




Research Article

Multiround Transfer Learning and Modified Generative Adversarial Network for Lung Cancer Detection

Kwok Tai Chui ¹, **Brij B. Gupta** ^{2,3,4,5,6}, **Rutvij H. Jhaveri** ⁷, **Hao Ran Chi**,⁸
Varsha Arya,^{4,9} **Ammar Almomani**,^{10,11} and **Ali Nauman**¹²

¹Department of Electronic Engineering and Computer Science, School of Science and Technology, Hong Kong Metropolitan University, Ho Man Tin, Hong Kong SAR, China

²International Center for AI and Cyber Security Research and Innovations, Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan

³Symbiosis Centre for Information Technology (SCIT), Symbiosis International University, Pune, India

⁴Lebanese American University, Beirut, 1102, Lebanon

⁵Center for Interdisciplinary Research at University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India

⁶Department of Computer Science, Dar Alheikma University, Jeddah, Saudi Arabia

⁷Department of Computer Science and Engineering, School of Technology, Pandit Deendayal Energy University, Gandhinagar, India

⁸Instituto de Telecomunicações, Aveiro, Portugal

⁹Asia University, Taichung 41354, Taiwan

¹⁰School of Information Technology, Skyline University College, P.O. Box 1797, UAE

¹¹Al-Balqa Applied University, Salt, Jordan

¹²Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, Republic of Korea

Correspondence should be addressed to Kwok Tai Chui; jktchui@hkmu.edu.hk and Brij B. Gupta; gupta.brij@gmail.com

Received 20 October 2022; Revised 31 January 2023; Accepted 14 February 2023; Published 3 March 2023

Academic Editor: Lianyong Qi

Copyright © 2023 Kwok Tai Chui et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lung cancer has been the leading cause of cancer death for many decades. With the advent of artificial intelligence, various machine learning models have been proposed for lung cancer detection (LCD). Typically, challenges in building an accurate LCD model are the small-scale datasets, the poor generalizability to detect unseen data, and the selection of useful source domains and prioritization of multiple source domains for transfer learning. In this paper, a multiround transfer learning and modified generative adversarial network (MTL-MGAN) algorithm is proposed for LCD. The MTL transfers the knowledge between the prioritized source domains and target domain to get rid of exhaust search of datasets prioritization among multiple datasets, maximizing the transferability with a multiround transfer learning process, and avoiding negative transfer via customization of loss functions in the aspects of domain, instance, and feature. In regard to the MGAN, it not only generates additional training data but also creates intermediate domains to bridge the gap between the source domains and target domains. 10 benchmark datasets are chosen for the performance evaluation and analysis of the MTL-MGAN. The proposed algorithm has significantly improved the accuracy compared with related works. To examine the contributions of the individual components of the MTL-MGAN, ablation studies are conducted to confirm the effectiveness of the prioritization algorithm, the MTL, the negative transfer avoidance via loss functions, and the MGAN. The research implications are to confirm the feasibility of multiround transfer learning to enhance the optimal solution of the target model and to provide a generic approach to bridge the gap between the source domain and target domain using MGAN.

1. Introduction

Cancer is the second leading cause of global death, according to the World Health Organization [1]. Among all types of cancers, lung cancer is ranked first that has caused 1.8 million deaths in each year. Lung cancer detection (LCD) in the early stage is important for medical staff to tailor-make the treatment plan and perform the prognostic estimation. LCD using artificial intelligence receives increasing attention in both academia and practice in view of the inadequacies of medical staff [2] and the heavy workload [3]. Reducing the time spent on medical diagnosis provides more time to medical doctors to concentrate on professional surgery and consultation and thus leveraging the healthcare quality. In this paper, we consider the traditional lung cancer screening via biomedical imaging, instead of an emerging approach using breath by the electronic nose [4, 5].

The traditional machine learning model is trained with a dataset that often reaches a bottleneck in achieving excellent model performance (e.g., in terms of sensitivity, specificity, and accuracy) to fulfil the mission-critical medical diagnosis. In addition, large-scale datasets may not be available for training an accurate deep learning-based model for all applications. These drive the emerging research trend in applying transfer learning, that performs knowledge transfer from the source domain to the target domain. In literature, it is well demonstrated the superiority and applicability of transfer learning in many research applications [6, 8]. Attention is drawn to a more general scenario, where the source domain and target domain are different but related (less difficult) or different and unrelated (more challenging). The issue of the negative transfer becomes more severe with the increase of dissimilarities between the source domain and target domain because there are more unrelated samples from the source domain [8]. The loss functions can be formulated to reduce the impact of negative transfer.

The rest of the paper is organized as follows. Section 1 is divided into three subsections to present a summary of the related works, a discussion of the research limitations of the related works, and the major research contributions of our work. Section 2 presents the design and formulations of the proposed algorithm for LCD. Section 3 summarizes the details of the 10 benchmark datasets and presents the performance evaluation and comparison. To investigate the contributions of the individual components of the proposed algorithm, ablation studies are conducted in Section 4. At last, in Section 5, a conclusion is drawn with future research directions.

1.1. Related Works. Although existing works [9–16] formulated the transfer learning problems with a single source domain and single target domain, the discussion has merit as these works fell into the same research area, i.e., transfer learning for LCD. In the following, two common types of formulations will be discussed: (i) transfer learning between the similar source domain and target domain [9–12] and (ii) transfer learning between the distant source domain and target domain [13–16].

The discussion begins with the transfer learning problem using a similar source domain and target domain. In [9], a hybrid residual and deep neural networks was proposed for the transfer learning from Luna16 to a small-scale dataset (125 chest computed tomography (CT) scans) collected by researchers in Shandong Provincial Hospital. The ablation study showed that the transfer learning strategy enhanced the accuracy of the LCD model from 79.5% to 85.7%. ImageNet was served as the source model in the transfer learning strategy to fine-tune the target model [10]. VGG16 and deep neural network were used to build the LCD model, which was evaluated using two benchmark datasets. Transfer learning enhanced the accuracy of the model from 87.5% to 90.8%. To transfer the knowledge from LUNA16 to the target domain of the Gangneung Asan Hospital for LCD, a YOLOX algorithm was used [11]. Results showed a slight enhancement of the model's accuracy from 89.7% to 90.9%. Some scenarios also suggested that improper settings in the fine-tuning of the target model may lead to deterioration on the model performance, which is a well-known issue of negative transfer. In [12], a nodule identification convolutional neural network was pretrained that would transfer knowledge to the target model (using data collected from some hospitals). Semi-supervised deep transfer learning was designed and implemented. Results showed that the sensitivity, specificity, and accuracy were improved from 90.2% to 92.2%, 66.3% to 78.6%, and 83.4% to 88.3%, respectively.

On the other hand, the transfer learning problems are formulated with distant sources and target domains. The work [13] conducted an exploratory analysis on 11 common feature extractors for the source domain (ImageNet), including NASNetLarge, NASNetMobile, DenseNet201, DenseNet169, InceptionResNetV2, ResNet50, InceptionV3, Xception, MobileNet, VGG19, and VGG16. The knowledge was transferred to build various classifiers, such as random forest, K-nearest neighbors, support vector machine, multilayer perceptron, and Naïve Bayes. Results revealed that ResNet50 with support vector machine achieved the best performance with sensitivity and accuracy of 85.4% and 88.4%, respectively. The work also demonstrated the effectiveness of the pretrained model using ImageNet to perform transfer learning on the target domain of chest CT [14]. Four common architectures, namely, DenseNet169, MobileNet, VGG19, and VGG16 were used to build the LCD model. The performance of the model was the best with VGG 16, yielding an accuracy of 91.3%. A recent work [15] has reported a difficulty in the transfer learning strategy without model overfitting. The model was with 98.8% and 83.4% of training accuracy and testing accuracy, respectively. ImageNet was served as the source domain for the knowledge transfer of a VGG19 pretrained model to the target domain of 150 patients with CT scans [16]. The model achieved sensitivity, specificity, and accuracy of 75%, 87%, and 82%, respectively.

1.2. Research Limitations of the Related Works. The major research limitations of the related works are summarized as follows:

- (i) Lack of studies in multiround transfer learning for LCD: existing works considered one-round transfer learning for LCD where only one source domain was involved. Although the target model receives a benefit in the enhancement of model's performance, the model is usually having room for further enhancement (not yet achieved global optimal solution). With more source datasets, it is expected that more unseen data and potential knowledge can be transferred (positive transfer) to further enhance the performance of the target model.
- (ii) Lack of studies in negative transfer between the source domain and the target domain: theoretically, one can formulate the transfer learning problem with the source dataset and target dataset with high similarities [9–12] or low similarities [13–16]. The negative transfer becomes more severe with the decrease in similarities because more unrelated samples can be found in the source dataset. If knowledge from unrelated samples is transferred to the target model, the model's performance becomes worsened. It is needed to avoid negative transfer to ensure the enhancement of performance of the target model, i.e., to guarantee the model moves towards the global optimal solution.
- (iii) Lack of studies in the creation of intermediate domains as a bridge between the source and target domains: controlling the knowledge transfer from the source domain to the target domain is important to enhance the chance of positive transfer. Intermediate domains should be used to break down the transfer learning problem into multiple subproblems. In this consideration, the similarities between the source domain and intermediate domain, as well as between intermediate domain and target domain, are higher than that in the original formulation, between the source domain and the target domain.

1.3. Research Contributions of Our Work. A multiround transfer learning and modified generative adversarial network (MTL-MGAN) algorithm is proposed to address the research limitations. The research contributions of our work are summarized as follows:

- (i) Enhancing the optimal solutions of the LCD model with multiround transfer learning: it has been demonstrated in many existing works for the benefits of transfer learning from the source model to the target model. Applying transfer learning multiple times (multiround transfer learning) with multiple source models is expected to enhance the optimal solutions of the LCD model (target model) where the performance of the target model in the next round is better than that in the current round. This strategy outperforms traditional single-round transfer learning. Ablation study reveals that multiround transfer learning (MTL) enhances the average sensitivity, specificity, and accuracy of the LCD model by 8.28%, 8.21%, and 8.26%, respectively.

- (ii) The loss functions are designed to minimize the impact of negative transfer: data heterogeneity is always existing between the source domain and target domain. Therefore, transfer learning is experienced discrepancies in the joint distributions between the source domain and the target domain. Reformulating the loss functions in domains, instances, and features for the reliable selection of relevant data and knowledge aims to enhancing the performance of the target model. Existing works did not fully consider the issue of negative transfer in the architecture of the transfer learning-based deep learning models. The ablation study shows that the proposed algorithm enhances the sensitivity, specificity, and accuracy of the LCD model by 1.57–2.23%, 1.42–2.26%, and 1.53–2.24%, respectively.
- (iii) A modified generative adversarial network (MGAN) is designed to create two intermediate domains as bridges between the source domain and target domain: bridging the gap between the source and target domains is important to maximize the enhancement of the performance of the LCD model, particularly when the distant source domain is selected. It is worth noting that the merit comes to the applicability of distant source domains where a wide variety of source domains can be selected to contribute to the target model. It could also serve as a generic formulation for distant transfer learning between various types of the source domain and target domain. The MGAN is designed to incorporate the advantages of various baseline generative adversarial network (GAN) algorithms. The rationale is to generate more relevant samples in source domains to enhance the model transferability. In other words, the unrelated samples become less dominant as more relevant samples are available with MGAN. Ablation study shows that the MGAN enhances the sensitivity, specificity, and accuracy of the LCD model by 3.07–4.61%, 2.92–4.33%, and 3.15–4.47%, respectively.

2. Methodology

The design and formulations of the MTL-MGAN are presented. This section is comprised of the overview of the MTL-MGAN, the prioritization algorithm, the loss functions, and the MGAN.

2.1. Overview of the MTL-MGAN. Before the illustration of the design and formulations of the proposed MTL-MGAN, an overview of the architecture is shown in Figure 1. For better visualization, it shows a scenario with multiple source datasets and one target dataset. Consider M source datasets (D_{s1}, \dots, D_{sM}) and one target dataset (TD). All source datasets are ranked in terms of the similarities between source datasets and target datasets using a prioritization algorithm (details in Subsection 2.2). The output of the algorithm provides prioritized source datasets in descending

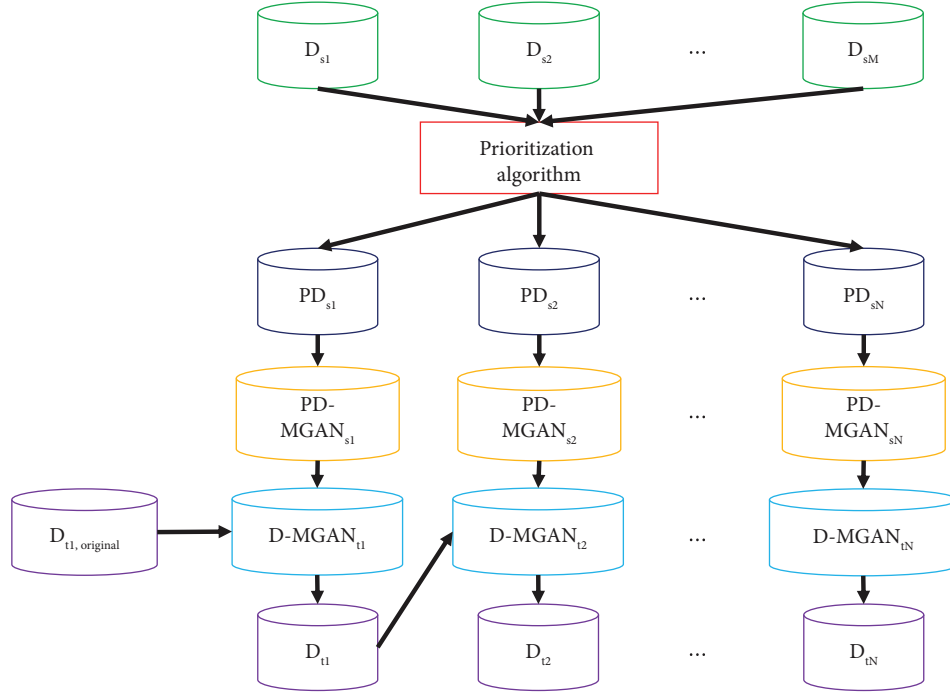


FIGURE 1: The overview of the architecture of the proposed MTL-MGAN.

order, where the highest similarity first, denoted by $(PD_{s1}, \dots, PD_{sN})$, with $N \leq M$ because some of the source datasets could be removed if they contain a significant portion of unrelated samples that may lead to negative transfer to the target domain. A threshold can be defined to filter source-target dataset pairs with low similarity. The removal of these pairs reduces the severity of negative transfer because more irrelevant knowledge could potentially be transferred to the target model. Both prioritized multiple source datasets and target datasets will perform MGAN to create intermediate domains as bridges. The trained target model D_t is updated using MTL with the repetitions of the abovementioned steps.

2.2. Prioritization Algorithm for Multiple Source Datasets.

Selecting appropriate source models to be transferred is important to avoid the waste of effort to transfer limited knowledge to the target domain. More importantly, the transfer of irrelevant knowledge to the target domain, as a well-known issue of negative transfer, should be avoided. Among relevant source models, for those carrying similarities (relevant samples) to the target domain, it is desired to prioritize the models to be transferred (one-to-one transfer learning) in descending order of similarities between the source and target domains. The rationale is due to the enhancement of the robustness of the target model during initial iterations to lower the impact of negative transfer from less similar source domains during later iterations. In addition, prioritization of multiple source datasets helps eliminate source-target domain pairs with low similarity (a threshold can be defined).

To design the prioritization algorithm for multiple source models, a hybrid approach is proposed to merge (i)

modified 2D dynamic warping (M2DW): traditional 2D dynamic warping (2DW) using bidirectional mapping optimally aligns between two images on a similarity basis. However, M2DW performs well only with even resolutions across multiple sensors [17]. The proposed M2DW fills the gap to enable uneven resolutions that are commonly used in practice; (ii) Silhouette coefficient: inspired by [18], where Silhouette coefficient was used to select the source domains using only with pretrained model and target domain. Our work extends the consideration with the aid of the characteristics of the source domains. To begin with, the design and formulations of the M2DW algorithm are presented.

The algorithm first runs through the classes of each dataset and then takes the mean of the image set for each class. Initializing the 2DW barycenter averaging with the medoid of the time series set. The iteration carries out for every pair of datasets using one-to-one mapping. The distance between any pair of datasets equals to the minimal 2DW distance between classes.

The total similarity score SS_{ij} for dataset D_i with N_i sequences and dataset D_j with N_j sequences is given by the following equation:

$$SS_{ij} = \sum_{m \in [1, N_i], n \in [1, N_j]} s_{mn}, \quad (1)$$

where s_{mn} is the similarity score between the m^{th} sequence in D_i and the n^{th} sequence in D_j .

Regarding the Silhouette coefficient, the target training datasets is first encoded with every source models. The average Silhouette coefficient \overline{SC} for each set of encodings is measured with the following formulations:

$$\begin{aligned}
SC_i &= \frac{(h-g)}{\max(g,h)}, \\
g &= \frac{[\sum_{x \in G, x \neq i} d(i,x)]}{(N_G-1)}, \\
h &= \frac{\min_{H \neq G} [\sum_{x \in H} d(i,x)]}{N_H}, \\
\overline{SC} &= \frac{[\sum_{i \in L} s(i)]}{N_L},
\end{aligned} \tag{2}$$

where SC_i is the Silhouette coefficient for a single encoding vector i , d is the distance between two encodings, G and H are some labels of i , L is the label for the final model, and N_G , N_H , and N_L are the number of encodings labeled for labels G , H , and L , respectively.

The total similarity scores for all pairs are normalized and weighted with the results using the Silhouette coefficient. As a result, the priorities of the source domains (to be transferred) are obtained.

2.3. Minimizing the Negative Transfer with Loss Functions. Transfer learning does not guarantee to improve the performance of the target model, that is a commonly known issue of negative transfer. A recent survey on negative transfer [19] summarizes the solutions into three types: (i) secure transfer: the objective function is defined to ensure positive transfer to the target model; (ii) distant transfer: low similarities between source dataset and target dataset may

happen when the datasets are in different domains (research topics). Some researchers demonstrated the effectiveness of setting up an intermediate domain to bridge between the source and target domains; (iii) transferability enhancement: enhancing the data quality in the source datasets leads to the improvement of the transfer learning to the target model.

The first approach is not chosen because of the requirement of the full understanding of all source domains and restrictions on the design and formulations of the transfer learning problem. It is not feasible based on the research initiative to allow distant transfer learning with a wide variety of dissimilar source datasets and target datasets. The second approach is also not appropriate that requires knowledge of source domains and experiences challenging to obtain or create an intermediate domain. Therefore, the last approach is considered to enhance the data transferability between the source domain and the target domain. To comprehensively enhance the data quality, we have formulated the optimization problems in the aspects of domains, instances, and features. The rationale is to fully consider the entire transfer learning process to ensure negative transfer avoidance in all phases. After the selection of useful samples (knowledge), unequal weighting factors are introduced to the first and second-order features. Penalization may also be performed for unrelated samples.

Regarding domains, we first consider the moment distance $d_{\text{moment}}(D_S, D_T)$ for the measurement of the similarity between every pair of domains. Denote the source domains as $D_S = \{D_1, \dots, D_{N_s}\}$ with a total number of source domains N_s and single target domain D_T . The moment distance is defined as

$$d_{\text{moment}}(D_S, D_T) = \frac{\left(\sum_{i=1}^{N_s} \left\| \overline{F(X_{s_i})^1} - \overline{F(X_t)^1} \right\|_2 + \sum_{i=1}^{N_s} \left\| \overline{F(X_{s_i})^2} - \overline{F(X_t)^2} \right\|_2 \right)}{N_s}, \tag{3}$$

where $\overline{F(X_{s_i})^1}$ and $\overline{F(X_{s_i})^2}$ are the average operation of the 1st order and 2nd order features with the s_i source domain, respectively. Likewise, $\overline{F(X_t)^1}$ and $\overline{F(X_t)^2}$ are the average operation of the 1st order and 2nd order features with the target domain, respectively.

Equation (3) assumes equal weighting factors for all source domains; however, this cannot precisely describe the fact that different extent of similarities exists between multiple source domains and target domain. Therefore, modified moment distance $d_{\text{modified}}(D_S, D_T)$ is proposed:

$$d_{\text{moment}}(D_S, D_T) = \frac{\left(\sum_{i=1}^{N_s} \alpha_i \left\| \overline{F(X_{s_i})^1} - \overline{F(X_t)^1} \right\|_2 + \sum_{i=1}^{N_s} \alpha_i \left\| \overline{F(X_{s_i})^2} - \overline{F(X_t)^2} \right\|_2 \right)}{N_s}, \tag{4}$$

where α_i is the normalized weight ($\sum_i \alpha_i = 1$) of the source domain s_i .

In the aspect of instances, the consideration is on the transfer of useful components in the source domain to the target domain. A minimization problem of the transfer learning based on component C_i can be formulated as

$$\min_{M_i} \beta_i (M_i^T M_i) + \gamma_i \|\alpha_i - \min\{\alpha_i\}\| + \delta_{T_i}, \tag{5}$$

where M_i is the Mahalanobis distance of C_i , β_i is the hyperparameter to control the generalization error of M_i , γ_i is the hyperparameter to control the regularization of the samples in C_i , and δ_{T_i} is the loss function (or error) to predict

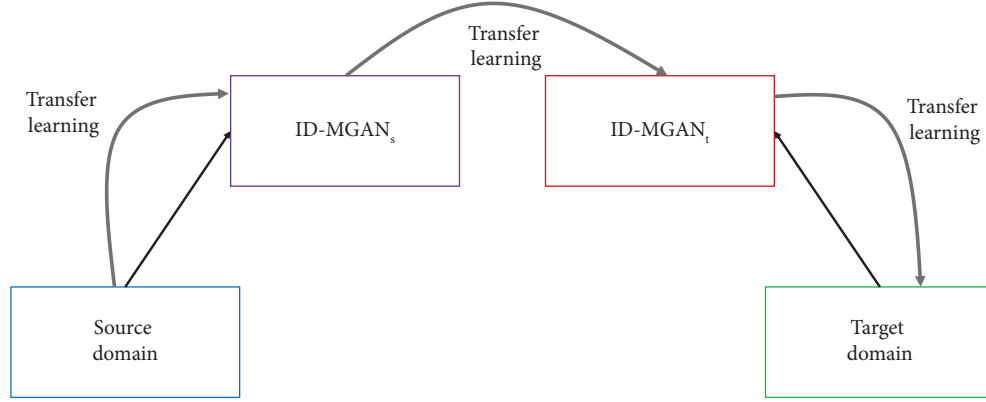


FIGURE 2: The architecture of the transfer learning process with two intermediate domains.

a sample in D_T . The loss function is calculated by the following equation:

$$\delta_{T_i} = \text{SWD}_{\text{within}} - \text{SWD}_{\text{across}}, \quad (6)$$

where $\text{SWD}_{\text{within}}$ and $\text{SWD}_{\text{across}}$ are the sum of the weighted differences within classes and across classes, respectively.

In the aspect of features, for those with small singular values can be penalized via singular value decomposition (SVD) with penalization. The feature matrix $F = [f_1, \dots, f_N]$ is denoted with size N . The representation of F using SVD is given by the following equation:

$$F = U\Sigma V^T, \quad (7)$$

where U is the left singular vector, Σ is the singular value matrix of F , and V is the right singular vector. Rearrange the singular values of Σ as $[\mu_1, \dots, \mu_N]$ in descending order. The idea for transferability enhancement in the feature layer is to penalize the smallest p singular values:

$$L_{\text{penalize}}(F) = \rho \sum_{i=1}^p (\mu_i)^2, \quad (8)$$

where ρ is the hyperparameter to control the strength of penalization and p equals to the number of penalized singular values.

2.4. MGAN for the Creation of Intermediate Domains.

Recall the rationale of the creation of intermediate domains between the source domain and the target domain, is to increase the similarities between the source domain and the target domain. In each round of MTL, two intermediate domains are created. One intermediate domain ID-MGAN_s is based on the source domain and another ID-MGAN_t is

based on the target domain using MGAN. The intermediate domains link closely with the source domain and the target domain to ensure they are based on the distribution of the original datasets (source dataset and target dataset). Figure 2 introduces the architecture of the transfer learning process with two intermediate domains. This has divided the original transfer learning process between the source domain and target domain into three subproblems: (i) subproblem 1: transfer learning between the source domain and ID-MGAN_s ; (ii) subproblem 2: transfer learning between ID-MGAN_s and ID-MGAN_t ; (iii) subproblem 3: transfer learning between ID-MGAN_t and target domain.

The baseline GAN is often not performing well in recent complex machine learning problems because of the fatal theory corruption with random noise vector [20]. Two popular (with high citations in the research publications) variants of GANs namely auxiliary classifier GAN [21] and conditional GAN [22] were thus proposed to solve the limitation. In this paper, we combine these variants of GANs, as the architecture of MGAN.

Figure 3 shows the architecture of the MGAN. Define the notations: noise vector n , conditional variable c , generator G , latent variable z , data distribution X , and discriminator D . MGAN is featured with (i) all generated samples are assigned with label and (ii) adding additional input, conditional variable to the discriminator. The idea of the algorithm is to use G to fool D , with c . G knows the mapping between latent space and data distribution whereas D classifies the generated samples from the ground truth distribution.

Define the loss functions L_{source} and L_{class} for the source and class, respectively. The objective functions of the MGAN are formulated as follows:

$$\begin{aligned} L_{\text{source}} &= E[\log P(\text{Source} = \text{fake} | X_{\text{fake}})] + E[\log P(\text{Source} = \text{real} | X_{\text{real}})], \\ L_{\text{class}} &= E[\log P(\text{Class} = c | X_{\text{fake}})] + E[\log P(\text{Class} = c | X_{\text{real}})], \\ \text{Generator: } &\max L_{\text{class}} - L_{\text{source}}, \\ \text{Discriminator: } &\max L_{\text{class}} + L_{\text{source}}. \end{aligned} \quad (9)$$

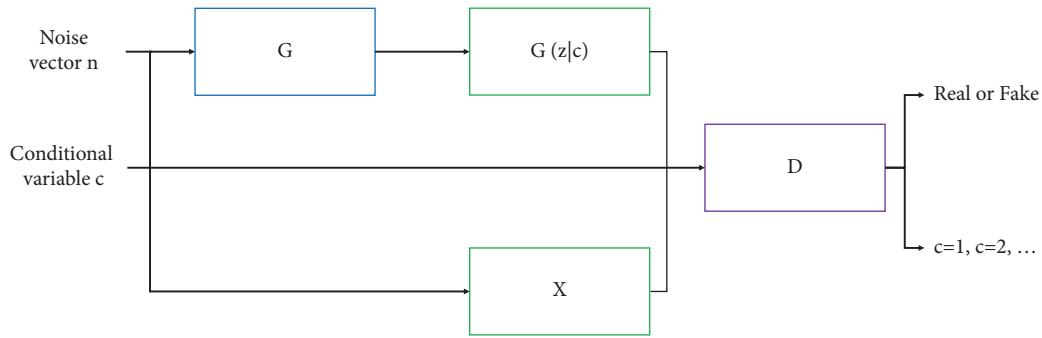


FIGURE 3: The architecture of the MGAN.

3. Performance Evaluation and Comparison

To evaluate the performance of the MTL-MGAN, 10 benchmark datasets are selected. The performance of the MTL-MGAN is analyzed. This is followed by the performance comparison between the MTL-MGAN and existing works.

3.1. Benchmark Datasets. 10 benchmark datasets are selected for which five of them are related to lung cancer datasets (with higher similarities given the application is LCD) and the remaining five of them are related to nonlung cancer datasets (with lower similarities). The five lung cancer datasets are NSCLC-Radiomics [23], NSCLC-Radiomics-Genomics [24], SPIE-AAPM Lung CT Challenge [25], LungCT-Diagnosis [26], and Lung CT Segmentation Challenge 2017 [27]. The nonlung cancer datasets are CIFAR-10 dataset [28], ImageNet dataset [29], Microsoft Common Objects in Context [30] of images for multidisciplinary research, prostate cancer dataset NaF Prostate [31], and breast cancer dataset QIN-Breast [32].

Trivially, it is expected that the similarities between lung cancer datasets [23–27] are high and thus the model experiences less severity of negative transfer. For image datasets of multidiscipline, the datasets [28–30] contain highly dissimilar samples which are more prone to negative transfer. For prostate cancer [31] and breast cancer datasets [32], there exist some similarities between datasets because of the nature of cancer images. These hypotheses will be examined in the following sections.

3.2. Performance Evaluation of the MTL-MGAN. This research study is intended to conduct research on the prioritization of source datasets, the negative transfer avoidance, generation of intermediate domains, and the multiple transfer learning so that the feature extraction and classification algorithms are not major research directions. Therefore, the convolutional neural network is employed as the basic architecture of the target model.

To examine the issue of model overfitting and better fine-tuning the models, 5-fold cross-validation is adopted that has been justified as a common setting of k -fold cross-validation (with $k=5$) [33, 34]. Since 10 benchmark datasets are chosen, at most, the target model performs 9-

round of MTL-MGAN from nine source datasets. The training will stop when negative transfer becomes severe, i.e., the performance (accuracy) of the target model is less than that of the target model using the preceding source dataset.

Figure 4 shows the accuracy of the 5 target models (lung cancer-related) in each round of MTL-MGAN. Several following observations are drawn:

- (i) The maximum number of rounds of MTL-MGAN varies across the target models. The ascending order is given by seven rounds in NSCLC-Radiomics-Genomics [24] and LungCT-Diagnosis [26], eight rounds in NSCLC-Radiomics [23] and Lung CT Segmentation Challenge 2017 [27], and nine rounds in SPIE-AAPM Lung CT Challenge [25].
- (ii) The rank in ascending order for the overall percentage improvement between the first and last round of iteration using MTL-MGAN is 6.85% in SPIE-AAPM Lung CT Challenge [25], 7.00% in NSCLC-Radiomics [23], 8.16% in NSCLC-Radiomics-Genomics [24], 8.70% in Lung CT Segmentation Challenge 2017 [27], and 9.92% in LungCT-Diagnosis [26].
- (iii) The percentage improvement per round using MTL-MGAN in ascending order is 0.761% in SPIE-AAPM Lung CT Challenge [25], 0.875% in NSCLC-Radiomics [23], 1.09% in Lung CT Segmentation Challenge 2017 [27], 1.17% in NSCLC-Radiomics-Genomics [24], and 1.42% in LungCT-Diagnosis [26].

3.3. Performance Comparison with Related Works. The performance comparison between our work and related works covered in Section 1.1 is shown in Table 1. We summarize the observations in each column as follows:

- (i) Source domain and target domain: the related works [9–12] formulated the transfer learning problem using a similar source domain and target domain whereas other works [13–16] considered the distant source and target domains. Our work considered 10 benchmark datasets to evaluate the MTL using similar and distant sources and target domains.

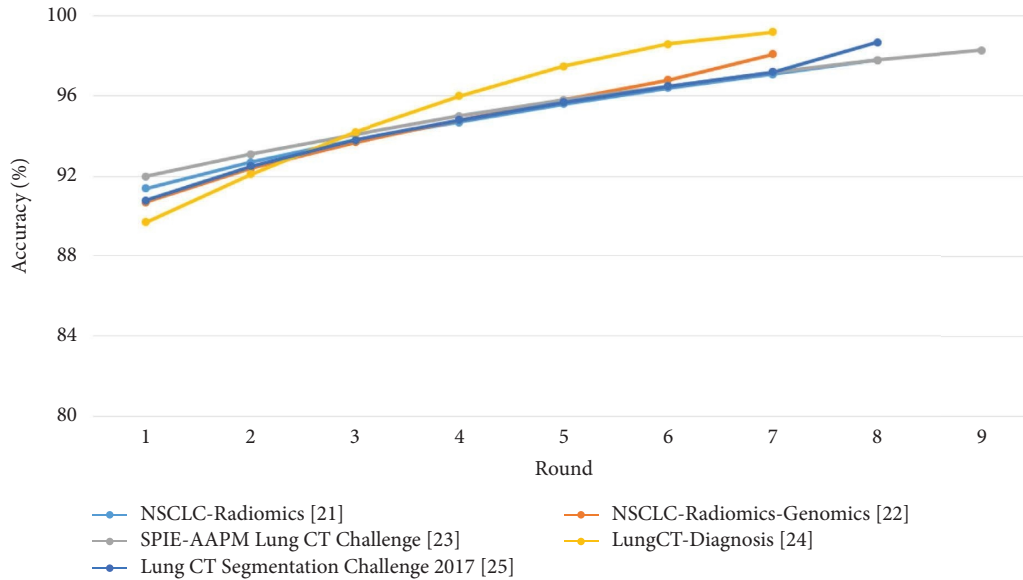


FIGURE 4: Accuracy of the five target models in each round of MTL-MGAN.

- (ii) Intermediate domains: related works [9–16] did not introduce any intermediate domains to bridge the gap between the source domain and target domain. Our work creates two intermediate domains using MGAN to reduce the level of dissimilarities between the source domain and target domain and thus enhancing the transferability. Particularly, it is important when the source domain and target domain are highly differed from each other.
- (iii) Methodology: the related works formulated the classification problems using traditional deep learning algorithms. In view of the research limitations, our work proposed the prioritization algorithm, the multiple transfer learning, the negative transfer avoidance algorithm by designing loss functions, and the MGAN.
- (iv) Cross-validation: related works [9–12, 14, 15] did not employ cross-validation. The performance evaluation possessed limitations in partial utilization of the dataset and lack of information on the evaluation of potential model overfitting when it comes to a deep learning environment. Related works [13, 16] adopted 10-fold cross-validation whereas our work used 5-fold cross-validation. Both 5-fold and 10-fold settings are commonly used in literature with comparable performance [35, 36].
- (v) Ablation study: related works [13–16] did not conduct an ablation study. It is an important element to evaluate the contributions of individual components of the transfer learning model on the performance enhancement of the target model. It is worth noting that negative transfer may exist that is equivalent to a worsened performance on the target model after transfer learning. Other related works [9–12] and our work carry out ablation studies and report the contributions of the transfer learning model in the enhancement of model performance.
- (vi) Sensitivity: related works [9–11, 14, 15] did not report the sensitivity. It is important to report both the sensitivity and specificity to ensure that biased classification is not observed. The works [13, 16] reported the sensitivity of the LCD model when transfer learning is applied. The work [12] revealed the improvement of sensitivity by 2.22% using the transfer learning model. Our work shows an improvement of sensitivity by 6.86–10.8% in the five target models.
- (vii) Specificity: similar to the sensitivity of the model, observation is made for the absence of reporting of the specificity and only the result after using the transfer learning model. The work [12] improved the specificity by 18.6%, nevertheless, model overfitting is observed. Our work shows an improvement of specificity by 6.70–10.4% in the five target models.
- (viii) Accuracy: all related works and our work report the accuracy. Related works [13–16] only reported the results after applying the transfer learning model. The percentage improvement of the accuracy is 7.80% [9], 3.77% [10], 1.34% [11], 5.88% [12], and 6.85–9.92% (our work).

4. Ablation Studies

To evaluate the benefits of the components of the MTL-MGAN, ablation studies are carried out on four key components namely prioritization algorithm, MTL, negative transfer avoidance with loss functions, and MGAN.

TABLE 1: Performance comparison between MTL-MGAN and related works.

Work	Source domain	Intermediate domains	Target domain	Methodology	Cross-validation	Ablation study	Sensitivity	Specificity	Accuracy
[9]	Luna16	No	Shandong provincial hospital	Residual neural network and deep neural network	No	Yes	N/A	N/A	From 79.5% to 85.7%
[10]	ImageNet	No	National lung screening trial and the national institute of allergy and infectious disease TB portal	VGG16 and deep neural network	No	Yes	N/A	N/A	From 87.5% to 90.8%
[11]	Luna16	No	Gangneung Asan hospital	YOLOX	No	Yes	N/A	N/A	From 89.7% to 90.9%
[12]	Nodule identification CNN	No	West China hospital of Sichuan university, Ruijin hospital of Shanghai Jiao Tong university School of medicine, and Changzheng hospital of second military medical university LIDC/IDRI	Semisupervised deep transfer learning	No	Yes	From 90.2% to 92.2%	From 66.3% to 78.6%	From 83.4% to 88.3%
[13]	ImageNet	No		ResNet50 and SVM VGG16, VGG19, MobileNet and DenseNet169	10-fold	No	85.4%	N/A	88.4%
[14]	ImageNet	No	Chest CT		No	No	N/A	N/A	91.3% with VGG16 (best)
[15]	VGG16	No	Iraq-oncology teaching hospital and national center for cancer diseases	Semisupervised deep transfer learning	No	No	N/A	N/A	98.8% (training) and 83.4% (testing)
[16]	ImageNet	No	CT	Deep convolutional neural network	10-fold	No	75%	87%	82%
Our work	Up to 9 of [24–32]	PD-MGAN _s and D-MGAN _t	[23]	MTL-MGAN	5-fold	Yes	From 91.8% to 98.1%	From 90.9% to 97.3%	From 91.4% to 97.8%
Our work	Up to 9 of [23, 25–32]	PD-MGAN _s and D-MGAN _t	[24]	MTL-MGAN	5-fold	Yes	From 90.2% to 97.6%	From 91.4% to 98.7%	From 90.7% to 98.1%
Our work	Up to 9 of [23, 24, 26–32]	PD-MGAN _s and D-MGAN _t	[25]	MTL-MGAN	5-fold	Yes	From 91% to 97.4%	From 92.5% to 98.7%	From 92% to 98.3%
Our work	Up to 9 of [23–25, 27–32]	PD-MGAN _s and D-MGAN _t	[26]	MTL-MGAN	5-fold	Yes	From 89.3% to 98.9%	From 90.0% to 99.4%	From 89.7% to 99.2%
Our work	Up to 9 of [23–26, 28–32]	PD-MGAN _s and D-MGAN _t	[27]	MTL-MGAN	5-fold	Yes	From 91.1% to 98.9%	From 90.3% to 98.3%	From 90.8% to 98.7%

TABLE 2: Performance comparison between MTL-MGAN with and without prioritization algorithm.

Target domain	Number of MTL-MGAN executions	
	With prioritization algorithm	Without prioritization algorithm
[23]	1	362880
[24]	1	181440
[25]	1	362880
[26]	1	181440
[27]	1	362880

TABLE 3: Performance comparison between MGAN and MTL-MGAN.

Target domain	MGAN/MTL-MGAN				
	Sensitivity (%)	Specificity (%)	Precision (%)	F-measure (%)	Accuracy (%)
[23]	91.8/98.1	90.9/97.3	91.2/97.6	91.2/97.5	91.4/97.8
[24]	90.2/97.6	91.4/98.7	90.4/98.0	91.0/98.3	90.7/98.1
[25]	91/97.4	92.5/98.7	91.5/97.8	91.5/97.9	92.0/98.3
[26]	89.3/98.9	90.0/99.4	89.8/99.1	90.1/99.3	89.7/99.2
[27]	91.1/98.9	90.3/98.3	90.6/98.5	91.0/98.9	90.8/98.7

4.1. Contribution of the Prioritization Algorithm. The prioritization algorithm helps ranking the similarities of the multiple source domains to the target domain. Table 2 compares the number of MTL-MGAN execution with and without the prioritization algorithm. The scenario without the prioritization algorithm is equivalent to the exhaustive search (the total number of executions can be found by permutation). The results are identical across different target domains.

4.2. Contribution of the MTL. The sensitivity, specificity, and accuracies of the target model with and without MTL are summarized in Table 3. Observations are drawn as follows:

- (i) Sensitivity: the improvement by MTL is 6.86% [23], 8.20% [24], 7.03% [25], 10.8% [26], and 8.56% [27]. The average sensitivity of the five target models is 8.28%.
- (ii) Specificity: the improvement by MTL is 7.04% [23], 7.99% [24], 6.70% [25], 10.4% [26], and 8.86% [27]. The average specificity of the five target models is 8.21%.
- (iii) Precision: the improvement by MTL is 7.02% [23], 8.41% [24], 6.89% [25], 10.4%, and 8.72% [27]. The average precision of the five target models is 8.29%.
- (iv) F-measure: the improvement by MTL is 6.91% [23], 8.02% [24], 6.99% [25], 10.0% [26], and 8.68% [27]. The average F-measure of the five target models is 8.12%.
- (v) Accuracy: the improvement by MTL is 7.00% [23], 8.16% [24], 6.85% [25], 10.6% [26], and 8.70% [27]. The average accuracy of the five target models is 8.26%.

4.3. Contribution of the Negative Transfer Avoidance with Loss Functions. Recall the loss functions are designed based on three aspects: domains, instances, and features. Table 4 summarizes the performance of the target model with and without the design of loss function in domains, instances, and features.

The comparisons are as follows:

- (i) Domains: the improvements of the sensitivity, specificity, precision, F-measure, and accuracy are ranged 2.09–2.40%, 1.97–2.47%, 2.07–2.40%, 2.07–2.52%, and 2.18–2.30%. The average improvements of the five target models are 2.23%, 2.26%, 2.27%, 2.31%, and 2.24% in sensitivity, specificity, precision, F-measure, and accuracy, respectively.
- (ii) Instances: the improvements of the sensitivity, specificity, precision, F-measure, and accuracy are ranged 1.67–2.06%, 1.84–2.40%, 1.55–2.41%, 1.75–2.29%, and 1.85–2.17%. The average improvements of the five target models are 1.97%, 2.05%, 2.06%, 2.01%, and 1.99% in sensitivity, specificity, precision, F-measure, and accuracy, respectively.
- (iii) Features: the improvements of the sensitivity, specificity, precision, F-measure, and accuracy are ranged 1.33–1.76%, 1.23–1.57%, 1.43–1.55%, 1.14–2.28%, and 1.34–1.66%. The average improvements of the five target models are 1.57%, 1.42%, 1.47%, 1.53%, and 1.53% in sensitivity, specificity, precision, F-measure, and accuracy, respectively.

4.4. Contribution of the MGAN. MGAN is applied to create two intermediate domains based on the source domain and target domain. Table 5 verifies the contributions of MGAN. The improvements of the sensitivity, specificity, precision, F-measure, and accuracy are ranged 3.07–4.61%, 2.92–4.33%, 3.06–4.81%, 2.18–4.24%, and 3.15–4.47%, respectively. The average improvements in sensitivity, specificity, precision, F-measure, and accuracy using with the inclusion of MGAN are 3.61%, 3.56%, 3.70%, 3.32%, and 3.58%, respectively.

TABLE 4: Performance comparison between MTL-MGAN with and without the loss functions in domains, instances, and features.

Loss function (with/without)	Target domain	Sensitivity (%)	Specificity (%)	Precision (%)	F-measure (%)	Accuracy (%)
Domains	[23]	98.1/95.8	97.3/95.2	97.6/95.4	97.6/95.2	97.8/95.6
	[24]	97.6/95.6	98.7/96.5	98.0/95.7	97.7/95.5	98.1/96.0
	[25]	97.4/95.4	98.7/96.4	97.8/95.7	98.5/96.3	98.3/96.2
	[26]	98.9/96.8	99.4/97.0	99.1/96.8	98.8/96.8	99.2/97.0
	[27]	98.9/96.6	98.3/96.4	98.5/96.5	98.5/96.2	98.7/96.5
Instances	[23]	98.1/96.3	97.3/95.3	97.6/95.5	98.0/96.2	97.8/95.9
	[24]	97.6/96.0	98.7/96.7	97.8/96.3	98.3/96.1	98.1/96.3
	[25]	97.4/95.2	98.7/96.9	97.8/95.5	98.8/96.0	98.3/96.3
	[26]	98.9/97.0	99.4/97.6	99.1/97.2	98.8/97.1	99.2/97.4
	[27]	98.9/96.9	98.3/96.0	98.5/96.4	98.5/96.3	98.7/96.6
Features	[23]	98.1/96.4	97.3/95.8	97.5/96.1	97.5/96.4	97.8/96.2
	[24]	97.6/96.1	98.7/97.5	97.9/96.5	98.3/97.0	98.1/96.8
	[25]	97.4/95.8	98.7/97.3	97.8/96.3	98.2/97.0	98.3/96.8
	[26]	98.9/97.6	99.4/97.9	99.1/97.7	98.8/97.2	99.2/97.7
	[27]	98.9/97.4	98.3/97.0	98.5/97.1	98.8/96.6	98.7/97.2

TABLE 5: Performance comparison between MTL and MTL-MGAN.

Target domain	MTL/MTL-MGAN				
	Sensitivity (%)	Specificity (%)	Precision (%)	F-measure (%)	Accuracy (%)
[23]	94.7/98.1	93.4/97.3	93.9/97.5	94.4/98.1	94.2/97.8
[24]	93.3/97.6	94.6/98.7	93.6/98.1	94.3/98.3	93.9/98.1
[25]	94.5/97.4	95.9/98.7	94.9/97.8	95.0/98.1	95.3/98.3
[26]	95.5/98.9	96.5/99.4	95.8/99.0	96.5/98.6	96.1/99.2
[27]	95.8/98.9	95.1/98.3	95.3/98.6	96.0/98.9	95.6/98.7

4.5. *Complexity of the Algorithms.* It can be seen from the results that the prioritization algorithm is important to significantly reduced the trials of the MTL-MGAN with different orders of multiple source datasets. This also reflects a significant reduction in the complexity of the model that avoids unnecessary computing power on exhaustive search. Regarding MTL, which is the strategy to perform multiple times of the transfer learning process. To avoid negative transfer, the loss functions are designed based on the aspects of domains, instances, and features. Although this increases the complexity of the optimization algorithm, the ablation study (Section 4.3) confirms the effectiveness of loss functions. Creating two intermediate domains using MGAN increases the time and computing power of the transfer learning process, however, they contribute to the avoidance of negative transfer.

5. Conclusion

The technological advancement of the machine learning algorithms has received attention in recent years to enhance the medical diagnosis of lung cancers. Responding to the research limitations of existing lung cancer detection models in multiround transfer learning, negative transfer, and lack of bridge between source and target domains, we have proposed a multiround transfer learning and modified generative adversarial network algorithm with a prioritization algorithm and modified loss functions in domains, instances, and features perspectives. 10 benchmark datasets are selected to evaluate the performance of the proposed

algorithm. It significantly enhances the performance of the lung cancer detection model, compared with related works. Ablation studies also provide convincing results to reveal the contributions of the components of the proposed algorithm in the aspects of prioritization algorithm, multiple transfer learning, customized loss functions in domains, instances, features, and modified generative adversarial network.

The implication of the proposed algorithm releases the constraints in the selection of source domains and target domains. Therefore, it can contribute to various research areas, such as sustainable development goals [37], green applications [38], cyber-physical systems [39, 40], smart homes [41], and medical diagnosis [6, 7, 42]. To enhance the efficiency of the optimization algorithm, future investigations could be conducted with various types of optimization approaches, which details can be referred to in review articles [46, 47].

Several future research directions are suggested such as (i) reducing the number of rounds of transfer learning by enhancing the negative transfer avoidance algorithm and generating more relevant samples; (ii) evaluating more baseline deep learning algorithms [43] such as recurrent neural networks, long short-term memory, gated recurrent network, self organization maps, and deep neural network; (iii) including more distant source datasets that are highly dissimilar to the target domain; (iv) modifying the transfer learning process with incremental learning [44] to gradually transfer knowledge between the source and target domains as well as reduce the impact of negative transfer.

List 1 Summary of the acronyms and symbols.

Acronyms

2DW:	2D dynamic warping
MTL:	Multiround transfer learning
c :	Conditional variable
MTL-MGAN:	Multiround transfer learning and modified generative adversarial network
CT:	Computed tomography
n :	Noise vector
d :	Distance between two encodings
N_G :	Number of encodings labeled for label G
$d_{\text{moment}}(D_S, D_T)$:	Moment distance
N_H :	Number of encodings labeled for label H
$d_{\text{modified}}(D_S, D_T)$:	Modified moment distance
:	
N_L :	Number of encodings labeled for label L
D :	Discriminator
PD_{s1}, \dots, PD_{sN} :	Prioritized source datasets with $N \leq M$
N :	
D_S :	Source domains
p :	Number of penalized singular values
D_{s1}, \dots, D_{sM} :	M source datasets
s_{mi} :	Similarity score between the m^{th} sequence in D_i and the n^{th} sequence in D_j
D_t :	Trained target model
SC_i :	Silhouette coefficient for a single encoding vector
D_T :	Single target domain
\overline{SC} :	Average Silhouette coefficient
$F = [f_1, \dots, f_N]$:	Feature matrix with size N
:	
SS_{ij} :	Total similarity score for dataset D_i with N_i sequences and dataset D_j with N_j sequences
$\overline{F(X_{s_i})^1}$:	Average operation of the 1 st order features with the s_i source domain
$\overline{F(X_{s_i})^2}$:	Average operation of the 2 nd order features with the s_i source domain
SWD_{across} :	Sum of the weighted differences across classes
$\overline{F(X_t)^1}$:	Average operation of the 1st order features with the target domain X_t
SWD_{within} :	Sum of the weighted differences within classes
$\overline{F(X_t)^2}$:	Average operation of the 2nd order features with the target domain X_t
TD:	Target dataset
G:	Generator
U :	Left singular vector
GAN:	Generative adversarial network
V :	Right singular vector
H :	Some labels of i
X :	Data distribution
ID-MGAN _s :	Intermediate domain based on the source domain using MGAN
z :	Latent variable
ID-MGAN _t :	Intermediate domain based on the target domain using MGAN

α_i :	Normalized weight ($\sum_i \alpha_i = 1$) of the source domain s_i
L :	Label for the final model
β_i :	Hyperparameter to control the generalization error of M_i
LCD:	Lung cancer detection
γ_i :	Hyperparameter to control the regularization of the samples in component C_i
M2DW:	Modified 2D dynamic warping
δ_{T_i} :	Loss function to predict a sample in D_T
M_i :	Mahalanobis distance of component C_i
ρ :	Hyperparameter to control the strength of penalization
MGAN:	Modified generative adversarial network
Σ :	Singular value matrix of F .

Data Availability

The data used to support the study are included in the paper.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] K. Ferlay, *Global Cancer Observatory: Cancer Today*, International Agency for Research on Cancer, Lyon, France, 2020.
- [2] World Health Organization, *Global Strategy on Human Resources for Health: Workforce 2030*, World Health Organization, Geneva, Switzerland, 2016.
- [3] K. Shankar, E. Perumal, M. Elhoseny, F. Taher, B. B. Gupta, and A. A. A. El-Latif, "Synergic deep learning for smart Health diagnosis of COVID-19 for connected living and smart cities," *ACM Transactions on Internet Technology*, vol. 22, no. 3, pp. 1–14, 2021.
- [4] B. N. Zamora-Mendoza, H. Sandoval-Flores, M. Rodríguez-Aguilar et al., "Determination of global chemical patterns in exhaled breath for the discrimination of lung damage in postCOVID patients using olfactory technology," *Talanta*, vol. 256, Article ID 124299, 2023.
- [5] N. Wijbenga, R. A. Hoek, B. J. Mathot et al., "Diagnostic performance of electronic nose technology in chronic lung allograft dysfunction," *The Journal of Heart and Lung Transplantation*, vol. 42, no. 2, pp. 236–245, 2023.
- [6] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [7] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, no. 1, pp. 9–40, 2016.
- [8] F. Zhuang, Z. Qi, K. Duan et al., "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2021.
- [9] A. Abdellatif, H. Abdellatif, J. Kanesan, C. O. Chow, J. H. Chuah, and H. M. Ghenni, "An effective heart disease detection and severity level classification model using machine learning and hyperparameter optimization methods," *IEEE Access*, vol. 10, pp. 79974–79985, 2022.
- [10] G. Suganya, M. Premalatha, S. Geetha, G. J. Chowdary, and S. Kadry, "Detection of COVID-19 cases from chest X-rays

- using deep learning feature extractor and multilevel voting classifier,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 30, no. 05, pp. 773–793, 2022.
- [11] J. W. Son, J. Y. Hong, Y. Kim et al., “How many private data are needed for deep learning in lung nodule detection on CT scans? A retrospective multicenter study,” *Cancers*, vol. 14, no. 13, pp. 3174–3219, 2022.
- [12] F. Shi, B. Chen, Q. Cao et al., “Semi-supervised deep transfer learning for benign-malignant diagnosis of pulmonary nodules in chest CT images,” *IEEE Transactions on Medical Imaging*, vol. 41, no. 4, pp. 771–781, 2022.
- [13] R. V. M. da Nóbrega, P. P. Rebouças Filho, M. B. Rodrigues, S. P. P. da Silva, C. M. J. M. Dourado Júnior, and V. H. C. de Albuquerque, “Lung nodule malignancy classification in chest computed tomography images using transfer learning and convolutional neural networks,” *Neural Computing & Applications*, vol. 32, no. 15, pp. 11065–11082, 2020.
- [14] P. Yadlapalli, D. Bhavana, and S. Gunnam, “Intelligent classification of lung malignancies using deep learning techniques,” *International Journal of Intelligent Computing and Cybernetics*, vol. 15, no. 3, pp. 345–362, 2022.
- [15] M. Humayun, R. Sujatha, S. N. Almuayqil, and N. Z. Jhanjhi, “A transfer learning approach with a convolutional neural network for the classification of lung carcinoma,” *Healthcare*, vol. 10, no. 6, pp. 1058–1115, 2022.
- [16] Y. Sasaki, Y. Kondo, T. Aoki, N. Koizumi, T. Ozaki, and H. Seki, “Use of deep learning to predict postoperative recurrence of lung adenocarcinoma from preoperative CT,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 17, no. 9, pp. 1651–1661, 2022.
- [17] R. McConnell, R. Kwok, J. C. Curlander, W. Kober, and S. S. Pang, “Psi-s correlation and dynamic time warping: two methods for tracking ice floes in SAR images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 29, no. 6, pp. 1004–1012, 1991.
- [18] A. Meiseles and L. Rokach, “Source model selection for deep learning in the time series domain,” *IEEE Access*, vol. 8, pp. 6190–6200, 2020.
- [19] W. Zhang, L. Deng, L. Zhang, and D. Wu, “A survey on negative transfer,” *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 2, pp. 305–329, 2023.
- [20] J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye, “A review on generative adversarial networks: algorithms, theory, and applications,” *IEEE Transactions on Knowledge and Data Engineering*, p. 1, 2022.
- [21] A. Odena, C. Olah, and J. Shlens, “Conditional image synthesis with auxiliary classifier GANs,” in *Proceedings of the 34th International Conference on Machine Learning*, pp. 2642–2651, JMLR, Sydney, Australia, August 2017.
- [22] M. Mirza and S. Osindero, “Conditional generative adversarial nets,” 2014, <https://arxiv.org/abs/1411.1784>.
- [23] H. J. W. L. Aerts, E. R. Velazquez, R. T. H. Leijenaar et al., “Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach,” *Nature Communications*, vol. 5, no. 1, pp. 4006–4009, 2014.
- [24] H. J. Aerts, E. R. Velazquez, R. T. Leijenaar et al., “Corrigendum: decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach,” *Nature Communications*, vol. 5, pp. 1–9, 2014.
- [25] S. G. Armato, L. Hadjiiski, G. D. Tourassi et al., “LUNGx Challenge for computerized lung nodule classification: reflections and lessons learned,” *Journal of Medical Imaging*, vol. 2, no. 2, pp. 020103–020104, 2015.
- [26] O. Grove, A. E. Berglund, M. B. Schabath et al., “Quantitative computed tomographic descriptors associate tumor shape complexity and intratumor heterogeneity with prognosis in lung adenocarcinoma,” *PLoS One*, vol. 10, no. 3, pp. 01182611–e118314, 2015.
- [27] J. Yang, H. Veeraraghavan, S. G. Armato et al., “Auto-segmentation for thoracic radiation treatment planning: a grand challenge at AAPM 2017,” *Medical Physics*, vol. 45, no. 10, pp. 4568–4581, 2018.
- [28] A. Krizhevsky and G. Hinton, *Learning Multiple Layers of Features from Tiny Images*, University of Toronto, Toronto, Canada, 2019.
- [29] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, “Imagenet: a large-scale hierarchical image database,” in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, IEEE, Miami, FL, USA, June 2009.
- [30] T. Y. Lin, M. Maire, S. Belongie et al., “Microsoft coco: common objects in context,” in *Proceedings of the European Conference on Computer Vision*, pp. 740–755, Springer, Zurich, Switzerland, September 2014.
- [31] K. A. Kurdziel, J. H. Shih, A. B. Apolo et al., “The kinetics and reproducibility of 18F-sodium fluoride for oncology using current PET camera technology,” *Journal of Nuclear Medicine*, vol. 53, no. 8, pp. 1175–1184, 2012.
- [32] X. Li, R. G. Abramson, L. R. Arlinghaus et al., “Multi-parametric magnetic resonance imaging for predicting pathological response after the first cycle of neoadjuvant chemotherapy in breast cancer,” *Investigative Radiology*, vol. 50, no. 4, pp. 195–204, 2015.
- [33] K. T. Chui, B. B. Gupta, H. R. Chi et al., “Transfer learning-based multi-scale denoising convolutional neural network for prostate cancer detection,” *Cancers*, vol. 14, no. 15, p. 3687, 2022.
- [34] D. Azar, R. Moussa, and G. Jreij, “A comparative study of nine machine learning techniques used for the prediction of diseases,” *International Journal of Artificial Intelligence*, vol. 16, no. 2, pp. 25–40, 2018.
- [35] A. Gaurav, B. B. Gupta, and P. K. Panigrahi, “A comprehensive survey on machine learning approaches for malware detection in IoT-based enterprise information system,” *Enterprise Information Systems*, pp. 1–25, 2022.
- [36] M. Kumar, S. Singhal, S. Shekhar, B. Sharma, and G. Srivastava, “Optimized stacking ensemble learning model for breast cancer detection and classification using machine learning,” *Sustainability*, vol. 14, no. 21, Article ID 13998, 2022.
- [37] J. Wu, S. Guo, H. Huang, W. Liu, and Y. Xiang, “Information and communications technologies for sustainable development goals: state-of-the-art, needs and perspectives,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2389–2406, 2018.
- [38] J. Wu, S. Guo, J. Li, and D. Zeng, “Big data meet green challenges: big data toward green applications,” *IEEE Systems Journal*, vol. 10, no. 3, pp. 888–900, 2016.
- [39] R. Atat, L. Liu, J. Wu, G. Li, C. Ye, and Y. Yang, “Big data meet cyber-physical systems: a panoramic survey,” *IEEE Access*, vol. 6, pp. 73603–73636, 2018.
- [40] A. Almomani, M. Alauthman, M. T. Shatnawi et al., “BB Phishing website detection with semantic features based on machine learning classifiers: a comparative study,” *International Journal on Semantic Web and Information Systems*, vol. 18, no. 1, pp. 1–24, 2022.

- [41] I. Cvitić, D. Peraković, M. Periša, and B. Gupta, "Ensemble machine learning approach for classification of IoT devices in smart home," *International Journal of Machine Learning and Cybernetics*, vol. 12, no. 11, pp. 3179–3202, 2021.
- [42] F. Gerges, F. Shih, and D. Azar, "Automated diagnosis of acne and rosacea using convolution neural networks," in *Proceedings of the 2021 4th International Conference on Artificial Intelligence and Pattern Recognition*, pp. 607–613, Xiamen, China, September 2021.
- [43] H. Guan and M. Liu, "Domain adaptation for medical image analysis: a survey," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 3, pp. 1173–1185, 2022.
- [44] S. Wang, L. Dong, X. Wang, and X. Wang, "Classification of pathological types of lung cancer from CT images by deep residual neural networks with transfer learning strategy," *Open Medicine*, vol. 15, no. 1, pp. 190–197, 2020.
- [45] H. Tan, J. H. T. Bates, and C. Matthew Kinsey, "Discriminating TB lung nodules from early lung cancers using deep learning," *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, pp. 161–167, 2022.
- [46] Z. H. Zhan, L. Shi, K. C. Tan, and J. Zhang, "A survey on evolutionary computation for complex continuous optimization," *Artificial Intelligence Review*, vol. 55, no. 1, pp. 59–110, 2022.
- [47] R. Abu Khurma, I. Aljarah, A. Sharieh, M. Abd Elaziz, R. Damaševičius, and T. Krilavičius, "A review of the modification strategies of the nature inspired algorithms for feature selection problem," *Mathematics*, vol. 10, no. 3, p. 464, 2022.
- [48] J. Manhas, R. K. Gupta, and P. P. Roy, "A review on automated cancer detection in medical images using machine learning and deep learning based computational techniques: challenges and opportunities," *Archives of Computational Methods in Engineering*, vol. 29, no. 5, pp. 2893–2933, 2022.
- [49] V. R. Gottumukkala, N. Kumaran, and V. C. Sekhar, "BLSNet: skin lesion detection and classification using broad learning system with incremental learning algorithm," *Expert Systems*, vol. 39, pp. 1–16, 2022.