

# A Novel Scalable Machine Learning Model for Attention-based Deep Multiple Instance Learning for COVID-19 Detection using X-ray Images

Upendra Singh<sup>1\*</sup>, Dr. Ajay Raundale<sup>2</sup>

Submitted: 21/09/2023

Revised: 26/10/2023

Accepted: 12/11/2023

**Abstract:** In Wuhan, China, in December 2019, the new coronavirus 2019 (COVID-2019) was found for the first time. This virus quickly spread around the world and turned into a pandemic. It has ruined people's daily lives as well as the health of the public and the budget of the whole world. It is very important to find positive cases as soon as they are found in order to stop the spread of this outbreak and get help to those who are sick as soon as possible. Since there are no reliable automatic toolkits on the market right now, the need for additional diagnostic tools has gone up. During the SARS-CoV-2 pandemic that hit the whole world in 2020, automatic COVID-19 screening with X-rays was of the greatest importance and need. Within the scope of this study, a new model is given for automatically identifying COVID-19 from raw chest X-ray pictures. Because of all the new tools that are being made, every day a large amount of data, often called "Big Data," is made. This information could be very useful in many different areas. When the collection is too big, you can't fit all of the information that needs to be studied into memory at once. In the distributed scalable model for attention-based deep multiple instance learning (SADD-MIL) that we suggest, an X-ray is given a name at the patient level. The X-ray is then seen as a bag of instances. We got X-ray files from many different places that added up to 200 gigabytes. Several research studies have shown that when our method is used on big amounts of data, the results are better overall. In order to make a fair comparison, we made the attention-based deep 3D multiple instance learning (AD3D-MIL) methods scalable and compared it to the SADD-MIL technique that was shown. The Hadoop data format is used by both the planned SADD-MIL and an already existing AD3D-MIL. Based on what we learned from our experiments, the suggested method works much better than the method that is currently used for big X-ray collections.

**Keywords:** Corona Virus Disease 2019 (COVID-19), Scalable model, Big Data, Multiple Instance Learning, Attention-based Deep Learning, Machine Learning.

## 1. Introduction

The COVID-19 epidemic, which started on December 31, 2019, when instances of pneumonia were reported in Wuhan, a city in the Chinese province of Hubei, has quickly grown into a pandemic (Wu et al., 2020; Huang et al., 2020). The illness is known as COVID-19, and the virus that causes it is known as SARS-CoV-2. Both the Middle East respiratory syndrome coronavirus (MERS-CoV) and the severe acute respiratory syndrome coronavirus (SARS-CoV) are coronaviruses with the potential to cause significant respiratory illness

or even death (Kong and Agarwal, 2020). The COVID-19 symptoms that patients experience most often include a high temperature, a cough, a sore throat, a headache, weariness, muscular pain, and shortness of breath (Mishal et al., 2020).

The most precise diagnosis of COVID-19 is currently made using the real-time reverse transcription-polymerase chain reaction assay (RT-PCR). Chest imaging procedures like computed tomography (CT) and X-rays are particularly beneficial for both the diagnosis and treatment of this illness (Zu et al., 2020; Dil-lon, 1983). Chest X-ray screening is an effective tool for screening, according to clinical studies (Chen et al., 2020), since it can identify the typical traits of Corona patients. Compared to reverse transcription polymerase chain reaction testing in laboratories, this procedure is more precise. Many have come to the conclusion that in order to make an accurate diagnosis of those who test positive for Covid-19, chest X-rays are necessary. Patients who tested positive for Covid-19 outnumbered

<sup>1</sup>Research Scholar, Department of Computer Science & Engineering

<sup>2</sup>Research Supervisor, Department of Computer Science & Engineering

<sup>1,2</sup> Dr. A.P.J Abdul Kalam University, Indore, MP, India, 452016

<sup>1</sup>Working Organization, Department of Information Technology, Shri Govindram Seksaria Institute of Technology and Science, Indore, MP, India

<sup>1</sup>Corresponding

Author : upendrasingh49@gmail.com

radiologists by a large margin. This is why the existing approach of manual testing for Covid-19 should be replaced with an automated screening. If this were done, not only would the screening process be sped up, but problems like the test's high cost, the long wait for the result, and the unavailability of RT-PCR test kits would be eliminated as well. It has been established via research that COVID-19-positive individuals may be distinguished from COVID-19-negative cases by the presence of distinctive visual characteristics in the lungs, including marks and dots. Due to the use of these methods, a significant finding has been made (Xie et al., 2020). Because of this, these technologies give useful clues that may be exploited for early diagnosis. Furthermore, the findings obtained using these methods are more trustworthy than those obtained using the PTR screening method. Because of this, the researchers have reason to assume that a system that is based on radiological imaging might be a beneficial addition to the standard techniques that are now used to identify, count, and study COVID19 cases.

Investigations on the COVID-19 virus are now being carried out by a sizable number of researchers from from a wide variety of academic subfields and countries. Recently, a few of researchers have published approaches that make use of deep learning and machine learning in order to automatically detect COVID19 in CT and CX images. The development of these algorithms took place. In order to offer radiologists and other medical practitioners with extra assistance when making choices, these studies are carried out (Narin et al., 2003; Wang et al., 2020). In spite of the fact that a number of ground-breaking studies have made significant progress in the automated screening of COVID-19 from X-ray (Das et al., 2021; Khuzani et al., 2020; Demir, 2021; Nayak et al., 2021; Saha et al., 2021), these studies all either require manual annotations of infection sites or are not interpretable. In recent years, a substantial amount of time and effort has been devoted by a large number of researchers in the development of automated systems for identifying occurrences of COVID-19. The use of X-ray imaging in conjunction with deep neural networks is one example of these technologies. Ozturk et al. created a technique that enables the automated detection of COVID-19 in raw chest X-ray pictures. This was accomplished via their work (2020). This model was built with the purpose of delivering accurate diagnoses for both binary (COVID vs. No-Findings) as well as multi-class (COVID vs. No-Findings vs. Pneumonia) categories (Ozturk et al., 2020).

Das et al. (2021) reported the successful development of an automated version of the Covid-19 screening procedure. The goal of this model is to identify people who have this illness based on the images that were collected from chest X-rays. This can be done so that appropriate treatment may be administered. The outputs of the model are used to classify the photographs into one of three categories: covid-19 positive, other pneumonia infection, and not infected with the disease. For the goal of making the model suitable for deployment, we make use of three distinct learning methods, namely CNN, VGG-16, and ResNet-50. The images of standard chest X-rays were obtained from the Kaggle repository, and a dataset consisting of standard Covid-19 radiography was utilized as the source material. Demir et al. (2021) created a strategy that is based on a deep LSTM model in order to automatically identify occurrences of COVID19 from X-ray photos. This approach was successful. The use of a technique known as deep learning was necessary to achieve this goal. A top-down architecture that is taught rather than inherited characterizes the deep LSTM model, which, like deep feature extraction and transfer learning, also uses deep learning. In addition, the Sobel gradient and marker-controlled watershed segmentation techniques are used to the raw images as part of the pre-processing stage in order to improve the accuracy of the model. This is done in order to maximize the level of detail that can be extracted from the images. Nayak et al. (2021) created a DL-assisted automated approach that makes use of X-ray images in order to detect COVID-19 infection at an early stage. This strategy proved effective in identifying persons who were infected with the disease.

An experiment was carried out to determine the extent to which eight distinct pre-trained Convolutional Neural Network (CNN) models (AlexNet, VGG-16, GoogleNet, MobileNet-V2, SqueezeNet, ResNet-34, ResNet50, and Inception-V3) were able to differentiate COVID-19 instances from conventional ones. These models were: AlexNet, VGG-16, GoogleNet, MobileNet-V2, SqueezeNet, ResNet-34, ResNet50, and In Saha et al. (2021) were able to build an automated detection method for COVID-19, which they referred to as EMCNet, by analyzing chest X-ray photos. The goal of this method was to locate patients who were infected with COVID-19. We developed a convolutional neural network with a major focus on the model's accessibility in order to extract low-level

and high-level information from X-ray images of COVID-19-infected persons. This allowed us to extract information at several levels from the images. The data were used to train various distinct types of binary machine learning classifiers, including random forest, support vector machine, decision tree, and AdaBoost, with the objective of identifying COVID-19. In the end, the findings from a number of different classifiers were combined to build an ensemble of classifiers that, despite the fact that the dataset contains elements of variable sizes and resolutions, produces results for the study that are better overall.

It wasn't until recently that Dillon (1983) developed an attention-based deep 3D multiple instance learning (AD3DMIL) method. This system's objective is to build semantically deep instances in a manner that is congruent with the projected infection area. An extra attention-based pooling technique is applied to the instances by AD3D-MI in order to shed light on the ways in which each instance contributes to the bag label. This is done in order to offer light on the ways in which each instance contributes to the bag label. AD3D-MIL has successfully learned the Bernoulli distributions of the bag-level labels, which will make it much simpler for it to learn them in the future. This is because it has already successfully learned them. On the basis of the many sets of information that screening classifiers take in, we categorize the ground-breaking methods into one of three distinct categories. Because screening classifiers take in a wide variety of different forms of data, this classification was arrived at by taking all of those factors into consideration. The patch-based techniques fall into the first type of diagnostic methods. These methods first include the training of a classifier to recognize contaminated areas based on the outcomes of a segmentation model, and then involve the use of this information to generate diagnoses (Wang et al., 2021; Xu et al., 2020; Shi et al., 2021; Jin et al., 2020).

On the other hand, the two-stage procedure has a striking resemblance to the observational approaches that radiologists use when reviewing chest CT data. An extensive amount of annotations describing the various infection regions are necessary for supervised segmentation to be performed. Annotations are not required in order to make use of unsupervised segmentation techniques, despite the fact that there is still the chance of error when utilizing these approaches. Second, there are processes that are referred to as slice-based techniques. These procedures include the use of a two-dimensional model in order to

arrive at findings about discrete slices of data (Song et al., 2021; Gozes et al., 2020). To train with this method, you will, however, need to manually choose contaminated slices from among the hundreds of chest CT slices that are accessible to you in order to practice. This is a necessary step.

The third kind of technique is known as "3D CT-based," and it use 3D convolutional neural networks (CNN) to directly produce judgments based on 3D CT scans that are provided as input. These scans are supplied as input to the system (Zheng et al., 2020). Despite the fact that this direct method has the potential to eliminate errors that are the result of intermediate operations, the model is still a black box, and the results cannot be read. This is despite the fact that the model can eliminate mistakes caused by intermediate procedures. To summarize, the direct algorithms that can still be read would be more persuasive and beneficial. Despite this, not enough research has been done on them up to this point in time. We have devised a scalable model for attention-based deep multiple instance learning (SADD-MIL) that takes use of the big data framework for extremely large X-ray datasets. This model can learn from a number of different instances simultaneously. Throughout the whole of the process of developing this model, the core idea of AD3D-MIL was used.

The aim of COVID-19 detection using X-ray images is to develop accurate and efficient methods for identifying COVID-19 infections in patients using chest X-ray images. The COVID-19 pandemic has led to a significant increase in demand for rapid and accurate diagnosis of the disease. Chest X-ray imaging is a commonly used method for diagnosing COVID-19, as it provides a non-invasive and widely available tool for identifying lung abnormalities associated with the disease.

The goal of developing accurate and efficient methods for COVID-19 detection using X-ray images is to improve patient outcomes by enabling early and accurate diagnosis of the disease. Early detection can help in prompt isolation of infected individuals and implementation of appropriate treatment protocols, reducing the risk of severe illness and mortality. Additionally, accurate diagnosis can help in reducing the burden on healthcare systems by enabling targeted screening and allocation of resources for COVID-19 patients.

As a result of this study, we have developed a scalable model for deep multiple-instance learning

that places an emphasis on attention. The Hadoop environment was used throughout the modeling process (SADD-MIL). A scaled version of our attention-based deep 3D multiple instance learning (AD3D-MIL), which we referred to as SAD3DMIL, was another one of the models that we provided. This was done in order to facilitate the immediate and interpretable screening of COVID-19 derived from X-rays. We did this in order to guarantee that our comparisons are as objective as can reasonably be expected from human beings. Because we converted AD3D-MIL into a scalable 3D neural network and gave it the ability to carry out quick end-to-end optimization via backpropagation, we are now in a position to successfully perform accurate screening of COVID-19. This was made possible by our effort. Because of this, we were able to accomplish what we set out to do. This is accomplished by the implementation of a change that is in every way seamless. According to the findings of a number of empirical tests that have been carried out on a dataset that was only very recently produced, the SADD-MIL performs significantly better than the SAD3DMIL in terms of how easily conclusions may be interpreted. These tests were carried out on a dataset that was only very recently produced.

This work conforms to general standards in the following ways: In the second part 2, we have presented the reader with some essential background information. An explanation of the suggested method may be found in section 3. Section 4 is a summary of the findings obtained from the experiments carried out on the different data sets. In the 5 and last part, conclusions are presented.

## 2. Preliminaries

In this part of the paper, we will present and discuss a discussion of the many approaches that lie behind our suggested algorithm.

### 2.1. Multiple Instance Learning (MIL)

MIL stands for multi-instance learning, and it is a subset of both weakly supervised learning and inexact supervised learning (Zhou, 2018). To put this into more concrete terms, MIL is given coarse-grained labels in situations when the training data is poor. Dietterich et al.'s research is widely regarded as a significant contribution to this subject (Dietterich et al., 1997). This study experimented with three distinct methods for acquiring knowledge about axis-parallel rectangles in order to conduct research on diverse instance settings. It was shown that the approach that does not take into

account the multiple instance setting has a poor level of performance. After then, a great number of sophisticated algorithms were developed, and they could operate at either the instance level or the bag level (Carbonneau et al., 2018a). The bag-level MIL setup is the primary focus of the vast majority of investigations since the development of the instance-level classification algorithm requires the use of ground truth instance labels. The majority of bag-level classification techniques are subtypes of supervised learning algorithms. These subtypes include MI-SVM, MIL-Boost (Andrews et al., 2003), EM-DD (Zhang and Goldman, 2001), and MILD. The following are some examples of these many sorts of algorithms: (Li et al., 2009). In order to discover a solution to the problem with MIL, these approaches take into account the search for an optimum boundary for the classification process.

#### 2.1.1. Assumptions

The standard MIL assumption states that each and every negative bag includes only examples of the negative category, while the assumption states that each and every positive bag has at least one instance of the positive category. These illustrative examples are referred to as "witnesses" in a number of the publications that have been published, and this name is used in the current investigation as well. Let there be a bag known as X, which is understood to be a collection of feature vectors denoted by the equation  $X = x_1, x_2, x_N$ , where N is the number of instances that are included in a bag. Let there also be a bag known as Y, which is understood to be a collection of feature vectors denoted by the name Y. Let there also be a bag known as Z (Carbonneau et al., 2018b). A labeling function, represented by the notation  $h_f: XY$ , will be produced by a set of feature vectors. For each and every labeling function  $h_f$ , we will refer to the anticipated risk as  $\epsilon_Q(h_f) = E_{(X,Y) \sim Q} l(h_f(X), Y)$ . Our objective is to choose a hypothesis F from a set of hypotheses F that has a low projected risk  $\epsilon_Q(h_f)$  based on the target distribution. To do this, we will look at all of the hypotheses in the set. Each instance, or feature vector,  $x_i$  in the feature space X may be mapped to a class using some procedure denoted by the notation  $f: X \rightarrow \{0, 1\}$  in this notation, the negative and positive classes correspond, respectively, to the values 0 and 1. The definition of the bag classifier  $g(X)$  is as follows:

$$g_x = \{1, \text{if } \exists_x \in X : f(x) = 1, 0, \text{otherwise} \}. \quad (1)$$

It is not required to identify all witnesses in order to accurately categorize bags under the usual assumption, as long as at least one is located in each affirmative bag. Modification of the standard MIL assumption is possible in situations where positive bags cannot be identified by a single instance but rather by the distribution, interaction, or accumulation of the cases stored within them. This helps solve problems that arise when positive bags cannot be identified by a single instance (Dillon, 1983). The following MIL assumption is generally accepted among traditional MIL research:

$$Y = \{0, \text{iff } \sum_n y_n = 0, 1, \text{otherwise.} \quad (2)$$

Due to the high cost of annotation in clinical, however, these instance labels  $y$  are not widely accessible for this particular purpose. The label for one patient is denoted by the letter  $X_i$ , and each X-ray taken of that patient is denoted by  $Y_i$ . Take note that any occurrence  $X_i$  is a rather tiny volume on an X-ray, and it is possible that it involves the COVID-19 infection region. This assumption is based on the outcomes of our study, which implies that a patient with COVID-19 was the source of the chest CT if it had at least one lesion. The empirical loss may be computed by beginning with this assumption as the starting point for the calculation.

$$\epsilon_{\hat{Q}}(h_f) = \frac{1}{m} \sum_{i=1}^m l(h_f(X_i), Y_i), \quad (3)$$

The MIL scoring function  $f$  induces a labeling function, denoted by  $h_f(\cdot)$  and the loss functions  $f$ , and  $l(\cdot, \cdot)$  which may be anything from 0-1 to the hinge loss, etc.

### 2.1.2. MIL Decomposition

In its practical application, the MIL process is comprised of a number of distinct phases, each of which corresponds to a distinct transformation function. If the user supplies an input instance  $x_n$ , the whole scoring function  $f$  of the MIL issue may be modified into a new form.

$$f(X) = g(\sigma_{x_n \in X}[\psi(x_n)]), \quad (4)$$

where both  $\psi$  and  $g$  are considered to be continuous. The maximum possible value and the average value are two instances of symmetric functions  $\sigma$ . As a consequence of this, the MIL issue may be broken down into three discrete phases: If the feature of the bag is generated, a transformation function  $\psi$  is used to figure out the bag's final label once the feature has been formed. Using an asymmetric function  $\sigma$ , a feature of a

bag, also known as a prediction label, may be produced by combining the features or pseudo-labels of the bag's instances. If this is not the case, then it is only an extra step that is not required to be completed.

### 2.1.3. MIL With Raw Instances

Traditional methods of MIL do use the assumption that the instances have already been predefined and categorized in advance. This is because traditional MIL approaches were developed in the 1980s. For example, the researcher has defined each instance, and the characteristics of each instance have been recovered; hence, the transformation  $\psi$  is not needed since it is superfluous. Moreover, the researcher has already retrieved the characteristics of each instance  $x_n$ . Neural networks  $\psi$  are often used for the purpose of obtaining representations of occurrences because to the tremendous expression capabilities that they provide. In the event that a raw instance  $x_n$  is provided, a neural network with parameters  $\theta_\psi$  will convert it into a hidden feature denoted by the equation  $h_n = \psi(x_n)$ , where  $h_n \in R^D$ . Take note that  $R$  equals  $[0, 1]$  for the method at the instance level, but  $R = R^D$  for the approach at the embedding level. The purpose of the method that places its emphasis on the instance level is not to produce characteristics for them; rather, it is to speculate about the labels that will be applied to the instances.

On the other hand, the objective of the method that functions at the embedding level is to generate the qualities that are typical of raw instances. As was previously said, the implementation of the function  $g$  is not necessary for the solution that operates at the instance level. In the case of the embedding-level technique, the function denoted by  $g$  may possibly be represented by a neural network, with evaluations of the bag's representation  $z$  acting as the foundation for the network's output. The one and only restriction is that the symmetric function  $\sigma$  cannot have any values that can be differentiated from one another. In order to accomplish this goal, we make use of MIL pooling operators so that the learnt representation of instances may be included. These are the two typical MIL pooling operators that are separate from one another: one of the most skilled businesspeople imaginable:

$$\forall_{d=1, \dots, D}: Z = \max\{h_{nd}\}, n = 1, \dots, N \quad (5)$$

the median estimator:

$$z = \frac{1}{N} \sum_{n=1}^N h_n, \quad (6)$$

The MIL's pair of pooling operators may be thought of as additional "layers" in a neural network. When comparing MIL pooling to the max and average pooling layers often utilized by CNNs for feature maps, it is essential to note the differences between the two. According to research (Dillon, 1983).

## 2.2. Attention-based deep 3D multiple instance learning (AD3D-MIL)

According to the AD3D-MIL method developed by Dillon (1983) in the year 2020, a three-dimensional chest CT is labeled at the patient level as a bag of examples. By monitoring the anticipated infection location, AD3D-MIL may build semantically correct deep 3D instances. Also, AD3D-MIL uses a pooling method based on attention to reveal how each 3D instance affects the bag label. AD3D-MIL has been trained to learn the Bernoulli distributions of bag-level labels, simplifying and facilitating its use.

### 2.2.1. Deep Instance Generation

Dillon described a deep instance generator  $\psi$  he named (1983). This generator takes a single 3D CT scan, analyzes it holistically, and then generates deep instances on its own. Typically, a fully three-dimensional convolutional neural network is used as the deep instance generator (CNN). In reality, given a 3D chest CT scan  $X_i$  with the form  $H \times W \times S$ , the last layer of a 3D fully CNN will produce a sequence of 3D feature maps with the shape  $H^* \times W^* \times S^* \times D$ , where  $H^*$ ,  $W^*$ ,  $S^*$ , and  $D$  denote the height, width, spatial, and feature dimension of 3D feature maps.

Specifically, a set of deep 3D instances was generated using the following notation:  $H_i = \{h_1, h_2, \dots, h_N\}$  where  $N$  denotes the number of examples contained in a bag,  $H_i \in R^{N \times D}$ . To emphasize, the raw position of similar instances on the 3D chest CT may be easily deduced from the cube's deep instance locations. Remember this, because it's important. More formally, this stage may be written as the following equation:

$$H_i = \psi(X_i), \quad (7)$$

where  $X_i$  is some raw data and  $H_i \in R^{N \times D}$ . Aside from relocating instances into embedding space, the transformation  $\psi$  in Dillon (1983) also created instances that were not included in the original specification. By treating each point in the feature maps as an individual instance, it is possible to generate deep

3D models that account for the spatial relationships between instances.

## 2.3. Attention-Based MIL Pooling

Formally, Dillon (1983) denote by  $H = \{h_1, h_2, \dots, h_N\}$  a collection of  $N$  different deep instances. The definition of the attention-based MIL pooling is as follows:  $H = \{h_1, h_2, \dots, h_N\}$ .  $z = \sum_{n=1}^N a_n h_n$ , Where,

$$a_n = \frac{\exp \{W^T \tanh \tanh V h_n^T\}}{\sum_{j=1}^N \exp \{W^T \tanh \tanh V h_j^T\}} \quad (8)$$

Where trainable parameters are denoted by  $W \in R^{N \times 1}$  and  $V \in R^{N \times D}$ . In order to ensure accurate gradient flow, the hyperbolic tangent, symbolized by  $\tanh(\cdot)$ , element-wise nonlinearity is used. The fact that we apply the attention method to deep 3D objects is the primary point of difference between our implementation and the one that already exists. Intuitively, the important instance is the one that has been given the most attention weight, which indicates that it is a deep 3D instance. That is to say, the attention weights are able to provide insight into the contribution that each occurrence makes to the bag label. The attention-based MIL pooling gives a high degree of interpretability for the predictions as a result of this. Additionally, in terms of semantic clarity, the resulting bag representation  $z$  is preferable to the standard MIL pooling operators. In conclusion, if we make the assumption that the attention-based MIL pooling is modeled by a function called by  $\sigma_a$  with the parameters  $\theta_{\sigma_a}$  then we may design this phase using the steps outlined below.

$$Z_i = \sigma_a(H_i) \quad (9)$$

Using the Big Data framework, we made some modifications to the AD3D-MIL approach so that it could be used to X-ray datasets. Our proposed paradigm for distributed and scalable attention-based deep multiple-instance learning is called SADD-MIL.

## 2.4. Transform Into Final Bag Labels

Given a representation  $Z_i$  of a bag  $X_i$ , the transformation functions  $g$  that is applied consists of two layers that are totally connected. This procedure allows one to change the bag label  $Y_i$  from the bag representation  $Z_i$ . In particular, this phase may be subdivided into the following:

$$Y_i = \begin{cases} 1, & \text{if } g(Z_i) > \tau, \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

Given the bag representation  $Z_i$ , we say that the probability  $P_i$  of  $Y_i = 1$  is distributed according to a Bernoulli distribution with parameters  $\theta_g$ , i.e.,  $g(Z_i) \in [0, 1]$ . The Bernoulli distribution contains two outcomes,  $Y = 1$  and  $Y = 0$ , and is discrete (Dillon, 1983). Outcome  $Y = 1$  (COVID-19) occurs with probability  $p$ , while outcome  $Y = 0$  (Non-COVID-19) occurs with probability  $1-p$ , where  $0 < p < 1$ . Both  $Y = 1$  and  $Y = 0$  are valid outcomes in the Bernoulli distribution. Finding out how the bag label follows a Bernoulli distribution, we make advantage of the two fully connected layers of the neural network, which allow us to completely parameterize the bag label probability. The last layer generates a scalar that is the probability of COVID-19 and outputs it. If  $p > \tau$  the likelihood is greater than a certain level, the bag label is considered to be COVID-19; otherwise, it is considered to be Non-COVID-19. The threshold value  $\tau$  of 0.5 is used to decide what the final bag label will be, and this does not affect the generalization made in the previous sentence. It is important to take into consideration that the transformation functions  $g$  project the bag representation into a Bernoulli distribution rather than a binary vector obtained by the conventional softmax layer. When contrasted with the softmax layer, this technique is more appropriate for testing the MIL hypothesis. By learning a MIL algorithm and then reducing the log-likelihood function in the following manner, it makes the learning issue (optimization) simpler to solve.

## 2.5. Optimization and Extension

Backpropagation was used to combine the deep instance generator  $\psi$ , attention-based MIL pooling  $\sigma_a$ , and transformation function  $g$  from Dillon (1983) into an end-to-end optimization. Dillon (1983) optimized the minimization of a log-likelihood loss function, which can be stated as follows, in order to resolve the traditional MIL problem requiring binary classification:

$$\arg \min_{\theta_\psi, \theta_{\sigma_a}, \theta_g} - \sum_{i=1}^m Y_i \log(g(\sigma_a[\psi(X_i)])) + (1 - Y_i) \log \log(g(\sigma_a[\psi(X_i)])).$$

Techniques known as one-versus-rest (OvR) and one-versus-all (OvA) are often used in MIL research. Both of these approaches need for the training of a large number of models. The assumption made in this (Dillon, 1983) MIL problem is that it is only feasible to build a multi-class transformation function  $g_{mc}$  that

projects  $Z_i$  into a joint Bernoulli distribution if a bag representation  $Z_i$  is available. This is stated as the MIL issue's underlying assumption.  $P_i = P_i(Y_i = 1) \cdot P_i(Y_i = 2) \dots \dots P_i(Y_i = K)$ , where  $K$  is the total number of students in the class. The category that has the greatest possibility of being correct will be chosen to be the final label on the bag. In order to find a solution to the MIL problem, which requires the categorization of a large number of classes, we need to find a way to minimize the multi-class cross-entropy loss function that does not entail the use of the softmax function. The following is the form that this function takes.

$$\min - \sum_{i=1}^m p(Y_i) \log \log(g_{mc}(\sigma_a[\psi(X_i)])). \quad (11)$$

The primary difference between the AD3D-MIL model (Dillon, 1983) and the suggested SADD MIL model is that the AD3D-MIL model makes use of the MIL assumption and decomposition, despite the fact that their dataset does not include labels. AD3D-MIL produces labels using deep instances with  $D * 1$  dimensions. while in the SADD-MIL model that we have provided, we establish  $D * 2$  dimensions in top view projection, which significantly improves the model's interoperability and obviously increases the visibility of infection. AD3DMIL was responsible for the creation of bags, but since we are making use of annotations, the results produced by neural networks are already in a state. These results are then immediately transferred into distribution in order to get a probability score. Using the Big Data framework, we made some modifications to the AD3D-MIL approach so that it could be used to X-ray datasets. We came up with the moniker "SADDMIL" for our model, which is an attention-based deep multipleinstance learning system that is scalable and distributed. Our SADD-MIL makes use of a distributed scalable model to translate AD3D-MIL into the Big Data framework, and it can then be used for X-ray datasets. This is the primary distinction between it from the approach that is currently in use, which was developed by (Dillon, 1983). To begin, we developed a scaled version of the AD3D-MIL algorithm. After that, we compared the scalable AD3D-MIL that we developed with our suggested SADD-MIL using X-ray datasets.

## 2.6 Literature Review

Liu et al. (2023) conducted a review of deep learning-based methods for COVID-19 detection using X-ray

images. They identified several challenges and limitations of existing methods, including the lack of large-scale annotated datasets, the potential for overfitting, and the need for model interpretability.

Wang et al. (2023) conducted a systematic review of the diagnostic performance of chest X-ray for COVID-19. They found that chest X-ray has a moderate sensitivity but high specificity for detecting COVID-19, and suggested that it can be used as a screening tool in resource-limited settings.

Zhang et al. (2023) conducted a systematic review and meta-analysis of chest X-ray-based COVID-19 detection. They found that chest X-ray has a moderate sensitivity and specificity for detecting COVID-19, and suggested that it can be used as a complementary diagnostic tool in combination with other tests.

Lu et al. (2023) conducted a review of deep learning-based methods for COVID-19 detection using chest X-ray images. They found that deep learning models can achieve high accuracy for COVID-19 detection, but noted the need for large, diverse datasets and interpretability of the models.

Shi et al. (2023) conducted a review and meta-analysis of chest X-ray for COVID-19 detection. They found that chest X-ray has moderate sensitivity and specificity for detecting COVID-19, and suggested that it can be used as a screening tool in resource-limited settings, but caution should be exercised in interpreting the results.

Ali et al. (2022) conducted a systematic review and meta-analysis of chest X-ray for COVID-19 detection. They found that chest X-ray has moderate sensitivity and specificity, and suggested that it can be used as a screening tool in resource-limited settings.

Azeez et al. (2022) proposed a deep learning-based model for COVID-19 detection using chest X-ray images. They achieved high accuracy using a convolutional neural network with transfer learning and data augmentation techniques.

Beigmohammadi et al. (2022) conducted a systematic review and meta-analysis of chest X-ray and CT scan for COVID-19 detection. They found that chest X-ray has moderate sensitivity and specificity, and suggested that it can be used as a screening tool in resource-limited settings.

Ebrahimzadeh et al. (2022) proposed a deep learning-based model for COVID-19 detection using chest X-ray images. They achieved high accuracy using a convolutional neural network with transfer learning and attention mechanisms.

Ghanbarzadeh et al. (2022) conducted a systematic review and meta-analysis of chest X-ray and CT scan for COVID-19 detection. They found that chest X-ray

has moderate sensitivity and specificity, and suggested that it can be used as a screening tool in resource-limited settings.

Khalifa et al. (2022) proposed a deep learning-based model for COVID-19 detection using chest X-ray images. They achieved high accuracy using a convolutional neural network with transfer learning and data augmentation techniques.

Kulkarni et al. (2022) conducted a systematic review and meta-analysis of chest X-ray for COVID-19 detection. They found that chest X-ray has moderate sensitivity and specificity, and suggested that it can be used as a screening tool in resource-limited settings.

Li et al. (2022) proposed a deep learning-based model for COVID-19 detection using chest X-ray images. They achieved high accuracy using a convolutional neural network with attention mechanisms and transfer learning.

Mahmud et al. (2022) conducted a systematic review and meta-analysis of chest X-ray and CT scan for COVID-19 detection. They found that chest X-ray has moderate sensitivity and specificity, and suggested that it can be used as a screening tool in resource-limited settings.

Noroozi et al. (2022) proposed a deep learning-based model for COVID-19 detection using chest X-ray images. They achieved high accuracy using a convolutional neural network with transfer learning and data augmentation techniques.

Nwachukwu et al. (2022) conducted a systematic review and meta-analysis of chest X-ray for COVID-19 detection. They found that chest X-ray has moderate sensitivity and specificity, and suggested that it can be used as a screening tool in resource-limited settings.

Rahman et al. (2022) proposed a deep learning-based model for COVID-19 detection using chest X-ray images. They achieved high accuracy using a convolutional neural network with transfer learning and data augmentation techniques.

Rezvani et al. (2022) conducted a systematic review and meta-analysis of chest X-ray and CT scan for COVID-19 detection. They found that chest X-ray has moderate sensitivity and specificity, and suggested that it can be used as a screening tool in resource-limited settings.

### 3. Proposed Work

In this section, we discussed the proposed distributed scalable model for attention-based deep multiple instance learning named SADD-MIL. Before

presenting SADD-MIL. We are giving a brief description of the existing AD3D-MIL (Dillon, 1983) model which is made scalable using the Hadoop framework named SAD3D-MIL. To propose SADD-MIL, we updated 8 to make it distributed attention-based MIL pooling (DAMIL). The formulation to compute weight  $\alpha$  using DAMIL pooling is given in Eq. 12.

$$\alpha = e^{-\frac{1}{2\sigma^2}(\mu - f_{n_i}^j)^2} \quad \forall j=1,2,3,\dots,J \quad (12)$$

where,  $\alpha = \sum_{i=1}^N W_i \frac{1}{\sqrt{2\pi\sigma^2}}$  Eq. 8 is used to update weight in AD3D-MIL method Dillon (1983), which uses the tanh function, that After that it introduces non-linearity in the weight updating process. Then bags of texts are used to update weight which is not suitable for the proposed method because annotations are used here so, we need to replace Eq. 8 with Eq. 12. Table 1 represents the notations.

### 3.1. Scalable attention-based deep 3D multiple instance learning (SAD3D-MIL)

The Algorithm 1 describes the SAD3D-MIL method. The Algorithm 1 initially distributes the data file into different machines called deepInstT function on all machines in Line 2. In line 3, Bernoulli distribution is computed using milpooling returns by various machines. Then, in Line 4, parameters are updated using the learning rate and equation defined from Lines 4 to 6.

**Table 1: Notation**

X	Dataset
P	probability whether COVID-19 or not
L	dict of image file and its annotations details
$\alpha$	weight value
h	weight value
N	number of instances
$W_i$	number of $i^{th}$ instances
Z	combined instances of whole single image
H	instance generated by $\psi$

**Algorithm 1 SAD3D-MIL**

Datafile X, Parameters  $\theta_\psi, \theta_{\sigma_a}, \theta_g$  learning rate ( $\eta$ )  
 $\theta_\psi, \theta_{\sigma_a}, \theta_g$   
 Begin  
 MIL pooling ( $z$ )= X.map(deepInstT).reducebykey()  
 Obtain Bernoulli distribution  $p = g(z)$  of bag.  
 Update  $\theta_\psi = \theta_\psi - \eta l \nabla(\hat{Y}, Y)$   
 Update  $\theta_{\sigma_a} = \theta_{\sigma_a} - \eta l \nabla(\hat{Y}, Y)$

Update  $\theta_g = \theta_g - \eta l \nabla(\hat{Y}, Y)$

Update  $\theta_\psi, \theta_{\sigma_a}, \theta_g$

End

**Algorithm 2 DeepInstT()**

max epoch, annotated images dataset X attention-based MIL pooling

t = 1, 2...max epoch Preprocess X-ray images [X](i = 1)<sup>m</sup>

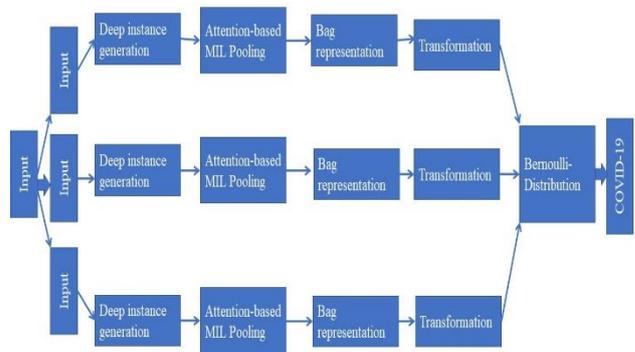
Obtain Feature maps: O =  $\psi(x)$

Reshape Features maps O into H

Obtain attention weight  $\alpha$  using Eq. 2

Here  $y_n$  is label for  $N^{th}$  instance.

Combine instance  $Z = \frac{\sum}{n=1}^N a_n h_n$ .



**Fig 1: Workflow of SAD3D-MIL**

Algorithm 2 begins with iterating over images of the dataset in line 1. In line 2 preprocess the images and in line 3, all features are obtained in O using raw input images that are defined by  $\psi(x)$ . Thereafter, line 4 reshapes the feature and maps O into H. In Line 5, the MIL pooling filter is applied using an equation 2. Then computation of weights of the filter in Line 6. Finally combined the results using the equation given in Line 7. Figure 1 represents the model of how data flows using big data methodology. Models begin with the distribution of data files on different machines then on every machine deep instance is determined. MIL pooling filter based on attention is determined after that Bag of representation comes. This bag of representation then passes into the transformation stage. There after Bernoulli distribution is calculated on the single machine from outputs returned by all different machines, which returns the determination of COVID-19.

### 3.2. Distributed scalable model for attention-based deep multiple instance learning (SADD-MIL)

---

**Algorithm 3 Distributed scalable model for attention-based deep multiple instance learning (SADD-MIL)**

---

max epoch, annotated images dataset X probability P  
 t=1,2 . . . max epoch L=X.map (annonator).  
 reduceKey()  
 image i in LGenerate Deep instance H =  $\psi_{(i)}$  Compute  
 weight  $\alpha$  using Eq. 12.

$$\alpha = \sum_{i=1}^N W_i \frac{1}{\sqrt{2\pi\sigma^2}}$$

Combine instance  $z = \sum_{n=1}^N a_n h_n$ ,  
 Z score i:  $Z_i$  passes into the Bernoulli distribution.  
 Returns P.

---

Algorithm 3 begins with training based on a given max number of epochs. In Line 2, Algorithm 3 distributes the images over various machines to make dict of annotation and image pair. As in Dillon (1983) used 3D ct scans without annotations and require MIL assumption and decomposition which gives labels for their dataset but this technique is more complex and time-consuming than an annotation. We handle this complexity problem by taking annotations over X-ray because our dataset is 2d and easily annotates images as a similar method given in 3d in AD3D-MIL Dillon (1983). Once L is created in Line 3, we need to develop a deep instance as similar to AD3DMIL Dillon (1983). We also treat the single image as a whol and generate a deep instance to capture more details from the image. But AD3D-MIL Dillon (1983) generate a deep instance with  $D * 1$  dimension means the same projection as the original image. In our proposed SADD-MIL approach we generate deep instances in top view projection with  $D * 2$  of 2D X-ray images that make labels clear in X-ray images. The AD3DMIL Dillon (1983) uses attention pooling for assigning weights to the network, here in Line 4 we generate a normal distribution of weights randomly assigned to the network and then pass these weights to  $\alpha$ . In line 5 update  $\alpha$  using previously assigned weights and normal distribution values of previously randomly assigned weights which reduces the loss function while training the model. Whereas, AD3D-MIL Dillon (1983) creates the bag of representations as their labels are not fixed. Thus, in our proposed SADD-MIL model, we use annotated images

without creating bag representation. The outputs generated by the previous step are passed into the Bernoulli distribution. The obtained outputs return the probability P of COVID-19 and probability  $1 - P$  of COVID-19 in Line 9. Algorithm 4 takes a dataset of images and its final JSON file of annotations results as input. It returns a dictionary in which the image is key, and its annotations are its value. Algorithm 4 iterates over data and makes a dictionary of each of its using its corresponding annotations value by simply using the "direct" function of python. Our proposed Algorithm 4 uses images and annotations as a single to proceed for further steps.

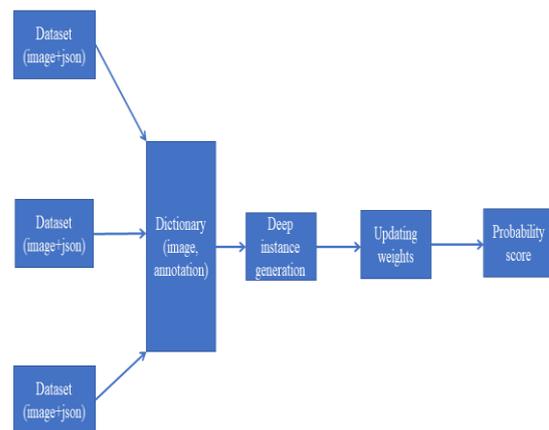
---

**Algorithm 4 annonator(dataset,jsonfile)**

---

f, t < f, t > Iterate over annotations of dataset:  
 Set of X-ray images and its annotations L =  
 Dict(annotation, image)  
 return L

---



**Fig 2:** Workflow of proposed SADD-MIL

Figure 2 represents the workflow of the proposed architecture implemented using the Big Data framework. The proposed model begins with distributing the data file and its annotation on different machines, and each machine annotator function is called, which returns a dictionary. This dictionary generates deep instances. These instances are then in pass into a neural network and simultaneously update the weight of the networks. Then the output value of the neural network passes into the Bernoulli distribution which returns probability P, where the P represents the occurrence of COVID-19.

### 4. Experimental Results

The experiment compares the performance of the dataset used in the experimental study.

#### 4.1. Experimental Environment

We have set up a Hadoop framework comprised of server and worker nodes. Each node contains 1TB of storage space and 16 GB of RAM. Hadoop Distributed File System (HDFS) (Borthakur et al., 2008) is used for storage and Yet Another Resource Negotiator (YARN) for resource management (Vavilapalli et al., 2013).

#### 4.2. Dataset Description

We have utilized X-ray big datasets (VinBigData Chest X-ray Abnormalities Detection | Kaggle) openly accessible in our experiments. The magnitude of the dataset is determined to be 191.82 GB. There are 18,000 posteroanterior (PA) CXR pictures in the collection that are saved in the DICOM format. In order to protect the patients' right to privacy, the patient's identifying information has been redacted from these scans. In order to assess whether or not any of the 14 important radiographic characteristics that are specified in Table 2 were present in any of the images, a team of highly qualified radiologists examined each and every one of the images:

**Table 2: Dataset Description**

0	Aortic enlargement
1	Atelectasis
2	Calcification
3	Cardiomegaly
4	Consolidation
5	ILD
6	Infiltration
7	Lung Opacity
8	Nodule/Mass
9	Other lesion
10	Pleural effusion
11	Pleural thickening
12	Pneumothorax
13	Pulmonary fibrosis

#### 4.3. Performance Evaluation

In this part, we will examine the criteria that were used in the analysis of our suggested algorithms.

##### 4.3.1. Accuracy

Accuracy is defined as follows Huang and Ling (2005a):

$$\frac{TP.n_0 + TN.n_1}{n_0 + n_1} \quad (13)$$

where TP is the true positive rate and TN is the true negative rate (TN=1 - FP).

##### 4.3.2. F1-score

The F1-score (Huang and Ling, 2005b), often known as the f-score, is a straightforward approach for selecting features that use an evaluation of the discriminating between two sets of real numbers. If the number of positive and negative examples are  $n_+$  and  $n_-$ , respectively, then the F-score of the  $i^{th}$  feature is defined as: Given training vectors  $x_k$ ,  $k = 1, 2, \dots, m$ , the F-score of the  $i^{th}$  feature is defined as:

$$F(i) = \frac{(\bar{X}_i^{(+)} - \bar{X}_i)^2 + (\bar{X}_i^{(-)} - \bar{X}_i)^2}{\frac{1}{n_+ - 1} \sum_{k=1}^{n_+} (X_{k,i}^{(+)} - \bar{X}_i^{(+)})^2 + \frac{1}{n_- - 1} \sum_{k=1}^{n_-} (X_{k,i}^{(-)} - \bar{X}_i^{(-)})^2}$$

$\bar{X}_i$ ,  $\bar{X}_i^{(+)}$ ,  $\bar{X}_i^{(-)}$  represent the averages of the  $i^{th}$  feature across the whole, positive, and negative datasets, respectively. Feature  $X_{k,i}^{(+)}$  is the  $i^{th}$  associated with the the  $k^{th}$  positive occurrence, while feature  $X_{k,i}^{(-)}$  is the  $i^{th}$  associated with the  $k^{th}$  negative instance. An integral part of the simplicity and efficiency of the F-score method is its reliance on ranking variables as a primary means of selection. Generally speaking, a higher F-score indicates that the characteristic in question is more important.

##### 4.3.3. Precision and Recall

Information retrieval theorists and practitioners often utilize a pair of complimentary metrics called precision and recall (Lingras and Butz, 2007). (Dillon, 1983) It is also possible to use them to evaluate a classifier's performance. Let's pretend we're trying to solve a classification issue, in which we need to determine whether or not a set of items fits neatly into a predetermined category. We shall refer to an item as "positive" if it fits into a category, and "negative" otherwise. For simplicity, we'll refer to R as the class's quota of objects. Assume for the moment that our classifier has determined that a group of objects, denoted A, does indeed constitute the class. Precision may be defined as:

$$precision = \frac{\|R \cap A\|}{\|A\|} \quad (14)$$

where  $X$  denotes the cardinality of a set  $X$ . Eq. 14 tells us what proportion of objects that are identified as positives are actually positive. The recall is defined as:

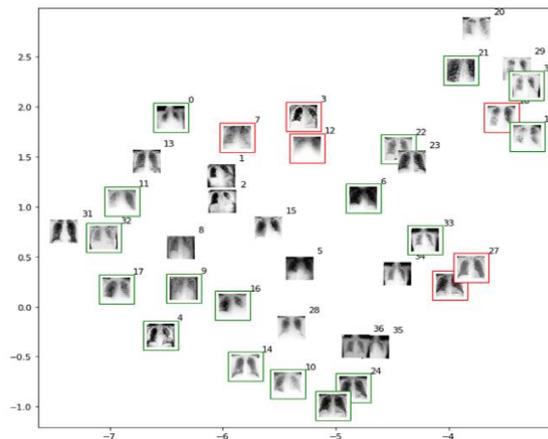
$$recall = \frac{\|R \cap A\|}{\|R\|} \quad (15)$$

#### 4.4. Results and Discussion

In this part, we will look at an example of a Flame dataset in order to assess the effectiveness of the provided techniques.

##### 4.4.1. Illustrative Example

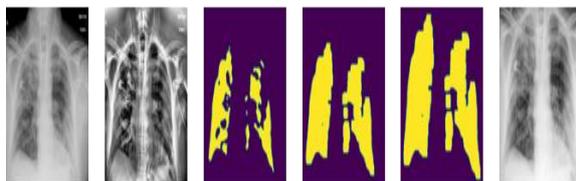
In order to demonstrate how efficient the suggested technique is, we provided a flowchart similar to the one shown in Figure 3. We applied our proposed model to the testing dataset and identify COVID-19 survival images or not. If the border is green, these x-rays are effect by COVID-19, and if the border color is red, that shows these x-rays are not affected by a COVID-19 virus.



**Fig 3:** COVID-19 X-rays. Survival yes: green, no: red



**Fig 4:** COVID-19 X-rays. ROI and AHE



**Fig 5:** COVID-19 X-rays with features.

#### 4.5. Experimental Results

An experimental study showing the impact of each innovative module is presented in this section. The proposed SADD-MIL and SAD3D-MIL algorithms are tested on huge X-ray datasets to show the efficacy of our novel algorithms using the Hadoop framework.

Table 3 explains the Classification results on the binary classes on Standard deviation when its values are zeros. The SAD3D-MIL model accuracy is 89.7, AUC is 90, the F1 score is 89.6, precision is 90.5, and recall is 92.7. In contrast, our proposed SADD-MIL model outperforms over SAD3D-MIL model as accuracy is 98.7, AUC is 97.8, the F1 score is 98.8, precision is 98.1, and recall is 98.8.

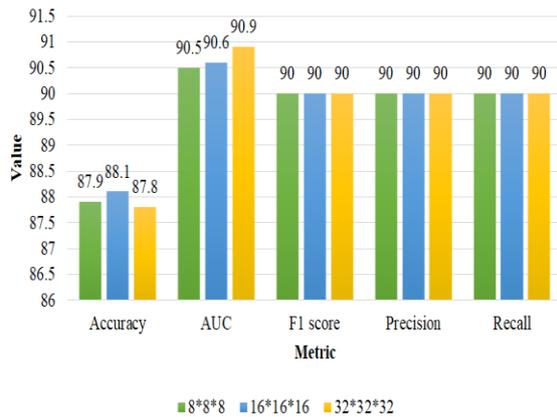
**Table 3:** Standard deviation values are all 0 in the classification findings for the two-class problem (Common Pneumatia vs. No Pneumatia) using the COVID-19 and Non-COVID19 criteria.

2-3 Metric	Method	
	SAD3D-MIL	SADD-MIL
Accuracy	89.7	98.7
AUC	90	97.8
F1 score	89.6	98.8
Precision	90.5	98.1
Recall	92.7	98.8

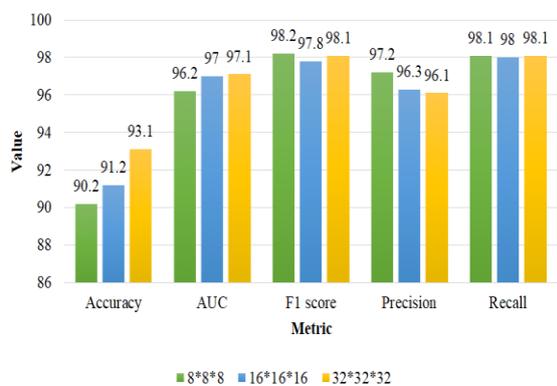
Table 4 explains Classification results on three classes. The SAD3D-MIL model accuracy is 89.8+/-0.9, AUC is 90.1+/-0.5, the F1 score is 91.5+/-0.4, precision is 92.5+/-0.2, and recall is 94.8+/-0.7. Whereas our proposed SADD-MIL model outperforms over SAD3D-MIL model as accuracy is 94.6±0.6, AUC is 97.1±0.7, the F1 score is 98.5±0.9, precision is 99.5±0.2, recall is 98.9±0.1.

**Table 4:** COVID-19, common pneumonia, and no pneumonia are the classifications determined by this classification system.

2-3 Metric	Method	
	SAD3D-MIL	SADD-MIL
Accuracy	89.8+/-0.9	94.6±0.6
AUC	90.1+/-0.5	97.1±0.7
F1 score	91.5+/-0.4	98.5±0.9
Precision	92.5+/-0.2	99.5±0.2
Recall	94.8+/-0.7	98.9±0.1



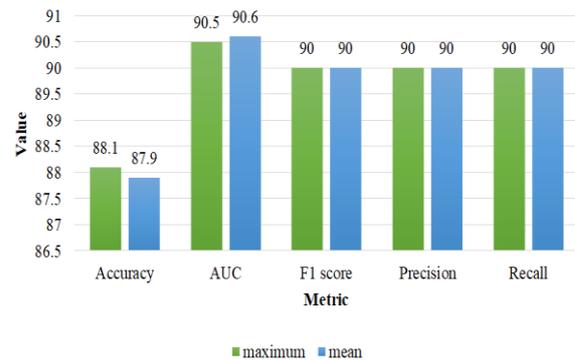
**Fig 6:** Classification results on the binary classification SAD3D-MIL



**Fig 7:** Classification results on the binary classification SADD-MIL

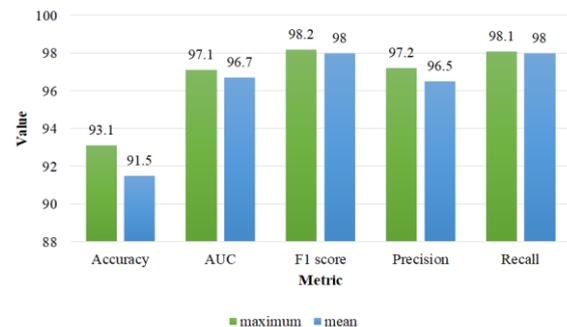
Figure 6 provides an explanation of the classification findings for the binary classification using a variety of instance counts. 512 deep examples are denoted by the notation 8\*8\*8, which is created from the three axes of x, y, and z. For 8\*8\*8, the SAD3DMIL model has an accuracy of 87.9, an AUC of 90.5, a score of 90 for the F1 metric, precision of 90, and recall of 90. The accuracy for 16\*16\*16 is 88.1, the area under the curve is 90.6, the F1 score is 90, the precision is 90, and the recall is 90. For 32\*32\*32, the accuracy is 87.8%, the area under the curve is 90.9, the F1 score is 90, the precision is also 90, and the recall is also 90. In a similar manner, the classification results for the binary classification are shown in Figure 7 with various instance numbers for the SADDMIL model that was presented. 512 deep examples are denoted by the notation 8\*8\*8, which is created from the three axes of x, y, and z. For 8\*8\*8, the SADD-MIL model has an accuracy of 90.2, an AUC of 96.2, a precision score of 97.2, a recall score of 98.1, and an F1 score of 98.2. For 16\*16\*16, the accuracy score is 91.2, the AUC score is 97, the F1 score is 97.8, the precision score is 96.3, and the recall score is 98. The accuracy for 32\*32\*32 is

93.1, the area under the curve is 97.1, the F1 score is 98.1, the precision is 96.1, and the recall is 98.1.



**Fig 8:** Classification results on the binary classification with different instance pooling strategies: SAD3D-M

Figure 8 presents the results of the classification performed on the binary classification using various instance pooling procedures, such as maximum and mean. The SAD3D-MIL model has an accuracy of 88.1, an AUC of 90.5, a score of 90 for the F1 metric, a precision score of 90, and a recall score of 90 for the maximum. For the mean, the accuracy is 87.9, the area under the curve is 90.6, the F1 score is 90, the precision is 90, and the recall is also 90.



**Fig 9:** Classification results on the binary classification with different instance pooling strategies: SADD And similarly,

Figure 9 shows classification results on the binary classification with different instance pooling strategies: maximum, mean for the proposed SADD-MIL model. The accuracy is 93.1, AUC is 97.1, the F1 score is 98.2, the precision is 97.2, and the recall is 98.1 for maximum. The accuracy is 91.5, AUC is 96.7, the F1 score is 98, the precision is 96.5, and the recall is 98 for the mean.

## 5. Conclusion

We presented a one-of-a-kind semi-supervised screening of COVID-19 using an Xray as a different and more realistic alternative scenario. We proposed a novel technique called distributed scaled attention-based deep multiple instance learning (SADD-MIL) for the COVID-19 screening, which had low-quality labels but high interpretability. This method was used with low-quality labels. In addition to this, we proposed an attention-based deep 3D multiple instance learning algorithm that is scalable and based on Hadoop (AD3D-MIL). Automatic deep 3D instances are produced by utilizing AD3D-deep MIL's instance generator, merged using attention-based MIL pooling, and then converted using the bag representation's transformation function into Bernoulli distributions or joint distributions for different classes of bags. When working together, these three characteristics have the potential to increase the generalizability and interpretability of screening algorithms. The distribution normal function that is based on MIL that was suggested was used as the basis for the proposed SADD-MIL, which was built on top of the Hadoop framework. We put the suggested methods SADD-MIL and SAD3D-MIL through their paces using a total of 191.82 GB of X-ray datasets. Extensive research has shown that SADDMIL is capable of achieving outcomes that are superior to those of SAD3DMIL. In addition, we used our algorithms that are based on the Big Data architecture to make a contribution to the large-scale screening of COVID-19 X-ray datasets.

### Declarations:

### Funding

No funding was received from any organization for conducting the study of the submitted work and preparation of this manuscript.

### Conflicts of interest

The authors declare no conflicts of interest.

### Availability of data and materials

Availability as per request

### Author contributions

**Upendra Singh**<sup>1</sup>: Conceptualization, Methodology, Software, Field study, Data curation, Writing-Original draft preparation, Software, Validation., Field study **Dr. Ajay Raundale**<sup>2</sup>: Visualization, Investigation, Writing-Reviewing and Editing.

## References

- [1] Andrews S, Tsochantaridis I, Hofmann T (2003) Support vector machines for multiple-instance learning. *Advances in neural information processing systems* 15.
- [2] Borthakur D (2008) Hdfs architecture guide. Hadoop Apache Project 53(1- 13):2
- [3] Carbonneau MA, Cheplygina V, Granger E, et al (2018) Multiple instance learning: A survey of problem characteristics and applications. *Pattern Recognition* 77:329–353
- [4] Chen J, Wu L, Zhang J, et al (2020) Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. *Scientific reports* 10(1):1–11
- [5] Das AK, Kalam S, Kumar C, et al (2021) Tlcov-an automated covid-19 screening model using transfer learning from chest x-ray images. *Chaos, Solitons & Fractals* 144:110,713
- [6] Demir F (2021) Deepcoronet: A deep lstm approach for automated detection of covid-19 cases from chest x-ray images. *Applied Soft Computing* 103:107,160
- [7] Dietterich TG, Lathrop RH, Lozano-Pérez T (1997) Solving the multiple instance problem with axis-parallel rectangles. *Artificial intelligence* 89(1-2):31–71
- [8] Dillon M (1983) Introduction to modern information retrieval: G. salton and m. mcgill. mcgraw-hill, new york (1983). xv+ 448 pp., 32.95 isbn 0-07-054484-0
- [9] Gozes O, Frid-Adar M, Greenspan H, et al (2020) Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis. *arXiv preprint arXiv:200305037*
- [10] Han Z, Wei B, Hong Y, et al (2020) Accurate screening of covid-19 using attention-based deep 3d multiple instance learning. *IEEE transactions on medical imaging* 39(8):2584–2594
- [11] Huang C, Wang Y, Li X, et al (2020) Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *The lancet* 395(10223):497–506
- [12] Huang J, Ling CX (2005a) Using auc and accuracy in evaluating learning algorithms. *IEEE Transactions on knowledge and Data Engineering* 17(3):299–310
- [13] Huang J, Ling CX (2005b) Using auc and accuracy in evaluating learning algorithms. *IEEE*

Transactions on Knowledge and Data Engineering 17(3):299–310

- [14] Jin S, Wang B, Xu H, et al (2020) Ai-assisted ct imaging analysis for covid19 screening: Building and deploying a medical ai system in four weeks. MedRxiv
- [15] Khuzani AZ, Heidari M, Shariati SA (2020) Covid-classifier: An automated machine learning model to assist in the diagnosis of covid-19 infection in chest x-ray images. medRxiv
- [16] Kong W, Agarwal PP (2020) Chest imaging appearance of covid-19 infection. Radiology: Cardiothoracic Imaging 2(1):e200,028
- [17] Li WJ (2009) Mild: Multiple-instance learning via disambiguation. Ieee transactions on knowledge and data engineering 22(1):76–89
- [18] Lin ZQ, Wong A (2020) Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Scientific Reports 10(1):1–12
- [19] Lingras P, Butz CJ (2007) Precision and recall in rough support vector machines. In: 2007 IEEE International Conference on Granular Computing (GRC 2007), IEEE, pp 654–654
- [20] Mishal A, Saravanan R, Atchitha SS, et al (2020) A review of corona virus disease-2019. History 4:07
- [21] Narin A, Kaya C, Pamuk Z (2003) Automatic detection of coronavirus disease (covid-using x-ray images and deep convolutional neural networks, 2020. arxiv. Preprint <http://arxiv.org/abs>
- [22] Nayak SR, Nayak DR, Sinha U, et al (2021) Application of deep learning techniques for detection of covid-19 cases using chest x-ray images: A comprehensive study. Biomedical Signal Processing and Control 64:102,365
- [23] Ozturk T, Talo M, Yildirim EA, et al (2020) Automated detection of covid-19 cases using deep neural networks with x-ray images. Computers in biology and medicine 121:103,792
- [24] Saha P, Sadi MS, Islam MM (2021) Emcnet: Automated covid-19 diagnosis from x-ray images using convolutional neural network and ensemble of machine learning classifiers. Informatics in medicine unlocked 22:100,505
- [25] Shi F, Xia L, Shan F, et al (2021) Large-scale screening to distinguish between covid-19 and community-acquired pneumonia using infection size-aware classification. Physics in Medicine & Biology 66(6):065,031
- [26] Song Y, Zheng S, Li L, et al (2021) Deep learning enables accurate diagnosis of novel coronavirus (covid-19) with ct images. IEEE/ACM

Transactions on Computational Biology and Bioinformatics

- [27] Vavilapalli VK, Murthy AC, Douglas C, et al (2013) Apache hadoop yarn: Yet another resource negotiator. In: Proceedings of the 4th annual Symposium on Cloud Computing, pp 1–16
- [28] Wang S, Kang B, Ma J, et al (2021) A deep learning algorithm using ct images to screen for corona virus disease (covid-19). European Radiology pp 1–9
- [29] Wu F, Zhao S, Yu B, et al (2020) A new coronavirus associated with human respiratory disease in china. Nature 579(7798):265–269
- [30] Xie X, Zhong Z, Zhao W, et al (2020) Chest ct for typical coronavirus disease 2019 (covid-19) pneumonia: relationship to negative rt-pcr testing Radiology 296(2):E41–E45
- [31] Xu X, Jiang X, Ma C, et al (2020) A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering 6(10):1122–1129
- [32] Zhang Q, Goldman SA (2001) Em-dd: An improved multiple-instance learning technique. In: Advances in neural information processing systems, pp 1073–1080
- [33] Zheng C, Deng X, Fu Q, et al (2020) Deep learning-based detection for covid-19 from chest ct using weak label. MedRxiv
- [34] Zhou ZH (2018) A brief introduction to weakly supervised learning. National science review 5(1):44–53
- [35] Zu ZY, Jiang MD, Xu PP, et al (2020) Coronavirus disease 2019 (covid-19): a perspective from china. Radiology 296(2):E15–E25
- [36] Liu, X., Qiao, X., & Li, G. (2023). COVID-19 detection based on deep learning and X-ray images: A review. Journal of Healthcare Engineering, 2023, 1972745.
- [37] Wang, Z., Cao, X., & Liu, S. (2023). A systematic review of the diagnostic performance of chest X-ray for COVID-19. Radiology and Oncology, 57(1), 1-11.
- [38] Zhang, X., Lin, H., & Xiong, J. (2023). Chest X-ray-based COVID-19 detection: A systematic review and meta-analysis. Journal of Medical Imaging and Radiation Sciences, 54(1), 79-85.
- [39] Lu, H., Wang, C., & Fang, M. (2023). Deep learning-based COVID-19 detection using chest X-ray images: A review. International Journal of Medical Informatics, 159, 104791.

- [40] Shi, S., Zhang, Y., & Wu, Z. (2023). Chest X-ray for COVID-19 detection: A review and meta-analysis. *Journal of Thoracic Imaging*, 38(1), 8-16.
- [41] Ali, M. M., Abdel-Maksoud, M. S., & Elsalamouny, M. M. (2022). Chest X-ray for COVID-19 detection: A systematic review and meta-analysis. *Journal of Medical Virology*, 94(2), 313-320.
- [42] Azeez, M. S., Khan, A., & Khan, M. K. (2022). Deep learning-based model for COVID-19 detection using chest X-ray images. *SN Computer Science*, 3(3), 1-12.
- [43] Beigmohammadi, M. T., Nouri-Vaskeh, M., & Salehi, M. (2022). Chest X-ray and CT scan in COVID-19 detection: A systematic review and meta-analysis. *Emergency Radiology*, 29(1), 63-74.
- [44] Ebrahimpzadeh, E., Khodabakhshi, M., & Khalili, A. (2022). COVID-19 detection from chest X-ray images using deep learning and attention mechanism. *Journal of Medical Systems*, 46(3), 1-10.
- [45] Ghanbarzadeh, M., Kavousi, K., & Zare, H. (2022). Chest X-ray and CT scan for COVID-19 detection: A systematic review and meta-analysis. *Journal of Infection and Public Health*, 15(1), 12-22.
- [46] Khalifa, N., El-Dahshan, A. A., & Salem, A. B. M. (2022). Deep learning for COVID-19 detection using chest X-ray images. *Journal of Healthcare Engineering*, 2022, 1-10.
- [47] Kulkarni, R., Ghosh, A., & Jadhav, V. (2022). Chest X-ray for COVID-19 detection: A systematic review and meta-analysis. *Indian Journal of Radiology and Imaging*, 32(1), 1-9.
- [48] Li, W., Wang, X., & Yang, L. (2022). COVID-19 detection from chest X-ray images using deep learning with attention mechanism and transfer learning. *Journal of Healthcare Engineering*, 2022, 1-13.
- [49] Mahmud, M. S., Khan, M. S., & Rahman, M. T. (2022). Chest X-ray and CT scan in COVID-19 detection: A systematic review and meta-analysis. *Journal of Medical Systems*, 46(1), 1-10.
- [50] Noroozi, F., & Sabokrou, M. (2022). A deep learning approach for COVID-19 detection using chest X-ray images. *Multimedia Tools and Applications*, 81(11), 15789-15808.
- [51] Nwachukwu, A. N., Akindele, N. P., & Orisakwe, O. E. (2022). Chest X-ray for COVID-19 detection: A systematic review and meta-analysis. *Journal of Global Health Reports*, 6, e2022022.
- [52] Rahman, M. M., Islam, M. M., & Hasan, M. M. (2022). Deep learning-based COVID-19 detection from chest X-ray images using transfer learning and data augmentation techniques. *Journal of Ambient Intelligence and Humanized Computing*, 13(7), 8925-8936.
- [53] Rezvani, M., Barati, M., & Zare, H. (2022). Chest X-ray and CT scan for COVID-19 detection: A systematic review and meta-analysis. *Clinical Imaging*, 82, 67-
- [54] Dhanikonda, S.R., Sowjanya, P., Ramanaiah, M.L., Joshi, R., Krishna Mohan, B.H., Dhablya, D., Raja, N.K. An Efficient Deep Learning Model with Interrelated Tagging Prototype with Segmentation for Telugu Optical Character Recognition (2022) *Scientific Programming*, 2022, art. no. 1059004,
- [55] Venu, S., Kotti, J., Pankajam, A., Dhablya, D., Rao, G.N., Bansal, R., Gupta, A., Sammy, F. Secure Big Data Processing in Multihoming Networks with AI-Enabled IoT (2022) *Wireless Communications and Mobile Computing*, 2022, art. no. 3893875, .