Modeling a deep transfer learning framework for the classification of COVID-19 radiology dataset

Michael Adebisi Fayemiwo [1], **Toluwase Ayobami Olowookere** [1], **Samson Afolabi Arekete** [1], **Adewale Opeoluwa Ogunde** [Corresp., 1], **Mba Obasi Odim** [1], **Bosede Oyenike Oguntunde** [1], **Oluwabunmi Omobolanle Olaniyan** [1], **Theresa Omolayo Ojewumi** [1], **Idowu Sunday Oyetade** [1], **Ademola Adegoke Aremu** [2], **Aderonke Anthonia Kayode** [1]

[1] Department of Computer Science, Redeemer's University, Ede, Osun, Nigeria

[2] Radiology Department, Ladoke Akintola University of Technology, Ogbomoso, Oyo, Nigeria

Corresponding Author: Adewale Opeoluwa Ogunde
Email address: ogundea@run.edu.ng

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-Coronavirus-2 or SARS-CoV-2), which came into existence in 2019, is a viral pandemic that caused coronavirus disease 2019 (COVID-19) illness and death. Research showed that relentless efforts had been made to improve key performance indicators for detection, isolation, and early treatment. This paper used Deep Transfer Learning Model (DTL) for the classification of real-life COVID-19 dataset of chest X-ray images in both binary (COVID-19 or Normal) and three-class (COVID-19, Viral-Pneumonia or Normal) classification scenarios. Four experiments were performed where fine-tuned VGG-16 and VGG-19 Convolutional Neural Networks (CNNs) with DTL were trained on both binary and three-class datasets that contain images of X-ray. The system was trained with X-ray images dataset for the detection of COVID-19. The fine-tuned VGG-16 and VGG-19 DTL were modelled by employing a batch size of 10 in 40 epochs, and Adam optimizer for weight updates, categorical cross-entropy loss function. The result showed that the fine-tuned VGG-16 and VGG-19 models produced an accuracy of 99.23% and 98.00%, respectively, in the binary task. In contrast, in the multiclass (three-class) task, the fine-tuned VGG-16 and VGG-19 DTL models produced an accuracy of 93.85% and 92.92%, respectively. Moreover, the fine-tuned VGG-16 and VGG-19 models have MCC of 0.98 and 0.96 respectively in the binary classification, and 0.91 and 0.89 for multiclass classification. These results showed strong positive correlations between the models' predictions and the true labels. In the two classification tasks (binary and three-class), it was observed that the fine-tuned VGG-16 DTL model had stronger positive correlations in the MCC metric than the fine-tuned VGG-19 DTL model. The VGG-16 DTL model has a Kappa value of 0.98 as against 0.96 for the VGG-19 DTL model in the binary classification task, while in the three-class classification problem, the VGG-16 DTL model has a Kappa value of 0.91 as against 0.89 for the VGG-19 DTL model. This result is in agreement with the trend observed in the MCC metric. Hence, it was

discovered that the VGG-16 based DTL model classified COVID-19 better than the VGG-19 based DTL model. Using the best performing fine-tuned VGG-16 DTL model, tests were carried out on 470 unlabeled images dataset, which was not used in the model training and validation processes. The test accuracy obtained for the model was 98%. The proposed models provided accurate diagnostics for both the binary and multiclass classifications, outperforming other existing models in the literature in terms of accuracy, as shown in this work.

Modeling a Deep Transfer Learning Framework for the Classification of COVID-19 Radiology Dataset

Michael Adebisi Fayemiwo[1], Toluwase Ayobami Olowookere[1], Samson Afolabi Arekete[1],
Adewale Opeoluwa Ogunde[1], Mba Obasi Odim[1], Bosede Oyenike Oguntunde[1], Oluwabunmi
Omobolanle Olaniyan[1], Theresa Omolayo Ojewumi[1], Idowu Sunday Oyetade[1], Ademola
Adegoke Aremu[2], Aderonke Anthonia Kayode[1]

[1] Department of Computer Science, Redeemer's University, Ede, Osun State, Nigeria
[2] Radiology Department, Ladoke Akintola University of Technology, Ogbomoso. Oyo State.
Nigeria

Corresponding Author:
Adewale Opeoluwa Ogunde[1]
Akoda, Ede, Osun State, 232103, Nigeria
Email address: ogundea@run.edu.ng

Abstract

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-Coronavirus-2 or SARS-CoV-2), which came into existence in 2019, is a viral pandemic that caused coronavirus disease 2019 (COVID-19) illness and death. Research showed that relentless efforts had been made to improve key performance indicators for detection, isolation, and early treatment. This paper used Deep Transfer Learning Model (DTL) for the classification of real-life COVID-19 dataset of chest X-ray images in both binary (COVID-19 or Normal) and three-class (COVID-19, Viral-Pneumonia or Normal) classification scenarios. Four experiments were performed where fine-tuned VGG-16 and VGG-19 Convolutional Neural Networks (CNNs) with DTL were trained on both binary and three-class datasets that contain images of X-ray. The system was trained with X-ray images dataset for the detection of COVID-19. The fine-tuned VGG-16 and VGG-19 DTL were modelled by employing a batch size of 10 in 40 epochs, and Adam optimizer for weight updates, categorical cross-entropy loss function. The result showed that the fine-tuned VGG-16 and VGG-19 models produced an accuracy of 99.23% and 98.00%, respectively, in the binary task. In contrast, in the multiclass (three-class) task, the fine-tuned VGG-16 and VGG-19 DTL models produced an accuracy of 93.85% and 92.92%, respectively. Moreover, the fine-tuned VGG-16 and VGG-19 models have MCC of 0.98 and 0.96 respectively in the binary classification, and 0.91 and 0.89 for multiclass classification. These results showed strong positive correlations between the models' predictions and the true labels. In the two classification tasks (binary and three-class), it was observed that the fine-tuned VGG-16 DTL model had stronger positive correlations in the MCC metric than the fine-tuned VGG-19 DTL model. The

40 VGG-16 DTL model has a Kappa value of 0.98 as against 0.96 for the VGG-19 DTL model in
41 the binary classification task, while in the three-class classification problem, the VGG-16 DTL
42 model has a Kappa value of 0.91 as against 0.89 for the VGG-19 DTL model. This result is in
43 agreement with the trend observed in the MCC metric. Hence, it was discovered that the VGG-
44 16 based DTL model classified COVID-19 better than the VGG-19 based DTL model. Using the
45 best performing fine-tuned VGG-16 DTL model, tests were carried out on 470 unlabeled images
46 dataset, which was not used in the model training and validation processes. The test accuracy
47 obtained for the model was 98%. The proposed models provided accurate diagnostics for both
48 the binary and multiclass classifications, outperforming other existing models in the literature in
49 terms of accuracy, as shown in this work.

51 **Keywords:** Convolutional neural networks, Coronavirus, COVID-19 test results, Deep transfer
52 learning, Machine learning, VGG-16, VGG-19

54 ## Introduction

55 Viral pandemics are usually a serious threat to the world, and Coronavirus disease 2019
56 (COVID-19) is not an exception. According to a COVID-19 report of the World Health
57 Organization (WHO), coronaviruses are from a large family of viruses that cause illness in
58 animals or humans (WHO, 2020). Numerous coronaviruses have been reported as the cause of
59 respiratory disease in humans, ranging from the common cold to more serious illnesses like
60 Middle East Respiratory Syndrome (MERS) and Extreme Acute Respiratory Syndrome (SARS)
61 (SARS). The newly discovered coronavirus led to the cause of the 2019-novel coronavirus
62 (COVID-19) disease. COVID-19 was initially observed in the Wuhan province of China and has
63 spread to all parts of the world (Nadeem, 2020). COVID-19 was recognized as a contributory
64 virus by Chinese authorities on January 7 2020. The Director-General of the WHO, on the 30th of
65 January 2020 reported that the epidemic constitutes a Public Health Emergency of International
66 Concern (PHEIC), based on the recommendations made by the Emergency Committee. WHO
67 activated the R&D Blueprint in reaction to the occurrence to speed up diagnostics, vaccines, and
68 therapeutics for this new coronavirus (WHO, 2020). The International Committee on Taxonomy
69 of Viruses named the novel coronavirus as "severe acute respiratory syndrome coronavirus 2
70 (SARS-Coronavirus-2 or SARS-CoV-2)". Globally, as of March 10, 2021, there have been
71 117,332,262 confirmed cases of COVID-19, including 2,605,356 deaths, reported to the World
72 Health Organization. Also, as of March 9, 2021, a total of 268,205,245 vaccine doses have been
73 administered (WHO, 2021). Specifically, in Nigeria, the Nigeria Centre for Disease Control
74 (NCDC) report on March 11, 2021, showed 394 new confirmed cases recorded. To date, 159,646
75 cases have been confirmed, 139,983 patients have recovered and discharged, and 1,993 deaths
76 have been recorded in 36 states and the Federal Capital Territory (NCDC, 2021).

77 Numerous researchers globally are putting their efforts together on collecting data and
78 developing solutions. The persistent focus has been on advancing key performance indicators,
79 for example, continually enhancing the speed of case detection, segregation, and early cure. The

80   execution of these containment procedures has been sustained and enabled by the pioneering and
81   aggressive use of cutting-edge technologies. Measures such as immediate case detection and
82   isolation, rigorous close contact tracing and monitoring/quarantine, and direct
83   population/community engagement have been considered in lessening COVID-19 illness and
84   death. This work, therefore, is aimed at using Artificial Intelligence (AI), specifically machine
85   learning, to identify those who are at risk of contracting COVID-19 to aid early diagnosis.
86   According to (BBC, 2020), a superhuman attempt is required to ease the deaths due to the global
87   epidemic. AI may have been overestimated - but in the case of medicine, it already has
88   established evidence. According to Arora, Bist, Chaurasia, and Prakash (2020), the role of AI is
89   going to be crucial for predicting the outcome based on symptoms, CT-Scan, X-ray reports, etc.

90   Laboratory checking of suspected cases is characterized by extended periods of testing and an
91   exponential rise in test requests (Kobia & Gitaka, 2020). Quick diagnostic tests with shorter
92   turnaround times of between 10 and 30 minutes have been developed to ease the problem.
93   However, many are presently going through clinical validation that is not in regular use (ECDC,
94   2020). In the process of result expectation, there is a need to continue to self-isolate. Once results
95   are received, there is a need to remain on self-isolation until the symptoms resolve after being in
96   seclusion for at least 14 days. If the symptoms worsen during the seclusion time or continued
97   after 14 days, the patient has to contact the accredited healthcare providers. The Rapid Test Kits
98   even deliver results after hours.

99   Pneumonia has been described as the most severe and frequent manifestation of COVID-19
100 infection (Huang *et al.* 2020); therefore, chest imaging which includes readily available and
101 affordable chest radiograph (X-rays), remains an essential factor in the diagnosis and evaluation
102 of COVID-19 patients (Rubin *et al.*, 2020). However, the availability of radiologists to report the
103 chest images is another obstacle. Therefore, there is the need to develop computer algorithms and
104 methods to optimize screening and early detection, which is the primary purpose of this research
105 in which deep learning, most especially Convolutional Neural Network (CNN), is deployed.
106 Deep learning provides the chance to increase the accuracy of the early discovery by automating
107 the primary diagnosis of medical scans (Madan, Panchal, & Chavan, 2019). The CNN belongs to
108 a category of a Deep Neural Networks (DNN), which comprise several layers that are hidden like
109 convolutional layers. The convolutional layers come with the non-linear activation function, a
110 rectified linear unit (ReLU layer), a Pooling layer, and a fully connected normalized layer. CNN
111 divides weights in the convolutional layer, thereby decreasing the memory footprint and
112 increasing the performance of the network (Sasikala, Bharathi & Sowmiya, (2018). The
113 objective of this paper is to classify a real-life COVID-19 dataset consisting of X-ray images
114 using a novel Deep Learning Convolutional Neural Network Model. Data from chest X-rays
115 were used because most hospitals have X-ray machines, and the COVID-19 X-ray dataset is now
116 available on the web. The remaining parts of this paper are organized as follows; literature
117 review, materials and methods, experimentation and results, evaluation of results , and
118 conclusion and future works.

119

Literature Review

120

121 COVID-19 is predominantly a respiratory illness, and pulmonary appearances constitute the
122 main presentation of the disease. SARS-CoV-2 infects the respiratory system but may also affect
123 other organs, as reported in some studies. Renal dysfunction (Chu, Tsang & Tang, 2005; Xu, Shi
124 & Wang, 2020), gastrointestinal complications (Pan, Mu & Ren, 2020), liver dysfunction
125 (Huang, Wang & Li, 2020), cardiac manifestations (Zhou, She, Wang & Ma, 2020), mediastinal
126 findings (Valette, du Cheyron & Goursaud, 2020), neurological abnormalities, and
127 haematological manifestations (Song & Shin, 2020) are among the reported extrapulmonary
128 features. Some of the clinical symptoms of COVID-19 are cough, expectoration, asthenia,
129 dyspnoea, muscle soreness, dry throat, pharyngeal dryness and pharyngalgia, fever, poor
130 appetite, shortness of breath, nausea, vomiting, nasal obstruction, and rhinorrhoea. A study on
131 COVID-19 credited to WHO, stated that the disease does not exhibit distinct symptoms, and
132 patients' symptoms can vary from fully asymptomatic to extreme pneumonia and death (WHO,
133 2020). Nevertheless, certain symptoms, such as dry cough, fever, dyspnea and fatigue, were
134 confirmed to be more prevalent in COVID-19 patients. Sore throat, nasal inflammation, fever,
135 arthralgia, chills, diarrhoea, hemoptysis, nausea, and conjunctival congestion are some of the
136 other clinical symptoms. (Shima, Leila, Amir & Ali, 2020). Other non-specific symptoms
137 include loss of smell and taste, dermatologic eruptions, delirium, and a general decline in health
138 (Recalcati, 2020). Because of the wide range of clinical indications and the growing global
139 burden of COVID-19, it is critical to promptly scale up diagnostic ability to diagnose the virus
140 and its risks.

141

142 Reverse transcription-polymerase chain reaction (RT-PCR) is the primary clinical instrument
143 currently in use detect COVID-19. It uses respiratory specimens for testing (Wang *et al.*, 2020a).
144 RT-PCR is used as a reference method for detecting COVID-19 patients; however, the technique
145 is expensive, manual, complicated, time-consuming, and requires specialized medical personnel.
146 Alternatively, X-ray imaging is an easily accessible tool that can be excellent in the COVID-19
147 diagnosis. Chest imaging is an essential part of evaluating respiratory complications, which
148 remain one of the most familiar presentations ranging from acute respiratory distress syndrome
149 to respiratory failure (Huang *et al.* 2020). Chest imaging has been defined as an efficient
150 screening method for detecting pneumonia, with a sensitivity of 97.5 percent for COVID-19.
151 (Nabila *et al.*, 2020; NHCPRC, 2020). Provided COVID-19's preference for the respiratory
152 system, chest radiography (X-ray), CT of the thorax, and/or Ultrasound have been verified not
153 only as case management and screening methods for COVID-19, but also as a way of reducing
154 infection transmission through early detection in initially False-negative RT-PCR tests. (Rubin *et*
155 *al.*, 2020). Imaging tests are useful for generating clinically actionable outcomes, which can be
156 used to determine a diagnosis or to guide management, triage, or treatment. Costs such as the
157 risk of exposure to radiation to the patient, the risk of COVID-19 transmission to uninfected
158 health care staff and other patients, the use of personal protective equipment (PPE), and the need

159   for sanitation and interruption of radiology rooms in resource-constrained settings reduce the
160   benefit. (Kooraki, Hosseiny, Myers & Gholamrezanezhad, 2020).
161   The role of imaging includes the detection of early parenchymal lung disease, disease
162   progression, complications, and alternative diagnoses, including acute heart failure from
163   COVID-19 myocardial injury and pulmonary thromboembolism (Driggin *et al.*, 2020). Although
164   CT is more sensitive than chest radiography, chest radiography remains the first-line imaging
165   modality in COVID-19 patients because of its availability, affordability, reduced radiation risk,
166   and ease of decontamination (Fatima *et al.*, 2020). Hence, this research work utilized chest
167   radiographs (X-ray) for identifying COVID-19 in patients. Bilateral, peripheral, lower zone
168   prevalent ground-glass opacities are common chest radiograph observations in COVID-19
169   patients. (Vancheri *et al.*, 2020). Other possible findings include normal, unilateral, or bilateral
170   reticular alterations, consolidations, ground-glass opacities, and pleural effusion (Hamid, Mir &
171   Rohela, 2020).
172

173   COVID-19 is transmitted through droplets from coughing or sneezing and on close contacts with
174   infected persons. Propagation of droplet based on infected surfaces are considered as the major
175   means of transmitting SARS-CoV-2'. However, patients going through screening are protected
176   and scanned through the use of treated tools (Kooraki *et al.*, 2020). The incubation period of
177   COVID-19 is usually about 14 days, during which it attacks the lung. Different countries
178   recommend personal protection equipment (PPE). According to the Centers for Disease Control
179   and Prevention, radiology personnel should use a face mask, glasses or face shield, sleeves, and
180   an isolation garment. A surgical cap and foot covers are required in countries with more rigorous
181   PPE guidelines, whereas a surgical mask, goggles, or face shield are recommended in countries
182   with less strict PPE guidelines. (CDCP, 2020). COVID-19 is highly contagious, and the
183   symptoms begin to appear 5-6 days after contracting it either from the droplet or close contact
184   with an infected person. According to Wang, Tang, and Wei (2020), the period between the
185   manifestation of symptom and demise ranges from 6 - 41 days, depending on the age and
186   immune system of the patient. The period is shorter among older people (Bai *et al.*, 2020;
187   Hamid, Mir & Rohela, 2020).
188

189   COVID-19 shows some unique clinical symptoms like targeting the lower airway that manifests
190   in sore throat, sneezing, and rhinorrhoea. In addition, the chest radiographs result in some cases
191   possess "infiltrate in the upper lobe of the lung" related to growing dyspnea with hypoxemia
192   (Hamid *et al.*, 2020). COVID-19 has become endemic and rapid diagnosis is imperative to
193   identify patients and carriers for possible isolation and treatment to curb the spread of the
194   disease. Attempts have been made to diagnose the disease, but many are slow and not accurate in
195   that they often give false-negative and false-positive results. This section examines some of the
196   work done in this area.
197

198   Over comprehensive detection of each aggressive spread chain, Lokuge *et al.* (2020) proposed an
199   accurate, fast, and flexible tracking approach for spotting all residual COVID-19 group
200   transmission. Using surveillance evaluation methods, they considered efficiency and sensitivity
201   in the classification of population transmission chains by testing primary care fever and cough
202   patients, hospital cases, or asymptomatic community members. They also varied the number of
203   duplications, monitoring capacities, and the prevalence of COVID-19 and non-COVID-19 fever
204   and cough. The findings of the study revealed that testing both syndromic fever and cough
205   primary care presentations, as well as precise and diligent case and touch monitoring and
206   assessment, allows for proper primary direct detection and elimination of COVID-19 population
207   transmission. Even with optimized test sensitivity, measures such as combining these approaches
208   could allow for increased case discovery if testing capacity is minimal. The impact analysis of
209   movement restriction as a result of emergence of COVID-19 was carried out by Hyafil and
210   Moria (2020). The study looked at the impact of the steps put in place in Spain to combat the
211   epidemic. The instances and the influence of the imposed restriction to movement on the
212   multiplicative quantity of hospitalization reports were estimated. The projected figure of
213   instances displayed a rapid rise towards total movement restriction as imposed. The primary
214   replication rate reduced meaningfully from 5.89 (95% CI: 5.46-7.09) before the lockdown to
215   0.48 (95% CI: 0.15-1.17) after the lockdown. The study found that managing a pandemic in the
216   magnitude of COVID-19 was very intricate and required timely decisions. The significant
217   modifications found in the infestation rate displayed that employing inclusive participation in the
218   first phase was vital in reducing the effect of a possible transferrable threat. This paper likewise
219   stressed the significance of dependable up-to-date epidemiological facts to precisely measure the
220   influence of Public Health guidelines on the virus-related outburst.

221   Yu *et al.* (2020) observed the mounting evidence that suggested that there remained a hidden
222   collection of COVID-19 asymptomatic but transferable cases and that approximating the count
223   of disease instances without symptoms was essential in knowing the virus and curtailing its
224   transmission; though, it was reported that it was difficult to precisely calculate the spread of the
225   infection. A machine learning-based fine-grained simulator (MLSim) was proposed to combine
226   many practical reasons, such as disease development during the maturation phase, cross-region
227   population movement, unobserved patients without symptoms, preventative measures and
228   confinement resilience, to estimate the number of asymptomatic infection, which is critical in
229   understanding the virus and accurately containing its spread. Digital transmission mechanisms
230   with many unspecified variables were used to simulate the relationships between the variables,
231   which were calculated from epidemic data using machine learning approach. When MLSim
232   learned to closely compare and contrast real-world data, it was able to simulate the instances of
233   patients without symptoms as well. The accessible Chinese global epidemic data helped the
234   MLSim to train better. The analysis indicated that fine-grained machine learning simulators
235   could improve the modelling of dynamic real-world infection transmission mechanism, which
236   can aid in the development of balanced mitigation steps. The simulator equally showed the
237   possibility of a great amount of undiscovered disease risk, posing a major threat to containing the

238   virus. COVID-19 was modelled using a composite stochastic and deterministic concept that
239   allowed for time-varying transmission capacity and discovery chances (Romero-Severson,
240   Hengartner, Meadors & Ke, 2020). Iterative particle sorting was used to adapt the model to a
241   historical data study of occurrence and casualty figures from fifty-one countries. The report
242   confirmed the fact that the spread rate is decreasing in forty-two of the fifty-one countries
243   surveyed. Out of the forty-two countries, thirty-four showed a big significant proof for
244   subcritical transmission rates, though the turndown in novel cases was moderately slow in
245   comparison to early development rates. The study concluded that attempts to reduce the
246   occurrence of COVID-19 by social distancing were successful. They could, however, be
247   improved and retained in various regions to prevent the disease from resurfacing. The study also
248   proposed other approaches to manage the virus before the relaxation of social distancing efforts.

249   The challenges associated with the storage and security of COVID-19 patients' data were
250   highlighted by ElDahshan, AlHabshy, and Abutaleb (2020). The authors pointed out that the
251   variety, volume, and variability of COVID-19 patients data required storage in NoSQL database
252   management systems (NoSQL DBMS). It was noted that available NoSQL DBMSs were
253   fraught with security challenges that rendered them unsuitable for storing confidential patient
254   data. Academic institutions, research centres, and enthusiasts find it difficult to select the most
255   suitable NoSQL DBMS because there are myriads of them without standard ways of
256   determining the best. Thus, the study presented an inventive approach to selecting and securing
257   NoSQL DBMS for medical information. The authors outlined the five most common NoSQL
258   database groups, as well as the most common NoSQL DBMS forms affiliated with every one of
259   them. In addition, their research included a comparison of the various types of NoSQL DBMS.
260   The paper provided an efficient solution to the myriads of security challenges, ranging from
261   authorization, authentication, encryption, and auditing in storing and securing medical
262   information utilizing a collection of web service-based functions.

263   Guerrero, Brito, and Cornejo (2020) used a mathematical model to depict a sneezing individual
264   in an urban setting with a meteorological wind of medium strength. The proliferation of airborne
265   route was demonstrated using a Lagrangian method and a wall-modeled Large Eddy Numerical
266   simulation. The results showed that the dimensions of two kinds of droplets differ in size: larger
267   droplets (400–900m) scatter between 2–5m in 2.3s, whereas smaller (100–200m) droplets are
268   transported in a larger, more impressive array between 8–11m by the windy conditions in 14.1s
269   on average.. Knowing the ambiguity of possible infection in this way aids in the development of
270   solutions for the possibility of adopting tougher self-care and distance policies.

271   To distinguish communicable acute abdomen patients suspected of COVID-19, Zhao et al.
272   (2020) proposed a forecasting model identified as a monogram and scale. The analytical
273   framework was built on the basis of a retrospective case study. In a training cohort, the model
274   was formulated using LASSO regression and multivariable logistic regression method. In the
275   training and testing cohorts, standard curve evaluated the efficiency of the monogram, receiver

276 operating characteristic (ROC) curves, decision curve analysis (DCA), and clinical effect curves.
277 According to the monogram, a simpler testing scale and management algorithm was developed.
278 In the testing cohort, the CIAAD monogram demonstrated strong differentiation and
279 standardization, which was approved. The CIAAD monogram was clinically useful, according to
280 decision curve research. The monogram was condensed even further into the CIAAD standard.
281 The estimated Bayesian Computation method was utilized by Vasilarou, Alachiotis, Garefalaki,
282 and Beloukas (2020) to determine the parameters of a demographic scenario engaging an
283 exponential growth of the size of the COVID-19 populations and revealed that rapid exponential
284 development in population size could sustain the monitored polymorphism patterns in COVID-
285 19 genomes.
286 Amrane *et al.* (2020) adopted a genetic approach using a rapid virological diagnosis on sputum
287 and nasopharyngeal samples from suspect patients. Two real-time RT-PCR systems employing a
288 "hydrolysis probe and the LightCycler Multiplex RNA Virus Master Kit" were used. The
289 primary technique probes the envelope protein (E)-encoding gene and used a synthetic RNA
290 positive control. The subsequent system targeted the spike protein-encoding gene (forward
291 primer, reverse primer, and probe) and used synthetic RNA positive control methods. Bai *et al.*
292 (2020) proposed the use of medical technology through the internet of things (IoT) to develop an
293 intelligent analysis and treatment assistance programme (nCapp). The conceptual cloud-based
294 IoT platform includes the basic IoT functions as well as a graphics processing unit (GPU). To aid
295 in deep mining and intelligent analysis, cloud computing systems were linked to existing
296 electronic health records, image cataloguing, picture cataloguing, and interaction. Li *et al.* (2020)
297 examined chest images for the analysis of COVID-19. High-resolution Computed tomography
298 (HRCT) was implemented for analysis of the virus infection. CT scans were taken with the
299 following parameters: 120 KV; 100-250 mAs; collimation of 5 mm; the pitch of 1-1.5; and 512
300 X 512 matrix. The images were reconstructed by high resolution and conventional algorithms.
301 The experiments were repeated several times, running into days for each patient. The High-
302 resolution CT objectively evaluated lung lesions giving a better understanding of the
303 pathogenesis of the disease.

304 Long *et al.* (2020) evaluated the suitability of Computed Tomography (CT) and real-time
305 reverse-transcriptase-polymerase Chain Reaction (rRT-PCR). A clinical experiment with life
306 data was executed, and the results presented showed that CT examination outperformed that of
307 rRT-PCR at 97.2% and 84.6%, respectively. In Vaishya, Javaid, Khan, and Haleem (2020),
308 seven critical AI applications for the novel COVID-19 were recognized to perform vital roles in
309 screening, analyzing, tracking, and predicting patients. Application areas identified comprised
310 early detection and diagnosis of the infection, treatment monitoring, individuals contact tracing,
311 projection of cases and mortality, drugs and vaccines development, lessening healthcare
312 workers' assignment, and deterrence of the disease by providing updated supportive information.

313 Zhang, Wang, Jahanshahi, Jia, and Schmitt (2020) conducted a survey that presented some
314 proofs of mental distress and associated predictors amongst adults in the current COVID-19

315 pandemic in Brazil. The data was composed of 638 adults from March 25 to 28, 2020, about one
316 month after the index case was confirmed in São Paulo. Female adults, who were young, more
317 trained, and practiced less, recorded higher levels of distress, with 52 percent experiencing mild -
318 to - moderate distress and 18.8 percent suffering extreme distress. The study's findings also
319 revealed that a person's distance from Sao Paulo, the epicenter, had a direct connection with the
320 psychological distress they were experiencing. For the older population who worked the least,
321 the "typhoon eye effect" was more potent. Adults who lived far off the worst hit geographic
322 area and did not go to work in the week preceding the survey were the most vulnerable. The
323 paper concluded that recognizing the predictors of suffering would allow mental health services
324 to improve target finding and assist the more mentally defenseless adults in the crisis. Ghafari *et
325 al.* (2020) assessed the challenges and indications of concern of the COVID-19 in Iran. The
326 heterogeneous COVID-19 casualty levels around the country in fourteen university hospitals in
327 Tehran were investigated, and it was revealed that the recorded cases on 13/03/2020 indicated
328 just under 10% of symptomatic patients in the population in adolescents. The finding indicated
329 that there was a major inaccurate reporting of cases in Iran. The study suggested that strict
330 measures be implemented throughout a period of widespread underreporting in order to prevent
331 the healthcare system from being exhausted within a month. Further studies on efficient
332 diagnosis, detection, and vaccine of the virus are continuing for two main reasons: partly because
333 the disease is new, and secondly because available research efforts have not been able to address
334 the concerns effectively. Therefore, in this paper, a deep learning modeling framework for
335 efficient identification, classification, and provision of new insights for the diagnosis of COVID-
336 19 was presented. Also, the prediction of probable patients of the novel COVID-19 using
337 radiology scanned images of suspected patients were shown.

338 Zivkovic *et al.* (2021) proposed a hybrid comprising machine learning Adaptive Neuro-Fuzzy
339 Inference System (ANFIS) and enhanced Bio-inspired Beetle Antennae Search (BAS) Algorithm
340 referred to as CESBAS-ANFIS. The hybrid study was carried out with a view to improving the
341 existing time-series prediction algorithms for forecasting COVID-19 new cases. An improved
342 BAS algorithm was adopted to update the parameters of ANFIS. A prediction model for the virus
343 outbreak was formulated using ANFIS trained by the improved BAS algorithm to enhance the
344 prediction accuracy of new cases of COVID-19. CESBAS was introduced to update ANFIS
345 parameters, thereby solving parameters' optimization problem of machine learning techniques
346 for prediction. Cauchy mutation operator was incorporated into the original BAS called
347 CESBAS (Cauchy Exploration Strategy BAS) to improve exploration ability and solution
348 diversity deficiencies observed. The proposed method consists of five layers; one input layer,
349 two hidden layers, one layer for conclusion parameters and the output layer that presented the
350 forecasted value. The proposed model with other hybrid models was tested under the same
351 conditions on two datasets; one dataset from WHO and the second from the "our world in data"
352 website, and their results were compared. CESBAS-ANFIS showed superior performance
353 compared to other hybrid techniques such as ABC-ANFIS, BAS-ANFIS, and FPA-ANFIS.

354   Irfan *et al.* (2021) explored the contributions of hybrid deep neural networks (HDNNs), chest X-
355   rays and computed tomography (CT) in the detection of COVID-19. The work employed X-ray
356   imaging and CT to develop the HDNNs for predicting the early infection of COVID-19. The
357   HDNNs were trained and tested on five thousand (5000) images collected from five different
358   sources (public and open), comprising 57% males and 32% females, and 3500 infected and 1500
359   healthy controls within an age group of 38-55 years. The proportion of the test dataset to the
360   training dataset was 20:80, and classification accuracy of 99% was achieved with the HDNNs on
361   the test dataset. The results of the performed experiments showed that the new multi-model and
362   multi-data approach achieved improved performance over the traditional machine learning
363   models.

364   Elzeki, Shams, Sarhan, Abd Elfattah, and Hassanien (2021) presented a new deep learning
365   computer-aided scheme for rapid and seamless classification of COVID-19. Consisting of three
366   separate COVID-19 X-ray datasets, the study presented the COVID Network (CXRVN) model
367   for assessing grayscale chest X-ray images. The scheme was implemented on three different
368   datasets using MatLab 2019b. A comparison was made with the pre-trained models of AlexNet,
369   GoogleNet, and ResNet using the mini-batch gradient descent and Adam optimizer to aid the
370   learning process. Performance evaluation results of the model using F1 score, recall, sensitivity,
371   accuracy, and precision, with generative adversarial network (GAN) data augmentation revealed
372   that the accuracy for the two-class classification was 96.7%. In comparison, that of the three-
373   class classification model reached 93.07%. However, the authors pointed out that increased
374   availability of datasets could improve the performance of future methodologies. In addition, it
375   was stated that the model could be enhanced by employing computed tomography (CT-images)
376   and studying different updated cases of the COVID-19 X-ray images.

377   From the reviews presented in this section, it was observed that some of the existing works
378   reported low accuracies, used imbalanced datasets, and there was no evidence of the use of some
379   standard evaluation metrics such as Matthew's Correlation Coefficient and Cohen's Kappa
380   Statistics in their approaches. The aforementioned issues were addressed in this paper. In
381   addition, the performance of the VGG-16 and VGG-19 networks, the forms of the VGGNet
382   Architecture, in predicting COVID-19 X-ray datatsets were compared in this study.

383

384   ***The VGGNet Architecture***
385   The VGGNet, proposed by Simonyan and Zisserman (2015), is a convolutional neural network
386   that performed very well in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC)
387   in 2014. The VGG-16 and VGG-19 networks are forms of the VGGNet Architecture. The
388   networks accept color images with size 224 x 224 and 3 channels (Red, Green, and Blue) as their
389   input data. The images pass through convolutional layers that are stacked on top of each other, in
390   which there is a limited reactive field of 3 x 3 and stride of 1 in the convolutional filter. The
391   convolutional kernel employs row and column padding such that the resolution before the

392   convolution is retained after the processing of the images. Max-pooling is then done over a max
393   pool window of size 2 x 2 with a stride of 2 (Simonyan & Zisserman, 2015).

394   The network of the VGG-16 CNN has 13 convolutional layers (that is, 3×3 convolutional layers
395   in blocks that are stacked on top of one another with growing depth). Two blocks house two 3×3
396   convolutional layers of the same setup in a sequential arrangement, while three blocks have three
397   3×3 convolutional layers of the same configuration in a sequential arrangement. As such, the
398   VGG-16 has two contiguous blocks of two convolutional layers, with each block accompanied
399   with a max-pooling. Also, it has three continuous blocks of three convolutional layers, with each
400   block accompanied with a max-pooling. In all, there are five max-pooling layers in the
401   architecture. Max pooling layer handles the reduction of the volume size after each block that
402   contains two convolutional layers and after each block that contains three convolutional layers.
403   The informative features are obtained by these max-pooling layers that are applied at the earlier
404   specified stages in the network. The VGG-16 further has two fully-connected layers, each with
405   4,096 nodes and one fully-connected layer with 1000 nodes, one node each for each of the 1000
406   categories of images in the ImageNet database on which the network was pre-trained, and is
407   followed by the SoftMax classifier (Simonyan & Zisserman, 2015), as presented in the VGG
408   Architecture, which can be found in Frossard *et al.* (2016).

409   The network of the VGG-19 CNN has 16 convolutional layers (that is, 3×3 convolutional layers
410   in blocks that are stacked on top of one another with growing depth). Two blocks house two 3×3
411   convolutional layers of the same setup in a sequential arrangement, while three blocks have four
412   3×3 convolutional layers of the same configuration in a sequential arrangement. In other words,
413   the VGG-19 has two continuous blocks of two convolutional layers, with each block
414   accompanied with a max-pooling. It also has three contiguous blocks of four convolutional
415   layers, with each block accompanied with a max-pooling. In all, there are five max-pooling
416   layers in the architecture. The Max-pooling layer handles the reduction of the volume size after
417   each block that contains two convolutional layers and after each block that contains four
418   convolutional layers. The informative features are obtained by these max-pooling layers that are
419   applied at the earlier specified stages in the network. The VGG-19 further has two fully-
420   connected layers, each with 4,096 nodes and one fully-connected layer with 1000 nodes, one
421   node each for each of the 1000 categories of images in the ImageNet database on which the
422   network was pre-trained, and is followed by the SoftMax classifier (Simonyan & Zisserman,
423   2015).

424

Materials & Methods

425

426   This study focused on diagnosing the COVID-19 chest X-ray dataset using a deep learning
427   convolutional neural network (CNN). The CNN comprises one or more convolution layers and
428   then followed by one or more fully connected layers as obtained in a standard multilayer neural
429   network. COVID-19 Radiology Dataset (chest X-ray) for Annotation and Collaboration was

430   collected from the Kaggle website (a database collated by researchers from Qatar University and
431   the University of Dhaka) with collaborators from Pakistan and Malaysia, and some medical
432   doctors), and Mendeley dataset repository. The data was preprocessed, and the median filter was
433   employed to restore the image undergoing evaluation by mitigating the severity of collection
434   degradations. Manikandarajan and Sasikala (2013) mentioned some preprocessing and
435   segmentation strategies that were used. Each data point was replaced by the average value of its
436   neighbors, and itself, in the median filter. As a result, data points that differed significantly to
437   their neighbors were removed. Following the preprocessing of the image dataset, the images
438   were sectioned by using a simulated annealing algorithm. Feature extraction and classification
439   were done using CNN. The neural network-based convolutional segmentation was implemented
440   in Jupyter Notebook using Python programming language and the model was built using sample
441   datasets for the system to recognize and classify the COVID-19. The model generated can be
442   used to develop a simple web-based application that medical personnel handling COVID-19 tests
443   could use to input new cases and quickly predict the presence of the COVID-19, with a very high
444   level of accuracy.
445
446   ***Dataset Description and Preprocessing***
447   Chest X-ray images were selected from a repository of COVID-19 positive cases' chest X-ray
448   images, as well as regular and viral pneumonia images, which were collated by researchers from
449   Qatar University and the University of Dhaka, along with collaborators from Pakistan and
450   Malaysia, and some medical doctors. There are 219 COVID-19 positive images in their current
451   release, 1341 normal images, and 1345 viral pneumonia images (Chowdhury *et al.*, 2020). For
452   multiple representations, the dataset of chest X-ray images for both COVID-19 and normal cases
453   were also selected from the Mendeley dataset repository (El-Shafai, 2020), which contains 5500
454   Non-COVID X-ray images and 4044 COVID-19 X-ray images. This study, therefore, adopted
455   these multisource datasets. Due to limited computing resources, in this study, 1,300 images were
456   selected from each category for model building and validation. In other words, 1,300 images of
457   COVID-19 positive cases, 1,300 Normal images, and 1,300 images of viral pneumonia cases,
458   totalling 3,900 images in all. Also, a different set of 470 images (containing 70 COVID-19
459   images, 200 ViralPneumonia, and 200 Normal) were selected and used for testing to obtain an
460   impartial evaluation of a final model. The dataset used in this study can be found at Fayemiwo *et*
461   *al.* (2021a). It should be noted here that further descriptions of the datasets were not provided by
462   the authors of the dataset's sources.

463   OpenCV (Bradski and Kaehler, 2008) was used for loading and preprocessing images in the
464   dataset. Each image was loaded and preprocessed by performing a conversion to RGB channel
465   and changing the size of the images to 224×224 pixels to be ready for the Convolutional Neural
466   Network. Pixel intensities were then scaled to the range [0, 1] and converted the data and labels
467   to NumPy array format. Labels were then encoded using a one-hot encoder while training/testing
468   splits were created. To ensure that the model generalizes well, data augmentation was performed
469   by setting the random image rotation to 15 degrees, random range zooming to 0.15, random shift

470   of width and height to 0.2, random shear range to 0.15, randomly flipping half of the images
471   horizontally by setting horizontal_flip = True and fill_mode to nearest. One thousand, nine
472   hundred and fifty (1,950) images were allocated initially to train the model in the binary
473   classification task, resulting in 78,000 training images after augmentation. Two thousand, nine
474   hundred and twenty-five (2,925) images were allocated initially for training the model in the
475   multiclass classification task, which resulted in 117,000 training images after augmentation.

476   ***The Deep-Transfer Learning Model***
477   This study employed the VGG-16 and VGG-19 Convolutional Neural Networks (CNN) with
478   Deep Transfer Learning (DTL) approach for COVID-19 detection. The Deep Transfer Learning
479   (DTL) approach focused on the storage of weights that have been grown while unraveling some
480   image classification tasks and then engaging them on a related task. Several DTL networks have
481   been proposed, some of which include VGGNet (Simonyan & Zisserman, 2015), GoogleNet
482   (Szegedy *et al.*, 2015), ResNet (He, Zhang, Ren & Sun, 2016), DenseNet (Huang, Liu, Van Der
483   Maaten & Wein, 2017) and Xception (Chollet, 2017). In this paper, VGG-16 CNN and VGG-19
484   CNN, forms of the VGGNet, were trained on the popular ImageNet images dataset. The VGG-16
485   and VGG-19 CNN were pre-trained deep neural networks on the ImageNet for computer vision
486   (image recognition) problems, having 16 weight layers and 19 weight layers, respectively. They
487   were used as pre-trained models to help learn the distinguishing features in COVID-19 X-ray
488   images with the aid of a transfer learning approach, thus, trained DTL models for the
489   identification of COVID-19 from X-ray images.

490   As shown in the workflow in Fig. 1 and Fig. 2, respectively, to train the VGG-16 based DTL
491   model and VGG-19 based DTL model for the detection of COVID-19, the VGG-16 CNN and
492   VGG-19 CNN was used as pre-trained models and were fine-tuned for COVID-19 detection
493   based on the principles of transfer learning. The weights of the lower layers of the network,
494   which train very common characteristics from the pre-trained model, were used as feature
495   extractors for the implementation of transfer learning with fine-tuning. Therefore, the pre-trained
496   model's lower layers weights were frozen and therefore not updated through the training process,
497   thus not participating in the transfer-learning process. The higher layers of the pre-trained model
498   were used for learning task-specific features from the COVID-19 images dataset. In this case, the
499   higher layers of the pre-trained model were unfrozen, made trainable or fine-tuned in which the
500   weights of the layers were updated. Consequently, the layers were allowed to participate in the
501   transfer-learning process. Each of these models ends with the SoftMax layer, which produces the
502   outputs. The weights for the VGG-16 and VGG-19 networks were pre-trained on ImageNet, and
503   the Fully Connected (FC) layer head was removed. From there, new fully-connected layer heads
504   were built comprising POOL => FC = SOFTMAX layers and attached on top of VGG-16 and
505   VGG-19. The Convolutional weights of VGG-16 and VGG19 were then frozen in such a way
506   that only the FC layer head was trained.

507 Both binary class scenario and three-class classification scenario are considered in the workflow,
508 in which the DTL model determines the class of the chest X-ray images as either "COVID-19"
509 category or "Normal" category in the binary class scenario or as either "COVID-19" category,
510 "Viral-Pneumonia" category, or "Normal" category in the three-class classification scenario.

511 The flowchart of the experimental algorithm for the deep transfer learning models based on the
512 VGG-16 and VGG-19 networks proposed in this paper is presented in Fig. 3.

513

514 **Figure 1:** The Architectural Workflow of the Proposed Fine-tuned VGG-16 based Deep Transfer
515 Learning Model for COVID-19 Detection

516

517 **Figure 2:** The Architectural Workflow of the Proposed Fine-tuned VGG-19 based Deep Transfer
518 Learning Model for COVID-19 Detection

519

520

521 **Figure 3:** The Flowchart of the experimental algorithm for the DTL models based on the VGG-
522 16 and VGG-19 networks

523 ## Experimentation and Results
524 Four different experiments were performed to classify radiological X-ray images using Deep
525 Transfer Learning approaches. Two of the experiments (Experiments A and B) include each of
526 the trained two Deep Transfer Learning models (VGG-16 and VGG-19 based) on binary class X-
527 ray image dataset (with COVID-19 and Normal classes). The other two experiments
528 (Experiments C and D) include each of the trained two Deep Transfer Learning models (VGG-16
529 and VGG-19 based) on a multiclass X-ray image dataset (with COVID-19, ViralPneumonia, and
530 Normal classes). For these experiments, out of a total of 3,900 images used, 2,925 images (75%)
531 were used for training the models, while 975 images (25%) were used for validation and to
532 perform hyper-parameter tuning. A different set of 470 images (containing 70 COVID-19
533 images, 200 ViralPneumonia, and 200 Normal) were used for testing to obtain an impartial
534 evaluation of the final model. The code for this experiment can be found at Fayemiwo *et al.*
535 (2021b).

536 The performances of the proposed models were obtained from the models' generated confusion
537 matrices, using standard metrics such as accuracy, specificity, precision, recall (sensitivity), F1-
538 Score, Matthews Correlation Coefficient and Cohen's Kappa statistics. Equations 1, 2 and 3

539   show the formulas for computing the three key metrics used in this article, namely Accuracy,
540   Matthews Correlation Coefficient and Cohen's Kappa statistics, respectively:

541   $$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

542   $$MCC = \frac{TP + TN - FP + FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{2}$$

543   $$Kappa = \frac{(TN + TP/TN + TP + FP + FN) - ((TN + FP) \times (TN + FN)/TN + TP + FP + FN)}{1 - ((TN + FP) \times (TN + FN)/TN + TP + FP + FN)} \tag{3}$$

544   where TP, TN, FP and FN denote True Positive, True Negative, False Positive and False
545   Negative, respectively.

546   ***Experiment A***

547   The first experiment was performed on the binary class dataset. A DTL model based on a pre-
548   trained VGG-16 model was trained to classify the X-ray images into the two classes of COVID-
549   19 or Normal; and to detect if X-ray images are simply of the class COVID-19 or Normal. The
550   VGG-16 based DTL model summary detailing the layers and parameters in each layer of the
551   model is shown in Table 1. The fine-tuned VGG-16 based DTL model consists of 14,747,650
552   total parameters, with 32,962 of them made trainable while 14,714,688 were non-trainable. The
553   VGG-16 DTL model was modelled by employing a batch size of 10 in 40 epochs, using Adam
554   optimizer specifically for updates of weights, certain cross-entropy loss function with a learning
555   rate of 1e^(-2). Performance of the proposed fine-tuned VGG-16 based DTL model was
556   evaluated on 25% of the X-ray images.

557   **Table 1:** The layers and layer parameters of the proposed fine-tuned VGG-16 based DTL model
558   (Binary Classification)

559

560   The output of the confusion matrix for the binary classification as obtained from the VGG-16
561   based DTL model is shown in Table 2. Figure 4 illustrates the training loss and accuracy, the
562   validation loss and accuracy graphs of the proposed fine-tuned VGG-16 based DTL model. The
563   validation accuracy, recall, specificity, precision, F1-Score, Matthews Correlation Coefficient
564   and Cohen's Kappa statistics of the proposed fine-tuned VGG-16 based DTL model were also
565   obtained.

566   **Table 2:** The Confusion Matrix of the Binary classification task obtained from the fine-tuned
567   VGG-16 based DTL

568

569

570   **Figure 4:** The Training loss and accuracy with the validation loss and accuracy curves obtained
571   for the fine-tuned VGG-16 based Deep Transfer Learning Model (For Binary Classification)

572   It was observed from Fig. 4 that the validation and training losses were slightly high in the earlier
573   epochs and then almost flattened as the training occurs in subsequent epochs. The decrease in the
574   loss values at around the 40th epoch was attributed to the fact that the fine-tuned VGG-16 based
575   DTL model was exposed to all the available X-ray images time and again during each of the
576   epochs considered during training.

577   The validation accuracy obtained for the fine-tuned VGG-16 based DTL model was 99.23%, its
578   recall was 100%, while its specificity stands at 98.48%. The obtained values for the precision,
579   recall, F1-Score, Matthews Correlation Coefficient and Cohen's Kappa statistics metrics for the
580   binary classification task using the VGG-16 based DTL model are given in Table 3.

581   **Table 3:** The Precision, Recall, F1-Score, Matthews Correlation Coefficient and Cohen's Kappa
582   statistics obtained for the classification task using the fine-tuned VGG-16 based DTL model
583   (Binary Classification)

584

585   ***Experiment B***

586   The second experiment was performed on the binary class dataset. A DTL model based on a pre-
587   trained VGG-19 model was trained to classify the X-ray images into the two classes of COVID-
588   19 or Normal; and also, to detect if X-ray images are simply of the class COVID-19 or Normal.
589   Also, the VGG-19 based DTL model summary detailing the layers and the parameters in each
590   layer of the model is shown in Table 4. The fine-tuned VGG-19 based DTL model consists of
591   20,057,346 total parameters, with 32,962 of them made trainable while 20,024,384 were non-
592   trainable. The VGG-19 DTL model was modelled by employing a batch size of 10 in 40 epochs,
593   using Adam optimizer for the updates of weights, categorical cross-entropy loss function with a
594   learning rate of $1e^{(-1)}$. Performance of the proposed fine-tuned VGG-19 based DTL model was
595   evaluated on 25% of the X-ray images.

596   **Table 4:** The layers and layer parameters of the proposed fine-tuned VGG-19 based DTL model
597   (Binary Classification)

598

599   The output of the confusion matrix for the binary classification as obtained from the fine-tuned
600   VGG-19 based DTL model are shown in Table 5. Figure 5 illustrates the training loss and
601   accuracy along with the validation loss and accuracy graphs of the proposed fine-tuned VGG-19

602   based DTL model. The validation accuracy, recall, specificity, precision, and F1-Score of the
603   proposed fine-tuned VGG-19 based DTL model were also obtained.

604   **Table 5:** The Confusion Matrix of the Binary classification task obtained from the fine-tuned
605   VGG-19 based DTL

606

607   **Figure 5:** The Training loss and accuracy with the validation loss and accuracy curves obtained
608   for the fine-tuned VGG-19 based Deep Transfer Learning Model (For Binary Classification)

609   It was observed from Fig. 5 that the validation and training losses were slightly moderate in the
610   earlier epochs and then decreases as the training occurs in more subsequent epochs (with a sharp
611   increase at about epochs 12 and 20). The decrease in the loss values at around the 40th epoch
612   were attributed to the fact that the fine-tuned VGG-19 based DTL model was exposed to all the
613   available X-ray images repeatedly on all the epochs considered during training.

614   The validation accuracy obtained for the fine-tuned VGG-19 based DTL model was 98.00%, its
615   recall was 95.95%, while its specificity stands at 100%. The values obtained for the precision,
616   recall, F1-Score, Matthews Correlation Coefficient and Cohen's Kappa statistics metrics for the
617   binary classification task using the VGG-19 based DTL model are given in Table 6.

618   **Table 6:** The Precision, Recall, F1-Score, Matthews Correlation Coefficient and Cohen's Kappa
619   statistics obtained for the classification task using the fine-tuned VGG-19 based DTL model
620   (Binary Classification)

621

622   *Experiment C*

623   The third experiment was performed on the multiclass (three-class) dataset, in which a DTL
624   model based on a pre-trained VGG-16 model was trained to classify the X-ray images into the
625   three classes of COVID-19, Viral Pneumonia or Normal; and to detect if X-ray images were
626   simply of the class COVID-19 or Viral Pneumonia or Normal. The VGG-16 based DTL model
627   summary, detailing the layers and parameters in each layer of the model, is shown in Table 7.
628   The fine-tuned VGG-16 based DTL model consists of 14,747,715 total parameters, with 33,027
629   of them made trainable while 14,714,688 were non-trainable. The VGG-16 DTL model was
630   modelled by employing a batch size of 10 in 40 epochs, using Adam optimizer specifically for
631   updates of weights, certain cross-entropy loss function with a learning rate of $1e^{-2}$, the
632   performance of the proposed fine-tuned VGG-16 based DTL model was evaluated on 25% of the
633   X-ray images.

634   **Table 7:** The layers and layer parameters of the proposed fine-tuned VGG-16 based DTL model
635   (Three-class Classification)

636

637   The output of the confusion matrix for the binary classification as obtained from the VGG-16
638   based DTL model are shown in Table 8. Figure 6 illustrates the training loss and accuracy and
639   the validation loss and accuracy graphs of the proposed fine-tuned VGG-16 based DTL model.
640   The validation accuracy, recall, specificity, precision, and F1-Score of the proposed fine-tuned
641   VGG-16 based DTL model were also obtained.

642   **Table 8:** The Confusion Matrix of the Three-class classification task obtained from the fine-
643   tuned VGG-16 based DTL

644

645

646   **Figure 6:** The Training loss and accuracy with the validation loss and accuracy curves obtained
647   for the fine-tuned VGG-16 based Deep Transfer Learning Model (For Three-class Classification)

648   It was observed from Fig. 6 that the validation and training losses were significantly high in the
649   earlier epochs and then slowly decrease as the training occurs in more subsequent epochs. The
650   decrease in the loss values at around the 40th epoch was attributed to the fact that the fine-tuned
651   VGG-16 based DTL model was exposed to all the available X-ray images repeatedly during each
652   of the epochs considered during training.

653   The validation accuracy obtained for the fine-tuned VGG-16 based DTL model was 93.85%, its
654   recall was 97.98%, while its specificity stands at 94.69%. The obtained values for the precision,
655   recall, F1-Score, Matthews Correlation Coefficient and Cohen's Kappa statistics metrics for the
656   three-class classification task using the VGG-16 based DTL model are given in Table 9.

657   **Table 9:** The Precision, Recall, F1-Score, Matthews Correlation Coefficient and Cohen's Kappa
658   statistics obtained for the classification task using the fine-tuned VGG-16 based DTL model
659   (Three-class Classification)

660

661

662   ***Experiment D***

663   The fourth experiment performed on the multiclass (three-class) dataset, in which a DTL model
664   based on a pre-trained VGG-19 model was trained to classify the X-ray images into the three

665   classes of COVID-19, ViralPneumonia or Normal; and also, to detect if X-ray images are simply
666   of the class COVID-19 or ViralPneumonia or Normal. The VGG-19 based DTL model summary
667   detailing the layers and parameters in each layer of the model is shown in Table 10. The fine-
668   tuned VGG-19 based DTL model consists of 20,057,411 total parameters, with 33,027 of them
669   made trainable while 20,024,384 were non-trainable. The VGG-19 DTL model was modelled by
670   employing a batch size of 10 in 40 epochs, using Adam optimizer specifically for updates of
671   weights, certain cross-entropy loss function with a learning rate of $1e^{(-1)}$, the performance of
672   the proposed fine-tuned VGG-19 based DTL model was evaluated on the 25% of the X-ray
673   images.

674   **Table 10:** The layers and layer parameters of the proposed fine-tuned VGG-19 based DTL
675   model (Three-class Classification)

676

677   The output of the confusion matrix for the binary classification as obtained from the VGG-19
678   based DTL model are shown in Table 11. Figure 7 illustrates the training loss and accuracy and
679   the validation loss and accuracy graphs of the proposed fine-tuned VGG-16 based DTL model.
680   The validation accuracy, recall, specificity, precision, and F1-Score of the proposed fine-tuned
681   VGG-16 based DTL model were also obtained.

682   **Table 11:** The Confusion Matrix of the Three-class classification task obtained from the fine-
683   tuned VGG-19 based DTL.

684

685   **Figure 7:** The Training loss and accuracy with the validation loss and accuracy curves obtained
686   for the fine-tuned VGG-19 based Deep Transfer Learning Model (For Three-class Classification)

687   It was observed from Fig. 7 that the validation and training losses were significantly high in the
688   earlier epochs and then gradually decrease as the training occurs in more subsequent epochs.
689   This decrease in the loss values at around the 40th epoch was attributed to the fact that the fine-
690   tuned VGG-19 based DTL model was exposed to all the available X-ray images time and again
691   during each of the epochs considered during training.

692   The validation accuracy obtained for the fine-tuned VGG-19 based DTL model was 92.92%, its
693   recall was 95.95%, while its specificity stands at 89.68%. The obtained values for the precision,
694   recall, F1-Score, Matthews Correlation Coefficient and Cohen's Kappa statistics metrics for the
695   three-class classification task using the VGG-19 based DTL model are given in Table 12.

696   **Table 12:** The Precision, Recall, F1-Score, Matthews Correlation Coefficient and Cohen's
697   Kappa statistics obtained for the classification task using the fine-tuned VGG-19 based DTL
698   model (Three-class Classification)

699

700   It was noted from the obtained confusion matrices and the computed performance evaluation
701   metrics of the binary and three-class classification tasks that the fine-tuned VGG-16 based deep
702   transfer learning model outperformed the fine-tuned VGG-19 based deep transfer learning
703   model in the detection of COVID-19. Based on this, some tests were carried out on unlabeled
704   images using the developed fine-tuned VGG-16 multi-classification model. The test was carried
705   out to obtained an impartial evaluation of the final model. Some results of the tests are shown in
706   Fig. 8 to Fig. 10.

707   Being the best performing model in this study, fine-tuned VGG-16 DTL model was tested on the
708   test dataset of 470 images. The test accuracy obtained for the model was 98%. The output results
709   of the tests as shown in Fig. 8 to Fig. 10 show how the fine-tuned VGG-16 DTL model classified
710   and detected each of the images as either "COVID-19", "Viral Pneumonia", or "Normal." The
711   level of confidence in the model classification is also shown. Figure 8 shows the sample images
712   that were detected as "COVID-19" along with the model's classification confidence accuracy
713   values. Figure 9 shows the sample images that were detected as "Viral Pneumonia" and the
714   model's classification confidence accuracy values. In contrast, Fig. 10 shows the sample images
715   that were detected as "Normal" along with the model's classification confidence accuracy
716   values.

717

718   **Figure 8:** COVID-19 sample test results with the predicted level of confidence value

719

720   **Figure 9:** Viral Pneumonia sample test results with a predicted level of confidence value

721

722   **Figure 10:** Normal sample test results with the predicted level of confidence value

723   Out of the six sample images shown in Fig. 8, only one showed a lower confidence level of
724   76.37%, while others were above 94%. Similar results could be seen in Fig. 10 for Normal
725   classification, where the lowest confidence level is 78.92%. However, the lowest output for Viral
726   Pneumonia is 96.21%, as shown in Fig. 9. These test results showed that the developed models
727   could generalize and adapt to new data outside the training and validation dataset. These test
728   results are necessary to show the adaptability of the developed models when related data is
729   considered.

Evaluation of Results
731   The results obtained in this work were compared with thirteen other existing approaches in the

732   literature. Few studies conducted before this study had used twenty-five and fifty images in each
733   class (Sethy & Behra, 2020; Hemdan *et al.,* 2020; Narin *et al.,* 2020), while nine out of the
734   twelve approaches benchmarked used imbalanced data (Table 9). Generally, the problem in
735   modeling imbalanced data is that it could lead to the inability of the model to generalize, or the
736   model can be biased towards a class with a high number of data points. Hence, in this study, an
737   equal value of data (1,300 images) was used for each category, and this is believed to have
738   contributed to increasing accuracies of the proposed models. At the moment, creating an
739   automated diagnostic tool for the detection of COVID-19 suffers from the drawback of limited
740   number of cases. To ensure the generalization of the models developed in this work, data
741   augmentation was performed by setting the random image rotation setting to 15 degrees
742   clockwise. The proposed new models were based on fine-tuning VGG-16 and VGG-19 methods
743   by constructing a new fully-connected layer head consisting of POOL => FC = SOFTMAX
744   layers and append it on top of VGG-16 and VGG-19; the Convolution weights of VGG-16 and
745   VGG-19 were then frozen, such that only the FC layer head was trained. The fine-tuned models
746   gave better results than other models that used ordinary pre-trained VGG-16 and VGG-19
747   (Apostolopoulos & Mpesiana, 2020; Khalid & Youness, 2020). Complete results of the
748   comparison with thirteen other existing results from the literature are presented in Table 13, with
749   the proposed model recording the best performance accuracy. The closest performing model to
750   the proposed model is that of the DarkCovidNet model (Tulin *et al.*, 2020) with 98.08%
751   accuracy, while the proposed DTL-based VGG-16 model has 99.23% accuracy, both in the
752   binary classification task.

753   **Table 13:** Comparison of the proposed COVID-19 diagnostic methods with other deep learning
754   methods developed using radiology images.

755

Conclusions and Future Work

757   Several researchers around the world are combining their efforts to collect data and develop
758   solutions for the COVID-19 pandemic problem. Laboratory testing of suspected cases
759   characterized by long waiting periods and an exponential increase in demand for tests has
760   hitherto constituted a significant bottleneck globally. Hence, rapid diagnostic test kits are being
761   developed, most of which are currently undergoing clinical validation and are yet to be adopted
762   for routine use. This paper proposed a solution using the Deep Learning Convolutional Neural
763   Network Model to classify a real-life COVID-19 dataset of chest X-ray images into three-
764   classes: COVID-19, Viral-Pneumonia and Normal categories. Two experiments were performed
765   where the VGG-16 and VGG-19 Convolutional Neural Networks (CNN) with Deep Transfer
766   Learning (DTL) was implemented in Jupyter Notebook using Python programming language.
767   Experimental results showed that the pre-trained VGG-16 DTL model classified COVID-19 data
768   better than the VGG-19 based DTL model. The fine-tuned VGG-16 and VGG-19 models

769 produced classification accuracies of 99.23% and 98.00%, respectively, for binary classification
770 and 93.85% and 92.92% for multiclass classification. The proposed model, therefore,
771 outperformed existing methods in terms of accuracy. Moreover, the fine-tuned VGG-16 and
772 VGG-19 models have MCC of 0.98 and 0.96 respectively in the binary classification, and 0.91
773 and 0.89 for multiclass classification. These results showed that there are strong positive
774 correlations between the models' predictions and the true labels. In the two classification tasks
775 (binary and three-class), it was observed that the fine-tuned VGG-16 DTL model had stronger
776 positive correlations in the MCC metric than the fine-tuned VGG-19 DTL model. The VGG-16
777 DTL model has a Kappa value of 0.98 as against 0.96 for the VGG-19 DTL model in the binary
778 classification task, while in the three-class classification problem, the VGG-16 DTL model has a
779 Kappa value of 0.91 as against 0.89 for the VGG-19 DTL model. This result is in agreement with
780 the trend observed in the MCC metric. The findings of this study have a high potential of
781 increasing the prediction accuracy for COVID-19 disease, which would be of immense benefit to
782 the medical field and the entire human populace as it could help save many lives from untimely
783 death.

784 The researchers suspect that the better performance of the VGG-16 DTL model might be
785 attributed to the volume of data used in the experiments; that is, the depth of layers in the VGG-
786 19 architecture may not have any significant effect on the performance when the dataset is small.
787 This suspicion would be investigated, as a future work, when more COVID-19 data is available.
788 Finally, the COVID-19 images from Chest CT scans are not readily available, unlike X-ray
789 images, because of their high cost. Therefore, other future works would consider using Chest CT
790 images to develop a more sensitive diagnostic tool for detecting viral pneumonia and COVID-19
791 variants. Further hyper-parameter tweaking would also be done to get more accurate results.
792

Acknowledgements

794 The authors acknowledge the efforts of Dr. (Mrs.) F. O. Unuabonah in proofreading this article
795 and increasing its readability.

796

References

798 Amrane, S., Tissot D., Doudier H., Eldin B., Hocquart C., Mailhe M., Colson M. (2020). Rapid
799   viral diagnosis and ambulatory management of suspected COVID-19 cases presenting at
800   the infectious diseases referral hospital in Marseille, France, - January 31 to March 1,
801   2020: A respiratory virus snapshot. Travel Medicine and Infectious Disease (In Press).
802   DOI:10.1016/j.tmaid.2020.101632
803 Apostolopoulos, I.D., Mpesiana, T.A (2020). COVID-19: automatic detection from X-ray
804   images utilizing transfer learning with convolutional neural networks. Physical and
805   Engineering Sciences in Medicine, Vol 43, 635–640. https://doi.org/10.1007/s13246-020-
806   00865-4.
807 Arora, K., Bist, A., Chaurasia, S., & Prakash, R. (2020). Analysis of deep learning techniques for
808   COVID-19 detection. *International Journal of Scientific Research in Engineering and
809   Management (IJSREM), 4*(4), 1-5.

810   Bai, L., Dawei, Y., Wang, X., Tong, L., Zhu, X., Zhong, N., . . . Wang. (2020). Chinese experts'
811   consensus on the Internet of Things-aided diagnosis and treatment of coronavirus disease
812   2019 (COVID-19). *Clinical eHealth, 3*, 7-15. doi:10.1016/j.ceh.2020.03.001
813   BBC. (2020). British Broadcasting Corporation. Retrieved from
814   https://www.bbc.com/news/technology-52120747
815   Bradski, G., and Kaehler, A. (2008). Learning OpenCV: Computer vision with the OpenCV
816   library. O'Reilly Media, Inc.
817   Centres for Disease Control and Prevention [CDCP], (2020). Coronavirus Disease 2019
818   (COVID-19) Interim Infection Prevention and Control Recommendations.
819   https://www.cdc.gov/coronavirus/2019-ncov/infection-control/control-
820   recommendationshtml#adhere. Accessed October 1, 2020.
821   Chollet, F. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. 2017
822   IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (pp. 1800-
823   1807). Honolulu, HI.
824   Chowdhury, M. E., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M. A., Mahbub, Z. B., . . .
825   Reaz, M. B. (2020, March 29). Can AI help in screening Viral and COVID-19
826   pneumonia? Retrieved from https://arxiv.org/abs/2003.13145;
827   https://www.kaggle.com/tawsifurrahman/covid19-radiography-database
828   Chu, K. H., Tsang, W. K., Tang, C. S. (2005). Acute renal impairment in coronavirus-associated
829   severe acute respiratory syndrome. *Kidney Int., 67*(2), 698-705.

830   Driggin, E., Madhavan, M. V., Bikdeli, B., (2020). Cardiovascular Considerations for Patients,
831   Health Care Workers, and Health Systems During the Coronavirus Disease 2019
832   (COVID-19) Pandemic. J. Am. Coll. Cardiol.
833   ECDC. (2020). ECDC: An overview of the rapid test situation for COVID-19 diagnosis in the
834   EU/EEA. https://dx.doi.org/10.1101/2020.03.18.20038059.
835   ElDahshan, K. A., AlHabshy, A. A., & Abutaleb, G. E. (2020). Data in the time of COVID-19: a
836   general methodology to select and secure a NoSQL DBMS for medical data. PeerJ
837   Computer Science, 6, e297.
838   El-Shafai, Walid; Abd El-Samie, Fathi (2020), "Extensive COVID-19 X-ray and CT Chest
839   Images Dataset", Mendeley Data, V3, doi: 10.17632/8h65ywd2jr.3
840   Elzeki, O. M., Shams, M., Sarhan, S., Abd Elfattah, M., & Hassanien, A. E. (2021). COVID-19:
841   a new deep learning computer-aided model for classification. PeerJ Computer Science, 7,
842   e358.
843   Fatima, S., Ratnani I, Husain M, Salim Surani (2020) Radiological Findings in Patients with
844   COVID-19. Cureus 12(4): e7651. doi:10.7759/cureus.7651.
845   Fayemiwo, M., Olowookere, T., Arekete, S., Ogunde, A., Odim, M., Oguntunde, B., Olaniyan,
846   O., Ojewumi, T., Oyetade, I., & Aremu, A. (2021a). RAIG COVID19 compiled
847   dataset.zip (Version 3). figshare. https://doi.org/10.6084/m9.figshare.14151854.v3
848   ([u'https://www.kaggle.com/tawsifurrahman/covid19-radiography-database'])
849   Fayemiwo, M., Olowookere, T., Arekete, S., Ogunde, A., Odim, M., Oguntunde, B., Olaniyan,
850   O., Ojewumi, T., Oyetade, I., & Aremu, A. (2021b). RAIG: Classification of COVID-19
851   Radiology Dataset (Version 2). figshare. https://doi.org/10.6084/m9.figshare.14151971
852   Frossard, D. (2016, June 17). VGG in TensorFlow. Retrieved May 24, 2020, from
853   https://www.cs.toronto.edu/~frossard/post/vgg16/

854 Ghafari, M., Hejazi, B., Karshenas, A., Dascalu, S., Ferretti, L., Ledda, A., &Katzourakis, A.
855   (2020). Ongoing outbreak of COVID-19 in Iran:.medRxiv preprint.
856   doi:10.1101/2020.04.18.20070904
857 Guerrero, N., Brito, J. M., & Cornejo, P. (2020). COVID-19. Transport of respiratory droplets in
858   a microclimatologic urban scenario. medRxiv preprint.
859   doi:10.1101/2020.04.17.20064394
860 Hamid, S., Mir, M. Y., &Rohela, G. K. (2020). Novel coronavirus disease (COVID-19): A
861   pandemic (epidemiology, pathogenesis and potential therapeutics). New Microbe and
862   New Infect, 35. doi:10.1016/j_nmni.2020.100679
863 Hemdan E.E.D., Shouman M.A. and Karar M.E. (2020), COVIDX-Net: A Framework of Deep
864   Learning Classifiers to Diagnose COVID-19 in X-ray Images, arXiv preprint
865   arXiv:2003.11055.
866 Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T,
867   Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang
868   G, Jiang R, Gao Z, Jin Q, Wang J, Cao B (2020). Clinical features of patients infected
869   with 2019 novel coronavirus in Wuhan, China. *The Lancet., 395*(10223), 497-506.
870 Huang, G., Liu, Z., Van Der Maaten, L., & Wein, K. Q. (2017). Densely connected
871   convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern
872   Recognition (CVPR), (pp. 2261-2269). Honolulu, HI.
873 Hyafil, A., &Moriña, D. (2020). Analysis of the impact of lockdown on the evolution of
874   COVID-19 epidemics in Spain. medRxiv preprint, 1-20.
875   doi:10.1101/2020.04.18.20070862
876 Irfan, M., Iftikhar, M. A., Yasin, S., Draz, U., Ali, T., Hussain, S., . . . Glowacz, A. (2021). Role
877   of Hybrid Deep Neural Networks (HDNNs), Computed Tomography, and Chest X-rays
878   for the Detection of COVID-19. International Journal of Environmental Research and
879   Public Health, 18(6), 3056.
880 Kobia, F., &Gitaka, J. (2020). COVID-19: Are Africa's diagnostic challenges blunting response
881   effectiveness? AAS Open Research, 1-11.
882 Kooraki, S., Hosseiny, M., Myers, L., &Gholamrezanezhad, A. (2020). Coronavirus (COVID-
883   19) Outbreak: What the Department of Radiology Should Know. *J Am Coll
884   Radiol,17*(4),447–451.
885 Li, M., Lei, P., Zeng, B., L. Z., Yu, P., Fan, B., . . . Liu, H. (2020). Coronavirus disease (COVID-
886   19): Spectrum of CT findings and temporal progression of the disease. *AcadRadiol.,
887   27*(5), 603-608. doi:10.1016/j.acra.2020.03.003.
888 Lokuge, K., Banks, E., Davies, S., R. L., Street, T., . . . Glass, K. (2020). Exit strategies:
889   Optimizing feasible surveillance for detection, elimination and ongoing prevention of
890   COVID-19 community transmission. medRxiv preprint.
891   Doi:10.1101/2020.04.19.20071217.
892 Long, C., Xu, H., Shen, Q., Zhang, X., Fan, B., Wang, C., . . . Li, H. (2020). Diagnosis of the
893   coronavirus disease (COVID-19): rRT-PCR or CT? Eur J Radiol.
894   doi:10.1016/j.ejrad.2020.108961
895 Madan, B., Panchal, A., & Chavan, D. (2019). Lung cancer detection using deep learning. 2nd
896   International Conference on Advances in Science & Technology (ICAST-2019).

897   Manikandarajan, A., &Sasikala, S. (2013). Detection and Segmentation of Lymph Nodes for
898   Lung Cancer Diagnosis. National Conference on System Design and Information
899   Processing.
900   Nabila, C., El Kettani, C., Amal, H., Abdelkrim, B., Mohamed, M., Said, E., Nezha, D. (2020).
901   The epidemiological and clinical profile of COVID-19 in children: Moroccan experience
902   of the Cheikh Khalifa University Center. The Pan African Medical Journal. 35(Supp
903   2):57, 1-5, DOI: 10.11604/pamj.2020.35.2.23571.
904   Nadeem, S. (2020). Coronavirus COVID-19: Available free literature provided by various
905   companies, Journals and Organizations around the World. *Journal of Ongoing Chemical*
906   *Research, 5*(1), 7-13. doi:10.5281/zenodo.3722904
907   Narin A., Kaya C., and Pamuk Z. (2020), Automatic Detection of Coronavirus Disease (COVID-
908   19) Using X-ray Images and Deep Convolutional Neural Networks, 2020 arXiv preprint
909   arXiv:2003.10849.
910   National Health Commission of People's Republic of China [NHCPRC], (2020). Diagnosis and
911   Treatment Protocol for Novel Coronavirus Pneumonia (Trial Version 7). 2020. Accessed
912   September 30 2020
913   Nigeria Centre for Disease Control [NCDC], (2021). Coronavirus (COVID-19) Highlights.
914   Abuja: Nigeria Centre for Disease Control (NCDC). Retrieved March 11, 2021, from
915   https://covid19.ncdc.gov.ng/
916   Pan, L., Mu, M., Ren, H. G. (2020). Clinical characteristics of COVID-19 patients with digestive
917   symptoms in Hubei, China: a descriptive, cross-sectional, multicenter study. *Am J*
918   *Gastroenterol, 115*(5),766-773.
919   Recalcati, S. (2020). Cutaneous manifestations in COVID-19: a first perspective. *J Eur Acad*
920   *Dermatol Venereol*, *34*(5),e212.
921   Romero-Severson, E., Hengartner, N., Meadors, G., &Ke, R. (2020). A decline in global
922   transmission rates of COVID-19. medRxiv preprint. doi:10.1101/2020.04.18.20070771
923   Rubin, G. D., Christopher, J. R., Linda, B. H., Nicola, S., Jeffrey, P. K., Suhail, R., … Ann, N. L.
924   (2020). The Role of Chest Imaging in Patient Management during the COVID-19
925   Pandemic: A Multinational Consensus Statement from the Fleischner Society. *Radiology*,
926   *296*(1), 172–180. doi.org/10.1148/radiol.2020201365
927   Sasikala, S., Bharathi, M., &Sowmiya, B. R. (2018). Lung cancer detection and classification
928   using deep CNN. *International Journal of Innovative Technology and Exploring*
929   *Engineering (IJITEE), 8*(25), 259-262.
930   Sethy P.K. and Behera S.K. (2020), Detection of Coronavirus Disease (COVID-19) Based on
931   Deep Features.
932   Shima, B., Leila, A., Amir, R. R., & Ali, G. (2020). Extrapulmonary manifestations of COVID-
933   19: Radiologic and clinical overview. *Clinical Imaging, 66*, 35-41.
934   doi:10.1016/j.clinimag.2020.05.013
935   Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional for Large-Scale Image
936   Recognition. International Conference on Learning Representations. San Diego.
937   Song Y., Zheng S., Li L., Zhang X., Zhang X., Huang Z., and Chong Y. (2020). Deep learning
938   enables accurate diagnosis of novel coronavirus (COVID-19) with CT images, medRxiv.
939   Song, Y.G., Shin, H.-S. (2020). COVID-19, a clinical syndrome manifesting as hypersensitivity
940   pneumonitis. *Infection Chemotherapy, 52*, 110-112.

941   Vaishya, R., Javaid, M., Khan, I. H., & Haleem, A. (2020). Artificial Intelligence (AI)
942   applications for COVID-19 pandemic. Diabetes and Metabolic Syndrome: *Clinical*
943   *Research and Review, 14*, 337 - 339. doi:https://doi.org/10.1016/j.dsx.2020.04.012.
944   Valette, X., du Cheyron, D., Goursaud, S. (2020). Mediastinal lymphadenopathy in patients with
945   severe COVID-19. *Lancet Infect Dis*., pii: S1473-3099(20)30310-8.
946   Vancheri, S. G., Giovanni, S., Francesco, B., Alessia, M., Costanza, C., Chandra, B., Adele, V.,
947   … Lorenzo, P. (2020). Radiographic findings in 240 patients with COVID-19
948   pneumonia: time-dependence after the onset of symptoms. Eur Radiol. 1–9. doi:
949   10.1007/s00330-020-06967-7
950   Vasilarou, M., A. N., Garefalaki, J., & Beloukas, A. (2020). Population genomics insights into
951   the recent. bioRxiv. doi:10.1101/2020.04.21.054122.
952   Wang L. and Wong A. (2020), COVID-Net: A Tailored Deep Convolutional Neural Network
953   Design for Detection of COVID-19 Cases from Chest Radiography Images, 2020 arXiv
954   preprint arXiv:2003.09871.
955   Wang, W., Tang, J., & Wei, F. (2020). Updated understanding of the outbreak of 2019 novel
956   coronavirus (2019-nCoV) in Wuhan, *China. Journal of Medical Virology, 92*(4), 441-
957   447. doi:10.1002/jmv.25689
958   Wang, W., Xu, Y., Gao, R., Lu, R., Han, K., Wu, G., & Tan, W. (2020a). Detection of SARS-
959   CoV-2 in different types of clinical specimens. *Jama, 323*(18), 1843-1844.
960   Wang S., Kang B., Ma J., Zeng X., Xiao M., Guo J., and Xu B. (2020b)., A deep learning
961   algorithm using CT images to screen for Corona Virus Disease (COVID-19), medRxiv.
962   WHO. (2020). Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-
963   19). Geneva: World Health Organization.
964   WHO. (2021). *WHO Coronavirus (COVID-19) Dashboard*. Retrieved March 11, 2021, from
965   World Health Organization: https://covid19.who.int/
966   Xu X., Jiang X., Ma C., Du P., Li X., and Lv S. (2020), Deep Learning System to Screen
967   Coronavirus Disease 2019 Pneumonia, arXiv preprint arXiv:200209334.
968   Xu, Z., Shi, L., Wang, Y. (2020). Pathological findings of COVID-19 associated with acute
969   respiratory distress syndrome. *Lancet Respir Med.,* 420-422.
970   Yu, Y., Liu, Y., Luo, F., Tu, W., Zhan, D., Yu, G., & Zhou, Z. (2020). COVID-19
971   Asymptomatic Infection Estimation. medRxiv preprint.
972   doi:10.1101/2020.04.19.20068072
973   Zhang, S. X., Wang, Y., Jahanshahi, A. A., Jia, J., & Schmitt, V. G. (2020). First study on mental
974   distress in Brazil during the COVID-19 crisis. medRxiv preprint. doi:doi:
975   https://doi.org/10.1101/2020.04.18.20070896
976   Zhao, B., Wei, Y., Sun, W., Qin, C., Zhou, X., Wang, Z., . . .Wang, Y. (2020). Distinguish
977   Coronavirus Disease 2019 Patients in General Surgery Emergency by CIAAD Scale:
978   Development and Validation of a Prediction Model Based on 822 Cases in China.
979   medRxiv preprint. doi:10.1101/2020.04.18.20071019
980   Zheng C., Deng X., Fu Q., Zhou Q., Feng J., Ma H., and Wang X. (2020), Deep learning-based
981   detection for COVID-19 from chest CT using a weak label, medRxiv,
982   https://doi.org/10.1101/2020.03.12.20027185.
983   Zhou, B., She, J., Wang, Y., Ma, X. (2020). The clinical characteristics of myocardial injury 1 in
984   severe and very severe patients with 2019 novel coronavirus disease. *J Infect*., pii: S0163-
985   4453(20)30149-3.

986   Zivkovic, M., Bacanin, N., Venkatachalam, K., Nayyar, A., Djordjevic, A., Strumberger, I., and
987   Al-Turjman, F. (2021). COVID-19 cases prediction by using hybrid machine learning
988   and beetle antennae search approach. Sustainable Cities and Society, 66(102669).
989   doi:10.1016/j.scs.2020.102669

Figure 1

A visualization of the VGG architecture (Source: Frossard, 2016)
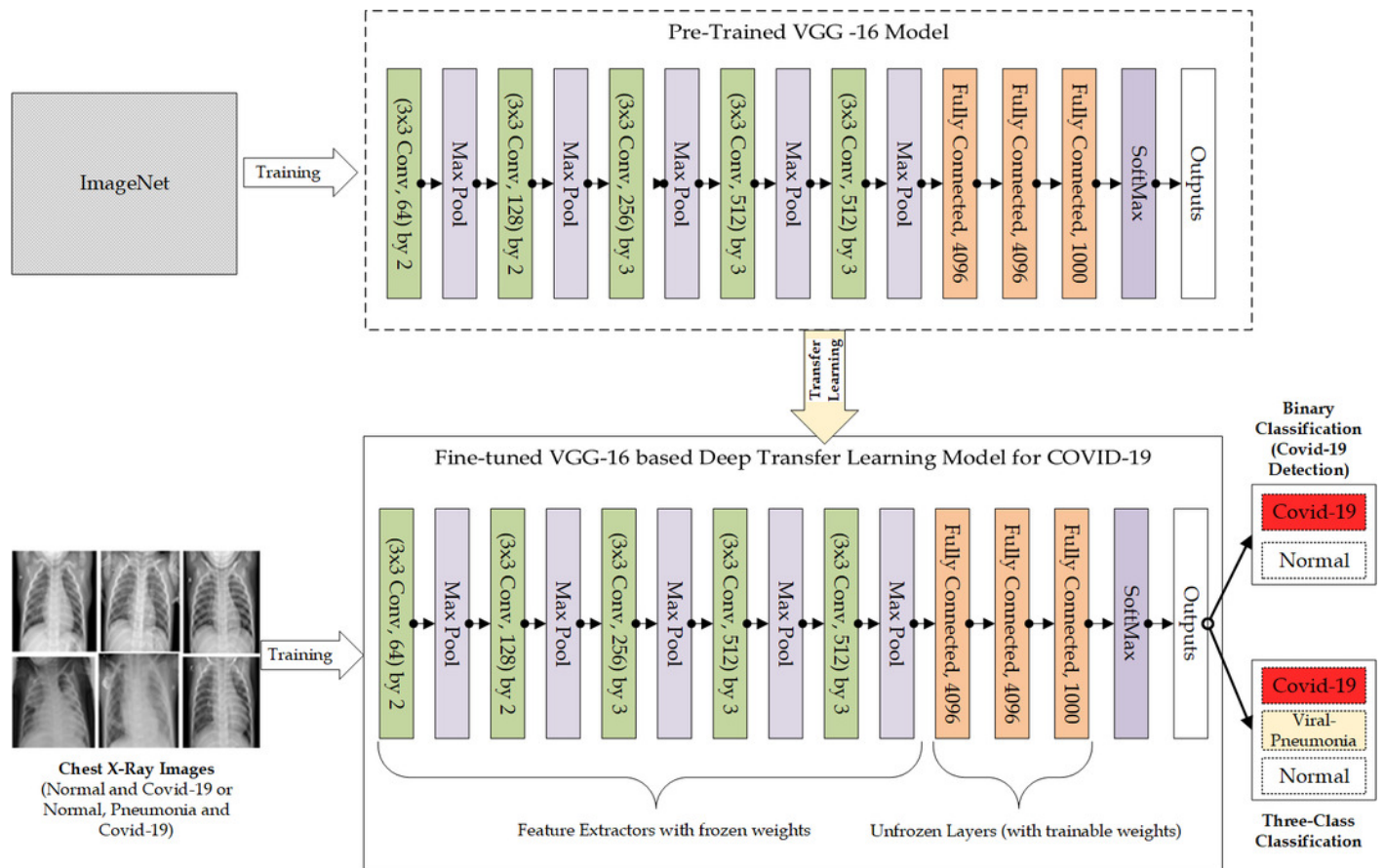
Figure 2

The Architectural Workflow of the Proposed Fine-tuned VGG-16 based Deep Transfer
Learning Model for COVID-19 Detection
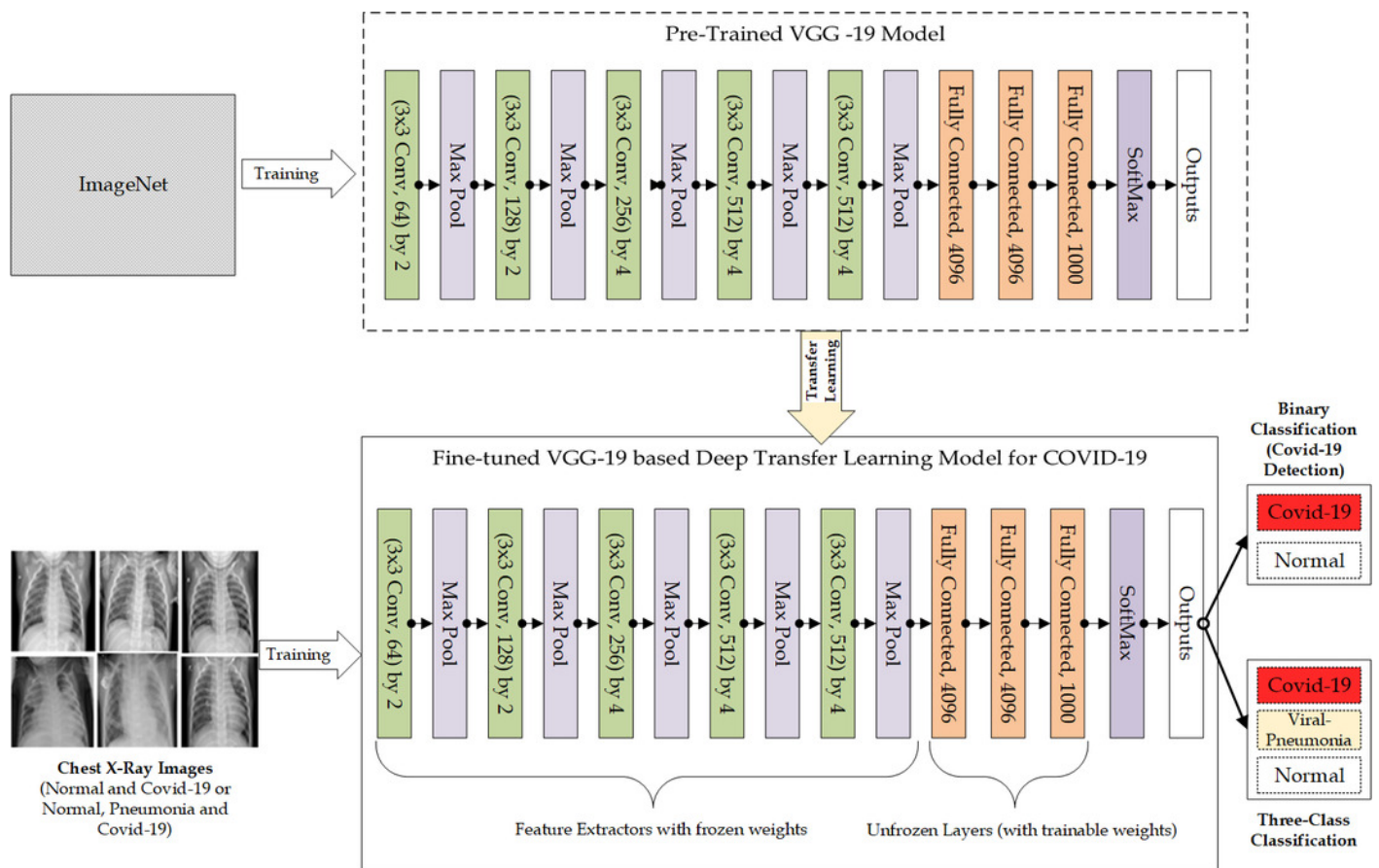
Figure 3

The Architectural Workflow of the Proposed Fine-tuned VGG-19 based Deep Transfer Learning Model for COVID-19 Detection
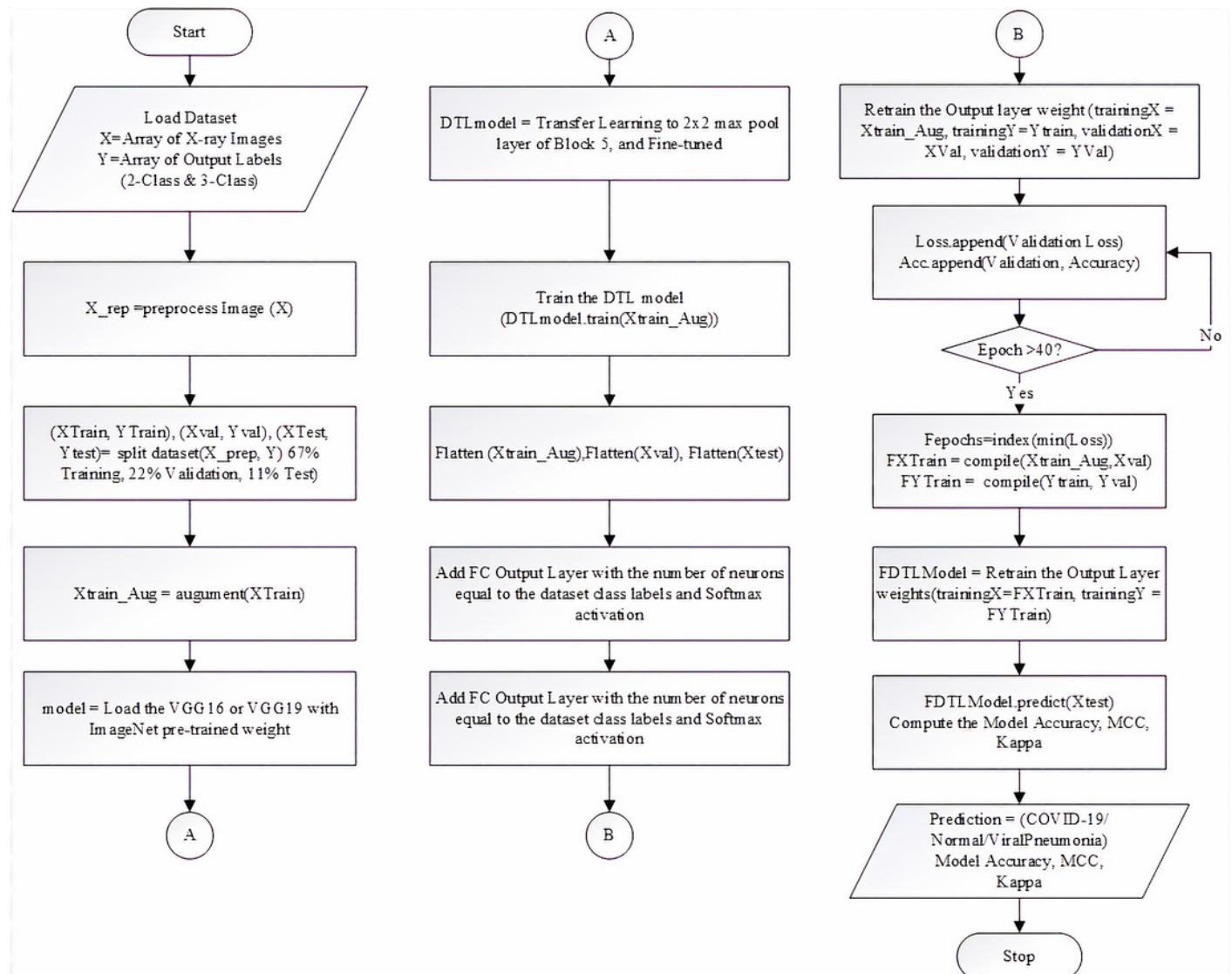
Figure 4

The Training loss and accuracy with the validation loss and accuracy curves obtained for the fine-tuned VGG-16 based Deep Transfer Learning Model (For Binary Classification)
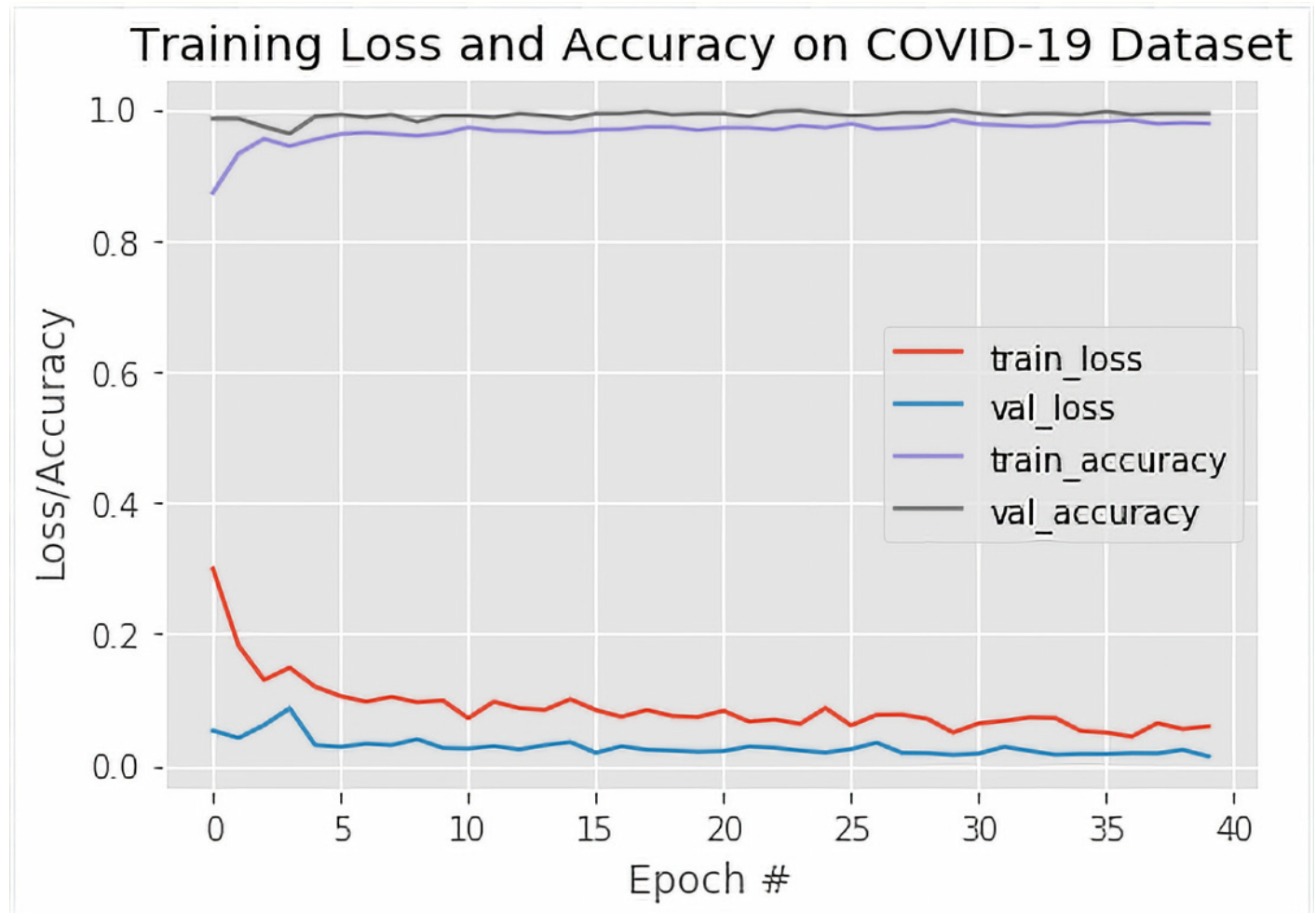
Figure 5

The Training loss and accuracy with the validation loss and accuracy curves obtained for the fine-tuned VGG-19 based Deep Transfer Learning Model (For Binary Classification)
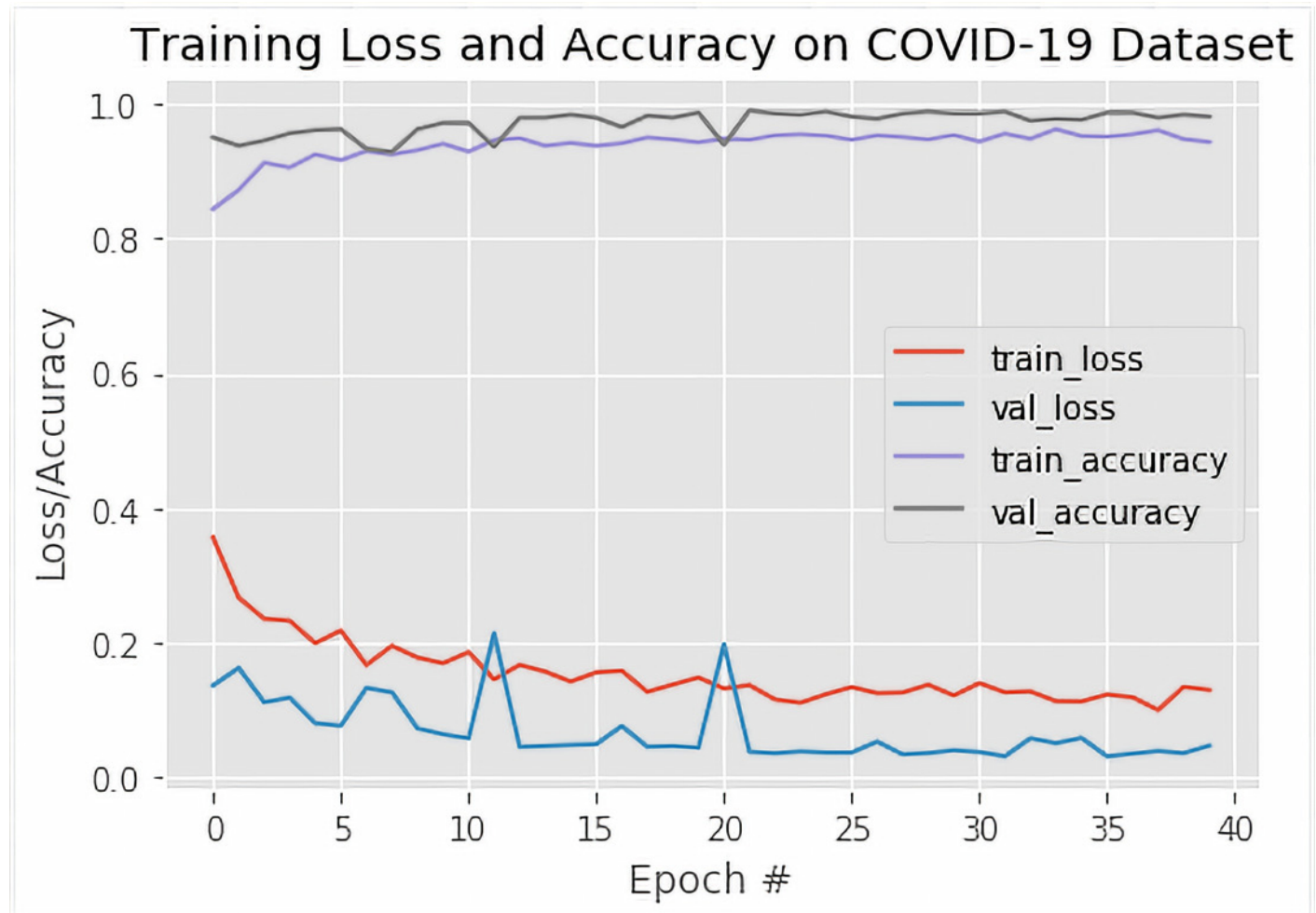
Figure 6

The Training loss and accuracy with the validation loss and accuracy curves obtained for the fine-tuned VGG-16 based Deep Transfer Learning Model (For Three-class Classification)
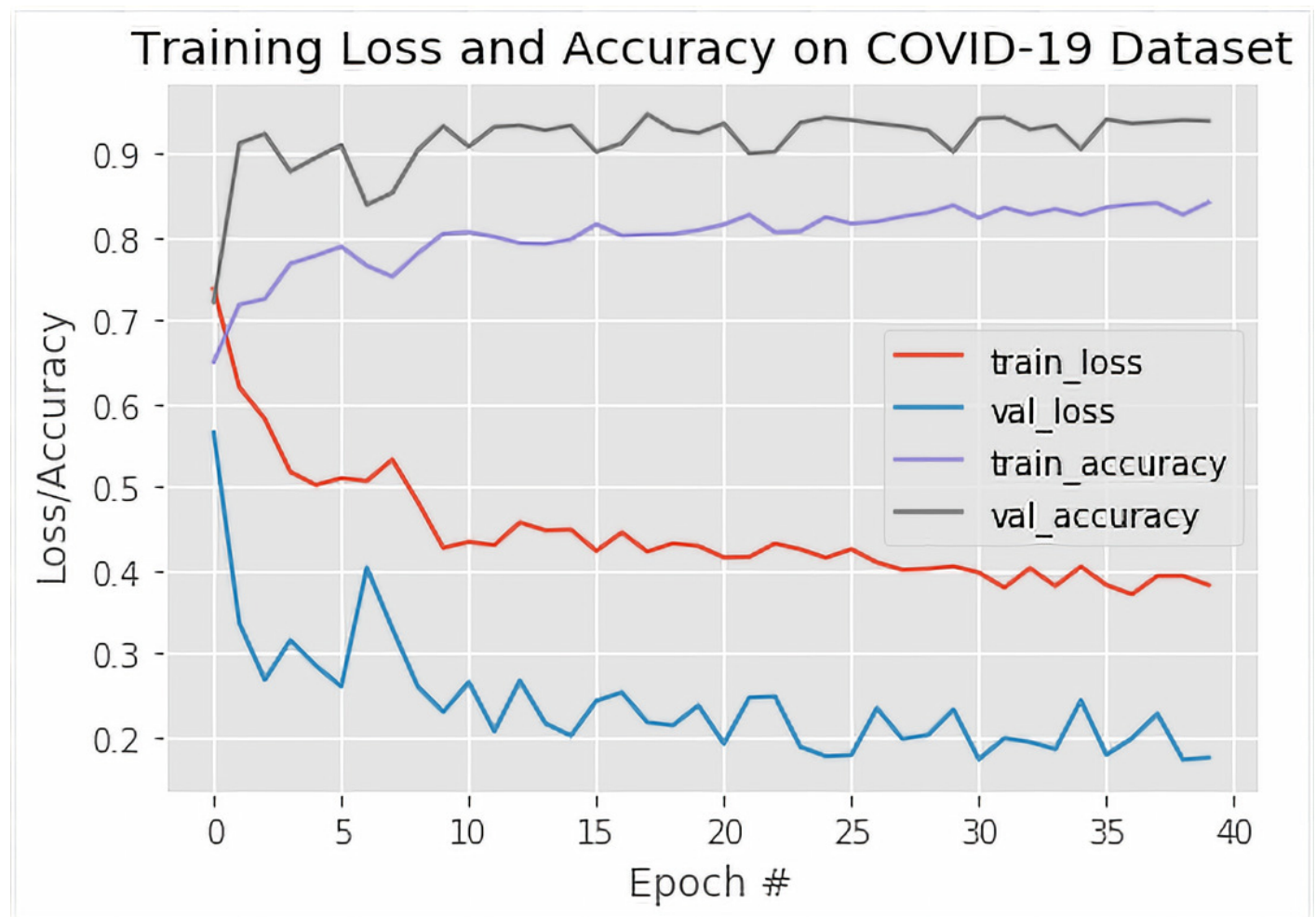
Figure 7

The Training loss and accuracy with the validation loss and accuracy curves obtained for the fine-tuned VGG-19 based Deep Transfer Learning Model (For Three-class Classification)
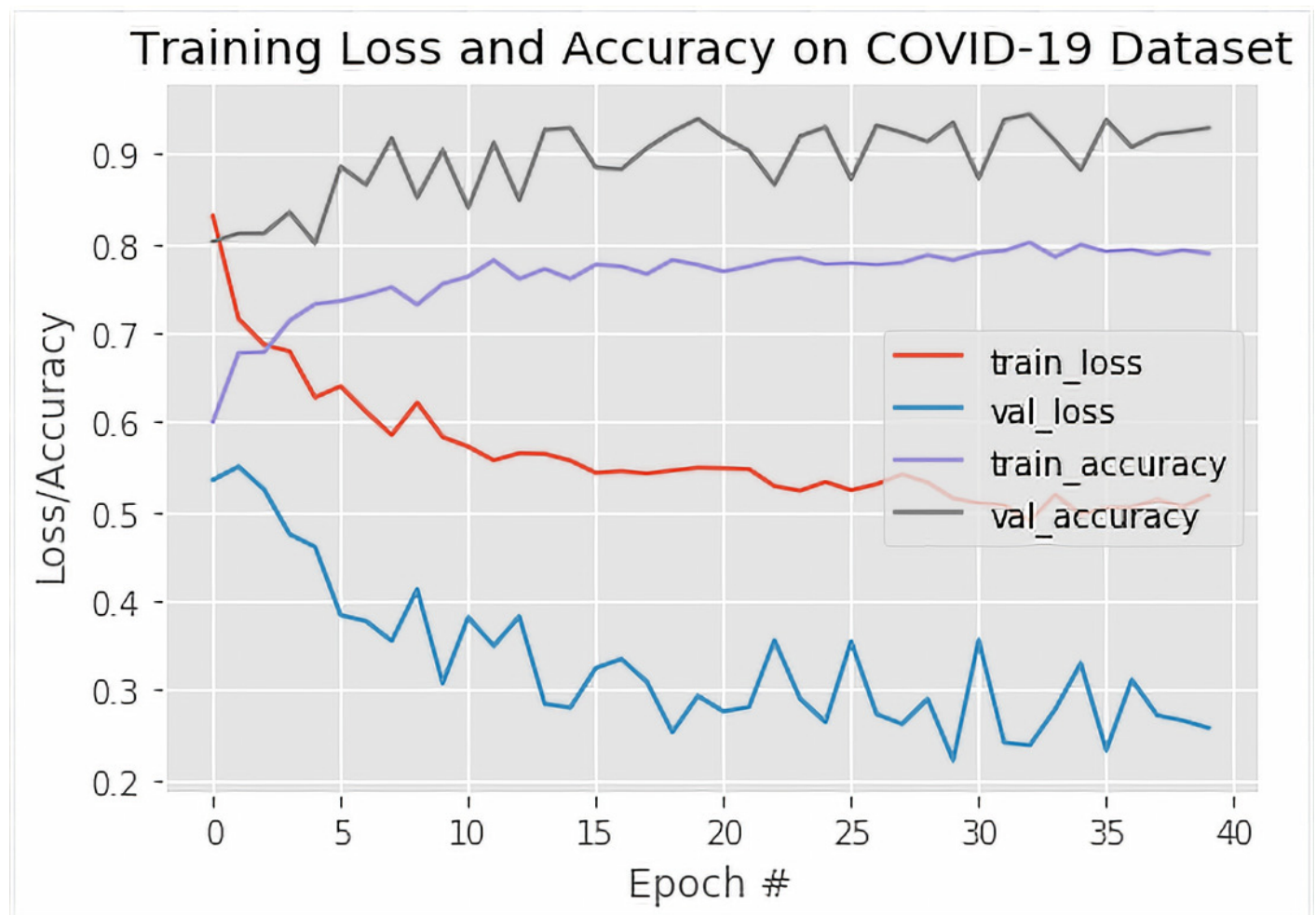
Figure 8

-19 sample test results with the predicted level of confidence value
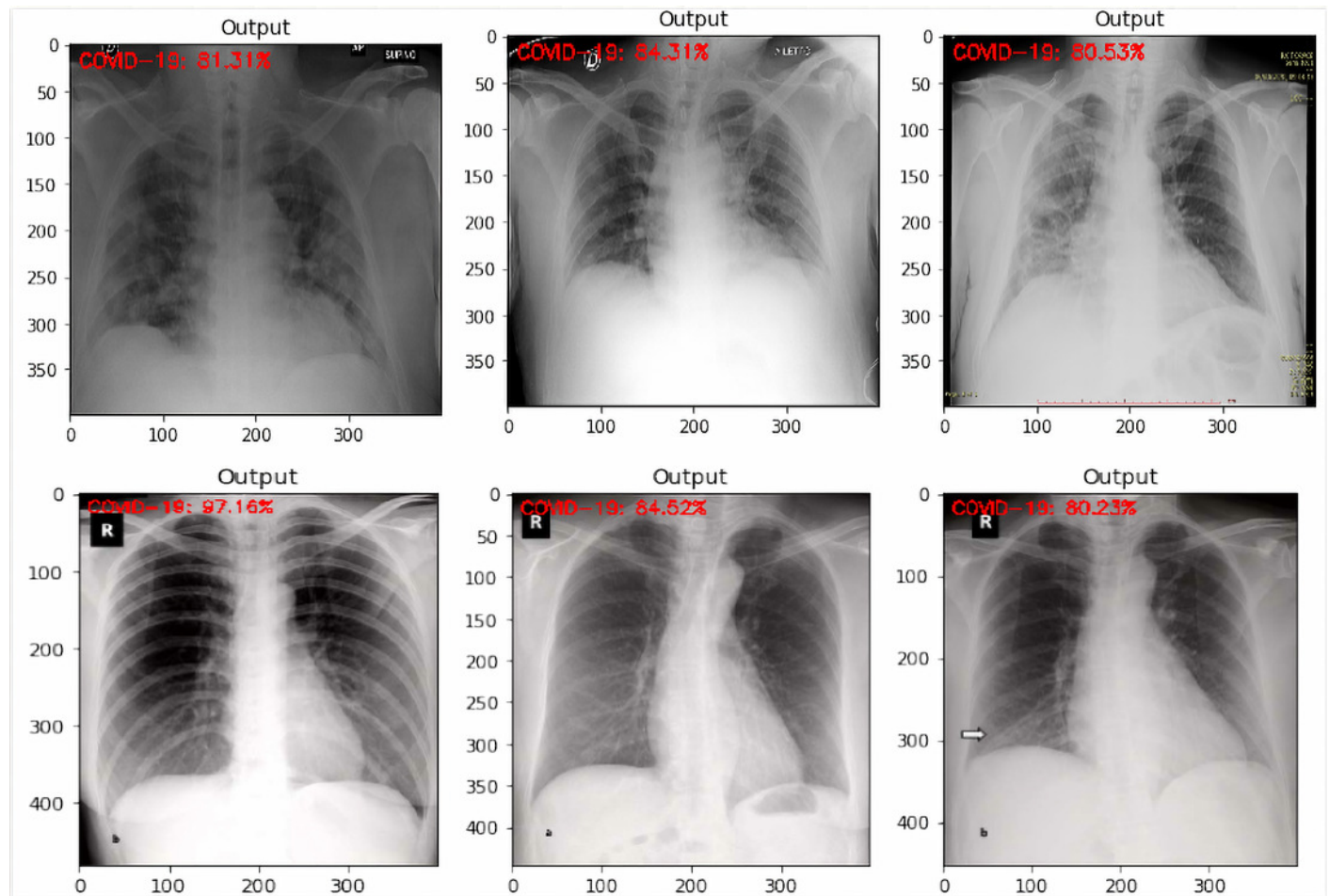
Figure 9

Viral Pneumonia sample test results with a predicted level of confidence value

Figure 10

Normal sample test results with the predicted level of confidence value

The Confusion Matrix of the Binary classification task obtained from the fine-tuned VGG-16 based DTL

1   **Table 1:** The layers and layer parameters of the proposed fine-tuned VGG-16 based DTL model
2   (Binary Classification)

	Layers	Layer's Type	Shape of Output	Num of Trainable Param
1	Convolution_1 of Block_1	Convolution 2D	[64, 224, 224]	1792
2	Convolution_2 of Block_1	Convolution 2D	[64, 224, 224]	36928
3	Convolution_1 of Block_2	Convolution 2D	[128, 112, 112]	73856
4	Convolution_2of Block_2	Convolution 2D	[128, 112, 112]	147584
5	Convolution_1 of Block_3	Convolution 2D	[256, 56, 56]	295168
6	Convolution_2of Block_3	Convolution 2D	[256, 56, 56]	590080
7	Convolution_3of Block_3	Convolution 2D	[256, 56, 56]	590080
8	Convolution_1 of Block_4	Convolution 2D	[512, 28, 28]	1180160
9	Convolution_2of Block_4	Convolution 2D	[512, 28, 28]	2359808
10	Convolution_3of Block_4	Convolution 2D	[512, 28, 28]	2359808
11	Convolution_1 of Block_5	Convolution 2D	[512, 14, 14]	2359808
12	Convolution_2of Block_5	Convolution 2D	[512, 14, 14]	2359808
13	Convolution_3of Block_5	Convolution 2D	[512, 14, 14]	2359808
14	Flatten	Flatten	[512]	0

| 15 | Dense | Dense | [64] | 32832 |
| 16 | Dense_1 | Dense | [3] | 130 |

The Precision, Recall and F1-Score obtained for the classification task using the fine-tuned VGG-16 based DTL model (Binary Classification)

1 **Table 2:** The Confusion Matrix of the Binary classification task obtained from the fine-tuned
2 VGG-16 based DTL

		PREDICTED	
		COVID-19	NORMAL
ACTUAL	COVID-19	TP = 321	FN = 0
	NORMAL	FP = 5	TN = 324

The Confusion Matrix of the Binary classification task obtained from the fine-tuned VGG-19 based DTL

1   **Table 3:** The Precision, Recall, F1-Score, MCC and Cohen's Kappa obtained for the
2   classification task using the fine-tuned VGG-16 based DTL model (Binary Classification)

	Precision	Recall	F1-score	MCC	Cohen Kappa
COVID-19	0.98	1.00	0.99	0.98	0.98
NORMAL	1.00	0.98	0.99		

The Precision, Recall and F1-Score obtained for the classification task using the fine-tuned VGG-19 based DTL model (Binary Classification)

1   **Table 4:** The layers and layer parameters of the proposed fine-tuned VGG-19 based DTL model
2   (Binary Classification)

	Layers	Layer's Type	Shape of Output	Num of Trainable Param
1	Convolution_1 of Block_1	Convolution 2D	[64,224, 224]	1792
2	Convolution_2 of Block_1	Convolution 2D	[64,224, 224]	36928
3	Convolution_1 of Block_2	Convolution 2D	[128, 112, 112]	73856
4	Convolution_2 of Block_2	Convolution 2D	[128, 112, 112]	147584
5	Convolution_1 of Block_3	Convolution 2D	[256, 56, 56]	295168
6	Convolution_2 of Block_3	Convolution 2D	[256, 56, 56]	590080
7	Convolution_3 of Block_3	Convolution 2D	[256, 56, 56]	590080
8	Convolution_4 of Block_3	Convolution 2D	[256, 56, 56]	590080
9	Convolution_1 of Block_4	Convolution 2D	[512, 28, 28]	1180160
10	Convolution_2 of Block_4	Convolution 2D	[512, 28, 28]	2359808
11	Convolution_3 of Block_4	Convolution 2D	[512, 28, 28]	2359808
12	Convolution_4 of Block_4	Convolution 2D	[512, 28, 28]	2359808
13	Convolution_1 of Block_5	Convolution 2D	[512, 14, 14]	2359808
14	Convolution_2 of Block_5	Convolution 2D	[512, 14, 14]	2359808
15	Convolution_3 of Block_5	Convolution 2D	[512, 14, 14]	2359808
16	Convolution_4 of Block_5	Convolution 2D	[512, 14, 14]	2359808
17	Flatten	Flatten	[512]	0
18	Dense	Dense	[64]	32832
19	Dense_1	Dense	[3]	130

The Confusion Matrix of the Three-class classification task obtained from the fine-tuned VGG-16 based DTL

1 **Table 5:** The Confusion Matrix of the Binary classification task obtained from the fine-tuned
2 VGG-19 based DTL

		PREDICTED	
		COVID-19	NORMAL
ACTUAL	COVID-19	TP = 308	FN = 13
	NORMAL	FP = 0	TN = 329

The Precision, Recall and F1-Score obtained for the classification task using the fine-tuned VGG-16 based DTL model (Three-class Classification)

1   **Table 6:** The Precision, Recall, F1-Score, MCC and Cohen's Kappa obtained for the
2   classification task using the fine-tuned VGG-19 based DTL model (Binary Classification)

	Precision	Recall	F1-score	MCC	Cohen Kappa
COVID-19	1.00	0.96	0.98	0.96	0.96
NORMAL	0.96	1.00	0.98		

The Confusion Matrix of the Three-class classification task obtained from the fine-tuned VGG-19 based DTL.

1   **Table 7:** The layers and layer parameters of the proposed fine-tuned VGG-16 based DTL model
2   (Three-class Classification)

	Layers	Layer's Type	Shape of Output	Num of Trainable Param
1	Convolution_1 of Block_1	Convolution 2D	[64, 224, 224]	1792
2	Convolution_2 of Block_1	Convolution 2D	[64, 224, 224]	36928
3	Convolution_1 of Block_2	Convolution 2D	[128, 112, 112]	73856
4	Convolution_2 of Block_2	Convolution 2D	[128, 112, 112]	147584
5	Convolution_1 of Block_3	Convolution 2D	[256, 56, 56]	295168
6	Convolution_2 of Block_3	Convolution 2D	[256, 56, 56]	590080
7	Convolution_3 of Block_3	Convolution 2D	[256, 56, 56]	590080
8	Convolution_1 of Block_4	Convolution 2D	[512, 28, 28]	1180160
9	Convolution_2 of Block_4	Convolution 2D	[512, 28, 28]	2359808
10	Convolution_3 of Block_4	Convolution 2D	[512, 28, 28]	2359808
11	Convolution_1 of Block_5	Convolution 2D	[512, 14, 14]	2359808
12	Convolution_2 of Block_5	Convolution 2D	[512, 14, 14]	2359808
13	Convolution_3 of Block_5	Convolution 2D	[512, 14, 14]	2359808
14	Flatten	Flatten	[512]	0

15	Dense	Dense	[64]	32832
16	Dense_1	Dense	[3]	195

The Precision, Recall and F1-Score obtained for the classification task using the fine-tuned VGG-19 based DTL model (Three-class Classification)

1   **Table 8:** The Confusion Matrix of the Three-class classification task obtained from the fine-
2   tuned VGG-16 based DTL

		PREDICTED			
		COVID-19	NORMAL	ViralPneumonia	TOTAL
ACTUAL	COVID-19	339	1	6	346
	NORMAL	0	321	18	339
	ViralPneumonia	4	31	255	290
	TOTAL	343	353	279	975

Comparison of the proposed COVID-19 diagnostic methods with other deep learning methods developed using radiology images.

1 **Table 9:** The Precision, Recall, F1-Score, MCC and Cohen's Kappa obtained for the
2 classification task using the fine-tuned VGG-16 based DTL model (Three-class Classification)

	Precision	Recall	F1-score	MCC	Cohen Kappa
COVID-19	0.99	0.98	0.98		
NORMAL	0.91	0.95	0.93	0.91	0.91
ViralPneumonia	0.91	0.88	0.90		

The layers and layer parameters of the proposed fine-tuned VGG-19 based DTL model (Three-class Classification)

1   **Table 10:** The layers and layer parameters of the proposed fine-tuned VGG-19 based DTL model
2   (Three-class Classification)

	Layers	Layer's Type	Shape of Output	Num of Trainable Param
1	Convolution_1 of Block_1	Convolution 2D	[64, 224, 224]	1792
2	Convolution_2 of Block_1	Convolution 2D	[64, 224, 224]	36928
3	Convolution_1 of Block_2	Convolution 2D	[128, 112, 112]	73856
4	Convolution_2 of Block_2	Convolution 2D	[128, 112, 112]	147584
5	Convolution_1 of Block_3	Convolution 2D	[256, 56, 56]	295168
6	Convolution_2 of Block_3	Convolution 2D	[256, 56, 56]	590080
7	Convolution_3 of Block_3	Convolution 2D	[256, 56, 56]	590080
8	Convolution_1 of Block_4	Convolution 2D	[512, 28, 28]	1180160
9	Convolution_2 of Block_4	Convolution 2D	[512, 28, 28]	2359808
10	Convolution_3 of Block_4	Convolution 2D	[512, 28, 28]	2359808
11	Convolution_1 of Block_5	Convolution 2D	[512, 14, 14]	2359808
12	Convolution_2 of Block_5	Convolution 2D	[512, 14, 14]	2359808
13	Convolution_3 of Block_5	Convolution 2D	[512, 14, 14]	2359808
14	Flatten	Flatten	[512]	0

15	Dense	Dense	[64]	32832
16	Dense_1	Dense	[3]	195

The Confusion Matrix of the Three-class classification task obtained from the fine-tuned VGG-19 based DTL

1   **Table 11:** The Confusion Matrix of the Three-class classification task obtained from the fine-
2   tuned VGG-19 based DTL

		PREDICTED			
		COVID-19	NORMAL	ViralPneumonia	TOTAL
ACTUAL	**COVID-19**	332	1	13	346
	NORMAL	2	304	33	339
	ViralPneumonia	1	19	270	290
	TOTAL	335	324	316	975

The Precision, Recall, F1-Score, Matthews Correlation Coefficient and Cohen's Kappa statistics obtained for the classification task using the fine-tuned VGG-19 based DTL model (Three-class Classification)

1   **Table 12:** The Precision, Recall, F1-Score, MCC and Cohen's Kappa obtained for the
2   classification task using the fine-tuned VGG-19 based DTL model (Three-class Classification)

	Precision	Recall	F1-score	MCC	Cohen Kappa
COVID-19	0.99	0.96	0.98		
NORMAL	0.94	0.90	0.92	0.89	0.89
ViralPneumonia	0.85	0.93	0.89		

Comparison of the proposed COVID-19 diagnostic methods with other deep learning methods developed using radiology images.

1 **Table 13:** Comparison of the proposed COVID-19 diagnostic methods with other deep learning
2 methods developed using radiology images.

S/No	Study	Type of Images	Number of Cases	Method Used	Accuracy (%)
1	Apostolopoulos and Mpesiana (2020)	Chest X-ray	224 COVID-19(+) 700 Pneumonia 504 Healthy	VGG-19	93.48
2	Wang and Wong (2020)	Chest X-ray	53 COVID-19(+) 5526 COVID-19 (-) 8066 Healthy	COVID-Net	92.4
3	Sethy and Behra (2020)	Chest X-ray	25 COVID-19(+) 25 COVID-19(-)	ResNet50+ SVM	95.38
4	Hemdan *et al.* (2020)	Chest X-ray	25 COVID-19(+) 25 Normal	COVIDX-Net	90.0
5	Narin *et al.* (2020)	Chest X-ray	50 COVID-19(+) 50 COVID-19(-)	Deep CNN ResNet-50	98.0
6	Ying *et al.* (2020)	Chest CT	777 COVID-19(+) 708 Healthy	DRE-Net	86.0
7	Wang *et al.* (2020b)	Chest CT	195 COVID-19(+) 258 COVID-19(-)	M-Inception	82.9
8	Zheng *et al.* (2020)	Chest CT	313 COVID-19(+) 229 COVID-19(-)	UNet+3D Deep Network	90.8
9	Xu *et al.* (2020)	Chest CT	219 COVID-19(-) 224 Viral pneumonia 75 Healthy	ResNet + Location Attention	86.7
10	Khalid and Youness (2020)	Chest X-ray and CT Scan	2780 bacterial pneumonia 1493 corona-virus 231 Covid19 1583 normal	Inception_Resnet_V2	92.18
				DensNet201	88.09
				Resnet50	87.54
				Mobilenet_V2	85.47
				Inception_V3	88.03

				VGG-16	74.84
				VGG-19	72.52
11	Tulin *et al.* (2020)	Chest X-ray	**2-Class:** 125 COVID-19(+) 500 No-Findings	DarkCovidNet	98.08
			3-Class: 125 COVID-19(+) 500 Pneumonia 500 No-Findings	DarkCovidNet	87.02
12	Elzeki *et al.* (2021)	Chest X-ray	**2-Class:** 221 COVID-19(+) 234 Normal	CXRVN Network	96.7
			3-Class: 221 COVID-19(+) 148 Pneumonia 234 Normal	CXRVN Network	93.07
13	Irfan *et al.* (2021)	Chest X-ray and Chest CT	3500 COVID-19 and 1500 healthy	CNN and Long-short term memory	99%
14	**Proposed Study**	Chest X-ray	**2-Class:** 1,300 COVID-19(+) 1,300 Normal	VGG-16	**99.23**
				VGG-19	98.00
			3-Class: 1,300 COVID-19(+) 1,300 Viral Pneumonia 1,300 Normal	VGG-16	**93.85**
				VGG-19	92.92