaps are apparent among VS Classes 1, 2, and 4 and between those classes and VS Classes 0 and 3. Lungs with little or no evidence of abnormal features tend to be associated with relatively low average and variance grayscale values. As abnormal features materialize in the lung-image extracts, grayscale average and variance values progressively increase. At the stage where abnormal features become extensive in the lung-image extracts, the percentage of image pixels at the average grayscale for the image tends to decrease. This tendency forms a distinctive characteristic in the VS Class 4 images.

The images displaying the severest abnormal lung features plot towards the lower-right portion of the ternary plot (Figure 3). Such images show a prevalence of lighter-gray pixels. This causes the grayscale variance value to decline and the grayscale average value to increase. From VS = 0 to VS = 3 the image positions tend to move from south-west to north-east position in Figure 3. Whereas, from VS = 3 to VS = 4 the image positions tend to move in a southerly direction in Figure 3.

Three-variable plots (Figures S4 and S5) reveal other progressive and distinctive trends that exist involving grayscale percentage statistics. The involvement of grayscale P90 in Supporting Information: Figure S4 displays a continuous trend from the base-right portion (VS = 0) to top-left left portion (VS = 3 or 4) with substantial overlap between VS Classes 3 and 4. By involving both grayscale P10 and P90 with grayscale average (Figure S5) the overlap between classes VS = 3 and VS = 4 is somewhat reduced.

The distributions of grayscale statistics from CT-extract images (Supporting Information: Table S2, Figures S4 and S5, and Figure 3) include substantial dispersion and inter-relationships. These can be exploited graphically (and by DL/ML methods) to classify the degree of abnormalities in lungs affected by COVID-19. The three-statistical-attribute combinations displayed in Figures 3, Supporting Information: S4 and S5 discriminate quite well between some of the VS classes. However, overlapping positions, particularly between VS Classes 0 and 1 and VS Classes 3 and 4 make a graphical approach unsuited to definitive VS classification. Involving DL/ML



methods to consider multiple grayscale attributes as input variables offers the potential to improve upon the graphical approach.

### 3.3 | Feature selection for DL/ML analysis

Thirteen grayscale-attribute statistics were extracted and computed for each of the (513) CT-extract images (Supporting Information: Table S2). All of these attributes could be employed as independent (input) variables with VS class as the single dependent variable for DL/ML model evaluation. However, as some of these attributes show relatively low correlations with VS class (Table 2 and Supporting Information: Figure S3) trial and error sensitivity analysis was conducted to determine their impact on the DL/ML prediction performance. The following sensitivity tests were conducted:

Test #1: 9 input variables (pixel#, standard error, minimum, variance/average ratio excluded)

Test #2: 10 input variables (pixel#, standard error, variance/average ratio excluded)

Test #3: 11 input variables (pixel# and variance/average ratio excluded)

Test #4: 12 input variables (pixel# excluded)

Test #5: 13 input variables

The results showed that Test #4 with 12-input variables included achieved the highest VS prediction accuracy. This suggests that all 12 of those attributes can provide useful input to the DL/ML models. Those 12 input features were, therefore, selected for detailed DL/ML model analysis.

### 3.4 | K-fold cross-validation analysis

K-fold cross-validation is applied to evaluate, select and justify the percentage splits of the data records between training and testing subsets to be applied in ML/DL analysis. Four specific folds or cases are evaluated for five algorithms. Those cases are: fourfold (75% training: 25% testing), fivefold (80% training: 20% testing), 10-fold (90% training: 10% testing), and 15-fold (93% training: 7% testing). Three runs are executed for each k-fold analysis case, thereby generating meaningful statistics for each case. Twelve distinct splits are evaluated for fourfold case (3 times 4). The number of splits evaluated increases to 45 for the 15-fold analysis (3 times 15). The mean MAE and the standard deviation MAE are recorded for all the splits evaluated for each fold (Table 3).

All four K-folds evaluated generate credible prediction results for the VS and algorithmic scale (AS) classifications (low MAE means and low MAE standard deviations). However, the fivefold (80%:20%) split is selected as this generates the lowest mean MAE and a lower standard

deviation than the 10- and 15-fold analysis. Also, the 80%:20% split involves a substantial number (103) of randomly selected images in each testing subset.

### 3.5 | DL/ML model VS class prediction performance

Table 4 shows the VS class prediction accuracy ( $VS_{pred}$  vs.  $VS_{act}$ ) for the 11 DL/ML models developed. The results refer to the trained and validated models applied to the entire data set. The number and percentages of misclassified data records is also listed in Table 4. The algorithms are ranked, with the best performing shown as rank 1, and the worst-performing shown as rank 11. The rankings are made based on RMSE and classification error numbers (Table 4). The CNN model is ranked 1 in the VS classification task incurring only 18 misclassifications from the 513 images (96.49% accuracy). Moreover, the CNN model achieved  $RMSE=0.19$  (with reference to the VS scale of 0–4) and  $R^2=0.98$ . The RF model was ranked second in its VS classifications (26 misclassifications;  $RMSE=0.26$ ;  $R^2=0.96$ ; 94.93% accuracy). Also, the ADA, DT, ELM, and KNN models performed well in classifying VS. On the other hand, the MLP, NBC, and QDA did not perform well in classifying VS. The percentage accuracies achieved for this multiclass VS scale prediction analysis compare favorably with the percentage accuracies achieved by recent deep learning studies published for the binary (COVID negative vs. COVID positive) predictions from CT and X-ray scans. For instance, Arora et al.<sup>6</sup> analyzed 812 CT images and achieved a binary prediction accuracy of 98%; Polsinelli et al.<sup>24</sup> analyzed 100 images using CNN and achieved a binary prediction accuracy of 85%; Dansana et al.<sup>23</sup> analyzed 360 CT images using CNN and achieved a binary prediction accuracy of 91%; and Bharati et al.<sup>22</sup> analyzed 3411 CT images using DL models and achieved a binary prediction accuracy of 82.42%. On the other hand, Bharati et al.<sup>25</sup> analyzed 5606 X-ray images and achieved a binary prediction accuracy of 73%.

Table 4 reveals that certain ML models achieve different rankings based on RMSE compared to rankings based on error numbers. For example, the ELM model has an RMSE-rank of 4 but error rank of 6, whereas ADA has RMSE-rank of 5 but error rank of 4 (Table 4). This is a consequence of the nature of the misclassification errors involved. Misclassifications involving adjacent VS classes have a smaller impact on RMSE than misclassifications that are more than one class away from the correct class.

### 3.6 | Confusion matrix assessment of VS classification

Classification errors made by the DL/ML models may increase, while for the same analysis, RMSE values might decrease (Table 4). Confusion matrices deliver a detailed

**TABLE 3** K-fold cross-validation results highlighting the visual scale (VS) and algorithmic scale (AS) prediction performances of five algorithms for training and testing splits varying from 75%: 25% to 93%: 7%, respectively.

K-fold cross-validation results					Visual scoring		Algorithmic scoring	
Model	K-fold	Training: testing percentage splits	Splits evaluated in each run	Distinct cases for three repeated runs	MAE mean	MAE standard deviation	MAE mean	MAE standard deviation
ADA	4-fold	75:25	4	12	0.3841	0.0597	0.0286	0.0151
	5-fold	80:20	5	15	0.3248	0.0709	0.0256	0.0281
	10-fold	90:10	10	30	0.3588	0.0845	0.0287	0.0230
	15-fold	93:7	15	45	0.3607	0.0946	0.0338	0.0339
DT	4-fold	75:25	4	12	0.3613	0.0459	0.0176	0.0132
	5-fold	80:20	5	15	0.3346	0.0475	0.0135	0.0172
	10-fold	90:10	10	30	0.3527	0.0898	0.0150	0.0192
	15-fold	93:7	15	45	0.3470	0.0872	0.0150	0.0267
KNN	4-fold	75:25	4	12	0.3191	0.0372	0.0721	0.0197
	5-fold	80:20	5	15	0.2985	0.0512	0.0608	0.0216
	10-fold	90:10	10	30	0.3112	0.0819	0.0663	0.0235
	15-fold	93:7	15	45	0.3193	0.0988	0.0677	0.0377
RF	4-fold	75:25	4	12	0.2743	0.0322	0.0240	0.0098
	5-fold	80:20	5	15	0.2582	0.0479	0.0222	0.0173
	10-fold	90:10	10	30	0.2652	0.0846	0.0231	0.0200
	15-fold	93:7	15	45	0.2614	0.0893	0.0234	0.0267
SVM	4-fold	75:25	4	12	0.5563	0.0954	0.1170	0.0257
	5-fold	80:20	5	15	0.4532	0.0709	0.1063	0.0276
	10-fold	90:10	10	30	0.5392	0.1038	0.1092	0.0482
	15-fold	93:7	15	45	0.5361	0.1153	0.1040	0.0535

Note: Model abbreviations are listed in Table 1.

Abbreviations: ADA, Adaptive Boosting “Adaboost”; AS, algorithmic scale; DT, Decision Tree; KNN, K-nearest Neighbor; MAE, mean absolute error; RF, Random Forest; SVM, support vector machine.

assessment of multiclass misclassifications on a VS class by class basis. Confusion matrices reveal VS classes predicted with most and least accuracy. The three models considered (CNN; KNN; ADA; Figure S6) perform quite distinctly. CNN (Figure S6A) performs best for VS Class 0 (just 1 misclassification) and, in percentage error terms worst for VS Class 1, although it misclassifies the most data records (6 out of 129) associated with VS Class 3, CNN predicts VS Classes 0, 1, and 2 more convincingly than VS classes 3 and 4. CNN misclassifications only involve erroneous predictions involving adjacent VS classes. Confusion matrices for KNN (Figure S6B) and ADA (Figure S6C) reveal that both models result in 28 misclassifications associated with the 513 CT-extract images evaluated. Nevertheless, the distribution of these misclassifications differs considerably for those two models. KNN has five misclassifications that are beyond the adjacent VS classes, whereas ADA results in eight such extreme misclassifications. Consequently, KNN (RMSE = 0.2887) outperforms ADA (RMSE = 0.3175) overall in its VS

class prediction. KNN and ADA struggle most to correctly classify VS Class 2 (11 misclassifications; about 40% of total misclassifications). KNN performs best in classifying VS Classes 0 and 1 (better than CNN). ADA does not deal as well with VS Class 0 as the CNN and KNN models but performs almost as well as the CNN model in its classification of VS Classes 3 and 4.

Analysis of confusion matrices provides a complementary perspective on the RMSE/error count analysis of each model in assessing their overall multiclass classification prediction strengths and weaknesses.

## 4 | DISCUSSION

### 4.1 | Formulaic/algorithmic scoring approach to classifying CT-extract images

The relatively good discriminatory performances among the VS classes of the graphical relationships (Figure 3,

**TABLE 4** VS Classification results for trained and validated deep learning/machine learning (DL/ML) models using 12-input-variable grayscale statistical attributes applied 513 CT-extract images.

Prediction accuracy visual score (VS: 0–4) using 12 grayscale input variables										
ML/DL Model	RMSE (VS)	MAE (VS)	APD (%)	AAPD (%)	Sdev (VS)	$R^2$	Number of errors	% correct	Rank by RMSE	Rank by errors
ADA	0.3175	0.0698	-0.6686	3.0459	0.3186	0.9445	28	94.54	5	4
CNN	0.1868	0.0349	0.3165	1.2209	0.1869	0.9805	18	96.49	1	1
DT	0.3294	0.0775	-0.3424	2.9199	0.3305	0.9404	32	93.76	6	5
ELM	0.3144	0.0872	-1.3372	3.7984	0.3156	0.9453	42	91.81	4	6
GPC	0.4176	0.1589	0.2778	4.8643	0.4175	0.9041	78	84.80	7	8
KNN	0.2887	0.0640	0.3230	2.1641	0.2890	0.9542	28	94.35	3	4
MLP	0.9139	0.5291	12.9167	18.7629	0.8225	0.6949	196	61.79	11	11
NBC	0.7587	0.4205	-9.3346	21.8992	0.7609	0.7024	177	65.50	10	10
QDA	0.5171	0.2054	-5.1583	9.8934	0.5184	0.8524	90	82.46	9	9
RF	0.2604	0.0562	-0.6944	2.3094	0.2614	0.9620	26	94.93	2	2
SVM	0.4313	0.1202	0.7526	4.6932	0.4285	0.9005	46	91.03	8	7

Note: VS system involves classes 0 to 4 assigned by a clinician.

Abbreviations: AAPD, absolute average percentage deviation; ADA, Adaptive Boosting “Adaboost”; APD, average percentage deviation; CNN, Convolutional Neural Network; DT, Decision Tree; ELM, Extreme Learning Machine; KNN, K-nearest Neighbor; MAE, mean absolute error; RMSE, root mean squared error; Sdev, standard deviation; SVM, support vector machine. Model abbreviations are listed in Table 1.

Supporting Information: Figures S4 and S5), involving just grayscale attributes, makes them exploitable for automated formulaic classification. Such an algorithmic/formulaic scoring system (AS) has the potential to provisionally classify the extent of abnormal lung features in CT-extract images before inspection by a clinician and/or DL/ML analysis. AS scores can complement the VS score assigned by a clinician (i.e., VS based on clinician's expert evaluation; AS based on a defined algorithm). Automating CT-extract-image assessments has the attraction of speed and objectivity. This can provide provisional guidance for a subsequent more detailed expert visual assessment.

An example AS approach involves just five grayscale attributes (P10, average, P90, variance and pixel% at the average value; attributes 10, 2, 12, 5, and 4 in Table 2 and Supporting Information: Figure S3, respectively) corresponding to those used in Figure 3, Supporting Information: Figure S4 and S5 for VS analysis.

AS distinguishes only four classes of lung abnormality. Class 1 corresponds to the least abnormality whereas Class 4 refers to the most extreme abnormality. There is no Class 0, which is included in the VS system to distinguish individuals with COVID-19-negative test results. The more generic AS is defined to address the extent of abnormal lung features, irrespective of the cause. The CT-extract images classified as either VS Class 0 or 1 tend to be classified as AS Class 1.

1. The AS formulas involve just four groups of rules: AS = 4 (extensive presence of abnormal lung features)

is established for those images that satisfy three grayscale attribute levels: P10 grayscale  $\geq 100$ ; average grayscale  $\geq 150$ ; P90 grayscale  $\geq 200$ .

2. AS = 1 (few if any abnormal lung features) is then established (after AS = 4 images have been allocated and removed from the selection) for those remaining images that satisfy four grayscale attribute levels: P10 grayscale  $< 80$ ; P90 grayscale  $< 125$ ; variance  $< 1000$ ; Pixel% at the average value  $> 1.5\%$ .
3. AS = 2 (minor abnormal lung features) is then established (after AS = 4 and AS = 1 images have been allocated and removed from the selection) for those remaining images that satisfy two grayscale attribute levels: Average grayscale  $< 125$ ; P90 grayscale  $< 150$ .
4. AS = 3 (substantial abnormal lung features) is then established (after AS = 4, AS = 1, and AS = 2 images have been allocated and removed from the selection) with all remaining CT-extract images not assigned to other groups.

Table 5 and Supporting Information: Figure S7 displays the Pearson and Spearman correlation coefficients between the grayscale attribute variables and the AS classes for all CT-extract images. As should be expected, the  $R$  and  $p$  values are slightly higher than those for correlations between the same variables and the VS classes (Table 2). This is particularly the case for those grayscale variables involved in the formulaic definition of the AS classes.

**TABLE 5** Correlation coefficients for 513 CT-scan-extract images grayscale statistical attributes versus AS.

Image grayscale statistical attributes	Attribute identifier	Pearson correlation coefficient ( <i>R</i> ) with algorithmic score (AS)	Spearman correlation coefficient ( <i>p</i> ) with AS
Number of Pixels in Extract	1	-0.0980	-0.1344
Grayscale Average	2	0.8202	0.8531
Number of Pixels at Grayscale Average	3	-0.5560	-0.5682
Percentage of Pixels at Grayscale Average	4	-0.7786	-0.7876
Grayscale Variance	5	0.6559	0.7340
Grayscale Variance/Average	6	0.3661	0.4144
Grayscale Standard Deviation	7	0.6988	0.7340
Grayscale Standard Error	8	0.4364	0.4871
Grayscale Minimum	9	0.2656	0.3304
Grayscale P10	10	0.6338	0.6516
Grayscale P50	11	0.7459	0.7897
Grayscale P90	12	0.9015	0.9174
Grayscale Maximum	13	0.3821	0.2892

The 513 CT-extract images assigned to AS classes are shown in Figure 4. An increased degree of segregation of the AS classes is discernible in Figure 4 compared to that achieved for the VS classes in Figure 3. This outcome confirms that the value of AS is a preliminary, automated, objective CT-extract image assessment. The five-attribute AS classification system can be further improved upon by applying DL/ML algorithms to consider all the available attributes.

Table 6 summarizes AS classification prediction accuracy determined by DL/ML models using the same 12 input variables (grayscale attributes) recorded for each of 513 CT-extract images. The results confirm that the DL/ML methods find it easy (fewer misclassifications and much lower RMSE/MAE values) to classify the CT-extract images using the AS classes than the VS classes.

Two of the ML algorithms, RF and DT, make only one classification error (out of 513 CT-extract assignments) in the AS classification exercise. Those two models achieved 99.81% accuracy in classifying AS correctly. That level of classification accuracy is better than that achieved for the VS scale (Section 3.5). It also compares very favorably with the accuracy achieved by published DL studies of CT scans focused on the binary (COVID negative vs. COVID positive) classifications,<sup>6,22-25</sup> as already mentioned in Section 3.5.

The CNN (DL model) ranks 6 out of 11 for its AS classification performance. CNN is outperformed by the RF, DT, ADA, GPC, and KNN models for AS classification (Table 6). Nevertheless, the CNN model

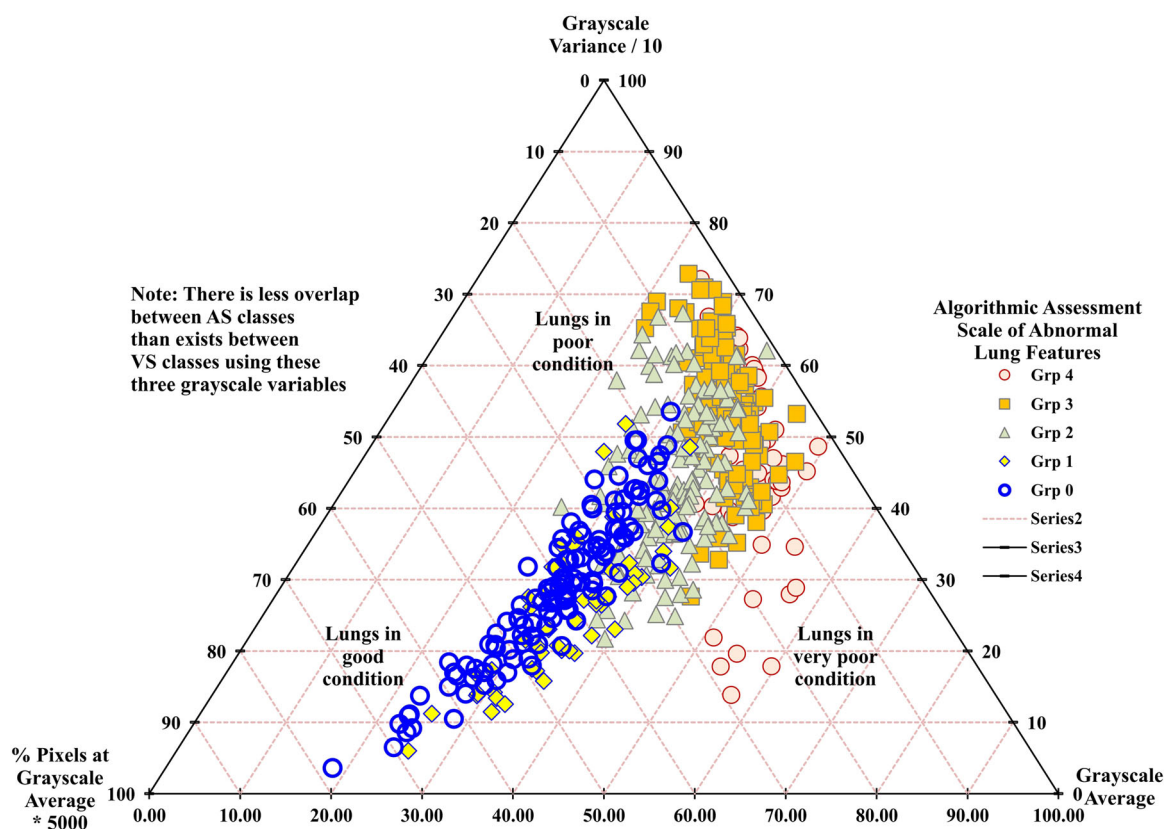
still achieves and AS classification accuracy of >98%. The confusion matrix analysis (not shown) reveals that DT misclassifies just one CT-extract image of AS = 2 as belonging to Class 1, with RF misclassifying a CT-extract image of AS = 4 as belonging to Class 3.

Alternative combinations of grayscale attributes (distinct from the five attributes used to define AS), and/or alternative classification scales including more or less than four classes, could also be explored. It is possible that such alternative algorithmic systems could also provide useful algorithmic segregation of the CT-extract images in terms of their severity of lung abnormalities. Further research is justified to evaluate such possibilities. Notwithstanding possible alternative algorithmic approaches, the AS system, as defined, is considered to provide a viable algorithmic CT-extract image classification system based on easy-to-measure grayscale attributes.

## 4.2 | Value of DL/ML models in predicting VS and AS classification systems

VS class definitions benefit from the input of a clinician's expertise, potentially taking account of a wider set of information than just image grayscale attributes. However, human judgement, irrespective of the expertise involved, makes VS class assignments by visual inspection somewhat subjective. This means that classifications are likely to vary, in detail, among clinicians. The magnitude of abnormal lung features associated with COVID-19 clearly varies over a broad range, as reflected by the grayscale attribute

## Lung CT-scan-extract Image Selected Grayscale Attribute Relationships With Algorithmic Classification Displayed



**FIGURE 4** Ternary plot of computed tomography (CT)-extract image grayscale average, variance, and pixel% at average grayscale with four algorithmic score (AS) classes distinguished. To be compared with Figure 3 highlighting that the AS classes can be more easily separated using these attributes than the classes related to the visual scores scale.

dispersions recorded (Figures 3, Supporting Information: Figures S4 and S5). Unfortunately, the measured grayscale attributes do not crisply segregate into clusters (e.g., to facilitate simple COVID-19-negative vs. COVID-19-positive classification). This makes the distinction of arbitrarily imposed visual-inspection classes unreliable.

In summary, the potential disadvantages of the VS method are (1) the risk of inconsistency and subjectivity in the VS class assignments of different clinicians; and (2) the slightly overlapping nature of the grayscale attribute clusters relating to the VS groups. Nevertheless, the classification performance of the CNN DL method applied to the prediction of VS classes is extremely good (just 18 misclassifications out of 513 CT-extract images evaluated; 96.49% accuracy). Based on that performance, evaluations with larger datasets may be able to further improve VS classification accuracy. However, it would be difficult to consistently achieve zero VS misclassifications due to the element of subjectivity involved the visual assessment process. By removing that subjectivity and applying a formulaic grayscale assessment, such as the

AS system defined and evaluated, more reliable image classification can be reached (Table 6 and Figure 4), achieving up to 99.81% accuracy by DT and RF models. Although the AS method overcomes the disadvantages of the VS method, a remaining disadvantage of the technique is that the extraction of the grayscale image slices from the CT scans currently lacks automation. Future developments planned for the technique are to automate the image slice extraction process.

A comparison of VS Class 0 (COVID-19-negative individuals) and VS Class 1 (COVID-19-positive individuals displaying minimal visual signs in terms of abnormal lung features) classification performance by DL/ML models is worth reflecting upon. CNN and KNN 12-variable models are able to distinguish between these classes with very few errors (Supporting Information: Figure S6). On the other hand, clinicians can only do this, in most cases, with the aid of a rRT-PCR COVID-19 test. This outcome highlights the power of the DL/ML models to use the CT-extract grayscale attributes to achieve valuable distinctions between the most difficult-to-separate VS classes based on visual assessment alone.

**TABLE 6** AS classification results for trained and validated deep learning/machine learning (DL/ML) models using 12-input-variable grayscale statistical attributes applied 513 CT-extract images.

Prediction accuracy algorithmic score (AS: 1-4) using 12 grayscale input variables										
ML/DL model	RMSE (VS)	MAE (VS)	APD (%)	AAPD (%)	Sdev (VS)	R <sup>2</sup>	Number of errors	% correct	Rank by RMSE	Rank by errors
ADA	0.0623	0.0039	0.1292	0.1292	0.0624	0.9955	2	99.61	3	3
CNN	0.1245	0.0155	0.1906	0.3844	0.1248	0.9821	8	98.44	6	6
DT	0.0440	0.0019	0.0646	0.0646	0.0442	0.9978	1	99.81	2	2
ELM	0.1460	0.0213	0.0258	0.6072	0.1465	0.9753	11	97.86	7	7
GPC	0.0984	0.0097	0.1421	0.2390	0.0986	0.9888	5	99.03	4	4
KNN	0.1165	0.0136	0.1098	0.3682	0.1168	0.9843	7	98.64	5	5
MLP	0.3294	0.1085	-0.2584	3.5853	0.3304	0.8760	56	89.08	11	11
NBC	0.3294	0.1085	-0.9496	3.7016	0.3306	0.8777	56	89.08	11	11
QDA	0.2567	0.0659	-0.6008	2.2416	0.2577	0.9243	34	93.37	9	9
RF	0.0440	0.0019	0.0388	0.0388	0.0442	0.9978	1	99.81	2	2
SVM	0.1761	0.0310	0.0032	1.0368	0.1766	0.9641	16	96.88	8	8

Note: AS system involves Classes 1-4 defined by formulaic/logical rules. Model abbreviations are listed in Table 1.

Abbreviations: AAPD, absolute average percentage deviation; APD, average percentage deviation; MAE, mean absolute error; RMSE, root mean squared error; Sdev, standard deviation; SVM, support vector machine.

There is clearly valuable grayscale attribute information available in the CT-scan images. The DL/ML methods offer a rapid and highly effective way of extracting that information in a way that could be used positively to better assess lung damage severity in COVID-19 patients and improve early decision-making on the most suitable care and treatment to administer. Although the use of DL methods for CT-image analysis and pattern recognition is expanding, ML methods should not be discounted as shown by the results of the DT and RF models with the AS data set described.

## 5 | CONCLUSIONS

Based upon grayscale attribute analysis of 513 CT-scan-extract images taken from 57 hospitalized individuals (49 with positive COVID-19 tests; 8 with negative COVID-19 tests) the following conclusions can be drawn: (1) Grayscale attributes can be used to classify, with reasonable accuracy using graphical methods, based a clinician's VS scale of VS = 0 (COVID-19-negative) to VS = 4 (COVID-19-positive with severe lung abnormalities visible) involving five classes. (2) Due to overlap among the VS classes, graphical analysis of grayscale image attributes does not on its own generate highly reliable VS classifications. (3) Deep- and machine-learning (DL/ML) methods can achieve much better VS classifications based on 12 grayscale image attributes than graphical analysis. (4) Of 11 DL/ML models evaluated, the CNN provided the best VS classifications; just 18 misclassifications

from the 513-image data set. This represents an accuracy of 96.49%. That multiclass prediction accuracy compares favorably with the range of published prediction accuracies (73%-98%) achieved for DL models configured to address the binary classification issue (COVID negative vs. COVID positive) using CT-scan and/or X-ray images. (5) A simple formula/logic-based AS system, based on just five grayscale-image attributes is able to separate all the CT-extract images evaluated into four classes; AS = 1 (minimal abnormal lung features) to AS = 4 (severe abnormal lung features) involving four classes without considering COVID-19 test information. (6) Applying DL/ML models with 12 grayscale attribute input variables to predict the AS classes, decision tree (DT) and random forest (RF) models outperformed the other models. Those two ML models did so by classifying all but one of the images into the correct AS classes. This represents a multiclass classification accuracy of 99.81%. That multiclass prediction accuracy exceeds the range of published prediction accuracies (73%-98%) achieved for DL models configured to address the binary classification issue (COVID negative vs. COVID positive) using CT-scan and/or X-ray images. (7) The CNN model performed less well classifying using AS scale and grayscale attributes than several of the ML methods evaluated but still achieved 98.44% accuracy. (8) The algorithmic approach, AS coupled with DL/ML models, offers the potential of providing an accurate, automated expert system for classifying into multiple classes the severity of lung abnormalities from CT-scan information based on grayscale image attributes.

Based on these findings, the methods proposed in this study, and grayscale attributes of CT scans in general, are worthy of further evaluation with larger datasets of CT images.

## AUTHOR CONTRIBUTIONS

*Clinical assessment of CT images and reviewing manuscript:* Sara Ghashghaei. *Methodology, CT image statistical analysis, machine and deep learning analysis, data interpretation, writing, reviewing, and editing subsequent drafts:* David A. Wood. *Reviewing and editing manuscript drafts:* Erfan Sadatshojaei. *Procurement and initial assessment of CT images and reviewing analysis:* Mansooreh Jalilpoor.

## ACKNOWLEDGMENT

No external funding was received in association with this research.

## CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT

There is no shared data available with this article.

## ETHICS STATEMENT

The data handling procedures of Namazi Medical Centre (Shiraz, Iran) were observed at all times. The images used for this study with procured with the patients' permissions with patient confidentiality and anonymity maintained.

## ORCID

Sara Ghashghaei  <https://orcid.org/0000-0001-5491-739X>

David A. Wood  <https://orcid.org/0000-0003-3202-4069>

Erfan Sadatshojaei  <https://orcid.org/0000-0002-1828-1694>

Mansooreh Jalilpoor  <https://orcid.org/0000-0001-7647-0541>

## REFERENCES

- Bai HX, Hsieh B, Xiong Z, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. *Radiology*. 2020;296(2):E46-E54. doi:10.1148/radiol.202000823
- Hare SS, Tavare AN, Dattani V, et al. Validation of the British Society of Thoracic Imaging guidelines for COVID-19 chest radiograph reporting. *Clin Radiol*. 2020;75(9):710.e9-710.e14. doi:10.1016/j.crad.2020.06.005
- Ai T, Yang Z, Hou H, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology*. 2020;296(2):E32-E40. doi:10.1148/radiol.202000642
- Fang Y, Zhang H, Xie J, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology*. 2020;296(2):E115-E117. doi:10.1148/radiol.202000432
- Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing. *Radiology*. 2020;296:200343. doi:10.1148/radiol.202000343
- Arora V, Ng EY, Leekh RS, Darshan M, Singh A. Transfer learning-based approach for detecting COVID-19 ailment in lung CT scan. *Comput Biol Med*. 2021;135:104575. doi:10.1016/j.combiomed.2021.104575
- Jiang Y, Guo D, Li C, Chen T, Li R. High-resolution CT features of the COVID-19 infection in Nanchong City: initial and follow-up changes among different clinical types. *Radiol. Infect. Dis*. 2020;7:71-77. doi:10.1016/j.jrid.2020.05.001
- Ng MY, Lee EYP, Yang J, et al. Imaging profile of the COVID-19 infection: radiologic findings and literature review. *Radiology*. 2020;2(1):e200034. doi:10.1148/ryct.2020200034
- Li K, Wu J, Wu F, et al. The clinical and chest CT features associated with severe and critical COVID-19 pneumonia. *Invest Radiology*. 2020;55(6):327-331. doi:10.1097/RLI.0000000000000672
- Farhat H, Sakr GE, Kilany R. Deep learning applications in pulmonary medical imaging: recent updates and insights on COVID-19. *Mach Vis Appl*. 2020;31:53. doi:10.1007/s00138-020-01101-5
- López-Cabrera JD, Orozco-Morales R, Portal-Díaz JA, Lovelle-Enríquez O, Pérez-Díaz M. Current limitations to identify COVID-19 using artificial intelligence with chest X-ray imaging. *Health Technol*. 2021;11(2):411-424.
- Chung M, Bernheim A, Mei X, et al. CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology*. 2020;295:202-207. doi:10.1148/radiol.202002030
- Duzgun SA, Durhan G, Demirkazik FB, Akpınar MG, Ariyurek OM. COVID-19 pneumonia: the great radiological mimicker. *Insights Imaging*. 2020;11:118. doi:10.1186/s13244-020-00933-z
- Yu Z, Li X, Sun H, et al. Rapid identification of COVID-19 severity in CT scans through classification of deep features. *Biomed Eng Online*. 2020;19:63. doi:10.1186/s12938-020-00807-x
- Hussain L, Nguyen T, Li H, et al. Machine-learning classification of texture features of portable chest X-ray accurately classifies COVID-19 lung infection. *Biomed Eng Online*. 2020;19:88. doi:10.1186/s12938-020-00831-x
- Shi F, Wang J, Shi J, et al. Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for COVID-19. *IEEE Rev Biomed Eng*. 2021;14:4-15. doi:10.1109/RBME.2020.2987975
- Dong D, Tang Z, Wang S, et al. The role of imaging in the detection and management of COVID-19: a review. *IEEE Rev Biomed Eng*. 2021;14:16-29. doi:10.1109/RBME.2020.2990959
- Li L, Qin L, Xu Z, et al. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. *Radiology*. 2020;296(2):E65-E71. doi:10.1148/radiol.202000905
- Song Y, Zheng S, Li L, et al. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *IEEE/ACM Trans Comput Biol Bioinform*. 2021;18:2775-2780. doi:10.1109/TCBB.2021.3065361
- Haque IRI, Neubert J. Deep learning approaches to biomedical image segmentation. *Inform Med Unlocked*. 2020;18:100297. doi:10.1016/j.imu.2020.100297
- Zhou T, Canu S, Ruan S. An automatic COVID-19 CT segmentation based on U-Net with integrated spatial and channel attention mechanism. *Int J Imaging Syst Technol*. 2021;31:16-27. doi:10.1002/ima.22527
- Bharati S, Podder P, Mondal MRH, Gandhi N. Optimized NASNet for diagnosis of COVID-19 from lung CT images. In: Abraham A, Piuri V, Gandhi N, Siarry P, Kaklauskas A, Madureira A, eds. *Intelligent Systems Design and Applications. ISDA 2020. Advances in Intelligent Systems and Computing*. Springer; 2021:1351. doi:10.1007/978-3-030-71187-0\_59
- Dansana D, Kumar R, Bhattacharjee A, et al. Early diagnosis of COVID-19-affected patients based on X-ray and computed tomography images using deep learning algorithm. *Soft Comput*. 2020;1-9. doi:10.1007/s00500-020-05275-y

24. Polsinelli M, Cinque L, Placidi G. A light CNN for detecting COVID-19 from CT scans of the chest. *Pattern Recognit Lett.* 2020; 140:95-100. doi:10.1016/j.patrec.2020.10.00
25. Bharati S, Podder P, Rubaiyat M, Mondal H. Hybrid deep learning for detecting lung diseases from X-ray images. *Inform Med Unlocked.* 2020;20:100391. doi:10.1016/j.imu.2020.100391
26. Amber Diagnostics. Different types of CT machines; 2022. Accessed March 27, 2022. <https://www.amberusa.com/blog/types-of-ct-machines/>
27. OpenCV. Open source computer vision library; 2021. Accessed March 27, 2022. <https://docs.opencv.org/master/d1/dfb/intro.html>
28. Barstugan M, Rahime Ceylan R, Sivri M, Erdogan H. Automatic liver segmentation in abdomen CT images using SLIC and AdaBoost algorithms. Proceedings of the 8th International Conference on Bioscience, Biochemistry and Bioinformatics: 2018;129-133. doi:10.1145/3180382.3180383
29. Kuwahara M, Kido S, Shouno H. Classification of patterns for diffuse lung diseases in thoracic CT images by AdaBoost algorithm. Proceedings Volume 7260, Medical Imaging: Computer-Aided Diagnosis. 2009;726037. doi:10.1117/12.811497
30. Agarwal C, Sharma A. Image understanding using decision tree based machine learning. Proceedings of the 5th International Conference on Information Technology & Multimedia. 2011. doi:10.1109/ICIMU.2011.6122757
31. Rajendran P, Madheswaran M. Hybrid medical image classification using association rule mining with decision tree algorithm. *J Comput.* 2020;2(1):127-136.
32. Yoo SH, Geng H, Chiu TL, et al. Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest X-ray imaging. *Front Med.* 2020;7:427. doi:10.3389/fmed.2020.00427
33. Albadr MAA, Tiun S, Ayob M, Al-Dhief FT, Omar K, Hamzah FA. Optimised genetic algorithm-extreme learning machine approach for automatic COVID-19 detection. *PLoS One.* 2020; 15(12):e0242899. doi:10.1371/journal.pone.0242899
34. Dogantakin A, Ozyurt F, Avci E, Koc M. A novel approach for liver image classification: PH-C-ELM. *Measurement.* 2019;137:332-338. doi:10.1016/j.measurement.2019.01.060
35. Huang GB, Zhou H, Ding X, Zhang R. Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics).* 2011;42: 513-529. doi:10.1109/TSMCB.2011.2168604
36. Zhu F, Carpenter T, Gonzalez DR, Atkinson M, Wardlaw J. Computed tomography perfusion imaging denoising using gaussian process regression. *Phys Med Biol.* 2012;57(12): N183-N198. doi:10.1088/0031-9155/57/12/N183
37. Hussein S, Gillies R, Cao K, Song Q, Bagci U. TumorNet: lung nodule characterization using multi-view Convolutional Neural Network with Gaussian Process. IEEE 14th International Symposium on Biomedical Imaging. 2017. doi:10.1109/ISBI.2017.7950686
38. Ramteke RJ, Khachane Monali Y. Automatic medical image classification and abnormality detection using K-nearest neighbour. *Int J Adv Comput Res.* 2012;2(4-6):190-196.
39. Kuruville J, Gunavathi K. Lung cancer classification using neural networks for CT images. *Comput Methods Programs Biomed.* 2014;113(1):202-209. doi:10.1016/j.cmpb.2013.10.011
40. Park DC. Image classification using naïve Bayes classifier. *Int J Comput Sci Eng.* 2016;4(3):135-139.
41. Suzuki K. Machine learning in computer-aided diagnosis of the thorax and colon in CT: a Survey. *IEICE Trans Inf Syst.* 2013; E96(D4):772-783. doi:10.1587/transinf.e96.d.772
42. Huidrom R, Chanu YJ, Singh KM. Pulmonary nodule detection on computed tomography using neuro-evolutionary scheme. *SIViP.* 2019;13:53-60. doi:10.1007/s11760-018-1327-4
43. Nedjar I, El Habib Daho M, Settouti N. Random forest based classification of medical X-ray images using a genetic algorithm for feature selection. *J Mech Med Biol.* 2015;15(2):1540025. doi:10.1142/S0219519415400254
44. Cyran KA, Kawulok J, Kawulok M, et al. Support vector machines in biomedical and biometrical applications. In: Ramanna S, Jain LC, Howlett RJ, eds. *Emerging Paradigms in ML. Smart Innovation, Systems and Technologies.* Springer; 2012:13. doi:10.1007/978-3-642-28699-5\_15
45. Zhang Z, Sejdic E. Radiological images and machine learning: trends, perspectives, and prospects. *Comput Biol Med.* 2019; 108:354-370. doi:10.1016/j.compbimed.2019.02.017
46. SciKit Learn. GridSearchCV: exhaustive search over specified parameter values for an estimator in Python 2022. Accessed March 27, 2022. [https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.GridSearchCV.html](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html)

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Ghashghaei S, Wood DA, Sadatshojaei E, Jalilpoor M. Grayscale image statistics of COVID-19 patient CT scans characterize lung condition with machine and deep learning. *Chronic Dis Transl Med.* 2022;8: 191-206. doi:10.1002/cdt3.27