JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 06 Issue 01
2023

Research Article
https://doi.org/10.56979/601/2023

Deep Learning-based Classification of COVID-19 Variants and Lung Cancer Using CT Scans

Hassaan Malik[1*], and Tayyaba Anees[2]

[1]Department of Computer Science, School of Systems and Technology, University of Management and Technology, Lahore 54770, Pakistan.
[2]Department of Software Engineering, School of Systems and Technology, University of Management and Technology, Lahore 54770, Pakistan.
[*]Corresponding Author: Hassaan Malik. Email: f2019288004@umt.edu.pk

Received: August 11, 2023 **Accepted:** November 08, 2023 **Published:** December 05, 2023

Abstract: As the pandemic continues to spread, the number of confirmed cases of coronavirus disease (COVID-19) is constantly increasing, and new variants are continually appearing. Reverse transcription-polymerase chain reaction (RT-PCR) testing is the gold standard for COVID-19 detection. A modern approach is taking shape to address these limitations by automating analysis using deep learning (DL) techniques. These approaches can achieve good diagnostic outcomes, particularly when utilizing different imaging methods such as chest X-rays, sonography, and computed tomography (CT) images. This work aims to create a multi-classification model utilizing deep learning method to classify three different types of lung cancer: adenocarcinoma (ADC), large cell carcinoma (LCC), and squamous cell adenocarcinoma (SCC), as well as COVID-19 variants such as COVID-19 delta (COD) and COVID-19 omicron (COO). This study set out to address this gap in the literature by developing a CNN-based deep learning COVID-19 and lung cancer classification network (DLCL_Net). The model was then tested on two benchmark datasets for CT scan-based COD, COO, and lung cancer classification. We used an SMOTE Tomek approach with the DLCL_Net model to deal with the minority class problem in these two datasets. We evaluate the suggested DLCL_Net model's classification performance with six baseline deep networks such as DenseNet-201 (B_1), EfficientNet-B0 (B_2), Inception-V3 (B_3), MobileNet-V3 (B_4), Vgg-16 (B_5), and Vgg-19 (B_6) to diagnose lung diseases. The proposed DLCL_Net model with SMOTE Tomek obtained a 99.93% AUC, along with an accuracy (ACC) of 98.92%, a recall (REC) of 98.92%, a precision (PRE) of 98.92%, and an F1-score of 98.89% in categorizing the five different types of lung diseases. While the rates of ACC for B_1, B_2, B_3, B_4, B_5, and B_6 are 92.40%, 94.12%, 88.78%, 95.80%, 93.85%, and 89.93%, respectively. The findings show that the DLCL_Net model with SMOTE Tomek outperforms baseline approaches (i.e., B_1, B_2, B_3, B_4, B_5, and B_6), hence providing great assistance to radiologists and health professionals in diagnosing lung illnesses.

Keywords: COVID-19; lung disease; lung cancer; deep learning; CNN; CT scans.

1. Introduction

Since its first appearance in late 2019, in Wuhan, Hunan Province, China, the coronavirus illness known as COVID-19 has been responsible for causing a pandemic. Its influence was far-reaching and pervasive, touching on many areas of contemporary life. Extremely Serious Respiratory Illness After confirming its link to a severe pneumonia epidemic in China, the International Committee (IC) on Classification of Viruses designated the novel coronavirus as SARS-CoV-2 [1-3]. According to accounts from the World Health Organization (WHO), the COVID-19 virus can be passed from person to person. The designation of COVID-19 as a Public Health Emergency of International Concern (PHEIC) was announced by the WHO on January 30, 2020 [4]. This was done due to the virus's accelerated proliferation. It is necessary to have

diagnostic techniques that are both sensitive and easy to access given the seriousness of COVID-19 and the potential consequences that it could have, as well as the quick proliferation of affected patients if they are not separated as soon as possible. Chest radiographs and CT scans are two types of medical imaging examinations that can help to detect, it is possible to determine which symptoms are most commonly experienced by individuals who are diagnosed with COVID-19. These symptoms include a dry cough, shortness of breath, a clogged pharynx, and other symptoms that are functionally identical to those described above [5]. Viruses, such as the SARS-CoV-2 virus, progress over time and will continue to do so even as their prevalence grows in the population [6]. Sometimes different strains of the infectious agent will appear [7]. When compared to the parent virus, a virus is said to have evolved into a variety when it incorporates at least one unique change [8]. When it comes to the ease with which some coronavirus strains, such as Delta and Omicron, can be transmitted from one individual to another, there has been an increase. A nucleic acid amplification analysis, also known as an RT-PCR (reverse transcription-polymerase chain reaction), is carried out on samples obtained from the respiratory system or blood [7-9]. This confirms the presence of COVID-19. Despite this, its sensitivity is restricted, and the results are extremely reliant on the individual operating it. In addition to this, it takes an extremely long amount of time, which is a major concern in light of the speed with which innovative COVID-19 variations are spreading. Due to this particular reason, it is essential to manufacture replacements for the RT-PCR instrument. The automated assessment of medical pictures might serve as a substitute for the real-time RT-PCR test. This means that it could be used as an alternative. Chest X-rays (CXRs) and computed tomography (CT) scans are two examples of such pictures. Additionally, CT scans and CXR are beginning to exhibit a higher degree of dependability and are becoming less dependent on human controllers [9]. Multiple studies have demonstrated that COVID-19 has a large and negative effect on the lungs from the moment it is visible by a CT scan, which is in its early stages.

A study [10] concluded that after infection with COVID-19, lung carcinoma typically developed three months later. These findings support the hypothesis that COVID-19 causes lung cancer. In a similar vein, numerous studies have shown that, in addition to clinical complaints, monitoring of the blood and biochemicals as well as CT imaging of the lungs are helpful diagnostic tools for identifying the condition [9]. Unfortunately, widespread utilization of CT scans and CXR image-based computer-aided diagnostic systems is not feasible at this time. This is because there are no automated scoring methods that can assist in early COVID-19 diagnosis [11]. Despite this, several studies have demonstrated that CXR can be utilized in conjunction with RT-PCR to verify the presence of issues [12]. Cancer is an illness that is characterized by the development of cells out of control [13]. Lung cancer (LC) is the name given to cancer that starts in the passages of the lungs [14]. The lungs are the primary site of genesis for lung cancer, but the disease can spread to other tissues including the lymph glands and the brain [15]. In addition, cancer that has already spread to other systems might spread to the lungs [16]. The process by which cancer cells expand from one tissue to another is referred to as metastasis [17]. Pulmonary malignancies are generally classified into small-cell and non-small-cell categories (including adenocarcinoma and squamous cell carcinoma) [18]. These various types of lung cancer both exhibit distinctive patterns of development and respond differentially to treatment [19]. Cancer affecting lung cells that are not small is far more common than cancer affecting small lung cells. The symptoms of LC can differ greatly from one person to the next. Symptoms relating to the lungs can be experienced by some people. Some people whose LC has expanded to other areas of the body (metastasized) display symptoms that are specific to the afflicted region [20]. These symptoms cannot be found anywhere else in the body. Indications of lung cancer can include a cough that is continuous or getting worse, irritation in the chest, shortness of breath, coughing, and an ongoing feeling of exhaustion [21]. Patients afflicted with COVID-19 are most likely to exhibit this cluster of symptoms. Because of this, it can be difficult for medical professionals to differentiate between COVID-19 and LC [22].

Automatic diagnostic technologies, such as a healthcare system that is driven by artificial intelligence, are of the highest urgency in light of the COVID-19 epidemic. These technologies are necessary to speed up the search for positive patients among the general population [15-19]. Numerous studies have shown, through the application of traditional machine learning (ML) [20] and deep learning (DL) [21] algorithms, that AI has been significantly utilized in the process of COVID-19 prediction. In recent years, a variety of techniques for AI have been established, which has drastically changed the environment of a lot of scholarly fields [13]. AI-based solutions are already posing a threat to long-standing beliefs that have been

prevalent within the healthcare business. Applying DL algorithms to medical images like CXR and CT scans yields remarkable and accurate results [10-13].

Despite concerning increases in death rates caused by COVID-19 and lung cancer (LC), patients have a better than 95% probability of survival if they are diagnosed and treated promptly [13-21]. As a result, we are driven to develop a model that can identify COVID-19 variants such as COVID-19 omicron (COO), and COVID-19 delta (COD) as well as three distinct forms of lung cancer, including adenocarcinoma (ADC), large cell carcinoma (LCC), and squamous cell adenocarcinoma (SCC) using CT scans, to save the lives of people. This study develops a novel model, named a deep learning-based COVID-19 and lung cancer classification network (DLCL_Net) which is based on a CNN for the classification of COD, COO, ADC, LCC, and SCC using CT scans. This is the primary study that uses DL methods for the classification of COVID-19 variants (i.e., COD and COO) and three different lung cancer diseases. To handle the imbalance class problem of the datasets, we applied the synthetic minority oversampling technique (SMOTE) Tomek with the DLCL_Net model. Additionally, DLCL_Net is also compared with six baseline classifiers such as DenseNet-201 (B_1), EfficientNet-B0 (B_2), Inception-V3 (B_3), MobileNet-V3 (B_4), Vgg-16 (B_5), and Vgg-19 (B_6).

The following are the key contributions of this work:

1. Through the use of CT scans, the DLCL_Net model can discriminate between five distinct lung disorders, including COD, COO, ADC, LCC, and SCC. The DLCL_Net model can extract dominant CT scan characteristics to assist in identifying lung diseases.
2. We simplify the DLCL_Net model by lowering the total number of trainable parameters to construct a reliable classifier.
3. The accuracy of the CNN model is suffering as a direct consequence of the problem of class disparity that is upsetting medical datasets. This problem is directly connected to the fact that the CNN model is experiencing difficulties. We employ an upsampling method called SMOTE Tomek and obtain sample concoctions from the CT scan at each class to get a higher degree of precision in the DLCL_Net model.
4. Using the Grad-CAM heat map, we can identify all of the evident characteristics that are associated with the categorization of lung disease.
5. The proposed model outperformed six baseline classifiers, including B_1, B_2, B_3, B_4, B_5, and B_6, based on a variety of evaluation metrics, including accuracy (ACC), an area under the curve (AUC), precision (PRE), recall (REC), loss, and F1-score.
6. Additionally, when compared to the most recent state-of-the-art (SOTA) classifications, the DLCL_Net model led to the production of findings that were substantial and noteworthy.

The following sections make up this study: Section 3 presents the recent studies on chest disease classification including COVID-19, pneumonia, LC, etc. Section 3 explains the dataset description, proposed model, and performance evaluation metrics. Section 4 presents the findings and explanations of the experiments. Section 5 consists of the conclusion of the study.

2. Literature Review

In the field of DL, one of the most essential responsibilities is the classification of diseases that affect the respiratory system by utilizing a wide range of medical imaging modalities. Among these methods are magnetic resonance imaging (MRI), radiography, CT images, and sonography. According to the findings of several studies, CT scans have the potential to detect COVID-19 variations (such as COD and COO), as well as ADC, LCC, and SCC. This will result in a reduction in the amount of time and effort that medical personnel need to spend [23-27]. During the ongoing research, one of the most challenging tasks has been to detect COVID-19 variations and LC illnesses at an early stage [28-42]. A selection of the most important and pertinent articles [43-53] on artificial intelligence approaches to recognizing COVID-19, pneumonia, lung cancer, and other diseases is presented in Table 1. A comparison is made in Table 1 between the current studies on COVID-19 and other respiratory illnesses. The current study's authors have recently created a BDCNet model to detect COVID-19-contaminated CXR [26]. Three distinct categories of lung diseases—COVID-19, PN, and LC—were investigated using a single DL model. A comparison was made between the BDCNet model with three pre-trained [28-30] models by utilizing performance indicators. Since the BDCNet model's accuracy rate is higher than 94.10%, it is better than competing methodologies.

Table 1. Modern research on COVID-19 and other lung disorders diagnostic methods.

Ref	Year	Models	Diseases	Medical Imaging		Outcomes	Limitations
				CXR	CT Scans		
[35]	2023	CapsNet with IELM	COVID-19	×	✓	Accuracy = 95.39%	This study only focuses on patient data privacy. COVID-19 diagnosis was not the primary goal of the study.
[37]	2023	DMFL_Net	COVID-19 & Pneumonia	✓	×	Accuracy = 96.91%	This study only uses federated learning for sharing patient data. Shared data was imbalanced.
[26]	2022	BDCNet	COVID-19, Lungs Cancer & Pneumonia	✓	×	Accuracy = 94.10%	No pre-processing of the dataset was done before training the model.
[48]	2022	CNN	COVID-19	✓	✓	Precision = 96.0%	The imbalance class of COVID-19 was used for training the model.
[49]	2022	VGG-16	COVID-19	×	✓	Accuracy = 97.0%	Several limitations have been observed: pre-trained models were not fine-tuned, imbalance data, and no pre-processing was done before training.
[50]	2022	MobileNet-V2	COVID-19	×	✓	Accuracy = 96.40%	No right set of hyperparameters was used for augmentation.
[51]	2023	ResNet34	COVID-19	×	✓	Accuracy = 95.47%	Enhance the inference capabilities of POC devices.

[30]	2021	DNN	COVID-19	✓	×	Accuracy = 93.96%	A limited number of COVID-19 samples.
[32]	2021	Inception ResNet-V2	COVID-19 & Pneumonia	×	✓	Accuracy = 94.80%	Imbalance class of pneumonia CT scan images.
[52]	2021	Capsule Network	COVID-19	×	✓	Precision = 83.0%	This study focuses on segmentation, different CT scanners were used.

With the use of medical data, CT scans, and a DNN, Jiao et al. [30] were able to predict the severity of COVID-19 infections as well as what the binary outcomes would be. When compared to the performance of the model on the external testing dataset, the model's performance on the internal testing dataset was much better. A CNN algorithm was constructed by Oh et al. [31] for a few different datasets. Patching classifiers were given the directive to decide in favor of the majority opinion of their peers so that the desired result could be achieved. The research utilized a total of 15,043 X-ray pictures, including 8851 photographs of healthy individuals, 6012 photographs of pneumonia, and 180 photographs of infection with COVID-19. ACC of 88.9%, REC of 84.9%, F-score of 84.4%, PRE of 83.4%, and SPF of 96.4% were all attained by the proposed CNN model. A study [32] developed an Inception-ResNet-V2 model to classify and identify COVID-19. The ACC of their classification of patients as normal, COVID-19 or pneumonia was 94.8%.

Through the use of three significant contributions as building blocks, Zhang et al. [33] were able to design an 18-layer residual CNN model for CXR images. Before continuing to the phase where the features are classified, they completed the first phase by determining the most notable aspects of the image with the help of a CNN model. In the concluding portion of the analysis, the anomalous module was utilized to identify model scores. During the test, there were a total of 1531 CXR taken, of which 100 had positive results for COVID-19, while the remaining images showed signs of lung infection. Among the CXR, 100 showed positive results for COVID-19. Their recall was 96% but their precision was only 70.65% when it came to identifying COVID-19.

While testing their model on publicly available COVID-19 datasets, Kassani et al. [34] presented that feature extraction using the DenseNet-121 and ResNet-50 frameworks allowed them to achieve an exceptionally high level of accuracy. The small sample size of this study is the main flaw because it makes the model unsuitable for use in broader contexts. The study encompassed 137 positive and 137 negative incidents.

A novel solution was proposed by Malik et al. [35], which incorporates data from five distinct databases to construct a blockchain-based FL global model. FL utilizes BCT to verify the data that is required to train the model on a global scale while simultaneously maintaining the anonymity of the institutions. In the structure that was presented, three distinct components were dependent on one another. The initial phase in the procedure consisted of doing data normalization to reconcile the discrepancy of data that was acquired from five different sources by using a large number of CT machines. Following the completion of this step, CapsNet and IELMs were used to classify COVID-19 patients. Last but not least, train a global model using BCT and FL while ensuring that the participants' confidentiality is maintained. Instances of COVID-19 were classified with an accuracy score of 95.99%, and the anonymity of patients was maintained throughout the whole procedure. To distinguish between COVID-19 and normal cases, Kavitha et al. [36] proposed a ResNet-100 model utilizing logistic regression. The model performed admirably when fed images from CT scans of individuals who tested positive for COVID-19.

The researchers Malik et al. [37] developed a one-of-a-kind DMFL_Net model for medical diagnostic image processing to differentiate COVID-19 from other respiratory illnesses. To generate accurate projections, the DMFL_Net collects data from several different hospitals, creates the model with the assistance of the DenseNet-169, and makes use of confidential information that is only revealed to those individuals who have been authorized to do so. Comprehensive testing with CXR was done, and the results showed

that the suggested model not only achieves a precision of 92.45% but also successfully safeguards the anonymity of the data for a wide variety of customers.

Kermany et al. [38] assess the accuracy of a pre-trained Inception-V3 model by evaluating its performance parameter when trying to identify pneumonia in children using CXR. The diagnostic precision of the model was determined to be 92.8%, and its true negative (TN) value was determined to be 93.2%. In this study, Wang et al. [39] investigate the impact that different localization approaches have on CXR to recognize and identify pneumonia infections. The AUC for the categorization of pneumonia is determined to be 63.3% according to their methodology. Guan et al. [40] categorize respiratory illnesses using a CNN with two branches. On the 'ChestX-ray14' dataset, their proposed model for identifying pneumonia obtains an AUC of 0.776%. Both the surveillance of illness identification based on radiological pictures and the application of the segmentation technique to different corporal systems using chest X-ray imaging were conducted in [41] and [42], respectively.

Patients suffering from pneumonia were categorized by using the CNN model, which was developed by [43] using a collection of CT scan pictures. CNN algorithm was trained from the beginning to identify pneumonia and to retrieve significant and prominent characteristics from the CT scans. Impressive results were produced by the CNN-based suggested model through the processing of medical images derived from a wide variety of patients with pneumonia datasets. When it comes to achieving a high level of effectiveness, other ways rely only on classification algorithms that have been established and trained beforehand.

The detection of COVID-19 from CXR images was accomplished by Panwar et al. [44] by the use of the rapid nCOVnet model [45]. Instead of taking several hours to determine if a patient is positive or negative for COVID-19 using the RT-PCR test, the study attained an accuracy of more than 97% in under five seconds. This is in contrast to the RT-PCR test, which takes several hours.

Janizek et al. [46] claim that pneumonia is an infection and illness of the respiratory system that can be caused by fungi, bacteria, and viruses. Pneumonia is a condition that affects the respiratory system. When infected, this specific type of sickness can harm the fluid-filled air sacs in the human lungs. CXRs are routinely performed on children as part of the diagnostic process for pediatric pneumonia. Now that we are living in the modern era, it is possible to devise a self-sufficient system for diagnosing and treating pneumonia. Authors have made widespread use of the CNN method to detect juvenile pneumonia, which assists medical personnel in giving possible early treatment and has been widely used by researchers. When it comes to diagnosing pneumonia from medical images, the CNN model variant that uses pretrained layers like Vgg-19, ResNet-50, and GoogleNet has demonstrated remarkable accuracy. Because of its dense internal segment structure, the ResNet50 model was shown to obtain a high level of accuracy by Bukhari et al. [47], even though its execution time was much greater than that of other models.

Several distinct deep learning models were used in earlier investigations [28, 30, 35, 48-52], to accomplish the classification of COVID-19 cases based on CT scan images. Additionally, limited studies [35, 37] used the federated learning (FL) model for patient data privacy. Through the utilization of CXR and CT scan images, the DL model was applied in several studies [26, 32-36] to differentiate COVID-19 from pneumonia. Many studies believe that the symptoms of COVID-19 are similar to several other lung diseases such as lung cancer (i.e., ADC, LCC, and SCC), pneumonia, etc. It was a challenge for radiologists and health experts to classify and identify COVID-19 and other lung diseases using CT scans. Regarding the classification of COVID-19 variants like COO and COD, it was observed that there is not a single study that has been found that makes use of DL models. Consequently, it became apparent that a DL model-based automated framework was necessary for the diagnosis of COVID-19 variants and other lung illnesses, such as the ones listed above, using CT images.

3. Materials and Methods

The DLCL_Net model, along with six other well-known DL classifiers (such as B_1, B_2, B_3, B_4, B_5, and B_6), was subjected to exhaustive testing using the experimental procedure that is described in this section.

3.1. Study flow for the Classification of COO, COD, and Lung Cancer

The identification of COVID-19 variations and the utilization of LC images make it possible to initiate treatment at the earliest possible stage to forestall the propagation of the disease. There is a possibility that an inefficient and time-consuming manual classification of COVID-19 variations and LC will be performed.

This highlights the need for a more sophisticated and automated diagnostic approach for this illness. This will decrease the stress and workload experienced by health experts. The automatic detection of COVID-19 variants and LC at an early stage has been a challenging and effective area of study in recent years. This disease's diagnosis has been the focus of extensive study, which reduces time as well as the effort of medical professionals.

DLCL_Net is an automated system that we built for this research. It is capable of diagnosing diseases such as COD, COO, ADC, LCC, and SCC by utilizing CT scans. The five most common forms of lung disorders, namely SCC, LCC, COO, COD, ADC, and normal, were used to train and assess this system. The resolution of the input image is fixed at 150 × 150 pixels. To prevent the model from being overly accurate, the dataset was handled using the data normalization approach. To address the issue of dataset inequality and ensure a balanced number of samples in each category, we also utilized a method called the SMOTE Tomek [53]. The COVID-19 and LC dataset is designed with three discrete sections: training, testing, and validation. Furthermore, the DLCL_Net approach for these disorders is shown in Fig. 1. The size of the training parameter is less than [54-55]. During training, 30 epochs were used to carry out the experimental process. The DLCL_Net achieved the extremely high accuracy level anticipated during training and validation once all epochs were run. The confusion matrix, which contains measures such as ACC, PRE, REC, AUC, and F1-score, was used to evaluate the performance of the DLCL_Net to six B_1, B_2, B_3, B_4, B_5, and B_6 models. To highlight the visible components of COO, COD, ADC, LCC, and SCC, the Grad-CAM heat-map technique was utilized. This technique outlines the characteristics that determine the categorization of an image sample. To illustrate the aspects that contribute to the diagnosis of this condition, these characteristics have previously been utilized.

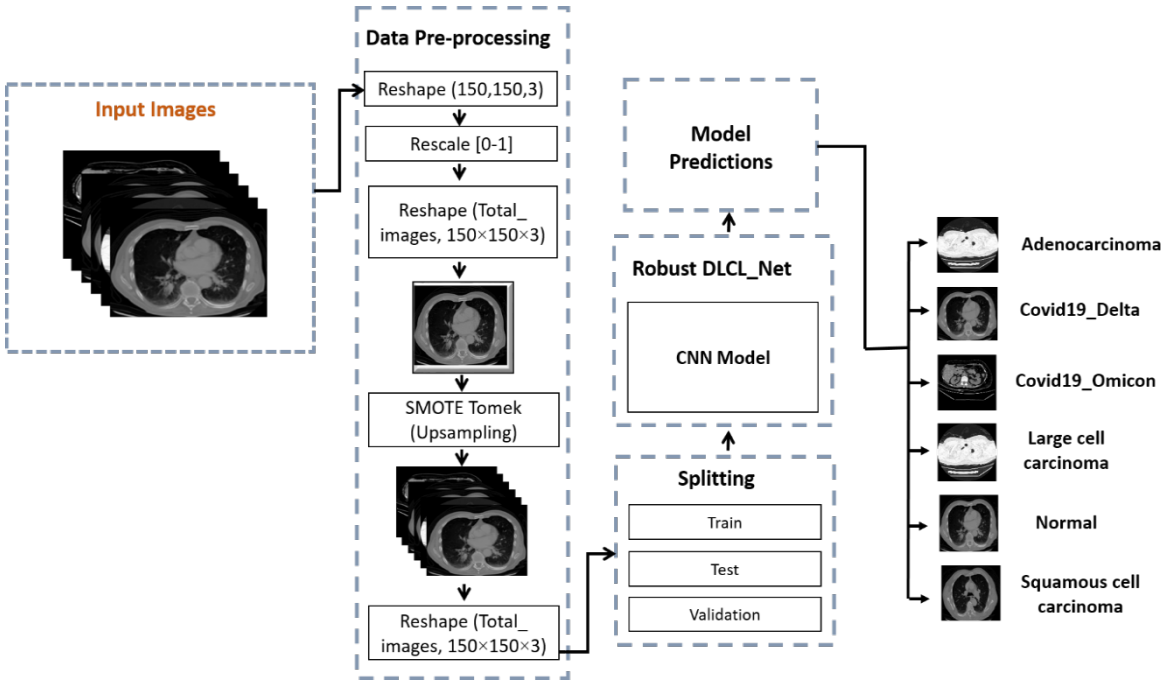


Figure 1. Work Flow of the proposed DLCL_Net.

3.2. Chest Diseases Datasets Description

A plethora of publicly available CT-Scan image datasets can be found online. The global prevalence of COVID-19 and LC prompted the current investigation into CT scans for both disorders. Fig. 2 depicts images of variants of COVID-19 such as COD, COO, and three types of LC. The DLCL_Net that was suggested was trained and assessed with datasets that were derived from two different repositories. Firstly, we collected 800 images of the COO and 863 images of the COD whereas 1540 images of normal cases from a Kaggle repository created by Gupta et al. [56]. This is a large dataset of lung CT scans that are publicly available for COVID-19 (i.e., Omicron and Delta variants). Through the analysis of a person's CT scans, this dataset aims to promote the study and creation of efficient and innovative techniques, such as deep

CNNs, that can determine whether a person is infected with COVID-19. A new sign of COVID-19 Omicron and COVID-19 Delta in the Middle East and Central Asia led to an enormous crisis and the failure of healthcare infrastructure while the COVID-19 pandemic was winding down in the majority of the world. Secondly, we collected 326 images of ADC, 163 images of LCC, and 252 images of SCC from a Kaggle repository created by Mohamed Hany [57]. Adenocarcinomas [58] of the lung are located in glands that secrete mucus and allow us to breathe in the lung's outer region. Coughing, sore throats, weight loss, and tiredness are some of the symptoms of ADC. LCC lung cancer expands and spreads rapidly, and it can occur anywhere in the lung. SCC lung cancer is located directly in the lung, in which the bigger bronchi link the airways to the lung, or in one of the major respiratory branches. Table 2 presents a detailed summary of the dataset.

Table 2. Distribution of image samples of COO, COD, and types of LC before the SMOTE Tomek process.

Chest Disease Classes	Disease Name	No. of CT scans
0	ADC	326
1	COD	863
2	COO	800
3	LCC	163
4	Normal	1540
5	SCC	252

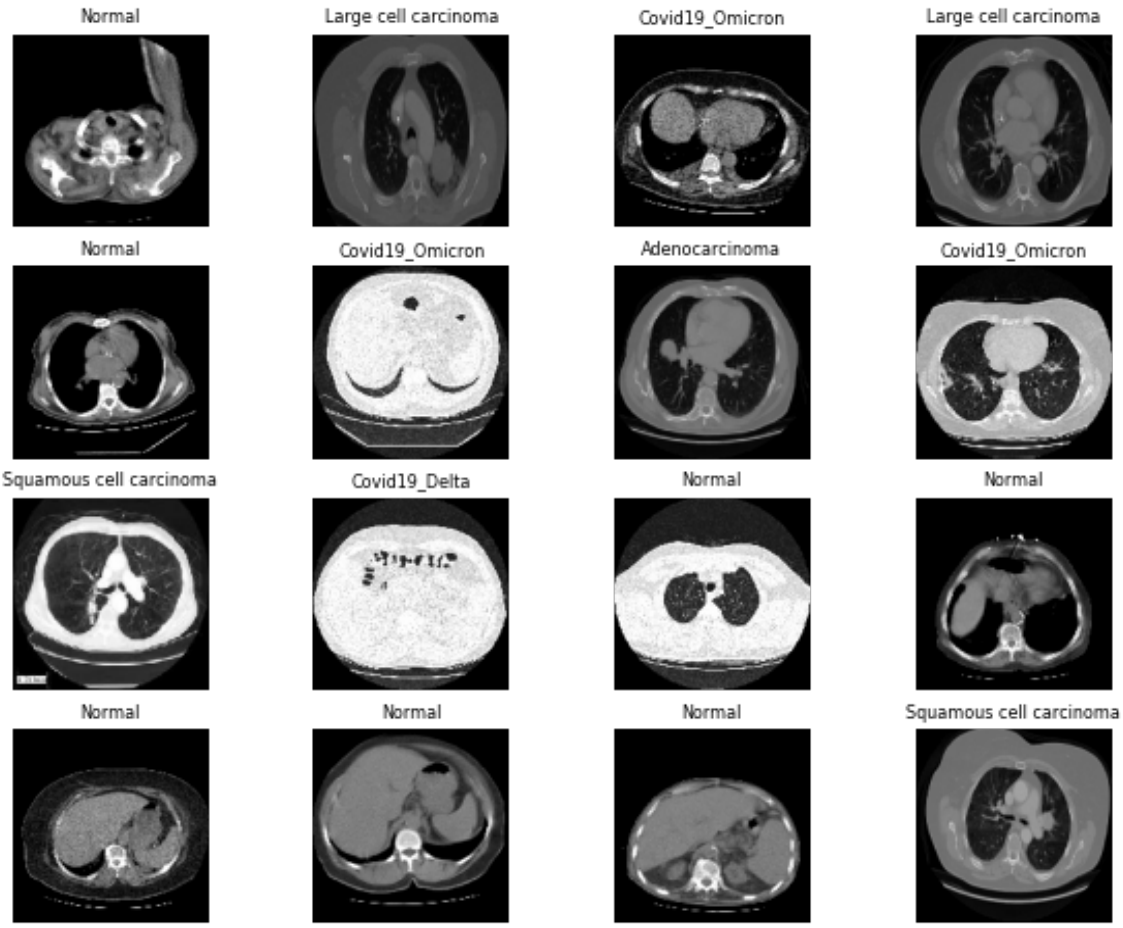


Figure 2. Real image samples of COVID-19 variants and three types of LC were taken out from two datasets.

Table 3. Distribution of image samples of COO, COD, and types of LC before the SMOTE Tomek process.

Chest Disease Classes	Disease Name	No. of CT scans
0	ADC	326
1	COD	863
2	COO	800

3	LCC	163
4	Normal	1540
5	SCC	252

3.3. Applying the SMOTE Tomek Technique

SMOTE is one of the most widely used and extremely well-oversampled approaches that researchers employ to generate false minority values in minority-class situations. SMOTE Tomek's approaches are going to be combined to achieve the goal of increasing the efficiency with which the unbalanced class may be dealt with. Through the utilization of the KNN algorithm, SMOTE can effectively produce synthetic points. As can be seen in Fig. 3, we can retrieve fusion examples for each class by employing the upsampling approach known as SMOTE Tomek [53].

The problem of classes being unequally distributed across the dataset is addressed by using the SMOTE Tomek technique. Before beginning the process of upsampling, the graphical representation of the sample distribution may be found in Table 2. An overview of the dataset that was obtained through the use of the upsampling method is presented in Table 3.

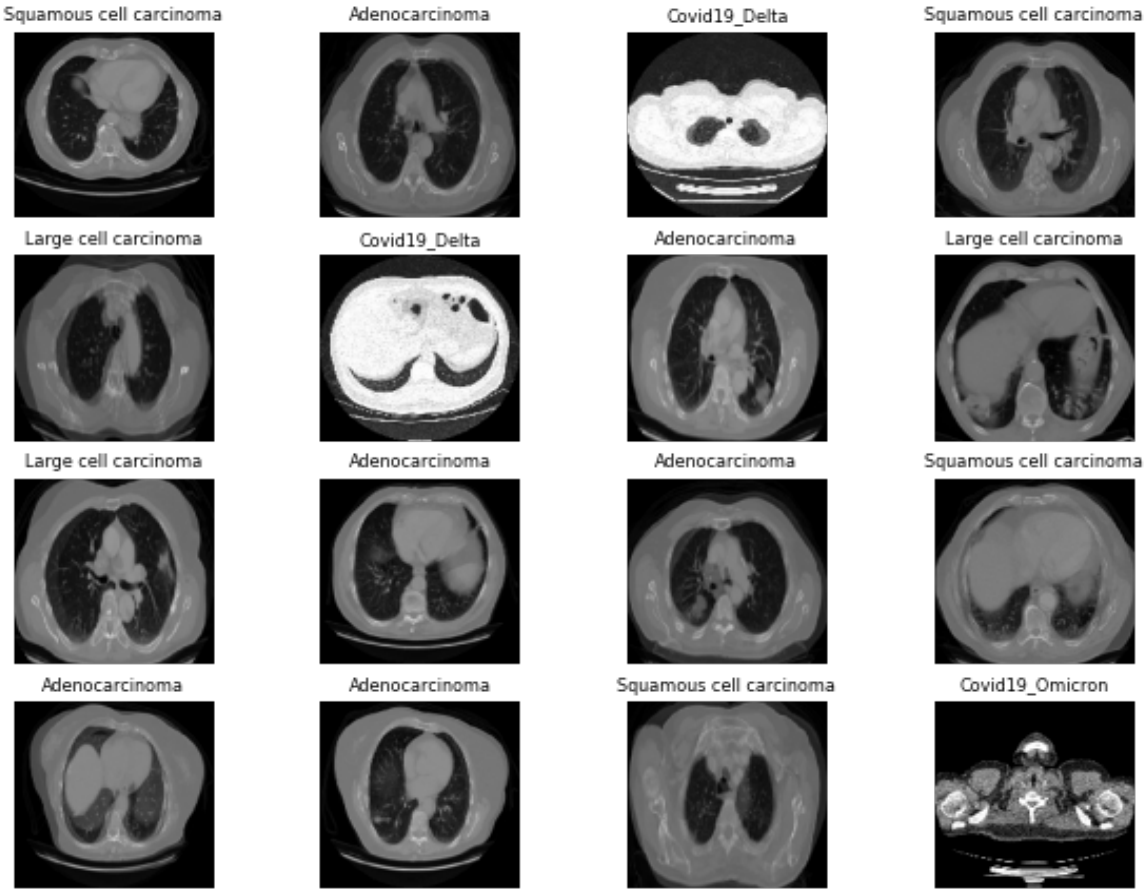


Figure 3. A solution to the issue of class imbalance is provided by SMOTE Tomek, which generates image samples.

Table 4. Image samples of COO, COD, and types of LC are distributed after SMOTE Tomek.

Chest Disease Classes	Disease Name	No. of CT scans
0	ADC	1497
1	COD	1587
2	COO	1534
3	LCC	1564
4	Normal	1540
5	SCC	1545

3.4. Proposed Model

This section provides an in-depth explanation of the suggested DLCL_Net model.

3.4.1. Architecture of the Proposed DLCL_Net

CNNs that are constructed based on the biological architecture of the human brain are extremely useful for computer vision applications such as face recognition, image segmentation, and object identification. Because of the concept of translation or spatial invariance, a CNN can recognize the same feature regardless of the number of CT images obtained. There is no variation in the location of the feature inside the CT scans; it is still able to be identified [59-60]. Within the scope of this investigation, we developed a CNN-based model that we referred to as DLCL_Net to accurately classify COD, COO, ADC, LCC, and SCC by utilizing CT images. Fig. 4 depicts the DLCL_Net model, which consists of 5 blocks of convolutional, a Rectified Linear Unit (ReLU) activation function, 2 dense layers, 1 dropout layer, and a SoftMax classification layer. A comprehensive explanation of the DLCL_Net model that has been developed for the classification of these diseases is presented in Table 4, which also includes the future layers.

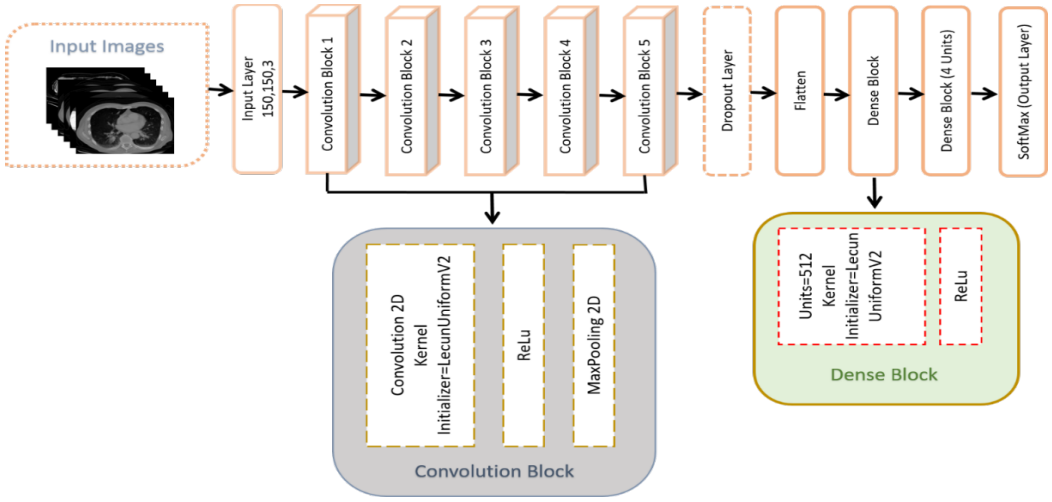


Figure 4. Architecture of the DLCL_Net model for classification of COO, COD, ADC, LCC, and SCC using CT scans.

Table 5. Breakage of layers, shape, and parameters of the DLCL_Net.

Layer_Type	Output_Shape	Parameters
DLCL_Net_InputLayer_000	(None, 150×150, 3)	0
DLCL_Net_Block_001	(None, 150×150, 8)	224
DLCL_Net_Block_002	(None, 75×75, 16)	1168
DLCL_Net_Block_003	(None, 37×37, 32)	4640
DLCL_Net_Block_004	(None, 18×18, 64)	18496
DLCL_Net_Block_005	(None, 9×9, 128)	73856
DLCL_Net_Dropout_006	(None, 4×4, 128)	0
DLCL_Net_Flatten_007	(None, 2048)	0
DLCL_Net_Dense_001	(None, 512)	1049088
DLCL_Net_ReLu_008	(None, 512)	0
DLCL_Net_Dense_002	(None, 6)	3078
DLCL_Net_Output: SoftMax_00	(None, 6)	0
Total Parameters:		1,150,554
Trainable Parameters:		1,150,554
Non-Trainable Parameters:		0

3.4.2. Convolutional Blocks

The entire convolutional block is made up of a convolutional 2D, a ReLU, and a pooling 2D that has a maximum value. Among the components that make up the overall architecture of the proposed work, the convolutional block is the most crucial component. The LecunUniformV2 initializer for the kernel layer is currently being developed to ensure that layer kernel weights are assigned appropriately. By utilizing the activation function of ReLU, the gradient vanishing problem may be resolved, and the network's capacity to figure out and carry out its responsibilities promptly can be simplified.

There is evidence of RGB channels in the image input stream. The convolutional layer is the name given to the first layer of our model for the time being. The process of applying filters is initiated by this layer, which is also frequently referred to as the kernel. As shown by Eq (1), the size of the kernel is determined by a relationship between two factors.

$$Filter\ Size(FS) = f_w \times f_h \quad (1)$$

where f_w indicates the width of the filter and f_h indicates its height. In our research, the filter size was adjusted to 3, hence Eq (1) became FS = 3 × 3. Feature identifiers are an alternative term for these filters that allow us to easily understand low-level visual characteristics, such as borders and shapes [61-62].

3.4.3. Flattened Layer

This layer follows the convolution layer and precedes the dense layer. Because dense layers require a one-dimensional architecture, while convolution layers can take tensor data types, the flattened layer was used to convert the two-dimensional image representation into the one-dimensional input data in Fig. 5.

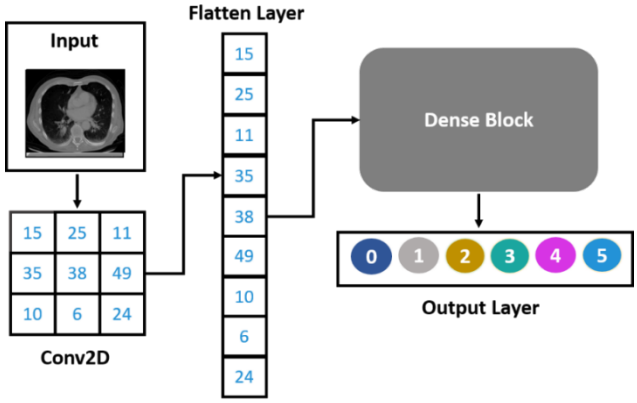


Figure 5. The basic structure that supports the flattened layer.

3.4.4. Dropout Layer

Our model made use of this layer, which had a dropout value of 0.2, to avoid overfitting with our proposed DLCL Net [64]. The goal of this layer was to activate or deactivate the units to decrease the overall training time and complexity, allowing the model to acquire the necessary characteristics.

3.5. Dense Blocks

For this investigation, we make use of two dense blocks, each of which has an activation function. The description of these activation functions is provided in the sections that occur after this one.

3.5.1. ReLU Function

A mathematical mechanism known as the activation function is responsible for determining whether or not the output of the neural network should be transmitted to the subsequent layer. Additionally, they are accountable for enabling and disabling network nodes in the majority of instances. Although DL classifiers make use of a variety of activation functions, we choose ReLU due to its simple and efficient processing. By swapping out all negative outcomes for zero, ReLU is activated. We used this activation function on the convolutional layer's outputs.

3.5.2. Dense Layer

Typically, this layer is referred to as the Fully connected layer (FCL). The dense layer only receives one matrix at a time and utilizes the attributes of that matrix to decide its output. During these stages, the classification and labeling of images take place. It is the combination of an activation function known as SoftMax and a dense layer of six neurons that results in the final output of the model. This final output categorizes the CT scans into one of six illness classes, which are as follows: ADC, COO, COD, LCC, SCC, and normal. Following the application of a couple of layers, an activation function known as SoftMax is utilized. SoftMax is a probability-based activation function, and it is an activation function in which the total number of classes is equal to the number of neurons [63]. There are a total of 1,150,554 parameters, and these parameters can be divided into two categories: 1,150,554 trainable parameters and zero non-trainable parameters.

3.6. Model Evaluations

Specifically, we address the multiclassification problems that arise when attempting to appropriately diagnose ADC, COO, COD, LCC, SCC, and normal in this work. As a result of this, a confusion matrix was used to evaluate the efficiency of the model. Before the training of the model, the dataset was first partitioned into two distinct sets: the training set and the test set. The test set was then used to evaluate the model. We use a broad range of measures to evaluate the efficiency of the model. To determine whether or not the proposed DLCL_Net is effective in identifying six illnesses, the following sets of assessment measures (see Eqs (2-5)) are used to an extreme degree.

$$Accuracy\ (ACC) = \frac{TP + TF}{TP+FN+FP+TN} \quad (2)$$

$$Precision\ (PRE)\ = \frac{TP}{TP+FP} \quad (3)$$

$$Recall\ (REC)\ = \frac{TP}{TP+FN} \quad (4)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

where TP stands for the True Positive, TN for the True Negative, and FP and FN for the False Positive and False Negative, respectively.

4. Results and Discussions

This section presents the results obtained by using the proposed DLCL_Net model and six other baseline models i.e., B_1, B_2, B_3, B_4, B_5, and B_6. Detailed results obtained by these models are presented in Table 5. To precisely evaluate the efficiency of deep neural networks, we used the same parameters for each.

4.1. Experimental Setup

There was a total of eight models that were successfully implemented with the help of Keras. These models included the six baseline models in addition to the DLCL_Net model with and without using the SMOTE Tomek method that was suggested. The programming language Python is also applied in the development of methods that are not immediately associated with CNN. For performing experimentation, a PC running Windows 10 with 32 gigabytes of memory and an NVIDIA graphics processing unit (GPU) with 11 gigabytes of memory was utilized.

4.2. Accuracy Comparison of Recent Deep Models with Proposed DLCL_Net

Applying SMOTE Tomek to the same dataset, we compared our suggested DLCL_Net to six different baseline networks. Additionally, before going ahead with the implementation of the SMOTE Tomek, we compared the proposed DLCL_Net. Upsampling is incorporated into the DLCL_Net model, which results in astonishing consequences for the model that is being offered. According to Table 5, the proposed DLCL_Net with upsampling, DLCL_Net without upsampling, B_1, B_2, B_3, B_4, B_5, and B_6 achieved respective accuracies of 98.92%, 88.73%, 92.40%, 94.12%, 88.78%, 95.80%, 93.85%, and 89.93%. Fig. 6 shows the significant improvement attained using the suggested DLCL_Net and implementing the upsampling.

Table 6. Evaluating the performance of the DLCL_Net model with baseline models such as B_1, B_2, B_3, B_4, B_5, and B_6 in classifying the COD, COO, and LC diseases using CT scan images.

Classifiers	Accuracy	Precision	Recall	F1-score	AUC
B_1	92.40%	92.71%	92.04%	92.41%	99.55%
B_2	94.12%	94.80%	94.00%	94.04%	99.70%
B_3	88.78%	89.95%	87.45%	88.62%	98.98%
B_4	95.80%	95.79%	95.44%	95.85%	99.72%
B_5	93.85%	94.38%	93.12%	93.88%	99.56%
B_6	89.93%	90.83%	89.09%	89.29%	99.23%
DLCL_Net Model (Without SMOTE Tomek)	88.73%	89.91%	87.89%	79.99%	98.54%

DLCL_Net Model (With SMOTE Tomek)	98.92%	98.92%	98.92%	98.89%	99.93%

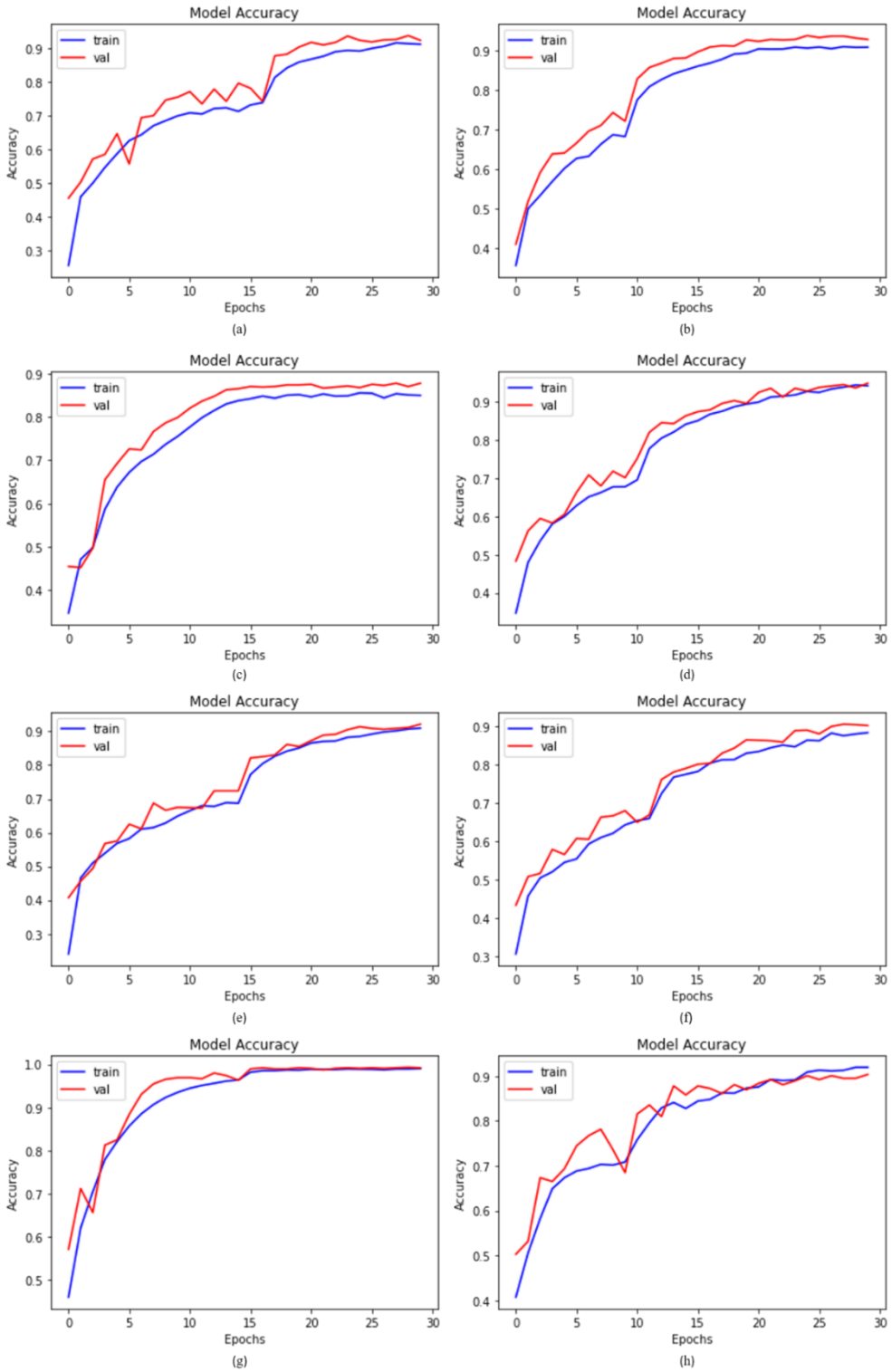


Figure 6. Training and validation accuracy; (a) B_1, (b) B_2, (c) B_3, (d) B_4, (e) B_5, (f) B_6, (g) and (h) represented DLCL_Net model with and without SMOTE Tomek, respectively.

4.3. Comparison of DLCL_Net with Baselines Models in Terms of AUC

As described previously in this study, our proposed model DLCL_Net is based on CNN with several units that are particularly effective at detecting the various COVID-19 and Lung Cancer classifications. To validate our proposed DLCL_Net, we compared it against six baseline networks. These baseline networks

such as B_1, B_2, B_3, B_4, B_5, and B_6 attained the AUC values of 99.55%, 99.70%, 98.98%, 99.72% 99.56%, and 99.23%, respectively. Using the datasets, Fig. 7 demonstrates that the DLCL_Net with and without SMOTE Tomek produced AUC values of 99.93% and 98.54 %, respectively. Based on the prior investigation, we have determined that the AUC results of the DLCL_Net model with SMOTE Tomek remain superior compared to the other six baseline CNN-based models i.e., B_1, B_2, B_3, B_4, B_5, and B_6.

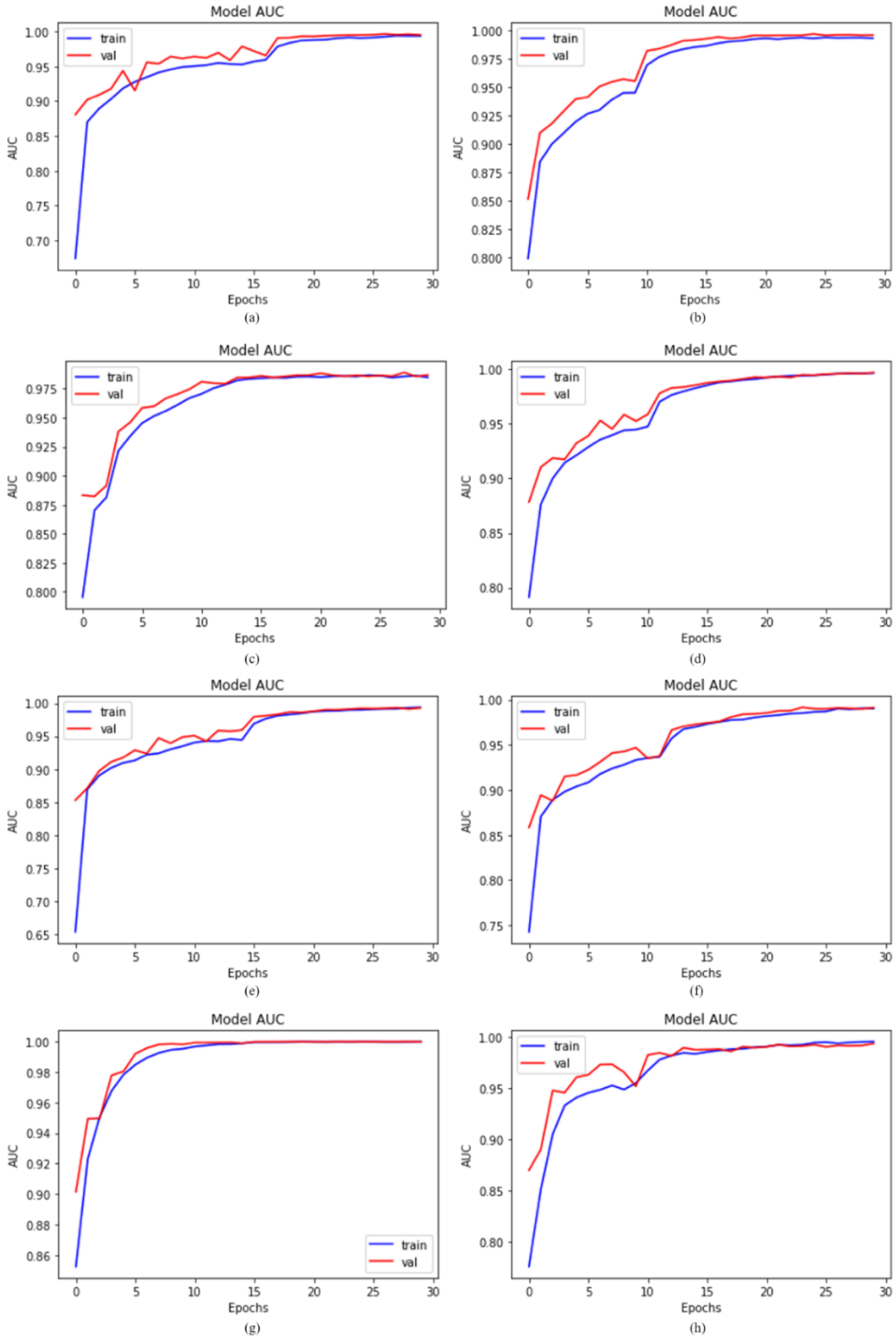


Figure 7. AUC outcomes achieved by the proposed model and baseline models; (a) B_1, (b) B_2, (c) B_3, (d) B_4, (e) B_5, (f) B_6, (g) and (h) represented DLCL_Net model with and without SMOTE Tomek, respectively.

4.4. Comparison of DLCL_Net with Baselines Models in Terms of Precision

For measuring the precision value, we compared the DLCL_Net model and six B_1, B_2, B_3, B_4, B_5, and B_6 models. The proposed DLCL_Net with and without SMOTE Tomek obtained precision values of 98.92% and 89.91%. The B_6 model achieves a precision value of 90.83%. The B_3 produces an inferior value of 89.95% as compared to other baseline models. Additionally, the B_1, B_2, B_5, and B_4 gained precision values of 92.71%, 94.80%, 94.38%, and 95.79%, respectively. The detailed results are presented in Fig. 8.

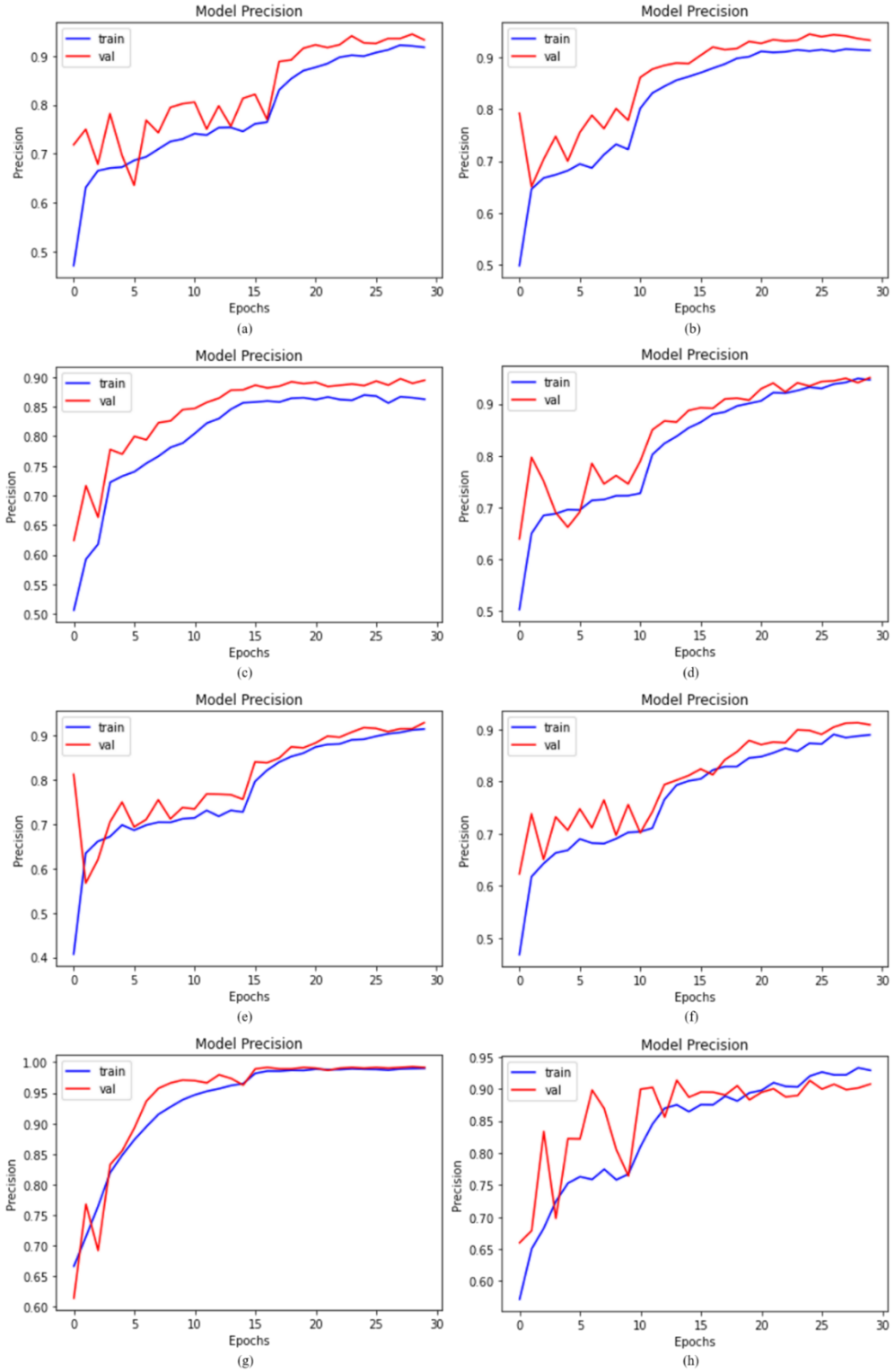


Figure 8. Visualizing the precision outcomes; (a) B_1, (b) B_2, (c) B_3, (d) B_4, (e) B_5, (f) B_6, (g) and (h) represented DLCL_Net model with and without SMOTE Tomek, respectively.

4.5. Comparison of DLCL_Net with Baselines Models in Terms of Recall

Recall is determined by dividing accurate positive predictions by the number of actual accurate positives as mentioned in Eq (4). When recall levels are high, it indicates that a greater number of positive samples have been identified. Through the utilization of a recall curve, the suggested DLCL_Net is compared to the baseline networks as shown in Fig. 9. The recall values for the proposed DCLC_Net with and without upsampling were 98.92% and 87.89%, respectively. The recall values for B_6, B_1, B_2, B_5, B_3, and B_4 were 89.09%, 92.04%, 94.00%, 93.12%, 87.45%, and 95.44%, respectively. It has been observed that the recall performance of the proposed model is remarkable.

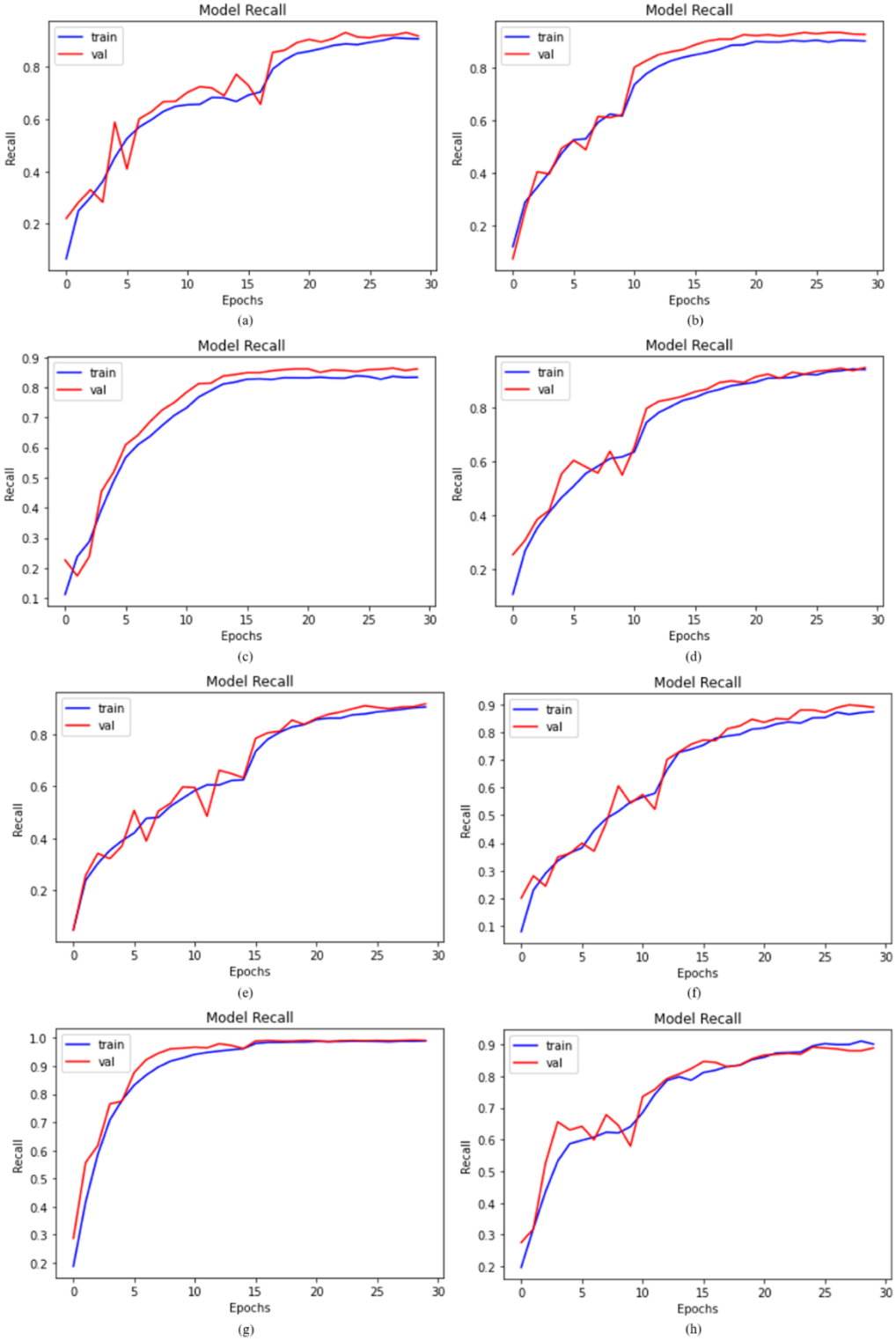


Figure 9. Measuring Recall; (a) B_1, (b) B_2, (c) B_3, (d) B_4, (e) B_5, (f) B_6 (g) DLCL_Net model with SMOTE Tomek (h) DLCL_Net model without SMOTE Tomek.

4.6. Comparison of DLCL_Net with Baselines Models in Terms of F1-score

The proposed model utilizing the SMOTE Tomek achieves the F1-score of 98.89%. The proposed DLCL_Net without using the class imbalance method attained the F1-score of 79.99%. Six baseline networks such as B_6, B_1, B_2, B_5, B_3, and B_4 obtained the F1-score values of 89.29%, 92.41%, 94.04%, 93.88%, 88.62%, and 95.85% respectively, as illustrated in Fig. 10. According to Fig. 10, the suggested DLCL_Net with SMOTE Tomek achieved the highest F1 score.

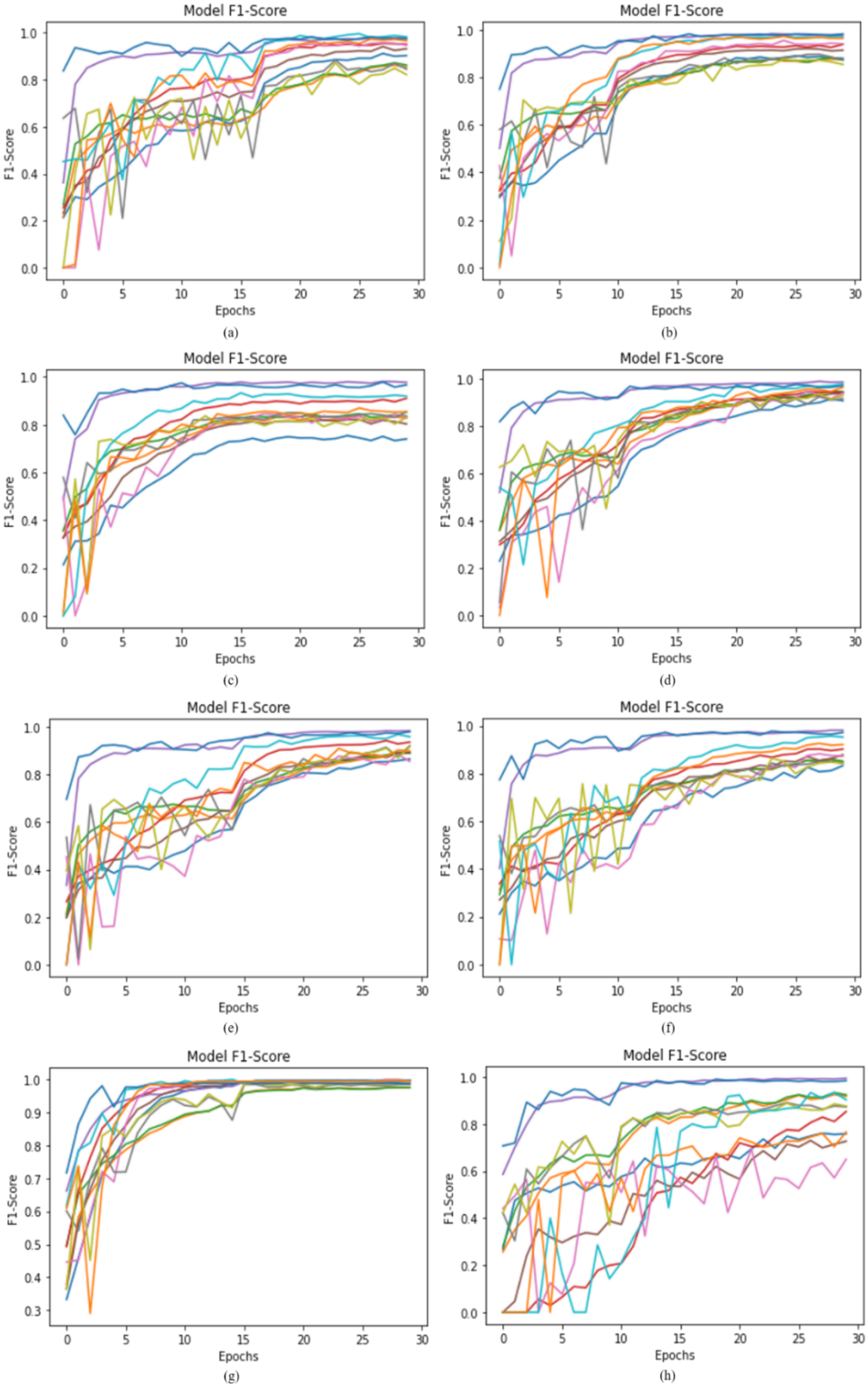


Figure 10. Computing the F1-score of the models; (a) B_1, (b) B_2, (c) B_3, (d) B_4, (e) B_5, (f) B_6, (g) and (h) represented DLCL_Net model with and without SMOTE Tomek, respectively.

4.7. Comparison of DLCL_Net with Baselines Models in Terms of Loss

It is the difference between the actual values and the expected values that is computed using loss functions. In this work, a categorical cross-entropy method was used to compute the loss. The outputs, on the other hand, are more spectacular when the network is constructed using up-sampling photos. The proposed DLCL_Net with and without upsampling achieved loss values of 0.0427 and 0.3417, whereas B_6, B_1, B_2, B_5, B_3, and B_4 obtained loss values of 0.2616, 0.2033, 0.1647, 0.1886, 0.3117 and 0.1244, respectively. Fig. 11 depicts this significant enhancement in DLCL_Net loss with SMOTE Tomek.

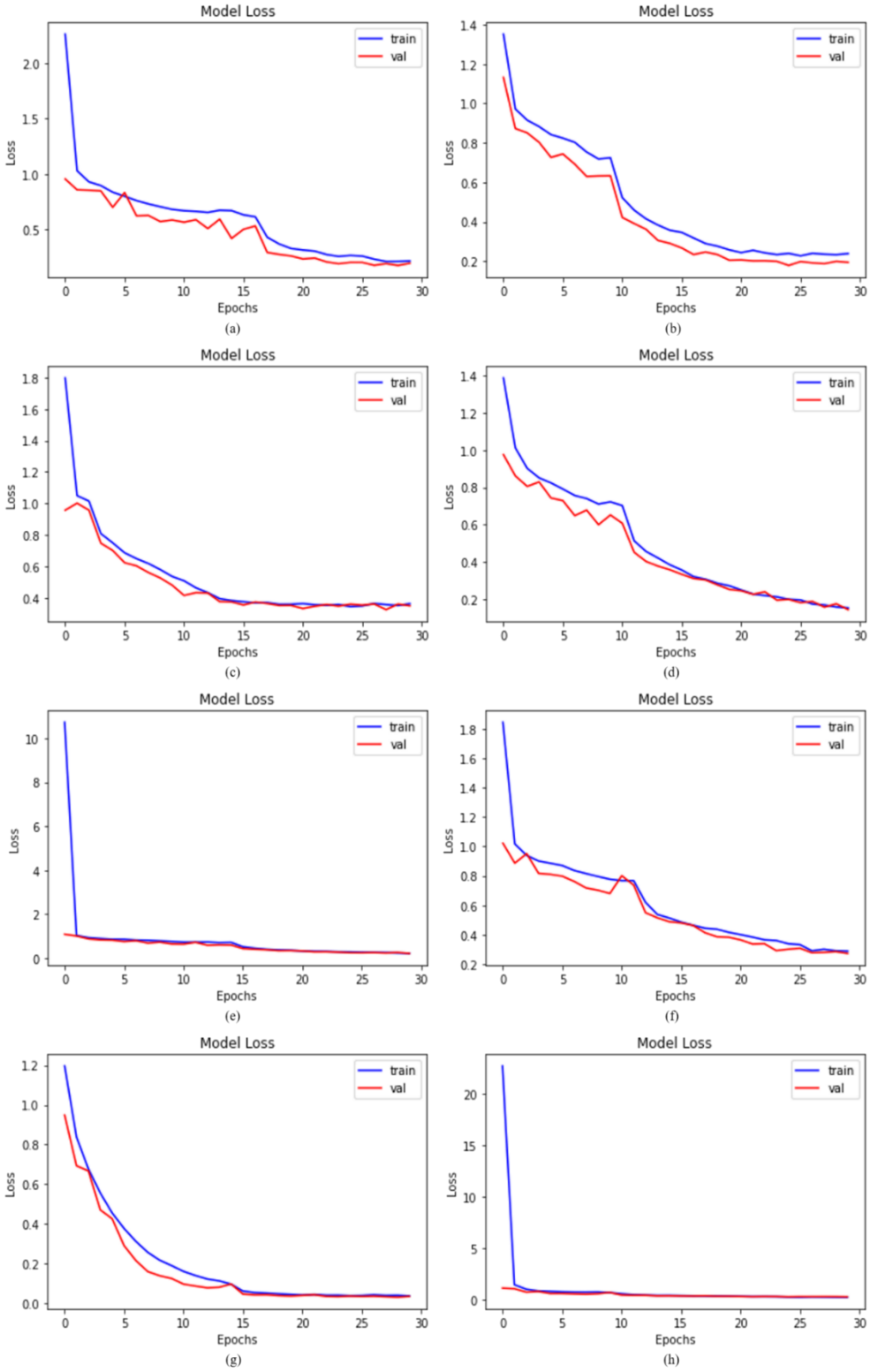


Figure 11. Evaluating the training and validation loss; (a) B_1, (b) B_2, (c) B_3, (d) B_4, (e) B_5, (f) B_6, (g) and (h) represented DLCL_Net model with and without SMOTE Tomek, respectively.

4.8. Comparison of DLCL_Net with Baselines Models in Terms of ROC

Specifically, it is implemented to investigate the efficacy of diagnostic tests, more specifically the predictability of a binary or multi-classifier. The AU(ROC) is utilized to evaluate the performance of a classifier; a higher AUC indicates a more effective classifier. We investigated the accuracy of our proposed DLCL_Net on the curve with and without upsampling by creating a dataset. Using the same dataset, this curve evaluates the proposed DLCL_Net with and without upsampling to six models such as B_1, B_2, B_3, B_4, B_5, and B_6. According to Fig. 12, the ROC values of the proposed DLCL_Net with and without upsampling, B_6, B_1, B_2, B_5, B_3, and B_4 were 0.9915, 0.9328, 0.9139, 0.9141, 0.9261, 0.9363, 0.9101, and 0.9715, respectively. Fig. 12 shows this significant improvement of the suggested DLCL_Net with SMOTE Tomek on the ROC curve.

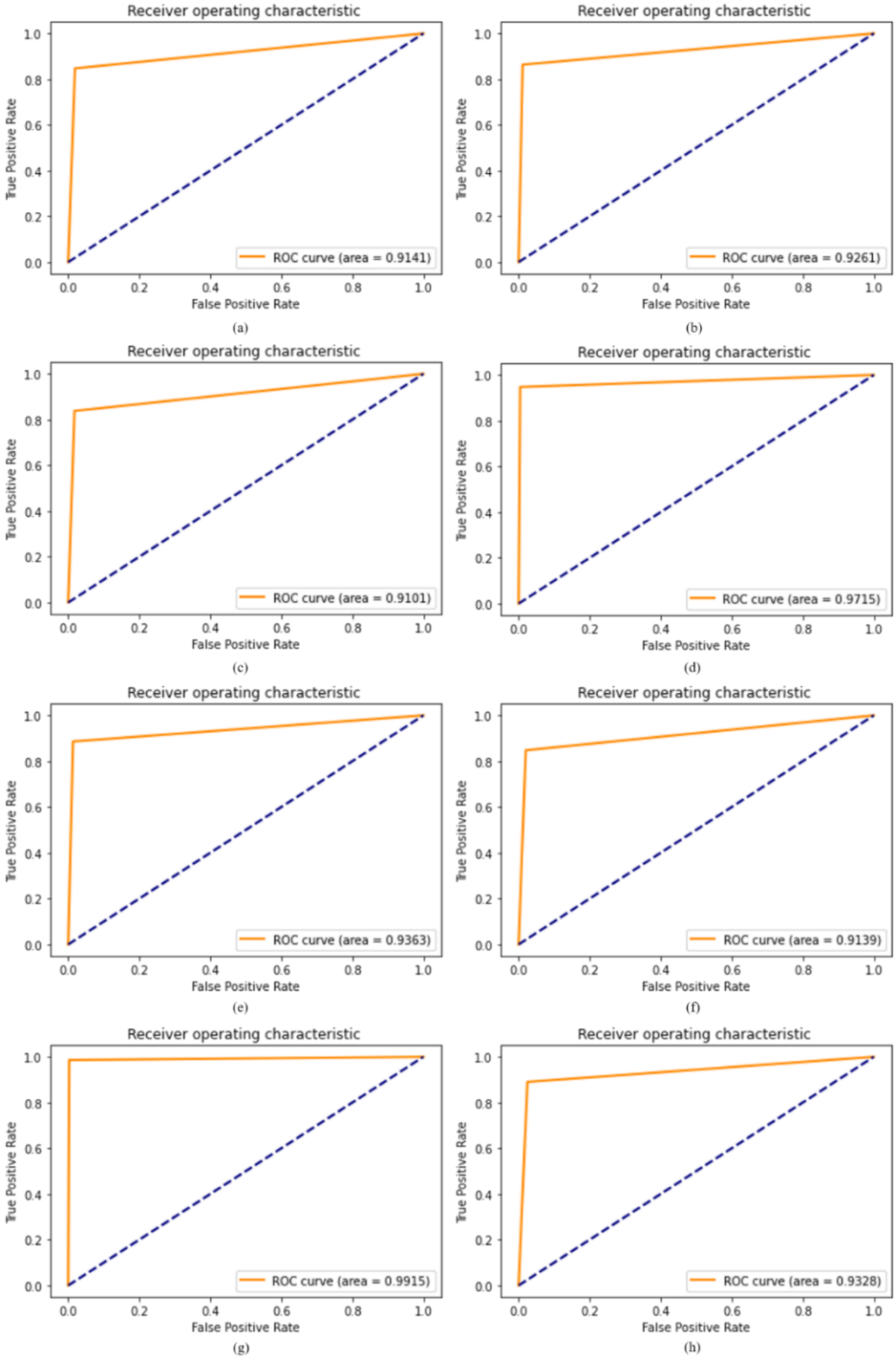


Figure 12. ROC curve evaluating the DLCL_Net model and baseline models; (a) B_1, (b) B_2, (c) B_3, (d) B_4, (e) B_5, (f) B_6, (g) and (h) represented DLCL_Net model with and without SMOTE Tomek, respectively.

4.9. Comparison of DLCL_Net with Baselines Models in Terms of AUC(ROC)

The extension of the ROC curve is used in Fig. 13 to display a comparison between the suggested DLCL_Net and six baseline networks. After balancing the dataset using the upsampling method, the proposed method improved greatly in contrast to the six models illustrated in Fig. 13. In the AUC, this major impact was also detected for class 0 (ADS), class 1 (COD), class 2 (COO), class 3 (LCC), class 4 (Normal), and class 5 (SCC) provided by the proposed DCLC_Net with and without the SMOTE Tomek. These improvements to AUC give evidence that the DLCL_Net selection of features is reliable and that the SMOTE Tomek method is also quite useful.

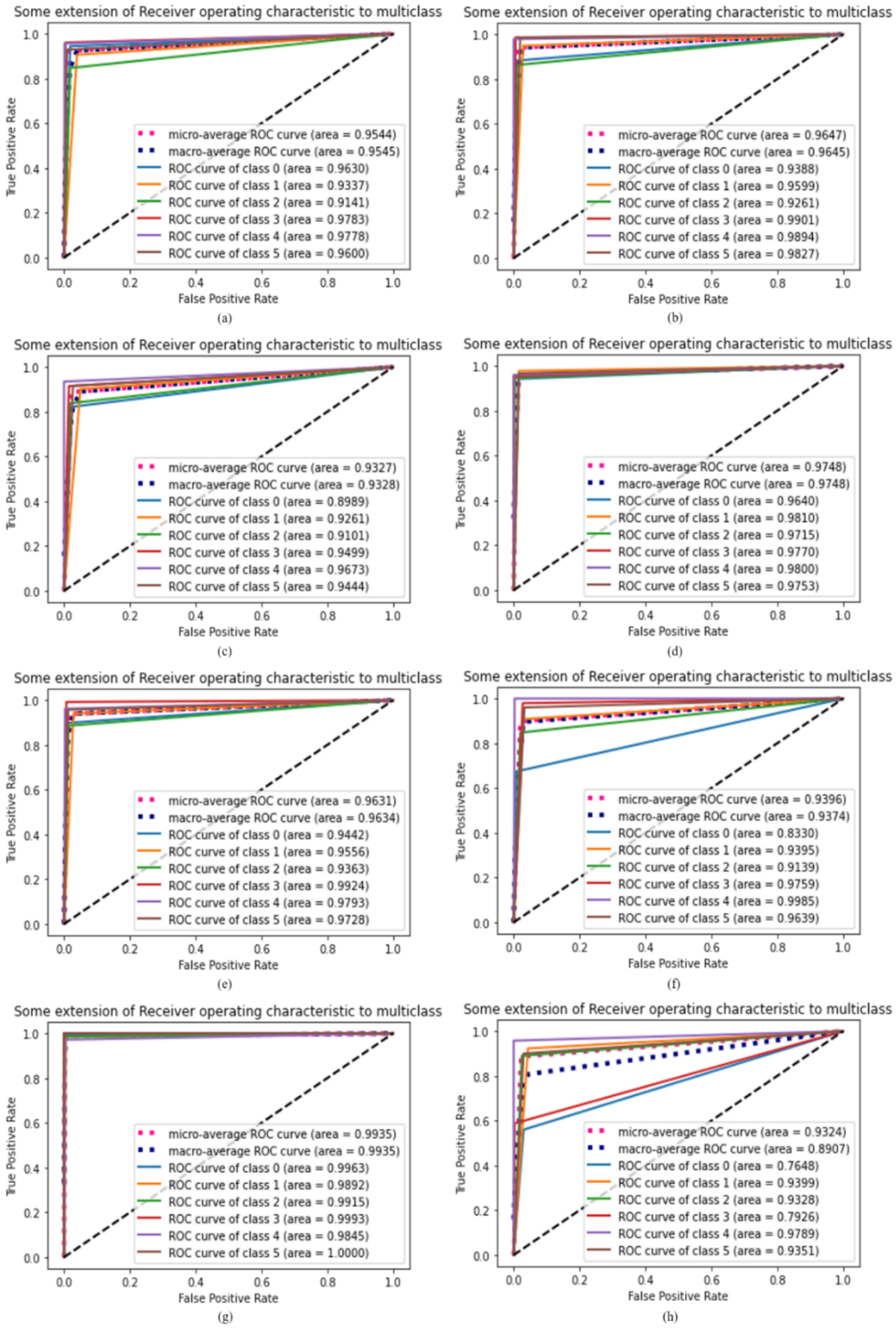
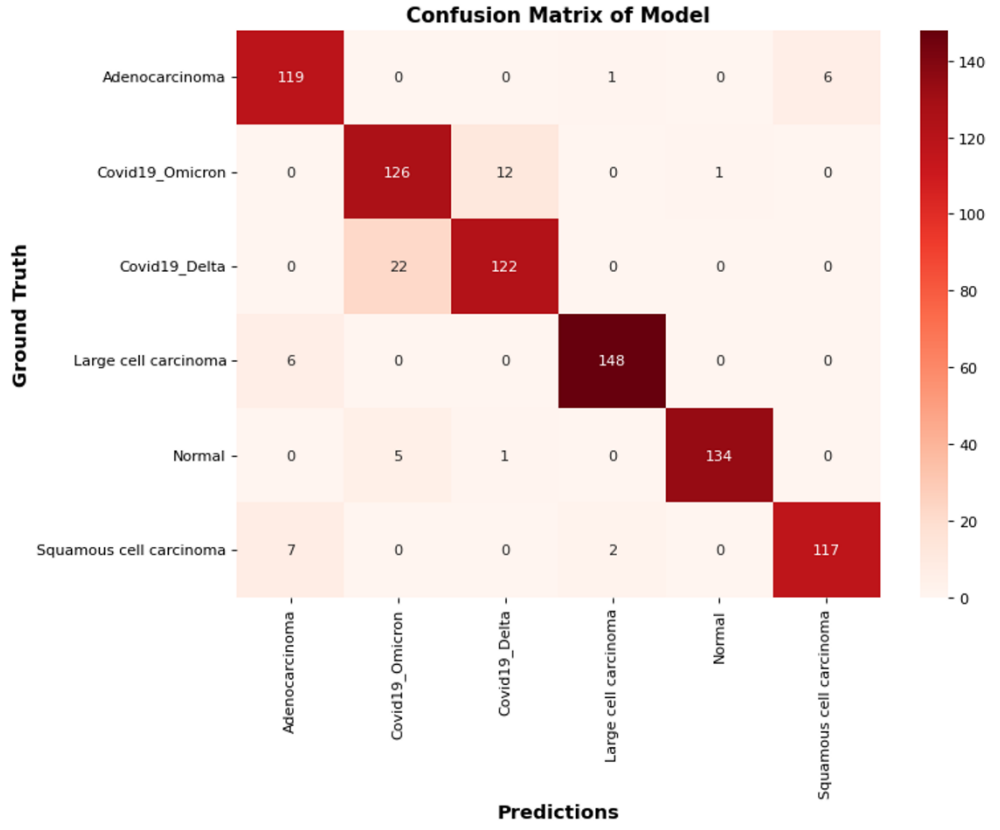


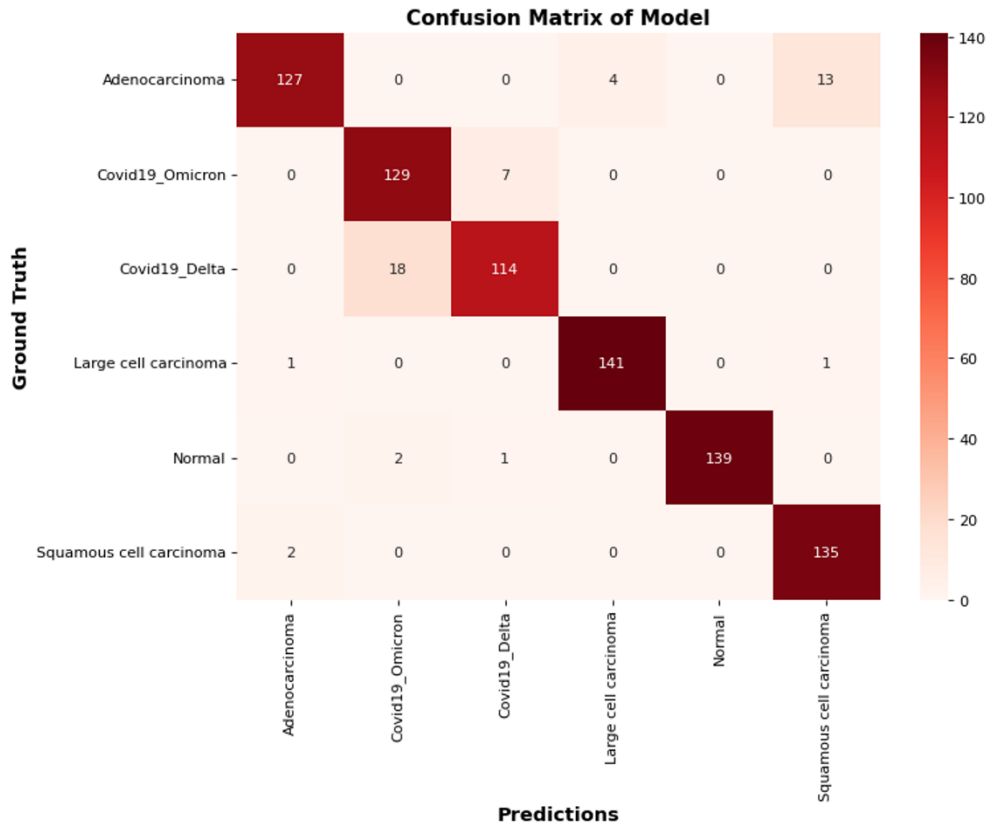
Figure 13. AUC(ROC) of the models used in this work; (a) B_1, (b) B_2, (c) B_3, (d) B_4, (e) B_5, (f) B_6, (g) and (h) represented DLCL_Net model with and without SMOTE Tomek, respectively.

4.10. Confusion Matrix Comparison of Proposed DLCL_Net with Other Networks

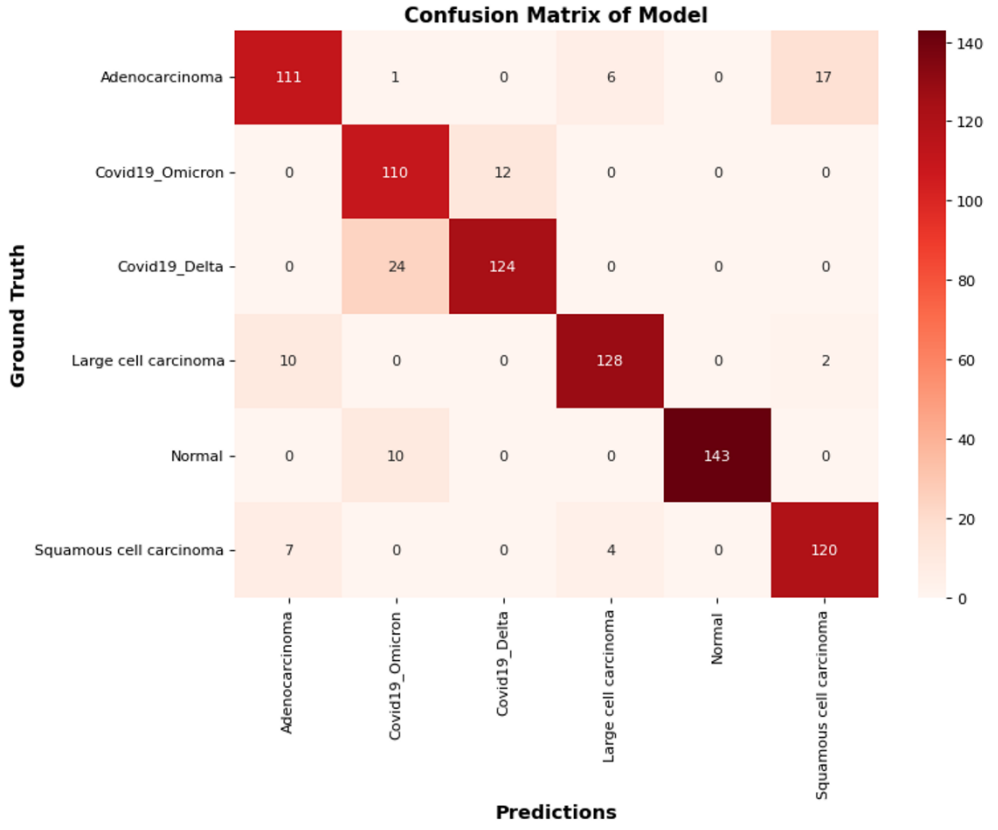
To evaluate the DLCL_Net model using the confusion matrix, we examined it with other networks. As seen in Fig. 14, the system applying to upsampled gives effective growth for DLCL_Net.



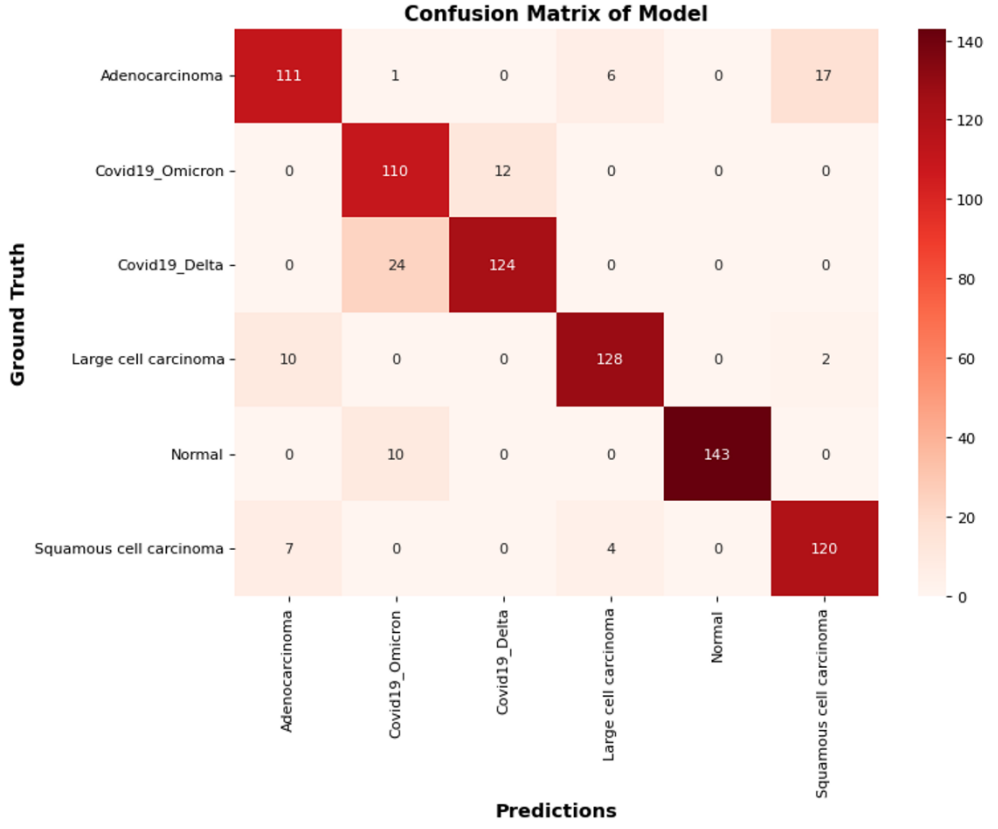
(a)



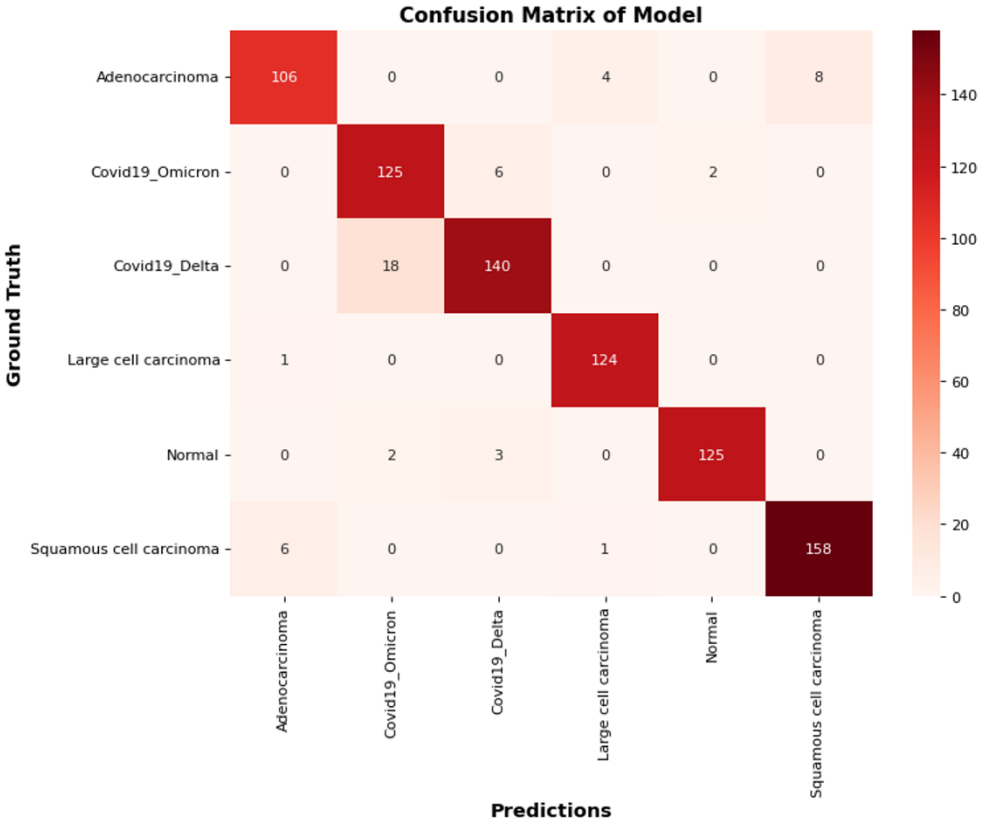
(b)


Confusion Matrix of Model
Adenocarcinoma
Covid19_Omicron
Covid19_Delta
Large cell carcinoma
Normal
Squamous cell carcinoma
Ground Truth
Predictions
111 1 0 6 0 17
0 110 12 0 0 0
0 24 124 0 0 0
10 0 0 128 0 2
0 10 0 0 143 0
7 0 0 4 0 120
140
120
100
80
60
40
20
0

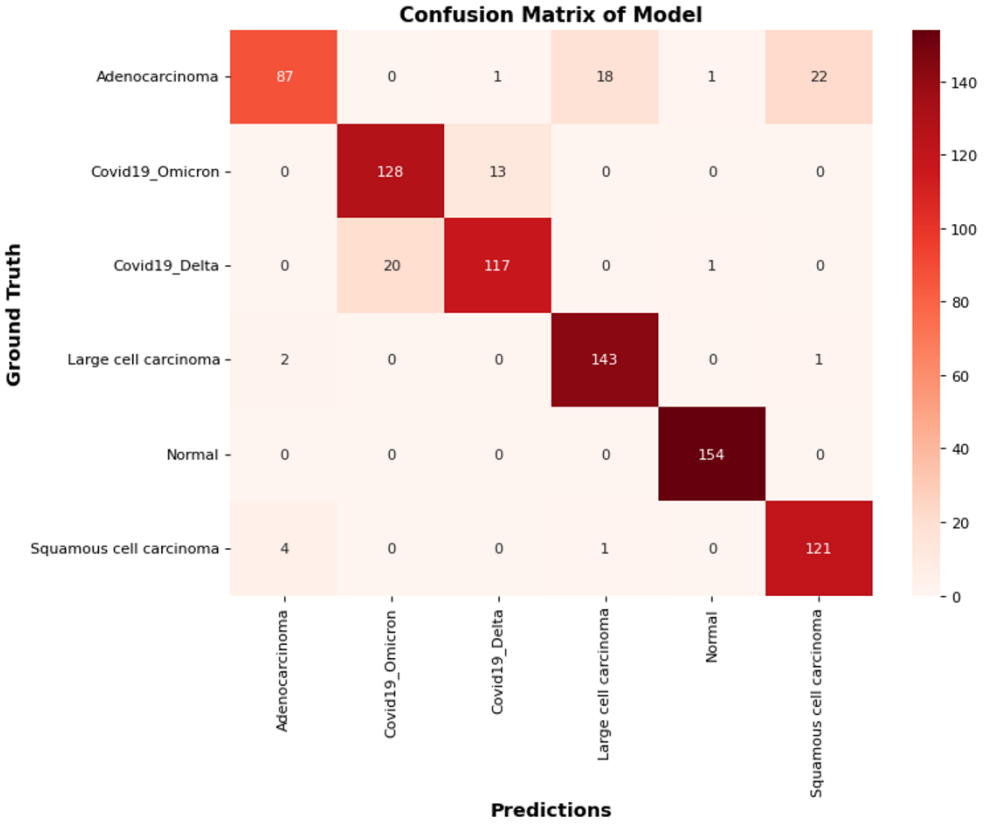
(c)


Confusion Matrix of Model
Adenocarcinoma
Covid19_Omicron
Covid19_Delta
Large cell carcinoma
Normal
Squamous cell carcinoma
Ground Truth
Predictions
111 1 0 6 0 17
0 110 12 0 0 0
0 24 124 0 0 0
10 0 0 128 0 2
0 10 0 0 143 0
7 0 0 4 0 120
140
120
100
80
60
40
20
0

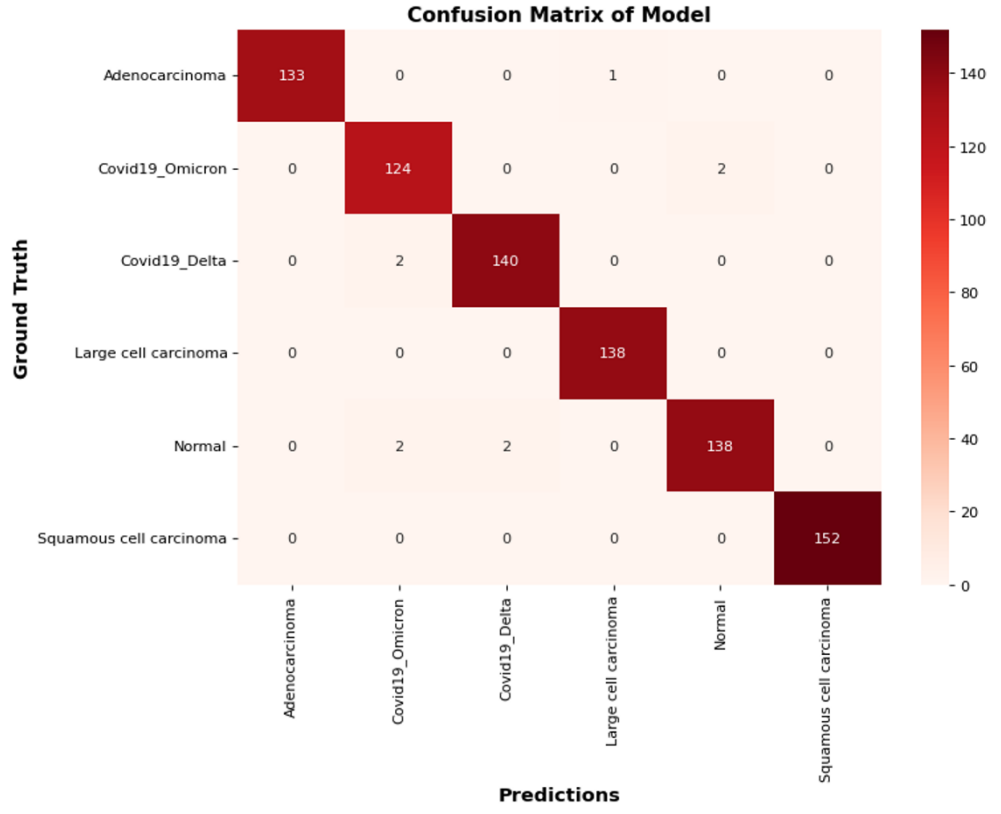
(d)


Confusion Matrix of Model
Adenocarcinoma
Covid19_Omicron
Covid19_Delta
Large cell carcinoma
Normal
Squamous cell carcinoma
Ground Truth
Predictions
106 0 0 4 0 8
0 125 6 0 2 0
0 18 140 0 0 0
1 0 0 124 0 0
0 2 3 0 125 0
6 0 0 1 0 158
140
120
100
80
60
40
20
0

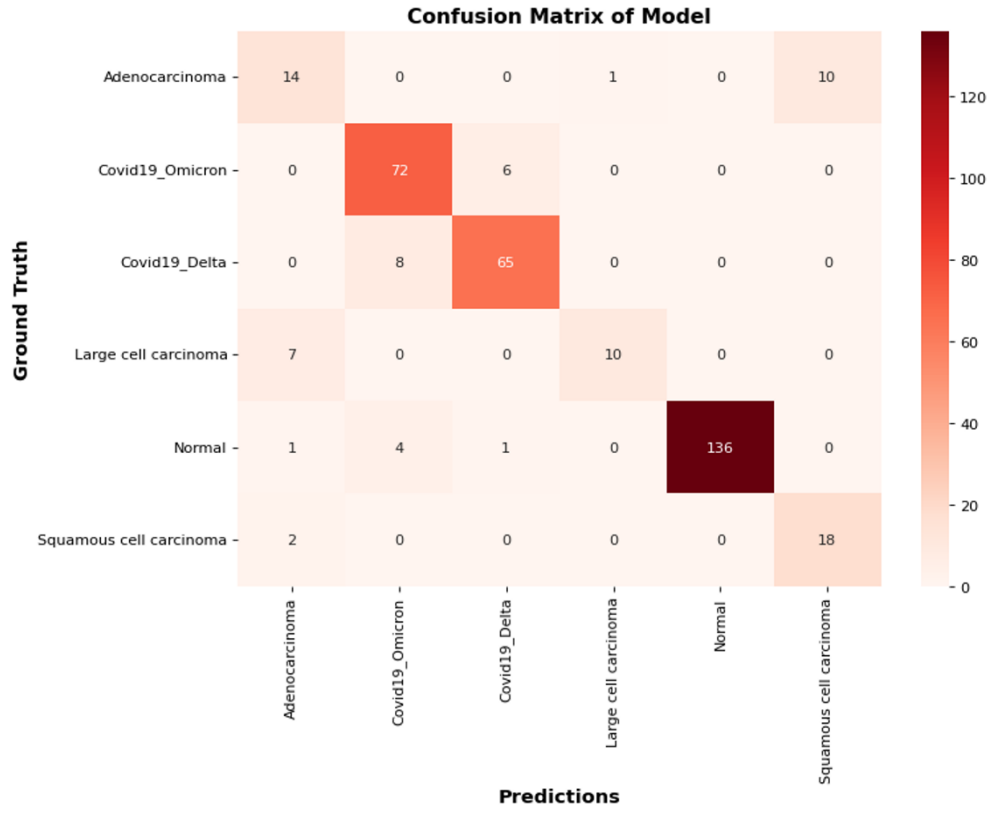
(e)


Confusion Matrix of Model
Adenocarcinoma
Covid19_Omicron
Covid19_Delta
Large cell carcinoma
Normal
Squamous cell carcinoma
Ground Truth
Predictions
87 0 1 18 1 22
0 128 13 0 0 0
0 20 117 0 1 0
2 0 0 143 0 1
0 0 0 0 154 0
4 0 0 1 0 121
140
120
100
80
60
40
20
0

(f)



(g)



(h)

Figure 14. Confusion matrix for measuring the performance metrics; (a) B_1, (b) B_2, (c) B_3, (d) B_4, (e) B_5, (f) B_6, (g) and (h) represented DLCL_Net model with and without SMOTE Tomek, respectively.

The DLCL_Net model with SMOTE Tomek properly classifies 133 out of a total of 134 images in ADS instances, while misclassifying just one image as LCC (see Fig. 14 (g)). Additionally, 124 COO images out of a total of 126 were classified correctly as COO, whereas 2 were incorrectly classified as normal. The DLCL_Net model with SMOTE Tomek accurately identifies 140 COD images out of a total of 142 images, while misclassifying 2 images as COO. The DLCL_Net detected accurately 138 LCC images out of 138 total images. In normal cases, 138 images of normal are correctly detected whereas 2 images as COO and 2 images as COD were incorrectly classified. The proposed DLCL_Net model with SMOTE Tomek accurately detected 152 SCC images out of a total of 152 images. In addition to this, we used Grad-CAM heatmaps to provide a graphical representation of the outcomes of our suggested model. The purpose of the heatmap is to display the relevant area on which the model concentrates as shown in Fig. 15.

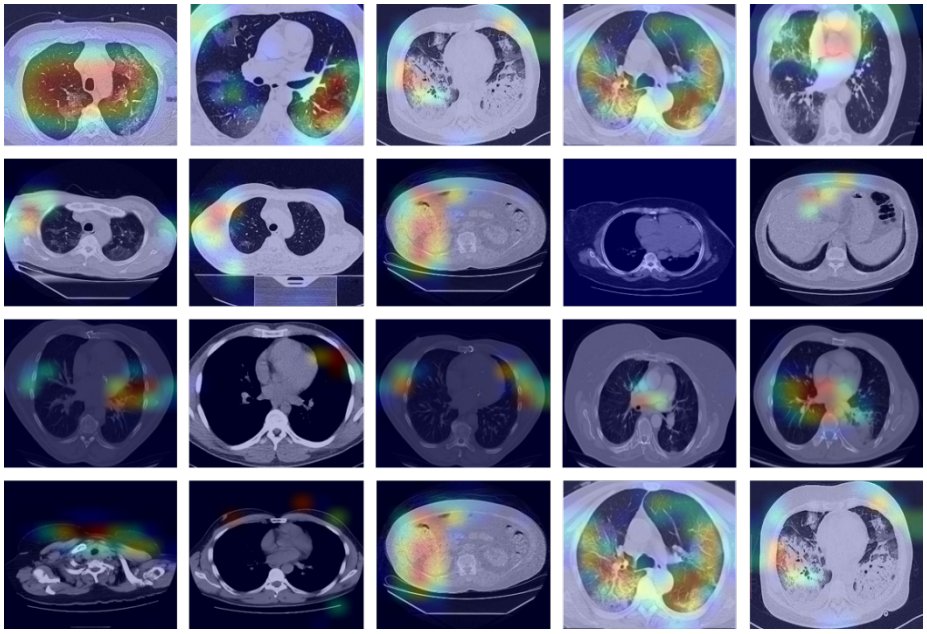


Figure 15. Grad-CAM analysis of the proposed DLCL_Net for the diseases of COVID-19 variants and LC.

4.11. Comparison with state-of-the-art classifiers

The results of a comparison between the proposed DLCL_Net model and other methods that are considered to be state-of-the-art (SOTA) are discussed in Table 6, which details the ACC, PRE, REC, and F1-score of each method.

Table 7. Evaluate the SOTA classifiers against the suggested DLCL_Net model using SMOTE Tomek.

Ref	Year	Models	Diseases	Accuracy	Precision	Recall	F1-score
[65]	2020	CNN	COVID-19	72.77%	73.83%	71.70%	×
[66]	2020	Inception-V3	COVID-19	76.01%	87.02%	93.20%	×
[67]	2021	Xception	COVID-19	92.97%	92.01%	78.00%	86.00%
[68]	2021	PAM-DenseNet	COVID-19	94.29%	93.75%	95.74%	×
[51]	2023	ResNet-34	COVID-19	95.47%	95.13%	95.05%	94.87%
[69]	2023	DarkNet-19	COVID-19	96.98%	96.89%	96.99%	97.04%
[50]	2022	MobileNet- V2	COVID-19	96.40%	96.45%	96.51%	96.58%
[70]	2023	Vgg-16	COVID-19 & Pneumonia	97.81%	95.01%	94.99%	95.00%
[71]	2023	CNN	COVID-19	94.36%	94.02%	94.04%	94.74%
[72]	2023	BCT + CNN	COVID-19	96.19%	96.09%	96.21%	96.14%
[73]	2023	ORB + CNN	COVID-19	97.39%	97.40%	97.15%	97.05%

Ours	**-**	**DLCL_Net (With SMOTE Tomek)**	**COVID-19 variants (COO & COD) and LC**	**98.92%**	**98.92%**	**98.92%**	**98.89%**

4.12 Discussions

Using CT scan images [11-14, 19, 29-32], it is possible to identify and classify a wide variety of lung illnesses. Our approach provides a comprehensive perspective of a given spot, allowing us to identify both the disease and contaminated internal parts. CT scan is the most accurate [26] and efficient [27-31] method for detecting whether an infection is a COD, COO, ADC, LCC, or SCC. Identifying COD, COO, ADC, LCC, and SCC requires a computerized diagnostic approach [56], as the number of confirmed COVID-19 cases continues to rise. Several studies [51-56] conclude that CT scan images may automatically distinguish between patients with lung cancer and those with other forms of lung illness, like COO and COD, with the use of DL approaches. As a result of this, we came up with a model that we named DLCL_Net. This model features DL as its foundation and can accurately diagnose a wide range of lung illnesses. Radiologists can begin therapy for their patients at an earlier stage when they make use of the DL models. Conditions such as COD, COO, ADC, LCC, and SCC are among those that are being treated at this time. A performance evaluation of the DLCL_Net model that was proposed was carried out with the use of two benchmark datasets that are accessible to the general public [63-66]. It was determined that the outcomes of the suggested model were assessed in conjunction with the outcomes of six distinct baseline models, which were B_1, B_2, B_3, B_4, B_5, and B_6 respectively. There is an imbalance in the CT scan pictures that were obtained from the datasets, as seen in Table 2. The model's performance is affected by the unequal class distribution of the CT scans when it is being trained [67-75]. Intending to address these concerns, we used the SMOTE Tomek method to boost the amount of CT scan pictures originating from the dataset's underrepresented minority [35]. Fig. 6 shows that the DLCL_Net model using SMOTE Tomek can accurately detect cases of infection with the five subtypes of lung illnesses: COD, COO, ADC, LCC, and SCC. This is because the model has received enough training on these subtypes. Table 5 presents that the proposed DLCL_Net with SMOTE Tomek performs significantly better in the classification of chest disorders when compared to the other classifiers. It was indicated that the DLCL_Net model, which makes use of the SMOTE Tomek approach, achieves an accuracy of 98.92% when it comes to the classification of CT scan images for COD, COO, ADC, LCC, and SCC. In addition, the DLCL_Net model achieves an accuracy of 88.73% even when it does not utilize the SMOTE Tomek approach. On the other hand, the B_1 model achieves a level of accuracy that is 92.40%. The B_2 and B_4 both obtained an accuracy of 94.12% and 95.80%, respectively, on their respective tests. When it comes to COVID-19 variations and LC classification, the performance of the B_3 is not very excellent when compared to the performance of all of the baseline models. The results of the GRAD-CAM evaluation of the proposed DLCL_Net model that makes use of SMOTE Tomek for the categorization of lung illnesses, such as COO, COD, and LC, are shown in Fig. 15.

As can be seen in Table 6, the classification performance of the DLCL_Net model is given by using SMOTE Tomek in combination with contemporary SOTA approaches. In [65], they suggested a CNN model that was able to obtain a classification accuracy of 72.77%. Tsiknakis et al. [66] developed a fine-tuned version of Inception-V3, which specifically targeted COVID-19 and normal cases. They arrived at a correct diagnosis 76.01% of the time for chest illnesses. The authors [67] develop a novel Xception model that focuses on the categorization of chest ailments into two distinct categories. With the help of this model, it is possible to differentiate between COVID-19 and the regular kinds of chest ailments. The multiclassification of chest disorders was carried out by Xue et al. [70] with the use of 2D superpixels using Vgg-16, and they achieved an accuracy of 97.81%. When it comes to the multiclassification of chest disorders, Gupta et al. [69] achieve a precision of 96.98%. When contrasted with other methods that are regarded as being at the cutting edge of technology, the DLCL_Net model using SMOTE Tomek that was proposed achieves a remarkable accuracy of 98.92%.

5. Conclusions

This study developed and assessed the DLCL_Net model for classifying the five different types of lung disorders such as COD, COO, ADC, LCC, and SCC. These lung disorders are currently increasing in

prevalence and harming people globally. Inaccurate, lengthy, and inadequate testing techniques, together with a lack of early detection of lung ailments, have directly contributed to the deaths of many people. In light of the substantial number of instances, it is essential to have a testing process that is not only speedy but also successful. To identify the five various types of lung diseases, the DLCL_Net model was made available. These blocks are then used for the classification of lung disorders in their early phases. Each convolutional block that constitutes the enhanced structure is created of many layers, and these blocks are subsequently employed. For this investigation, we make use of the SMOTE Tomek approach to produce samples. This strategy allows us to solve the concerns about the imbalance in the dataset as well as the need to keep a balanced number of samples for each class. By displaying a heat map of class activation, Grad-CAM illustrates the operation of the model that has been proposed. Our suggested DLCL_Net obtained 99.93% AUC, 98.92% precision, 98.92% recall, 98.89% F1-score, and 98.92% accuracy. This leads one to the conclusion that DLCL_Net can function as a substantial contributing component for the medical practitioner. One of the constraints of the research is that the DLCL_Net model with SMOTE Tomek that we have proposed is not suitable for X-ray information. In the future, we include BCT and FL together with a CNN model to obtain better classification results for lung disorders.

Funding: This research received no external funding.
Data Availability Statement: The datasets utilized for this study are discussed in the the datasets description section, and are publicly available.
Conflicts of Interest: The authors declare no conflict of interest.

References

1. Guo, Huaqiu, Yin Zhou, Xiaoqiang Liu, and Jianguo Tan. "The impact of the COVID-19 epidemic on the utilization of emergency dental services." Journal of dental sciences 15, no. 4 (2020): 564-567.
2. Chavez, Summer, Brit Long, Alex Koyfman, and Stephen Y. Liang. "Coronavirus Disease (COVID-19): A primer for emergency physicians." The American journal of emergency medicine 44 (2021): 220-229.
3. Bhargava, Anuja, and Atul Bansal. "Novel coronavirus (COVID-19) diagnosis using computer vision and artificial intelligence techniques: a review." Multimedia tools and applications 80 (2021): 19931-19946.
4. Kostoulas, Polychronis, Paolo Eusebi, and Sonja Hartnack. "Diagnostic accuracy estimates for COVID-19 real-time polymerase chain reaction and lateral flow immunoassay tests with bayesian latent-class models." American Journal of Epidemiology 190, no. 8 (2021): 1689-1695.
5. Fox, Tilly, Julia Geppert, Jacqueline Dinnes, Katie Scandrett, Jacob Bigio, Giorgia Sulis, Dineshani Hettiarachchi et al. "Antibody tests for identification of current and past infection with SARS-CoV-2." Cochrane database of Systematic reviews 11 (2022).
6. Afshar, Parnian, Shahin Heidarian, Nastaran Enshaei, Farnoosh Naderkhani, Moezedin Javad Rafiee, Anastasia Oikonomou, Faranak Babaki Fard, Kaveh Samimi, Konstantinos N. Plataniotis, and Arash Mohammadi. "COVID-CT-MD, COVID-19 computed tomography scan dataset applicable in machine learning and deep learning." Scientific Data 8, no. 1 (2021): 121.
7. Chen, Jun, Lianlian Wu, Jun Zhang, Liang Zhang, Dexin Gong, Yilin Zhao, Qiuxiang Chen et al. "Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography." Scientific reports 10, no. 1 (2020): 19196.
8. Kassani, Sara Hosseinzadeh, Peyman Hosseinzadeh Kassani, Michal J. Wesolowski, Kevin A. Schneider, and Ralph Deters. "Depthwise separable convolutional neural network for skin lesion classification." In 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 1-6. IEEE, 2019.
9. Grassi, Roberto, Vittorio Miele, and Andrea Giovagnoni. "Artificial intelligence: A challenge for third millennium radiologist." La radiologia medica 124, no. 4 (2019): 241-242.
10. Neri, Emanuele, Francesca Coppola, Vittorio Miele, Corrado Bibbolino, and Roberto Grassi. "Artificial intelligence: Who is responsible for the diagnosis?." La radiologia medica 125 (2020): 517-521.
11. Sadigov, Alizaman, Sakhavatdin Akhundov, and Aynur Agayeva. "Risk Factors for Lung Cancer in Individuals With COVID-19 Without Cancer History." (2021).
12. van Assen, Marly, Giuseppe Muscogiuri, Damiano Caruso, Scott J. Lee, Andrea Laghi, and Carlo N. De Cecco. "Artificial intelligence in cardiac radiology." La radiologia medica 125 (2020): 1186-1199.
13. Asif, S., Wenhui, Y., Jin, H., & Jinhai, S. (2020, December). Classification of COVID-19 from chest X-ray images using deep convolutional neural network. In 2020 IEEE 6th international conference on computer and communications (ICCC) (pp. 426-433). IEEE.
14. Alqudah, Ali Mohammad, Shoroq Qazan, and Ihssan S. Masad. "Artificial intelligence framework for efficient detection and classification of pneumonia using chest radiography images." Journal of Medical and Biological Engineering 41, no. 5 (2021): 599-609.
15. Apostolopoulos, Ioannis D., and Tzani A. Mpesiana. "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks." Physical and engineering sciences in medicine 43 (2020): 635-640.
16. Ayan, Enes, and Halil Murat Ünver. "Diagnosis of pneumonia from chest X-ray images using deep learning." In 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), pp. 1-5. Ieee, 2019.
17. Aydogdu, M., Ezgi Ozyilmaz, Handan Aksoy, G. Gursel, and Numan Ekim. "Mortality prediction in community-acquired pneumonia requiring mechanical ventilation; values of pneumonia and intensive care unit severity scores." Tuberk Toraks 58, no. 1 (2010): 25-34.
18. Berahmand, Kamal, Elahe Nasiri, and Yuefeng Li. "Spectral clustering on protein-protein interaction networks via constructing affinity matrix using attributed graph embedding." Computers in Biology and Medicine 138 (2021): 104933.

19. Berahmand, Kamal, Elahe Nasiri, and Yuefeng Li. "Spectral clustering on protein-protein interaction networks via constructing affinity matrix using attributed graph embedding." Computers in Biology and Medicine 138 (2021): 104933.
20. Canayaz, Murat. "C+ EffxNet: A novel hybrid approach for COVID-19 diagnosis on CT images based on CBAM and EfficientNet." Chaos, Solitons & Fractals 151 (2021): 111310.
21. El Asnaoui, Khalid. "Design ensemble deep learning model for pneumonia disease classification." International Journal of Multimedia Information Retrieval 10, no. 1 (2021): 55-68.
22. Hammoudi, Karim, Halim Benhabiles, Mahmoud Melkemi, Fadi Dornaika, Ignacio Arganda-Carreras, Dominique Collard, and Arnaud Scherpereel. "Deep learning on chest X-ray images to detect and evaluate pneumonia cases at the era of COVID-19." Journal of medical systems 45, no. 7 (2021): 75.
23. Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. "Dermatologist-level classification of skin cancer with deep neural networks." nature 542, no. 7639 (2017): 115-118.
24. Guan, Qingji, Yaping Huang, Zhun Zhong, Zhedong Zheng, Liang Zheng, and Yi Yang. "Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification." arXiv preprint arXiv:1801.09927 (2018).
25. Gulshan, Varun, Lily Peng, Marc Coram, Martin C. Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan et al. "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs." Jama 316, no. 22 (2016): 2402-2410.
26. Malik, Hassaan, and Tayyaba Anees. "BDCNet: Multi-classification convolutional neural network model for classification of COVID-19, pneumonia, and lung cancer from chest radiographs." Multimedia Systems 28, no. 3 (2022): 815-829.
27. Malik, Hassaan, Muhammad Shoaib Farooq, Adel Khelifi, Adnan Abid, Junaid Nasir Qureshi, and Muzammil Hussain. "A comparison of transfer learning performance versus health experts in disease diagnosis from medical imaging." IEEE Access 8 (2020): 139367-139386.
28. Cifci, Mehmet Akif. "Deep learning model for diagnosis of corona virus disease from CT images." Int. J. Sci. Eng. Res 11, no. 4 (2020): 273-278.
29. Grewal, Monika, Muktabh Mayank Srivastava, Pulkit Kumar, and Srikrishna Varadarajan. "Radnet: Radiologist level accuracy using deep learning for hemorrhage detection in ct scans." In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 281-284. IEEE, 2018.
30. Jiao, Zhicheng, Ji Whae Choi, Kasey Halsey, Thi My Linh Tran, Ben Hsieh, Dongcui Wang, Feyisope Eweje et al. "Prognostication of patients with COVID-19 using artificial intelligence based on chest x-rays and clinical data: a retrospective study." The Lancet Digital Health 3, no. 5 (2021): e286-e294.
31. Oh, Yujin, Sangjoon Park, and Jong Chul Ye. "Deep learning COVID-19 features on CXR using limited training data sets." IEEE transactions on medical imaging 39, no. 8 (2020): 2688-2700.
32. Alruwaili, Madallah, Abdulaziz Shehab, and Sameh Abd El-Ghany. "COVID-19 diagnosis using an enhanced inception-ResNetV2 deep learning model in CXR images." Journal of Healthcare Engineering 2021 (2021): 1-16.
33. Zhang, Jianpeng, Yutong Xie, Yi Li, Chunhua Shen, and Yong Xia. "Covid-19 screening on chest x-ray images using deep learning based anomaly detection." arXiv preprint arXiv:2003.12338 27 (2020): 141.
34. Kassania, Sara Hosseinzadeh, Peyman Hosseinzadeh Kassanib, Michal J. Wesolowskic, Kevin A. Schneidera, and Ralph Detersa. "Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: a machine learning based approach." Biocybernetics and Biomedical Engineering 41, no. 3 (2021): 867-879.
35. Malik, Hassaan, Tayyaba Anees, Ahmad Naeem, Rizwan Ali Naqvi, and Woong-Kee Loh. "Blockchain-Federated and Deep-Learning-Based Ensembling of Capsule Network with Incremental Extreme Learning Machines for Classification of COVID-19 Using CT Scans." Bioengineering 10, no. 2 (2023): 203.
36. Kavitha, M., T. Jayasankar, P. Maheswara Venkatesh, G. Mani, C. Bharatiraja, and Bhekisipho Twala. "COVID-19 disease diagnosis using smart deep learning techniques." Journal of Applied Science and Engineering 24, no. 3 (2021): 271-277.

37. Malik, Hassaan, Ahmad Naeem, Rizwan Ali Naqvi, and Woong-Kee Loh. "DMFL_Net: A Federated Learning-Based Framework for the Classification of COVID-19 from Multiple Chest Diseases Using X-rays." Sensors 23, no. 2 (2023): 743.
38. Kermany, D. S., M. Goldbaum, W. Cai, C. C. S. Valentim, H. Liang, and S. L. Baxter. "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell [Internet]. 2018; 172 (5): 1122-1131. e9."
39. Wang, Xiaosong, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M. Summers. "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2097-2106. 2017.
40. Guan, Qingji, Yaping Huang, Zhun Zhong, Zhedong Zheng, Liang Zheng, and Yi Yang. "Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification." arXiv preprint arXiv:1801.09927 (2018).
41. Melendez, Jaime, Bram Van Ginneken, Pragnya Maduskar, Rick HHM Philipsen, Klaus Reither, Marianne Breuninger, Ifedayo MO Adetifa, Rahmatulai Maane, Helen Ayles, and Clara I. Sánchez. "A novel multiple-instance learning-based approach to computer-aided detection of tuberculosis on chest X-rays." IEEE transactions on medical imaging 34, no. 1 (2014): 179-192.
42. Hermann, Simon. "Evaluation of scan-line optimization for 3D medical image registration." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3073-3080. 2014.
43. Stephen, Okeke, Mangal Sain, Uchenna Joseph Maduh, and Do-Un Jeong. "An efficient deep learning approach to pneumonia classification in healthcare." Journal of healthcare engineering 2019 (2019).
44. Panwar, Harsh, P. K. Gupta, Mohammad Khubeb Siddiqui, Ruben Morales-Menendez, and Vaishnavi Singh. "Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet." Chaos, Solitons & Fractals 138 (2020): 109944.
45. Khan, Asif Iqbal, Junaid Latief Shah, and Mohammad Mudasir Bhat. "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images." Computer methods and programs in biomedicine 196 (2020): 105581.
46. Janizek, Joseph D., Gabriel Erion, Alex J. DeGrave, and Su-In Lee. "An adversarial approach for the robust classification of pneumonia from chest radiographs." In Proceedings of the ACM conference on health, inference, and learning, pp. 69-79. 2020.
47. Bukhari, Syed Usama Khalid, Syed Safwan Khalid Bukhari, Asmara Syed, and Syed Sajid Hussain Shah. "The diagnostic evaluation of Convolutional Neural Network (CNN) for the assessment of chest X-ray of patients infected with COVID-19." MedRxiv (2020): 2020-03.
48. Ravi, Vinayakumar, Harini Narasimhan, Chinmay Chakraborty, and Tuan D. Pham. "Deep learning-based meta-classifier approach for COVID-19 classification using CT scan and chest X-ray images." Multimedia systems 28, no. 4 (2022): 1401-1415.
49. Kogilavani, S. V., J. Prabhu, R. Sandhiya, M. Sandeep Kumar, UmaShankar Subramaniam, Alagar Karthick, M. Muhibbullah, and Sharmila Banu Sheik Imam. "COVID-19 detection based on lung CT scan using deep learning techniques." Computational and Mathematical Methods in Medicine 2022 (2022).
50. Singh, Vipul Kumar, and Maheshkumar H. Kolekar. "Deep learning empowered COVID-19 diagnosis using chest CT scan images for collaborative edge-cloud computing platform." Multimedia Tools and Applications 81, no. 1 (2022): 3-30.
51. Choudhary, Tejalal, Shubham Gujar, Anurag Goswami, Vipul Mishra, and Tapas Badal. "Deep learning-based important weights-only transfer learning approach for COVID-19 CT-scan classification." Applied Intelligence 53, no. 6 (2023): 7201-7215.
52. Kumar, Rajesh, Abdullah Aman Khan, Jay Kumar, Noorbakhsh Amiri Golilarz, Simin Zhang, Yang Ting, Chengyu Zheng, and Wenyong Wang. "Blockchain-federated-learning and deep learning models for covid-19 detection using ct imaging." IEEE Sensors Journal 21, no. 14 (2021): 16301-16314.
53. Han, Hui, Wen-Yuan Wang, and Bing-Huan Mao. "Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning." In Advances in Intelligent Computing: International Conference on Intelligent Computing, ICIC 2005, Hefei, China, August 23-26, 2005, Proceedings, Part I 1, pp. 878-887. Springer Berlin Heidelberg, 2005.
54. Charisis, Vasileios, Alexandra Tsiligiri, Leontios J. Hadjileontiadis, Christos N. Liatsos, Christos C. Mavrogiannis, and George D. Sergiadis. "Ulcer detection in wireless capsule endoscopy images using bidimensional nonlinear analysis." In XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010: May 27–30, 2010 Chalkidiki, Greece, pp. 236-239. Springer Berlin Heidelberg, 2010.

55. Abra Ayidzoe, Mighty, Yongbin Yu, Patrick Kwabena Mensah, Jingye Cai, Kwabena Adu, and Yifan Tang. "Gabor capsule network with preprocessing blocks for the recognition of complex images." Machine Vision and Applications 32, no. 4 (2021): 91.
56. Gupta, Anand Kumar, Asadi Srinivasulu, Kamal Kant Hiran, Goddindla Sreenivasulu, Sivaram Rajeyyagari, and Madhusudhana Subramanyam. "Prediction of Omicron Virus Using Combined Extended Convolutional and Recurrent Neural Networks Technique on CT-Scan Images." Interdisciplinary Perspectives on Infectious Diseases 2022 (2022).
57. Lung Cancer Dataset, "https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images". Access Date:2nd February 2023.
58. Sünnetci, Kubilay Muhammed, and Ahmet Alkan. "Lung cancer detection by using probabilistic majority voting and optimization techniques." International Journal of Imaging Systems and Technology 32, no. 6 (2022): 2049-2065.
59. Wang, Sen, Yuxiang Xing, Li Zhang, Hewei Gao, and Hao Zhang. "Deep convolutional neural network for ulcer recognition in wireless capsule endoscopy: experimental feasibility and optimization." Computational and mathematical methods in medicine 2019 (2019).
60. Wen, Fushuan, and A. Kumar David. "A genetic algorithm based method for bidding strategy coordination in energy and spinning reserve markets." Artificial Intelligence in Engineering 15, no. 1 (2001): 71-79.
61. Aslani, S., and J. Jacob. "Utilisation of deep learning for COVID-19 diagnosis." Clinical Radiology 78, no. 2 (2023): 150-157.
62. Gajera, Himanshu K., Deepak Ranjan Nayak, and Mukesh A. Zaveri. "A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features." Biomedical Signal Processing and Control 79 (2023): 104186.
63. Gulakala, Rutwik, Bernd Markert, and Marcus Stoffel. "Rapid diagnosis of COVID-19 infections by a progressively growing GAN and CNN optimisation." Computer Methods and Programs in Biomedicine 229 (2023): 107262.
64. Wang, Ling, Xiuting Wang, Jingqi Fu, and Lanlan Zhen. "A Novel Probability Binary Particle Swarm Optimization Algorithm and its Application." J. Softw. 3, no. 9 (2008): 28-35.
65. Zhang, Jianpeng, Yutong Xie, Yi Li, Chunhua Shen, and Yong Xia. "Covid-19 screening on chest x-ray images using deep learning based anomaly detection." arXiv preprint arXiv:2003.12338 27 (2020): 141.
66. Tsiknakis, Nikos, Eleftherios Trivizakis, Evangelia E. Vassalou, Georgios Z. Papadakis, Demetrios A. Spandidos, Aristidis Tsatsakis, Jose Sánchez-García et al. "Interpretable artificial intelligence framework for COVID-19 screening on chest X-rays." Experimental and Therapeutic Medicine 20, no. 2 (2020): 727-735.
67. Jain, R., Gupta, M., Taneja, S., Hemanth, D.J.: Deep learningbased detection and analysis of COVID-19 on chest X-ray images. Appl Intell 51(3), 1690–1700 (2021).
68. Xiao, Bin, Zeyu Yang, Xiaoming Qiu, Jingjing Xiao, Guoyin Wang, Wenbing Zeng, Weisheng Li, Yongjian Nian, and Wei Chen. "PAM-DenseNet: A deep convolutional neural network for computer-aided COVID-19 diagnosis." IEEE Transactions on Cybernetics 52, no. 11 (2021): 12163-12174.
69. Gupta, Kapil, and Varun Bajaj. "Deep learning models-based CT-scan image classification for automated screening of COVID-19." Biomedical Signal Processing and Control 80 (2023): 104268.
70. Xue, Xingsi, Seelammal Chinnaperumal, Ghaida Muttashar Abdulsahib, Rajasekhar Reddy Manyam, Raja Marappan, Sekar Kidambi Raju, and Osamah Ibrahim Khalaf. "Design and Analysis of a Deep Learning Ensemble Framework Model for the Detection of COVID-19 and Pneumonia Using Large-Scale CT Scan and X-ray Image Datasets." Bioengineering 10, no. 3 (2023): 363.
71. Malik, Hassaan, Tayyaba Anees, Ahmad Sami Al-Shamaylehs, Salman Z. Alharthi, Wajeeha Khalil, and Adnan Akhunzada. "Deep Learning-Based Classification of Chest Diseases Using X-rays, CT Scans, and Cough Sound Images." Diagnostics 13, no. 17 (2023): 2772.
72. Malik, Hassaan, Tayyaba Anees, Ahmad Naeem, Rizwan Ali Naqvi, and Woong-Kee Loh. "Blockchain-Federated and Deep-Learning-Based Ensembling of Capsule Network with Incremental Extreme Learning Machines for Classification of COVID-19 Using CT Scans." Bioengineering 10, no. 2 (2023): 203.

73. Malik, Hassaan, Tayyaba Anees, Muhammad Umar Chaudhry, Radomir Gono, Michał Jasiński, Zbigniew Leonowicz, and Petr Bernat. "A Novel Fusion Model of Hand-Crafted Features with Deep Convolutional Neural Networks for Classification of Several Chest Diseases using X-ray Images." IEEE Access (2023).
74. Kareem, Amer, Haiming Liu, and Vladan Velisavljevic. "A federated learning framework for pneumonia image detection using distributed data." Healthcare Analytics (2023): 100204.
75. Prasad Koyyada, Shiva, and Thipendra P. Singh. "An explainable artificial intelligence model for identifying local indicators and detecting lung disease from chest X-ray images." Healthcare Analytics (2023): 100206.