Evolving Systems (2023) 14:1101–1115
https://doi.org/10.1007/s12530-023-09484-2

ORIGINAL PAPER

AMTLDC: a new adversarial multi-source transfer learning framework to diagnosis of COVID-19

Hadi Alhares[1] · Jafar Tanha[1] · Mohammad Ali Balafar[1]

Received: 8 June 2022 / Accepted: 2 January 2023 / Published online: 12 January 2023

Abstract
In recent years, deep learning techniques have been widely used to diagnose diseases. However, in some tasks, such as the diagnosis of COVID-19 disease, due to insufficient data, the model is not properly trained and as a result, the generalizability of the model decreases. For example, if the model is trained on a CT scan dataset and tested on another CT scan dataset, it predicts near-random results. To address this, data from several different sources can be combined using transfer learning, taking into account the intrinsic and natural differences in existing datasets obtained with different medical imaging tools and approaches. In this paper, to improve the transfer learning technique and better generalizability between multiple data sources, we propose a multi-source adversarial transfer learning model, namely AMTLDC. In AMTLDC, representations are learned that are similar among the sources. In other words, extracted representations are general and not dependent on the particular dataset domain. We apply the AMTLDC to predict Covid-19 from medical images using a convolutional neural network. We show that accuracy can be improved using the AMTLDC framework, and surpass the results of current successful transfer learning approaches. In particular, we show that the AMTLDC works well when using different dataset domains, or when there is insufficient data.

Keywords Diagnose diseases · COVID-19 diagnosis · Deep learning · Multi-source adversarial domain adaptation · Coronavirus pneumonia

1 Introduction

Nearly 251 million people worldwide officially have been infected with COVID-19, and more than 5 million death tolls until November 2021 (Worldometer 2021; Ghaderzadeh et al. 2021a) as of epidemic declaration in March 2020, signifies the rapid diagnosis of the COVID-19 with high reliability in the early stages; Not only to save human lives but also to reduce the social and economic burden on the communities involved. Although the RT-PCR (Real-time polymerase chain reaction) test is the standard reference for confirming COVID-19, some studies show that this laborious method cannot diagnose the disease in the early stages (Ai et al. 2020; Alshazly et al. 2021; Jokandan et al. 2007), and some studies report a high false-negative rate (Long et al. 2020; Ghaderzadeh and Aria 2020).

One standard way to identify morphological patterns of lung lesions associated with COVID-19 is to use chest scan images. There are two common techniques for scanning the chest: X-rays and computer tomography (CT). Detection of COVID-19 from chest images by a radiologist is time-consuming, and the accuracy of COVID-19 diagnosis depends strongly on the radiologist's opinion (Ng et al. 2020; Ghaderzadeh et al. 2021b). Also, manually checking every image might not be feasible in emergency cases. Recently, deep learning-based methods (Ghaderzadeh et al. 2021a; Hemdan et al. 2003; Farooq and Hafeez 2003; Luz et al. 2021; Li et al. 2020c; Wang et al. 2020a) have been applied to help the medical community diagnose COVID-19 quickly, accurately, and automatically.

The use of deep learning in various fields of machine vision has shown promising results. In particular, deep

✉ Jafar Tanha
tanha@tabrizu.ac.ir

Hadi Alhares
Hadixhares@gmail.com

Mohammad Ali Balafar
balafarila@tabrizu.ac.ir

[1] Department of Computer Engineering, Faculty of Electrical and Computer Engineering, University of Tabriz, 29th Bahman Blvd, Tabriz 5166616471, Iran

learning is widely used in medical imaging (e.g. for the diagnosis of diabetic retinopathy (Gulshan et al. 2016), skin cancer (Esteva et al. 2017), breast cancer (Wang et al. 2016), and other tasks (Bayani et al. 2022; Bayani et al. 2022; Aria et al. 2022a). However, deep learning faces many challenges. Some of these challenges are related to the intrinsic of deep models; for example, a lot of data is needed for the success of deep learning models. The reason for medical applications' success based on deep learning is the data that has been collected over the years. However, in most other applications, it is difficult to collect sufficient medical data to train the model. This is because of the cost of labeling them, which requires an expert in this field (Wang et al. 2019). The lack of a sufficient dataset is also a major challenge in the Covid-19 diagnostic task with medical imaging. To solve this problem, various methods have been proposed, such as (Altae-Tran et al. 2017; Christodoulidis et al. 2017; Dhungel et al. 2017; Bar et al. 2015). One of these methods is data augmentation. In this method, data can be increased by using some data transformation techniques, such as zooming, image rotation, horizontal or vertical shifting, and horizontal or vertical flipping (Dhungel et al. 2017). In some other methods, such as few-shot learning, data efficiency can be increased (Altae-Tran et al. 2017). Some other methods instill the knowledge learned on sufficient data into target deep models. The purpose of this approach is to train the model with insufficient data. This knowledge can be obtained by training the model on a semi-related dataset and then fine-tuning it with the target dataset (Bar et al. 2015). This technique is known as transfer learning.

The transfer learning technique is very appropriate due to the variety of medical datasets available. For example, we can transfer knowledge between different datasets. Also, multi-source transfer learning can be used to combine multiple sources and extract knowledge from them (Christodoulidis et al. 2016). Knowledge transfer between these datasets and learning common features can improve model generalizability. The advantage of using multi-source transfer learning is that it allows the use of several different datasets, each of which may not be sufficient to train and generalize the model alone. However, the nature of the source datasets can be very different, and the transmission efficiency strongly depends on the similarity between the source tasks and the target task. So in some cases, transfer learning hurts the model instead of helping it to train better. There is also a risk that the model will learn the specific features of each dataset instead of learning the common features between the datasets. Therefore, it harms the generalization of the learned model. This is especially true in the case of COVID-19 detection. Because medical datasets are often collected with different medical imaging tools and methods.

Problem statement: Most existing methods for classifying COVID-19 are trained and evaluated with images from the same dataset. Using only a dataset reduces the generalizability of these methods. So the results of training and testing the network on the same dataset are much better than the results of training and testing the network on different datasets. In other words, in the feature extraction stage, most of the proposed models are very dependent on the domain of the training dataset and do not perform well in the face of unseen datasets. For this reason, they are not trusted in real-world applications where the data used is new and independent of training data. Numerous studies demonstrate that the most recent approaches in the literature are unreliable (Tartaglione et al. 2020; Tabik et al. 2020). For example, two well-known studies (Wang et al. 2020a; Afshar et al. 2020) in this field show a performance close to random classification facing unseen data (i.e., datasets on which the model has not been trained). The classification accuracy in research (Silva et al. 2020) decreases from 98.5% on the test set to 59.12% on unseen datasets. The structural and inherent differences in the images from the available datasets, which arise from different tools and medical imaging methods, are the cause of this issue.

Method: To solve the above problem, in this research, we propose the Adversarial Multi-source Transfer Learning Framework framework for COVID-19 diagnosis from CT (Computed Tomography) images, namely AMTLDC. We use two separate datasets to learn common representations that are independent of the domain of each dataset. The source dataset is used for network training and the target dataset is used to increase the transferability and generalizability. Deep models used traditionally in transfer learning (e.g., convolutional neural networks) generally implement two modules: a feature extractor that extracts knowledge from the inputs, and a predictor that uses the knowledge to make the predictions. In the AMTLDC setting, a new module infers the source of the input data based on its extracted features. By making the features extractor compete against this objective, the learned feature representation generalizes better across the sources. Our hypothesis is that the feature representation, being more general, will then transfer better to an unknown target. This idea is particularly well suited for COVID-19 diagnosis because of the structural and inherent differences in the images from the available datasets, which arise from different tools and medical imaging methods. AMTLDC can perform the correct classification regardless of the specific features of each input data domain. In other words, the representations learned are shared among both data domains and do not depend on a particular dataset domain.

Contribution: The contributions of this research are threefold:

1. The effect of intrinsic and natural differences in existing datasets obtained with different medical imaging tools

and approaches is minimized as a result of the proposed adversarial multi-source transfer learning framework.
2. An efficient deep framework is developed to make Covid-19 detection more accurate.
3. Extensive experiments show that the AMTLDC has high generalizability on unseen data.

In the remainder of the paper: the related works are reviewed in Sect. 2; the proposed AMTLDC framework is introduced in Sect. 3; experiments performed are explained in Sects. 4 and 5; and in the last section, the conclusion is presented.

2 Related work

Various deep learning methods have been introduced to detect Covid 19. These methods can be divided into three general categories. The first category includes methods that have developed customized architectures for COVID-19 detection, such as COVID-Net (Wang et al. 2020a) and CVR-Net (Hasan et al. 2020). The second category includes methods that use common architectures [such as ResNet (Residual Network) (He et al. 2016)] and transfer learning. The last category includes very few studies that have employed handcrafted feature extraction approaches and conventional classifiers. In the following, each of these categories is reviewed.

2.1 Customized models

Some methods introduced a customized architecture for COVID-19 detection. COVID-Net (Wang et al. 2020a) is one of the pioneering methods that has introduced a new convolutional architecture for identifying COVID-19. This architecture is trained and evaluated on X-ray images. An improved version of the COVID-Net method has been developed in Wang et al. (2020b). The authors developed a novel joint learning model to detect COVID-19 by effectively learning with heterogeneous datasets with distribution discrepancies. In this model, the generated representations and the network performance are computationally improved.

In Hasan et al. (2020), a robust CNN (Convolutional Neural Network) based network, called CVR-Net was proposed. In this framework, both CT and X-ray images are used to train and test the model. The proposed end-to-end CVR-Net is a multi-scale multi-encoder ensemble model.

To increase the efficiency of coronavirus detection based on CT images, the authors proposed a set of deep models called CovidCTNet (Javaheri et al. 2005) that successfully detect Covid-19 from other lung diseases. CovidCTNet is designed to work with small sample sizes and heterogeneous datasets.

In Amyar et al. (2020) Multitask deep learning-based model was proposed. The proposed model can improve the state-of-the-art U-NET model by leveraging useful information contained in multiple related tasks. The main aim of this approach is on the one hand to leverage useful information contained in multiple related tasks to improve both segmentation and classification performances, and on the other hand to deal with the problems of small datasets.

2.2 *Pre-trained models based on transfer learning*

Various methods based on transfer learning are proposed to detect coronavirus from medical images. In Singh et al. (2020a), the authors used convolutional networks to detect negative and positive cases of coronavirus on CT scan samples. In Apostolopoulos and Mpesiana (2020), common convolutional architectures, such as VGG19, MobileNet v2, Inception, Xception, and Inception ResNet v2 have been used along with transfer learning to classify samples into three categories: normal, bacterial pneumonia, and COVID-19. A common transfer learning technique with fine-tuning is used in Minaee et al. (2020) to identify COVID-19. The authors used some convolutional neural network architectures, such as DenseNet-121, ResNet50, SqueezeNet, and ResNet18. These models were tested on a dataset of 5000 X-ray images. In Hasan et al. (2020), as in the methods mentioned, transfer learning on a trained VGG-16 model is used to diagnose COVID-19. In Brunese et al. (2020), like other similar methods, the pre-trained VGG-16 network is used to detect COVID-19. In Li (2020), an efficient 3D deep learning framework called CONVNet is introduced. CONVNet uses the pre-trained Resnet architecture to extract two-dimensional and three-dimensional features. In Song (2021), the authors have developed a new method called DeepPneumonia for the diagnosis of bacterial pneumonia, COVID-19, and healthy cases. This model has achieved 86.5% and 94% accuracy for detecting COVID-19 with bacterial pneumonia and healthy cases, respectively. Other similar methods are introduced in Zhou et al. (2020), Jaiswal et al. (2020).

2.3 *Methods based on handcrafted feature extraction*

Some COVID-19 detection methods used handcrafted feature extraction approaches. In Pereira et al. (2020), first, different texture features are extracted from the images by popular texture descriptors, and then these texture features are combined with the extracted features from the pre-trained InceptionV3 (Szegedy et al. 2016) model. In Al-Karawi et al. (2020a), a method for classifying the positive and negative cases of COVID-19 based on CT scan images was proposed. Different texture features were extracted from CT images using the Gabor filter, and then the SVM method was used to classify these images. In Hasan et al. (2020), to reduce intensity variations between CT slices, a preprocessing step was applied on CT slices. Then a long short-term memory (LSTM) classifier is used to discriminate between COVID-19, pneumonia, and healthy cases. Other related methods based on the combination of feature extraction approaches and deep learning models are introduced in Farid et al. (2020a).

Recently, some new methods for the segmentation or classification of Corona images have been introduced. In Abd Elaziz et al. (2021), the goal is to present an efficient image segmentation method for COVID-19 CT images. This method depends on improving the density peaks clustering (DPC) using generalized extreme value (GEV) distribution. The DPC is faster than other clustering methods, and it provides more stable results. However, it is difficult to determine the optimal number of clustering centers automatically without visualization. So, GEV is used to determine the suitable threshold value to find the optimal number of clustering centers that lead to improving the segmentation process. The proposed model is applied to a set of twelve COVID-19 CT images.

In Elaziz et al. (2020), the authors proposed a hybrid swarm intelligence (SI) based approach that combines the features of two SI methods, marine predators algorithm (MPA) and moth-flame optimization (MFO). This approach was called MPAMFO, in which, the MFO was utilized as a local search method for MPA to avoid trapping at local optima. The MPAMFO was proposed as an MLT approach for image segmentation, which showed excellent performance in all experiments. the authors tested the MPAMFO for a real-world application, such as CT images of COVID-19. Thirteen CT images were used to test the performance of MPAMFO.

To determine the COVID-19 case from other normal and abnormal cases, the authors in Yousri et al. (2021) proposed a method that extracted the informative features from X-ray images, leveraging on a new feature selection method to determine the relevant features. in this method, an enhanced cuckoo search optimization algorithm (CS) was proposed using fractional-order calculus (FO) and four different heavy-tailed distributions in place of the Lévy flight to strengthen the algorithm performance during dealing with the COVID-19 multi-class classification optimization task. The classification process included three classes, called normal patients, COVID-19 infected patients, and pneumonia patients. The distributions used are Mittag–Leffler distribution, Cauchy distribution, Pareto distribution, and Weibull distribution. Two datasets for COVID-19 X-ray images were considered for testing the proposed method.

Most of the mentioned methods are highly dependent on the image domain of datasets on which they were trained. If the test set is from the same domain of the training set, the model performance will be acceptable. However, when the domain of the evaluation dataset is different, model performance is significantly reduced. However, in real-world applications, the domain of the inference image is not always the same as the training set. In other words, unseen data is often independent of the training set, so the results would not be reliable.

3 Proposed framework for COVID-19 detection

The steps of the AMTLDC framework are shown graphically in Fig. 1. As shown in this figure, the COVID-19 detection model uses two separate datasets to learn common representations that are independent of the domain of each dataset. The source dataset is used for network training and the target dataset is used to increase the transferability and generalizability. The next step is the preprocessing step. In this step, the input data is decoded, resized, normalized, and finally transformed by data augmentation techniques. The next step is the AMTLDC architecture, which consists of three parts: CNN-based Feature Extractor, classifier, and discriminator. These blocks are responsible for extracting

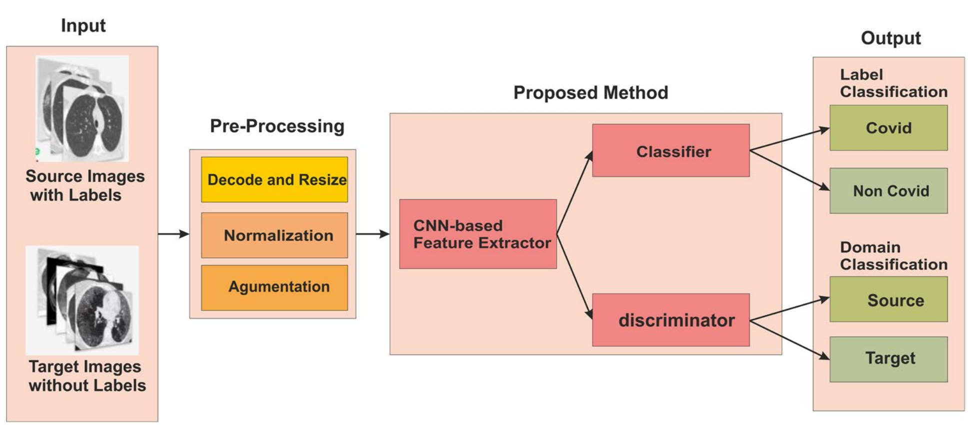

Fig. 1 AMTLDC framework

features, classifying data into two classes COVID-19 or Non-COVID-19, and distinguishing source data from target data, respectively.

The purpose of the AMTLDC is to learn general features that are useful for both datasets so that the correct classification can be done regardless of the input source and the specific aspects of each input distribution.

3.1 Preprocessing

The preprocessing steps are described below:

Step1: Decode: CT images are often saved in DICOM format. These files must be converted to the common image format. In this research, images are converted to *png* format.

Step2: Resize: CT images are collected from different sources, which may not be in the same size. Therefore, all images should be resized to be suitable for the proposed network input layer.

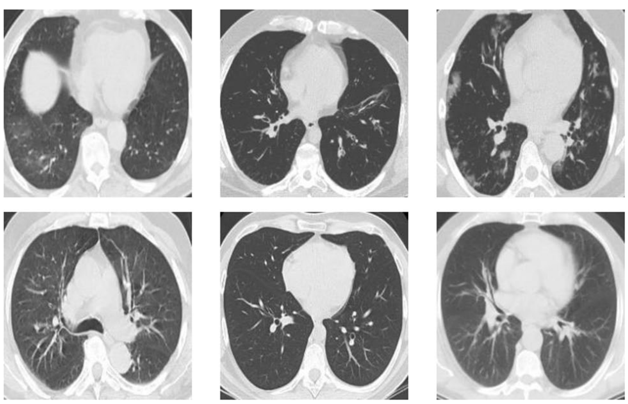

Fig. 2 Some images after preprocessing step

Step3: Normalize: These images are often in the range of 0 to 255, which should be normalized for network training. So we normalized these images in the [0, 1] range.

Step4: Data Augmentation: Due to insufficient data for network training, we use the data augmentation technique to generate new data. The transformations used on the images are brightness, contrast, rotation, and noise, which are applied in the range/type of 0.2, 0.2, [− 20°, + 20°], and horizontal respectively.

Some images after preprocessing step are shown in Fig. 2.

3.2 *AMTLDC framework*

The AMTLDC architecture consists of three parts: CNN-based Feature Extractor, classifier, and discriminator. Figure 3 shows these modules. As shown in this figure, in the feature extraction block, common convolutional architectures such as VGG16, and Resnet with transfer learning techniques can be used. In AMTLDC, we use the pre-trained ResNet50 (He et al. 2016) architecture on the imagenet dataset. To classify data into two classes, COVID-19 and Non_COVID-19, we pass the extracted representations from the feature extraction block into two consecutive modules consisting of Dense, Batch-normalization, Relu, and Dropout layers. Then, on top of these two blocks, apply the sigmoid activation function. The output of this model determines the probability of assigning each sample to each class. The domain classifier is responsible for identifying and distinguishing source data from target data. The goal is to learn representations that are common among datasets. In other words, representations specific to one dataset are not learned, which may have some structural and inherent differences from other datasets. The architecture of this block is the same as the classification block, except that the output

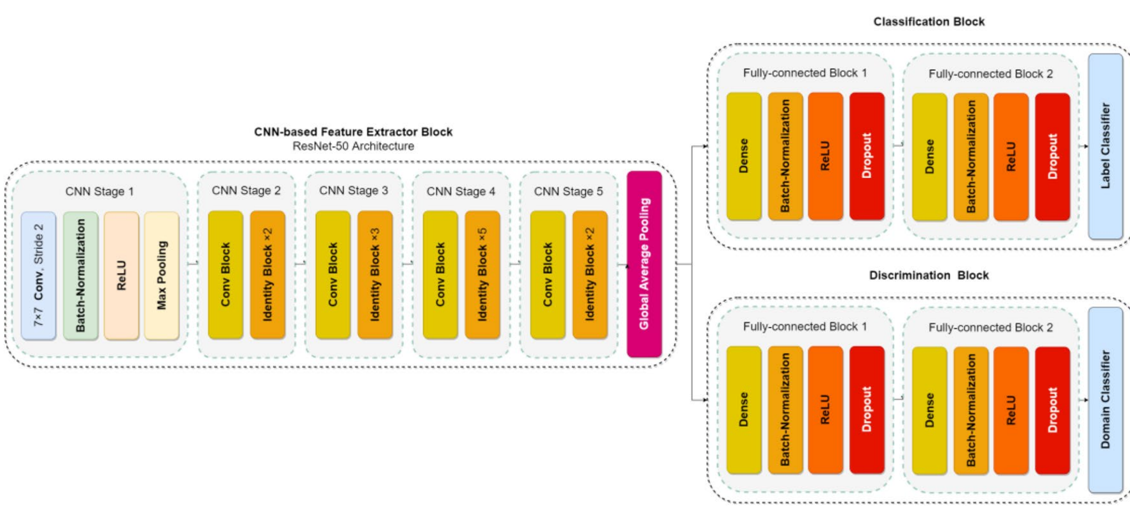

Fig. 3 CNN-based Feature Extractor, Classification, and Discrimination Blocks

estimates the probability of assigning each image to each dataset (source and target). The purpose of this AMTLDC architecture is to increase the model transferability and generalizability, simultaneously. So that the learned representations are independent of the input domain and general, i.e. they are suitable for both source and target datasets; therefore, the representations learned are based on general features, independent of the specific domain and dataset.

3.3 *Training phase*

To solve the issue of over-specialization of the trained model on multiple datasets, and increase the generalizability and transferability of the model inspired by Bois et al. (2021), we train the model with two loss functions. Figure 4 shows a graphical view of the training approach. This model is trained by an efficient adversarial training approach in a multi-source transfer learning environment. The classification and discrimination blocks use the features extracted by the feature extraction block to classify the input data and the domain the data come from, respectively. Both of these blocks are trained by backpropagating their respective losses. The binary cross-entropy loss function is used to calculate the losses of both blocks. When arriving at the feature extractor block, the loss related to the discrimination block is reversed by the inverse gradient layer. Thus, the feature extractor block learns common and general representations of both sources that are useful for classifying input, at the same time, these learned representations are indiscriminative of the domain the data come from.

AMTLDC is trained simultaneously with two loss functions: classification loss and discrimination loss. Equation (1) shows the used loss function in this method. This loss combines a classification loss ($\mathcal{L}_c$) and a discrimination loss ($\mathcal{L}_d$). λ_c and λ_d are coefficients, controlling the bias-vs-variance tradeoff of the generalization.

$$\mathcal{L} = \lambda_c \mathcal{L}_c + \lambda_d \mathcal{L}_d \tag{1}$$

We use the cross-entropy loss function to calculate the discriminator domain loss and the classification loss. The classification loss ($\mathcal{L}_c$) in this algorithm is defined by Eq. (2)

$$\mathcal{L}_c = -y log(\hat{y}) \tag{2}$$

where y indicates the correct class, $\hat{y}$ indicates the model prediction.

The discrimination domain loss ($\mathcal{L}_d$) in this algorithm is defined in Eq. (3).

$$\mathcal{L}_d = -y log(\hat{y}) \tag{3}$$

where y indicates the correct domain class, $\hat{y}$ indicates the model prediction.

4 Experiments

In this section, the efficiency of the AMTLDC method is evaluated and compared with the following groups of methods:

1. Methods based on customized network
2. Methods based on pre-trained networks and transfer learning

For the AMTLDC, the parameters and their values are described in Table 1. In the proposed method, the two

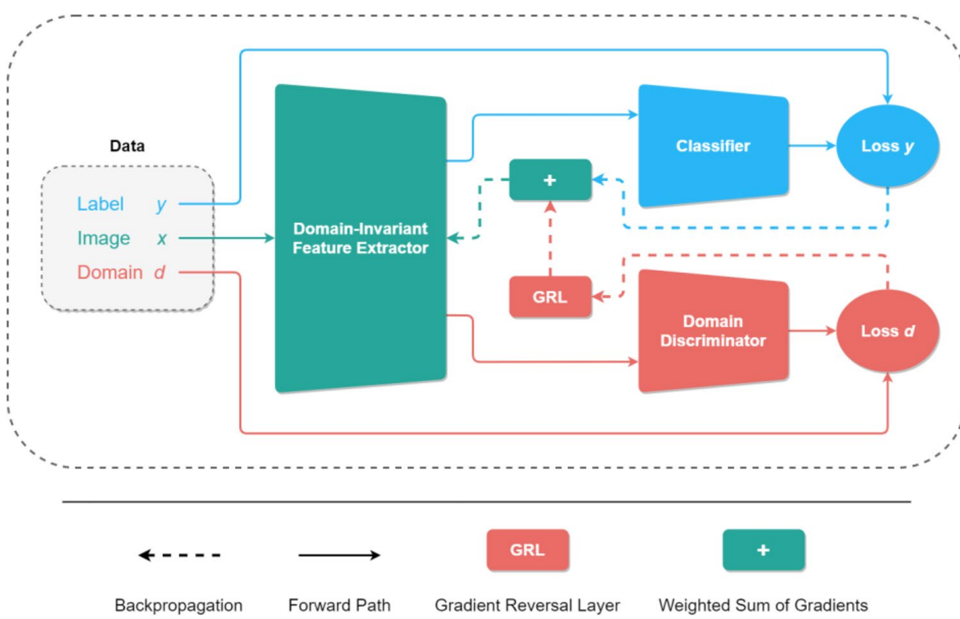

Fig. 4 Efficient adversarial training approach in a multi-source transfer learning environment

Table 1 AMTLDC parameters

Parameter	Value
dropout rate	0.5
Coefficients λ_d	1
Coefficients λ_c	4
Batch size	32
maximum number of iterations	2×10^4
Learning rate (Adam optimizer)	10^{-2}

parameters that have a significant effect on the results are the λ_d and λ_c coefficients. We tested these parameters in the range of Worldometer (2021), Hemdan et al. (2003). According to this test, the best results were obtained with the values of 1 and 4 for λ_d and λ_c coefficients, respectively.

In the proposed method and conventional architecture (the second group of compared methods), such as VGG-16 and ResNet, all common parameters such as learning rate and batch size are considered the same. Also, the number of layers and neurons in the classification module are similar. So the comparisons are quite fair. In the methods whose source code is not available or have their parameters, the best results are reported directly from the relevant papers. Most of these methods often introduce a customized architecture for Covid classification tasks.

4.1 *Evaluation criteria*

In the experiments, such as (Luz et al. 2021; Yousri et al. 2021; Hashemzadeh et al. 2019; Golzari Oskouei et al. 2021a, 2021b, 2022; Aria et al. 2022b; Golzari Oskouei and Hashemzadeh 2022; Wang et al. 2021; Ghaderzadeh et al. 2022), we use *Accuracy*, *Precision*, *Recall*, *F1*, and *Specificity* criteria to evaluate the algorithms. These evaluation criteria are shown in Eqs. (4–8), where TP, FN, TN, and FN represent True Positive, False Positive, True Negative, and False Negative, respectively.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{4}$$

$$Precision = \frac{TP}{Tp + FP} \times 100 \tag{5}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100 \tag{6}$$

$$\mathrm{F1} = 2 \times \frac{\mathrm{Recall} \times \mathrm{Precision}}{\mathrm{Recall} + \mathrm{Precision}} \times 100 \tag{7}$$

$$\mathrm{Specificity} = \frac{TN}{TN + FP} \times 100 \tag{8}$$

4.2 Dataset

In recent research, three datasets, SARS-CoV-2 CT (Angelov and Almeida Soares 2020), COVID19-CT (He et al. 2020), and COVID19-CT_v2 (Zhao et al. 2020) are often used to evaluate model performance. We also test the performance of the AMTLDC method on these datasets. The SARS-CoV-2 dataset contains 1252 corona images and 1230 non-corona images. The COVID19-CT dataset contains 349 corona images and 397 non-corona images. The COVID19-CT_v2 dataset contains 349 corona images and 463 non-corona images.

5 Results

Tables 2, 3, 4, 5, 6, 7 show the results of different evaluation criteria. The results of other methods are reported directly from the relevant articles.

5.1 Experiment 1: evaluation of the SARS-CoV-2 dataset

In this section, we evaluate the AMTLDC method on the SARS-CoV-2 dataset and compare it with other successful methods. The results of the AMTLDC and other methods are stated in Table 2.

Table 2 shows the performance of our method, compared to other methods, is the best. The results indicate the higher performance of our method compared to other advanced methods in this research field. After our method, EfficientNetB0, and xDNN methods have the best performance, respectively and the Decision Tree method has the worst result among all the methods. Also, for the AMTLDC, the confusion matrix of evaluation on the test set of the SARS-CoV-2 dataset is shown in Fig. 5. From Table 2 and Fig. 5, it is evident that AMTLDC performs better than the other methods. The average Accuracy, Precision, Recall, and F1 metrics of AMTLDC are 99.8%, 99.8%, 99.7%, and 99.7%, respectively. Recall 99.7% indicates that, on average, only two COVID-19 images is incorrectly predicted as non-COVID-19. Moreover, the AMTLDC correctly diagnoses all non-COVID-19 cases with only two false positive. After AMTLDC, the EfficientNetB0, xDNN, DenseNet201-Based, and ShuffleNet methods have relatively good performance, respectively. In EfficientNetB0 architecture, on average, three COVID-19 images are incorrectly predicted as non-COVID-19.

Table 2 Performance comparison of different models on the SARS-CoV-2 dataset

Model/method	Evaluation metrics			
	Accuracy	Precision	Recall	F_1
Decision tree	79.4	76.8	83.1	79.8
GoogleNet	91.7	90.2	93.5	91.8
AlexNet	93.7	94.9	92.2	93.6
ResNet50V2	94.2	92.8	96.7	94.1
ResNet50	94.9	93.0	97.1	95.0
VGG-16	94.9	94.0	95.4	94.9
AdaBoost	95.1	93.6	96.7	95.1
SqueezeNet	95.1	94.2	96.2	95.2
ShuffleNet	97.5	96.1	99.0	97.5
EfficientNetB0	98.9	99.1	98.9	99.0
Xception	98.8	99.0	98.6	98.8
COVID CT-Net (Yazdani et al. 2020)	90.7	88.5	85.0	90.0
Contrastive learning (Wang et al. 2020b)	90.8	95.7	85.8	90.8
Modified VGG19 (Panwar et al. 2020)	95.0	95.3	94.0	94.3
DenseNet201-based (Jaiswal et al. 2020)	96.2	96.2	96.2	96.2
xDNN (Soares et al. 2020)	97.3	99.1	95.5	97.3
AMTLDC	**99.8**	**99.8**	**99.7**	**99.7**

The best performace for the used dataset and methods are in bold

Table 3 Performance comparison of different models on the COVID19-CT dataset

Model/method	Evaluation metrics			
	Accuracy	Recall	Specificity	F_1
AlexNet	74.5	70.4	79.0	75.0
SqueezeNet	78.5	86.5	63.8	82.0
VGG-16	78.5	74.6	82.8	76.0
GoogleNet	78.9	75.9	82.3	79.0
VGG-19	83.2	90.7	74.7	85.0
NasNet-mobile	83.4	84.8	81.9	85.0
NasNet-large	85.2	79.3	91.9	84.0
Xception	85.6	88.3	80.6	87.7
ShuffleNet	86.1	83.5	89.0	86.0
Inception-ResNet-v2	86.3	88.1	84.2	87.0
MobileNet-v2	87.2	93.2	77.6	89.0
DenseNet-121	88.9	88.8	88.9	88.2
Inception-v3	89.4	90.0	88.9	88.8
ResNet-101	89.7	82.2	89.2	89.0
ResNet-18	90.1	89.4	90.9	91.0
ResNeXt-50	90.6	93.4	88.2	90.3
ResNeXt-101	90.9	93.1	88.9	90.6
DenseNet-169	91.2	93.3	88.9	90.8
DenseNet-201	91.7	88.6	94.1	91.9
Contrastive learning (Wang et al. 2020b)	78.6	78.0	77.0	78.8
ResNet-101-based (Saqib et al. 2020)	80.3	85.7	86.0	81.8
DenseNet-169-based (He et al. 2020)	83.0	84.8	85.5	81.0
DenseNet-121 + SVM (Jokandan et al. 2007)	85.9	84.9	86.8	86.2
DenseNet-169-based (Martinez 2009)	87.7	85.6	86.9	87.8
Decision function (Mishra et al. 2020)	88.3	87.0	87.9	86.7
AMTLDC	**95.1**	**94.6**	**95.8**	**94.1**

The best performace for the used dataset and methods are in bold

Table 4 Cross-dataset evaluation

Method	Training dataset	Test dataset	Evaluation metrics		
			Accuracy	Recall	Precision
Without multi-source transfer learning	SARS-CoV-2	COVID19-CT (train set)	64.11	65.05	64.10
	SARS-CoV-2	COVID19-CT (test set)	62.01	63.21	61.89
	SARS-CoV-2	COVID19-CT (all data)	64.21	66.45	65.04
	COVID19-CT	SARS-CoV-2	56.92	59.28	56.47
	SARS-CoV-2	COVID19-CT-v2 (train set)	60.34	61.15	60.10
	SARS-CoV-2	COVID19-CT-v2 (test set)	58.78	59.41	57.23
	SARS-CoV-2	COVID19-CT-v2 (all data)	60.13	62.91	61.16
	COVID19-CT-v2	SARS-CoV-2	52.87	55.41	52.32
With multi-source transfer learning	SARS-CoV-2	COVID19-CT (train set)	91.00	92.55	92.01
	SARS-CoV-2	COVID19-CT (test set)	90.44	90.89	91.81
	SARS-CoV-2	COVID19-CT (all data)	**92.37**	**93.42**	**92.37**
	COVID19-CT	SARS-CoV-2	82.97	87.37	84.09
	SARS-CoV-2	COVID19-CT-v2 (train set)	89.32	90.50	90.35
	SARS-CoV-2	COVID19-CT-v2 (test set)	88.12	88.10	89.68
	SARS-CoV-2	COVID19-CT-v2 (all data)	**90.41**	**91.42**	**90.75**
	COVID19-CT-v2	SARS-CoV-2	80.64	85.74	82.42

The best performace for the used dataset and methods are in bold

Comparing AMTLDC with other ResNet-based methods, it can be seen that the AMTLDC, while having a smaller network, has much better results than other versions. So that the Accuracy, Precision, Recall, and F1 rates of the AMTLDC approach with comparing other ResNet-based methods are improved by an average of 5.04%, 6.27%, 3.24%, and 5.90%, respectively.

5.2 Experiment 2: evaluation of the COVID19-CT dataset

In this section, we evaluate the AMTLDC on the COVID19-CT dataset and compare it with other successful methods. The results of the AMTLDC and other methods are stated in Table 3.

As Table 3 shows, the performance of AMTLDC—is the best among all methods. The results indicate the higher performance of AMTLDC compared to other successful methods in this field. After our method, ResNet-based methods have relatively good performance and the AlexNet method has the worst result among all the methods.

For AMTLDC, the average Accuracy, Recall, Specificity, and F1 metrics are 95.1%, 94.6%, 95.8%, and 94.1%, respectively. Also, after AMTLDC, the ResNet-based methods have relatively good performance. The average Accuracy, Recall, Specificity, and F1 metrics on second best method (DenseNet-169) are 91.2%, 93.3%, 88.9%, and 90.8%, respectively. In DenseNet-169, the average Recall of 93.3% indicates that, on average, eight images of COVID-19 are incorrectly predicted as non-COVID-19. Also, the average specificity of 88.9 indicates that all non-COVID-19 cases are detected with more than ten false-positive samples. In AMTLDC, the average Recall of 94.6% indicates that, on average, seven COVID-19 images are incorrectly predicted as non-COVID-19. Also, the average of Specificity 95.8% indicates that it detects all cases of non-COVID-19 with only six false-positive samples.

Comparing the AMTLDC with the other ResNet-based methods, such as ResNeXt-50, ResNeXt-101, and ResNet-50 models, it can be seen that the AMTLDC has much better results than these methods. The Accuracy, Recall, Specificity, and F1 rates of the proposed approach with comparing ResNeXt-101 architecture are improved by an average of 5.28%, 1.82%, 7.98%, and 4.85%, respectively. Similarly, the Accuracy, Recall, Specificity, and F1 rates of the proposed approach with comparing ResNeXt-50 architecture are improved by an average of 5.62%, 1.49%, 8.84%, and 5.20%, respectively.

5.3 Experiment 3: cross-dataset evaluation

In this section, we evaluate the transferability and generalizability of our AMTLDC. To investigate whether the AMTLDC prevents negative transfer or not, we test our AMTLDC once with proposed multi-source transfer learning and once without it. In both modes, we train the network once on the SARS-CoV-2 dataset and evaluate it on the COVID19-CT and COVID19-CT–v2 datasets, and vice versa. Table 4 shows the results of this evaluation. According to this table, it can be seen that the results of the AMTLDC have been improved by about 30% compared to without using the multi-source transfer technique. As can be

Table 5 AMTLDC vs. pre-trained models

References	Data sources	No. of samples	Model	Performance
Ardakani et al. (2020b)	Real-time data from the hospital environment	Total: 1,020 COVID-19: 510 Non-COVID-19: 510	AlexNet, VGG-16, VGG-19, …	Accuracy: 99.51, Recall: **100**, Specificity: 99.02
Chen et al. (2020)	Renmin Hospital of Wuhan University	Total: 35,355	UNet+ +	Accuracy: 98.85, Recall: 94.34, Specificity: 99.16
Cifci (2022)	Kaggle benchmark dataset (Kaggle 2020)	Total: 5,800	AlexNet, Inception-V4	Accuracy: 94.74, Recall: 87.37, Specificity: 87.45
Javaheri et al. (2005)	Five medical centers in Iran, SPIE-AAPM-NCI (Armato et al. 2015), LUNGx (Armato et al. 2016)	Total: 89,145 COVID-19: 32,230 Non-COVID-19: 56,915	BCDU-Net (U-Net)	Accuracy: 91.66, Recall: 87.5, Specificity: 94
Jin et al. (2020)	Wuhan Union Hospital, LIDC-IDRI (Armato et al. 2011), ILD-HUG (Depeursinge et al. 2012)	Total: 1881 COVID-19: 496 Non-COVID-19: 1385	ResNet152	Accuracy: 94.98, Recall: 94.06, Specificity: 95.47, F1: 92.78
Jin et al. (2020)	Five different hospitals of China	Total: 1,391 COVID-19: 850 Non-COVID-19: 541	DPN-92, Inception-v3, ResNet-50	Recall: 97.04, Specificity: 92.2
Dadário et al. (2020))	Multiple hospitals environment	Total: 4536 COVID-19: 1296 Non-COVID-19: 1325	ResNet50	Recall: 90, Specificity: 96
Wu et al. (2020)	China Medical University, Beijing Youan Hospital	Total: 495 COVID-19: 368 Non-COVID-19: 127	ResNet50	Accuracy: 76, Recall: 81.1, Specificity: 61.5
Xu et al. (2019)	Zhejiang University, Hospital of Wenzhou, Hospital of Wenling	Total: 618 COVID-19: 219 Non-COVID-19: 399	ResNet18	Accuracy: 86.7, Recall: 81.5, F1: 81.1
Yousefzadeh et al. (2020)	Real-time data from the hospital environment	Total: 2124 COVID-19: 706 Non-COVID-19: 1418	DenseNet, ResNet, Xception, EcientNetB0	Accuracy: 96.4, Recall: 92.4, Specificity: 98.3, F1: 95.3
AMTLDC	SARS-CoV-2 CT-scan dataset	Total: 2482 COVID-19: 1252 Non-COVID-19: 1229	ResNet50	Accuracy: **99.96**, Recall: 99.80, Specificity: **99.80**, F1: **99.90**

The best performace for the used dataset and methods are in bold

seen in this table, the AMTLDC improves generalizability. A closer look reveals that network training on COVID19-CT and COVID19-CT datasets are less generalizable than network training on SARS-CoV-2 datasets. The reason for this is quite clear, the SARS-CoV-2 dataset is richer (size, variety of data collected) than the two other datasets. Also, the data collected in this dataset are from different sources, in different contrasts, and with different visual features. Therefore, it is not a suitable dataset for model training.

The Accuracy, Recall, and Precision rates of the proposed approach (with the proposed multi-source transfer learning mode) with comparing without it are improved by an average of 42.37%, 41.15%, and 43.24%, respectively.

5.4 Experiment 4: AMTLDC vs. pre-trained models

We test the AMTLDC method with methods in which the models are pre-trained on the ImageNet dataset. As shown in Table 5, the results show that the proposed algorithm has a higher performance than other successful methods in this field. The critical point is that the proposed method is trained on the small SARS-CoV-2 CT-scan dataset, while the other methods are often trained on a large dataset. Therefore, apart from the qualitative contributions and the proposed innovations that offer a low-cost and practical solution to overcome the shortcut learning problem (Geirhos et al. 2020), the proposed method achieves significant improvements using only a few sets of training samples without suffering from overfitting problem.

Table 6 AMTLDC vs. customized models

Reference	Data Sources	No. of samples	Model	Performance
Liu et al. (2020)	Ten designated COVID-19 hospitals in China	Total: 1993 COVID-19: 920 Non-COVID-19: 1073	Modified DenseNet-264	Accuracy: 94.3, Recall: 93.1, Specificity: 95.1
Amyar et al. (2020)	COVID-CT (Zhao et al. 2020), COVID-19 CT segmentation dataset (2020), Henri Becquerel Center	Total: 1044 COVID-19: 449 Non-COVID-19: 595	Encoder-Decoder with multi-layer perceptron	Accuracy: 86.0, Recall: 94.0, Specificity: 79.0
Elghamrawy and Hassanien. (2020)	COVID-19 Database (Italian Society of Medical and Interventional Radiology : COVID-19 Database 2020), COVID-CT (Zhao et al. 2020)	Total: 583 COVID-19: 432 Non-COVID-19: 151	WOA-CNN	Accuracy: 96.40, Recall: 97.25, Precision: 97.3
Farid et al. (2020b)	Kaggle benchmark dataset (Kaggle Benchmark dataset 2020)	Total: 102 COVID-19: 51 Non-COVID-19: 51	CNN	Accuracy: 94.11, Precision: 99.4, F1: 94.0
Hasan et al. (2020)	COVID-19 (2020), SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset (Armato et al. 2015)	Total: 321 COVID-19: 118 Non-COVID-19: 203	QDE–DF	Accuracy: 99.68
Singh et al. (2020b)	COVID-19 patient chest CT images (Li et al. 2020a)	Total: 150 COVID-19: 75 Non-COVID-19: 75	MODE-CNN	Accuracy: 93.25, Recall: 90.70, Specificity: 90.72
Wang et al. (2020c)	Xi'an Jiaotong University, Nanchang University, Xi'anMedical College	Total: 1,065 COVID-19: 740 Non-COVID-19: 325	Modified Inception	Accuracy: 79.3, Recall: 83.0, Specificity: 67.0
Song et al. (2020)	Hospital of Wuhan University, Third Affiliated Hospital	Total: 1990 COVID-19: 777 Non-COVID-19: 1213	DRE-Net	Accuracy: 94.3, Recall: 93.0, Precision: 96.0
Zheng et al. (2020)	Union Hospital, Tongji Medical College, Huazhong University of Science and Technology	Total: 630	DeCoVNet	Accuracy: 90.1, Recall: 90.7, Specificity: 91.1
AMTLDC	SARS-CoV-2 CT-scan dataset	Total: 2482 COVID-19: 1252 Non-COVID-19: 1229	ResNet50	Accuracy: **99.86**, Recall: **99.80,** Specificity: **99.70**, F1: **99.70**

The best performace for the used dataset and methods are in bold

Table 7 Number of parameters and runtime of different methods

Model	No. of parameters	Runtime (second)
MobileNet-v2	6,444,417 (~6 M)	4 s
DenseNet-121	10,253,057 (~10 M)	11 s
VGG-16	16,324,609 (~16 M)	22 s
DenseNet-169	17,865,473 (~17 M)	23 s
VGG-19	21,634,305 (~21 M)	30 s
DenseNet-201	24,347,393 (~24 M)	34 s
InceptionV3	25,083,873 (~25 M)	36 s
Xception	27,288,297 (~27 M)	39 s
ResNet50V2	29,991,617 (~30 M)	43 s
ResNet-50	30,014,529 (~30 M)	43 s
ResNet101V2	49,053,377 (~49 M)	71 s
ResNet-101	49,084,993 (~49 M)	71 s
InceptionResNetV2	56,798,625 (~57 M)	83 s
ResNet152V2	64,758,465 (~64 M)	90 s
ResNet152	64,797,761 (~65 M)	93 s
AMTLDC	36,441,346 (~36 M)	49 s

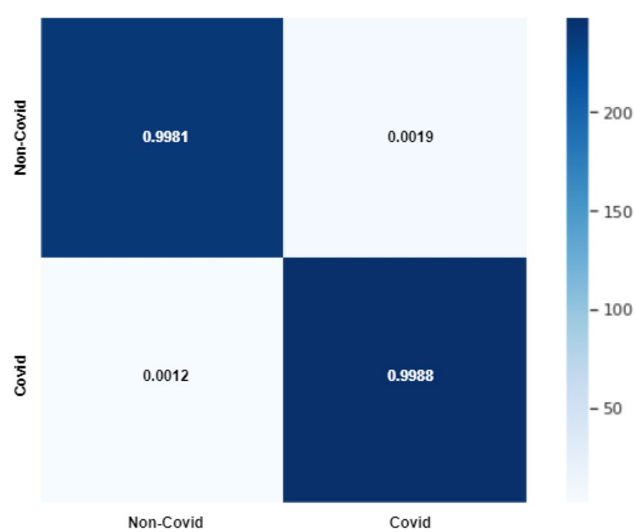

Fig. 5 Confusion matrix of evaluation on the test set of the SARS-CoV-2 dataset

The method presented by Ardakani et al. (2020a) has almost higher performance than AMTLDC in terms of Recall metric; however, it suffers from low reliability. In other words, in the face of unseen data, network performance decreases dramatically.

5.5 Experiment 5: AMTLDC vs. customized models

This section compares the proposed method with methods that have developed customized architectures specifically to detect COVID-19. In these methods, transfer learning is not used, and the network is trained from scratch.

Table 6 shows the results for AMTLDC and other compared approaches. As shown in this table, the AMTLDC method, for all metrics, has the best results. After AMTLDC, Elghamrawy and Hassanien (2020) has the second-best results. Moreover, Hasan et al. (2020) has relatively good performance. Among the reported results, Wang et al. (2020c) has the worst performance.

5.6 Experiment 6: AMTLDC runtime and comparison with the baseline method

In this section, the runtime of AMTLDC is compared with the baseline CNN-based models. The experiments are conducted on a computer with Intel corei7-4700HQ, CPU 2.40 GHz, and 8 GB RAM. Table 7 shows a time complexity analysis in terms of batch training time (secs). As shown in this table, the runtime time of the proposed method is less than the ResNet-based models, such as ResNet152, ResNet152V2, InceptionResNetV2, ResNet-101, and ResNet101V2. The lowest running time belongs to MobileNet-v2 and the highest running time belongs to ResNet152. The reason that the runtime time of the proposed method is longer than some methods is that the proposed method uses an additional (discriminator module) block for domain classification. Therefore, compared to a similar method such as ResNet-50, it has about 6 million more parameters. The discriminator block increases the training time by approximately 7 s. By analyzing and comparing the proposed method with methods that have a lower runtime, such as VGG-16, it can be seen that although the training time is not optimal in AMTLDC, the accuracy has improved significantly. So, in the AMTLDC, the accuracy has been improved by more than 15% on average compared to VGG-16. This claim is also true for other models with less runtime.

6 Conclusion

In this research, we proposed a multi-source adversarial transfer learning model for the diagnosis of COVID-19 disease, a task in which the generalizability of the model is greatly reduced due to a lack of data. Existing methods do not have good results in unseen data due to a lack of sufficient data. Therefore, they are not reliable in real-world applications. Thanks to the use of two different sources in the proposed COVID-19 detection framework, the AMTLDC ensures that the representations learned are common among datasets and are not specific to the domain of a particular dataset. In other words, in the AMTLDC, the generalizability and transferability of the model are improved and it has brilliant results for unseen data. The performance of the AMTLDC was compared with many advanced models. The results showed that the AMTLDC has high generalizability and transferability and has improved up to 50% of the results compared to similar methods. Also, The obtained results indicate that AMTLDC achieves classification improvements of at least 2%, 18%, 18%, and 9% in Accuracy, Precision, Recall, and F1, respectively, compared with the best results of competitors, even without directly training on the same data.

Although many methods have been proposed in the field of diagnosing COVID-19 using medical images, and even some methods have reported 100% accuracy, they are not highly efficient in real-world applications. The main reason is that the datasets used to train the network are not from a specific imaging device. Positive samples were mostly collected from one imaging device and negative samples were collected from different devices. The proposed methods learn the specific patterns of an imaging device, not the patterns and structures involved in corona images. Therefore, we believe that the main challenge in this field is still collecting quality and standardized data.

Common loss functions are used in the proposed method. As a future direction, it would be useful to improve AMTLDC by using efficient loss functions, such as triple loss or center loss. By using these loss functions for samples of a class, in the embedded space, the distance is reduced, and therefore better separating representations are learned for classification.

Author contributions All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.

Data availability The authors make all datasets underlying the findings described in their manuscript available without restriction.

Declarations

Conflict of interest The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

References

Abd Elaziz M, Al-qaness MAA, Abo Zaid EO, Lu S, Ali Ibrahim R, Ewees AA (2021) Automatic clustering method to segment COVID-19 CT images. PLoS ONE 16(1):e0244416. https://doi.org/10.1371/journal.pone.0244416

Afshar P, Heidarian S, Naderkhani F, Oikonomou A, Plataniotis KN, Mohammadi A (2020) COVID-CAPS: a capsule network-based framework for identification of COVID-19 cases from X-ray images. Pattern Recogn Lett 138:638–643. https://doi.org/10.1016/j.patrec.2020.09.010

Ai T et al (2020) Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology 296(2):E32–E40

Al-Karawi D, Al-Zaidi S, Polus N, Jassim S (2020a) Machine learning analysis of chest CT scan images as a complementary digital test of coronavirus (COVID-19) patients. MedRxiv. https://doi.org/10.1101/2020.04.13.20063479

Alshazly H, Linse C, Barth E, Martinetz T (2021) Explainable COVID-19 detection using chest ct scans and deep learning. Sensors 21(2):455

Altae-Tran H, Ramsundar B, Pappu AS, Pande V (2017) Low data drug discovery with one-shot learning. ACS Cent Sci 3(4):283–293

Amyar A, Modzelewski R, Li H, Ruan S (2020) Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: classification and segmentation. Comput Biol Med 126:104037

Angelov P, Almeida-Soares E (2020) SARS-CoV-2 CT-scan dataset: a large dataset of real patients CT scans for SARS-CoV-2 identification. MedRxiv. https://doi.org/10.1101/2020.04.24.20078584

Apostolopoulos ID, Mpesiana TA (2020) COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43(2):635–640. https://doi.org/10.1007/s13246-020-00865-4

Ardakani AA, Kanafi AR, Acharya UR, Khadem N, Mohammadi A (2020a) Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: results of 10 convolutional neural networks. Comput Biol Med 121:103795

Ardakani AA, Kanafi AR, Acharya UR, Khadem N, Mohammadi A (2020b) Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: results of 10 convolutional neural networks. Comput Biol Med 121:103795

Aria M, Hashemzadeh M, Farajzadeh N (2022a) QDL-CMFD: a quality-independent and deep learning-based copy-move image forgery detection method. Neurocomputing 511:213–236. https://doi.org/10.1016/j.neucom.2022.09.017

Aria M, Nourani E, Golzari Oskouei A (2022b) ADA-COVID: adversarial deep domain adaptation-based diagnosis of COVID-19 from lung CT scans using triplet embeddings. Comput Intell Neurosci 2022:2564022. https://doi.org/10.1155/2022/2564022

Armato SG III et al (2011) The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Med Phys 38(2):915–931

Armato S III et al (2015) SPIE-AAPM-NCI lung nodule classification challenge dataset. Cancer Imaging Arch 10:K9

Armato SG III et al (2016) LUNGx Challenge for computerized lung nodule classification. J Med Imaging 3(4):044506

Bar Y, Diamant I, Wolf L, Greenspan H (2015) Deep learning with non-medical training used for chest pathology identification. In: Medical Imaging 2015: Computer-Aided Diagnosis, 2015, vol. 9414: International Society for Optics and Photonics, p. 94140V

Bayani A et al (2022) Identifying predictors of varices grading in patients with cirrhosis using ensemble learning. Clin Chem Lab Med (CCLM) 60:1938–1945

Bayani A et al (2022) Performance of machine learning techniques on prediction of esophageal varices grades among patients with cirrhosis. Clin Chem Lab Med (CCLM). https://doi.org/10.1515/cclm-2022-0623

Brunese L, Mercaldo F, Reginelli A, Santone A (2020) Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. Comput Methods Prog Biomed 196:105608. https://doi.org/10.1016/j.cmpb.2020.105608

Chen J et al (2020) Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study. MedRxiv 10:1–11

Christodoulidis S, Anthimopoulos M, Ebner L, Christe A, Mougiakakou S (2016) Multisource transfer learning with convolutional neural networks for lung pattern analysis. IEEE J Biomed Health Inform 21(1):76–84

Christodoulidis S, Anthimopoulos M, Ebner L, Christe A, Mougiakakou S (2017) Multisource transfer learning with convolutional neural networks for lung pattern analysis. IEEE J Biomed Health Inform 21(1):76–84. https://doi.org/10.1109/JBHI.2016.2636929

Cifci MA (2020) Deep learning model for diagnosis of corona virus disease from CT images

COVID-19. [Online]. Available: https://radiopaedia.org/. Accessed: 9 Apr 2020

COVID-19 CT Segmentation Dataset. [Online]. Available: http://medicalsegmentation.com/covid19/. Accessed: 15 Apr 2020

Dadário AMV, Paiva JPQ, Chate RC, Machado BS, Szarf G (2020) Regarding Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT. Radiology, p. 201178

De Bois M, El Yacoubi MA, Ammi M (2021) Adversarial multi-source transfer learning in healthcare: application to glucose prediction for diabetic people. Comput Methods Prog Biomed 199:105874. https://doi.org/10.1016/j.cmpb.2020.105874

Depeursinge A, Vargas A, Platon A, Geissbuhler A, Poletti P-A, Müller H (2012) Building a reference multimedia database for interstitial lung diseases. Comput Med Imaging Graph 36(3):227–238

Dhungel N, Carneiro G, Bradley AP (2017) A deep learning approach for the analysis of masses in mammograms with minimal user intervention. Med Image Anal 37:114–128. https://doi.org/10.1016/j.media.2017.01.009

Elaziz MA et al (2020) An improved marine predators algorithm with fuzzy entropy for multi-level thresholding: real world example of COVID-19 CT image segmentation. IEEE Access 8:125306–125330. https://doi.org/10.1109/access.2020.3007928

El-Ghamrawy SM (2020) Diagnosis and Prediction Model for COVID19 Patients response to treatment based on convolutional neural networks and whale optimization algorithm using CT images. MedRxiv. https://doi.org/10.1101/2020.04.16.20063990

Esteva A et al (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639):115–118. https://doi.org/10.1038/nature21056

Farid AA, Selim GI, Khater HAA (2020a) A novel approach of CT images feature analysis and prediction to screen for corona virus disease (COVID-19)

Farid AA, Selim GI, Awad H, Khater A (2020b) A novel approach of CT images feature analysis and prediction to screen for corona virus disease (COVID-19). Int J Sci Eng Res 11(3):1–9

Farooq M, Hafeez A (2020) COVID-resnet: a deep learning framework for screening of covid19 from radiographs. arXiv preprint https://arXiv.org/2003.14395

Geirhos R et al (2020) Shortcut learning in deep neural networks. Nat Mach Intell 2(11):665–673

Ghaderzadeh M, Aria M (2021) Management of COVID-19 detection using artificial intelligence in 2020 pandemic. Presented at the 2021 5th International Conference on Medical and Health Informatics, Kyoto, Japan, 2021. https://doi.org/10.1145/3472813.3472820

Ghaderzadeh M, Asadi F, Jafari R, Bashash D, Abolghasemi H, Aria M (2021a) Deep convolutional neural network–based computer-aided detection system for COVID-19 using multiple lung scans: design and implementation study. J Med Internet Res 23(4):e27468

Ghaderzadeh M, Aria M, Asadi F (2021b) X-ray equipped with artificial intelligence: changing the COVID-19 diagnostic paradigm during the pandemic. BioMed Res Int 2021:9942873. https://doi.org/10.1155/2021/9942873

Ghaderzadeh M, Aria M, Hosseini A, Asadi F, Bashash D, Abolghasemi H (2022) A fast and efficient CNN model for B-ALL diagnosis and its subtypes classification using peripheral blood smear images. Int J Intell Syst 37(8):5113–5133

Golzari Oskouei A, Hashemzadeh M (2022) CGFFCM: a color image segmentation method based on cluster-weight and feature-weight learning. Softw Impacts 11:100228. https://doi.org/10.1016/j.simpa.2022.100228

Golzari Oskouei A, Hashemzadeh M, Asheghi B, Balafar MA (2021a) CGFFCM: cluster-weight and group-local feature-weight learning in Fuzzy C-means clustering algorithm for color image segmentation. Appl Soft Comput 113:108005. https://doi.org/10.1016/j.asoc.2021.108005

Golzari Oskouei A, Balafar MA, Motamed C (2021b) FKMAWCW: categorical fuzzy k-modes clustering with automated attribute-weight and cluster-weight learning. Chaos Solitons Fractals 153:111494. https://doi.org/10.1016/j.chaos.2021.111494

Golzari Oskouei A, Balafar MA, Motamed C (2022) EDCWRN: efficient deep clustering with the weight of representations and the help of neighbors. Appl Intell. https://doi.org/10.1007/s10489-022-03895-5

Gulshan V et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316(22):2402–2410. https://doi.org/10.1001/jama.2016.17216

Hasan AM, Al-Jawad MM, Jalab HA, Shaiba H, Ibrahim RW, Al-Shamasneh AAR (2020) Classification of COVID-19 coronavirus, pneumonia and healthy lungs in CT scans using Q-deformed entropy and deep learning features. Entropy 22(5):517

Hasan M, Alam M, Elahi M, Toufick E, Roy S, Wahid SR (2020) CVR-Net: a deep convolutional neural network for coronavirus recognition from chest radiography images. arXiv preprint https://arXiv.org/2007.11993

Hashemzadeh M, Golzari Oskouei A, Farajzadeh N (2019) New fuzzy C-means clustering method based on feature-weight and cluster-weight learning. Appl Soft Comput 78:324–345. https://doi.org/10.1016/j.asoc.2019.02.038

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp 770–778.

He X et al (2020) Sample-efficient deep learning for COVID-19 diagnosis based on CT scans. Medrxiv 3:034501

Hemdan EED, Shouman MA, Karar ME (2020) Covidx-net: a framework of deep learning classifiers to diagnose COVID-19 in X-ray images. arXiv preprint https://arXiv.org/2003.11055

Italian Society of Medical and Interventional Radiology : COVID-19 Database. [Online]. Available: https://www.sirm.org. Accessed: 28 Mar 2020

Jaiswal A, Gianchandani N, Singh D, Kumar V, Kaur M (2020) Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2020.1788642

Javaheri T et al. (2020) Covidctnet: an open-source deep learning approach to identify COVID-19 using ct image. arXiv preprint https://arXiv.org/2005.03059

Jin C et al (2020) Development and evaluation of an AI system for COVID-19 diagnosis. medRxiv 2:1

Jin S et al (2020) AI-assisted CT imaging analysis for COVID-19 screening: building and deploying a medical AI system in four weeks. medRxiv. https://doi.org/10.1101/2020.03.19.20039354

Jokandan AS et al (2020) An uncertainty-aware transfer learning-based framework for COVID-19 diagnosis. arXiv preprint https://arXiv.org/2007.14846

Kaggle Benchmark Dataset. [Online] Available: https://www.kaggle.com/andrewmvd/covid19-ct-scans. Accessed: 1 Mar 2020

Li X, Zeng X, Liu B, Yu Y (2020a) COVID-19 infection presenting with CT halo sign. Radiol Cardiothor Imaging 2(1):e200026

Li L et al (2020b) Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. Radiology

Li T, Han, Wei B, Zheng Y, Hong Y, Cong J (2020c) Robust screening of COVID-19 from chest X-ray via discriminative cost-sensitive learning. arXiv preprint https://arXiv.org/2004.12592

Liu et al. (2020) Assisting scalable diagnosis automatically via CT images in the combat against COVID-19. medRxiv,

Long C et al (2020) Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? Eur J Radiol 126:108961. https://doi.org/10.1016/j.ejrad.2020.108961

Luz E et al (2021) Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images. Res Biomed Eng 38:1–14

Martinez AR (2020) Classification of covid-19 in ct scans using multi-source transfer learning. arXiv preprint https://arXiv.org/2009.10474

Minaee S, Kafieh R, Sonka M, Yazdani S, Jamalipour-Soufi G (2020) Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning. Med Image Anal 65:101794. https://doi.org/10.1016/j.media.2020.101794

Mishra AK, Das SK, Roy P, Bandyopadhyay S (2020) Identifying COVID19 from chest CT images: a deep convolutional neural networks based approach. J Healthc Eng. https://doi.org/10.1155/2020/8843664

Ng M-Y et al (2020) Imaging profile of the COVID-19 infection: radiologic findings and literature review. Radiol Cardiothor Imaging 2(1):e200034

Panwar H, Gupta P, Siddiqui MK, Morales-Menendez R, Bhardwaj P, Singh V (2020) A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. Chaos, Solitons Fractals 140:110190

Pereira RM, Bertolini D, Teixeira LO, Silla CN, Costa YMG (2020) COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. Comput Methods Prog Biomed 194:105532. https://doi.org/10.1016/j.cmpb.2020.105532

Saqib M, Anwar S, Anwar A, and Blumenstein M (2020) COVID19 detection from Radiographs: Is Deep Learning able to handle the crisis? TechRxiv

Silva P et al (2020) COVID-19 detection in CT images with deep learning: a voting-based scheme and cross-datasets analysis. Inform Med Unlocked 20:100427. https://doi.org/10.1016/j.imu.2020.100427

Singh D, Kumar V, Vaishali V, Kaur M (2020a) Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks. Eur J Clin Microbiol Infect Dis 39(7):1379–1389. https://doi.org/10.1007/s10096-020-03901-z

Singh D, Kumar V, Kaur M (2020b) Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks. Eur J Clin Microbiol Infect Dis 39:1–11

Soares E, Angelov P, Biaso S, Froes MH, Abe DK (2020) SARS-CoV-2 CT-scan dataset: a large dataset of real patients CT scans for SARS-CoV-2 identification. medRxiv. https://doi.org/10.1101/2020.04.24.20078584

Song Y et al. (2020) Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. medRxiv

Song Y et al (2021) Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. IEEE/ACM Trans Comput Biol Bioinform. https://doi.org/10.1148/radiol.2020200905

Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the Inception Architecture for Computer Vision. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27–30 June 2016, pp 2818–2826, doi: https://doi.org/10.1109/CVPR.2016.308

Tabik S et al (2020) COVIDGR dataset and COVID-SDNet methodology for predicting COVID-19 based on chest X-ray images. IEEE J Biomed Health Inform 24(12):3595–3605

Tartaglione E, Barbano CA, Berzovini C, Calandri M, Grangetto M (2020) Unveiling COVID-19 from chest X-ray with deep learning: a hurdles race with small data. Int J Environ Res Public Health 17(18):6933

Wang D, Khosla A, Gargeya R, Irshad H, Beck AH (2016) Deep learning for identifying metastatic breast cancer. arXiv preprint https://arXiv.org/1606.05718

Wang F, Casalino LP, Khullar D (2019) Deep learning in medicine—promise, progress, and challenges. JAMA Intern Med 179(3):293–294. https://doi.org/10.1001/jamainternmed.2018.7117

Wang L, Lin ZQ, Wong A (2020a) COVID-net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. Sci Rep 10(1):1–12

Wang Z, Liu Q, Dou Q (2020b) Contrastive cross-site learning with redesigned net for COVID-19 ct classification. IEEE J Biomed Health Inform 24(10):2806–2813

Wang et al. (2020c) A deep learning algorithm using CT images to screen for Corona virus disease (COVID-19). MedRxiv

Wang S et al (2021) A deep learning algorithm using CT images to screen for corona virus disease (COVID-19). Eur Radiol 31(8):6096–6104. https://doi.org/10.1007/s00330-021-07715-1

Worldometer—real time world statistics. [Online]. Available: https://www.worldometers.info/. Accessed 13 Aug 2021

Wu X et al (2020) Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: a multicentre study. Eur J Radiol 128:109041

Xu X et al (2020) A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering 6:1122–1129

Yazdani S, Minaee S, Kafieh R, Saeedizadeh N, Sonka M (2020) COVID ct-net: Predicting covid-19 from chest ct images using attentional convolutional network. arXiv preprint https://arXiv.org/2009.05096

Yousefzadeh M et al (2020) AI-corona: radiologist-assistant deep learning framework for COVID-19 diagnosis in chest CT scans. MedRxiv

Yousri D, Abd Elaziz M, Abualigah L, Oliva D, Al-Qaness MAA, Ewees AA (2021) COVID-19 X-ray images classification based on enhanced fractional-order cuckoo search optimizer using heavy-tailed distributions. Appl Soft Comput 101:107052. https://doi.org/10.1016/j.asoc.2020.107052

Zhao J, Zhang Y, He X, Xie P (2020) COVID-CT-Dataset: a CT scan dataset about COVID-19. arXiv preprint https://arXiv.org/2003.13865, vol. 490

Zheng C et al (2020) Deep learning-based detection for COVID-19 from chest CT using weak label. MedRxiv 395:497

Zhou M et al (2020) Improved deep learning model for differentiating novel coronavirus pneumonia and influenza pneumonia. medRxiv 2:e191095