Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-021-02989-1

ORIGINAL RESEARCH

Multistage transfer learning technique for classifying rare medical datasets

R. V. Aswiga[1] · R. Aishwarya[1] · A. P. Shanthi[1]

Received: 12 October 2020 / Accepted: 1 March 2021

Abstract
It is said that about 8% of the people across the world are impacted by different kinds of rare diseases. Identifying such rare diseases accurately is a challenging task, as they exhibit common symptoms that may be incorrectly recognized as a common disease. As a result, the treatment is insufficient and by the time a diagnosis is made, it may be too late for the patient to survive. Therefore, any attempt towards early diagnosis of rare diseases becomes the need of the hour and several researchers are concentrating on machine learning techniques to do the same. This work proposes one such approach to diagnose rare disorders. A multi-level transfer learning (MLTL) framework with three models is designed to detect and classify rare diseases with very limited datasets, using the knowledge acquired from easily available datasets. The first model, for the source domain, uses the abundantly available non-medical images and learns the generalized features. The acquired knowledge is then transferred to the second model, for the intermediate and auxiliary domain, which is again commonly available and related to the target domain, and helps in learning the auxiliary or intermediate task. This information is then used to classify the final target domain, which consists of medical datasets that are very scarce. Experimental results with non-medical and the auxiliary CT abdomen images show good classification accuracies for identifying rare diseases related to the liver and kidneys. An area under receiver operating characteristic (ROC) curve of 0.90 and 0.89 are achieved for two different rare diseases, with just 2.08% of the source domain dataset, 6.6% of the intermediate domain dataset and less than 10% of the rare target domain dataset, when compared to the work reported in the literature.

Keywords Multistage transfer learning · Multitask learning · Deep learning · Medical imaging · Machine learning

1 Introduction

Accurate diagnosis of rare diseases is an important and challenging task. The symptoms of such diseases often appear unfamiliar and atypical to a clinician, because of the rarity of such cases. Machine assisted diagnostic methods (Zou et al. 2019; Li et al. 2019) can support the clinicians in such cases, and in recent times, machine learning approaches are gaining momentum. Such rare disease diagnosis is challenging to machine learning approaches as well. Machine learning algorithms often require a significant number of training examples to achieve a good generalization performance. However, by the very nature of rare diseases, the number of such datasets is going to be very scarce.

Transfer learning (Singh and Kisku 2018) is a deep learning concept generally used in applications that suffer from insufficient training data. Inspired from the human's ability to transfer the acquired knowledge across different tasks, transfer learning has emerged as an upcoming area of research. In transfer learning, the knowledge acquired by the model while solving a particular task, is used later to solve another new task by transferring the acquired knowledge in the form of weights of the model. The effect of knowledge transferred depends on the relatedness between the two tasks. The pre-trained features may not be sufficiently fine-tuned according to the target task, when the available training data of the target domain are limited, and the source and target domains are unrelated. This problem is overcome

✉ R. V. Aswiga
aswiga91@gmail.com

R. Aishwarya
aishwarya.rsv@gmail.com

A. P. Shanthi
shanthiap@gmail.com

[1] Department of Computer Science and Engineering, College of Engineering, Guindy, Anna University, Chennai, Tamilnadu, India

Published online: 08 March 2021

using multi-level transfer learning by adding intermediate stages of transfer learning from related auxiliary domains.

This work applies the concept of multistage transfer learning to the classification of rare diseases that have very limited datasets. The easily available non-medical images dataset, consisting of birds, cars, dogs, flowers and humans, is used as the source domain. Computed tomography (CT) abdomen images, which are commonly available and related to the target domain, are used as the intermediate domain and the limited datasets of rare diseases is used as the target domain.

This paper includes datasets related to two rare diseases collected from hospitals, viz. liver caroli syndrome and renal cell carcinoma syndrome. Identifying such liver and kidney disorders is challenging because of the shape variation, tissue linkage, and gray scale similarity and due to the overlapping of different organs. Caroli disease is a very rare inherited disease, which is caused due to the formation of cystic dilatation in the bile ducts within the liver. This disease is broadly classified into two types—focal caroli disease and caroli syndrome. The former is a simple disorder that affects the isolated left portion of the liver, containing abnormally widened bile ducts, and the latter is more diffuse, which leads to liver failure or polycystic kidney disease. Detecting this rare caroli syndrome associated with hepatic fibrosis and hypertension at the early stage can reduce the threat imposed on the quality of a patient's life to some extent.

In addition, scientists are not certain about what exactly causes Renal Cell Carcinoma (RCC), especially between ages 50 and 70. There are several subtypes of RCC, which are very rare as shown in Fig. 1, each making up less than 1% of patients having this kind of cancer. Among the rare RCCs, the proposed work includes neuroblastoma-associated carcinoma dataset, which belongs to the class of neuroendocrine tumour. Radiation therapy is the only treatment available when diagnosed at an advanced stage, which leads to various side effects, including skin problems, fatigue, diarrhoea, etc. Although it is a serious disease, finding and treating it earlier makes it more likely that the patient can be cured.

Therefore, diagnosing such rare diseases at the earliest is important and essential in the medical field. However, classifying these diseases is challenging and tricky because of the limited instances. Therefore, there is a need for computer-aided diagnosis, to aid the physicians in spotting such rare diseases.

This work proposes a multistage deep convolutional neural network (CNN) for the classification of rare medical images, based on transfer learning. The proposed work consists of a multi-level transfer learning (MLTL) framework, based on the work proposed (Samala et al. 2019). This MLTL framework consists of the source, intermediate and target domains. A widely available non-medical dataset is used as the source domain and the basic common features are transferred to the intermediate domain, consisting of the CT abdomen dataset. Thereafter, the features of the CT abdomen dataset are transferred to the target domain, consisting of a rare CT liver and renal cell carcinoma dataset and classification is performed. The results obtained are

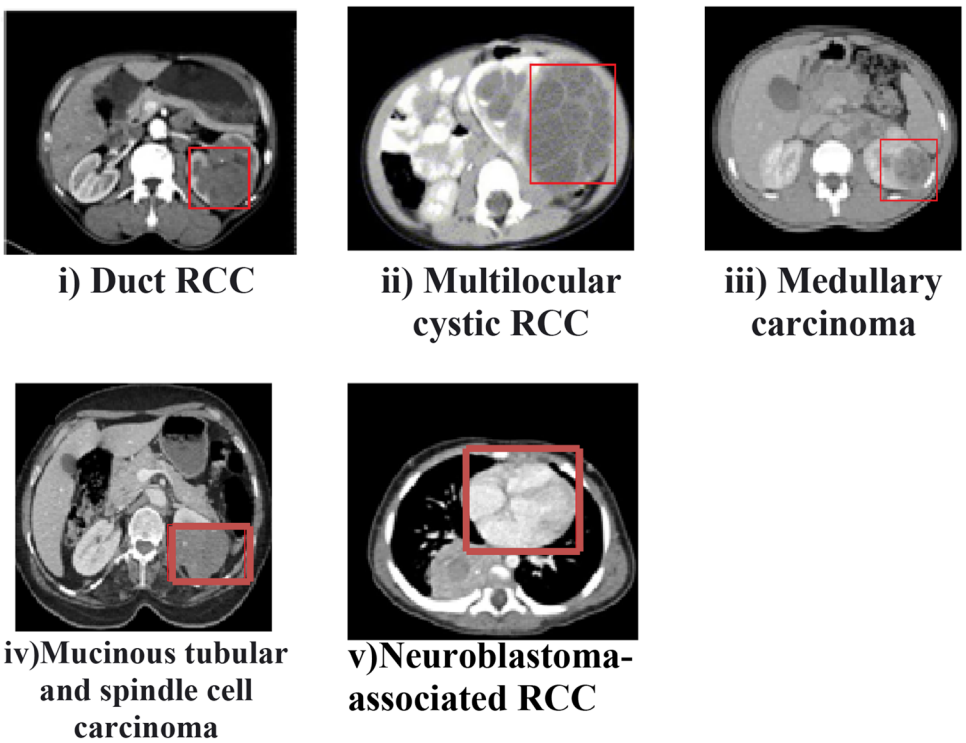


Fig. 1 Rare types of renal cell carcinoma

promising, paving the way of classifying rare diseases with limited datasets.

The major contribution of the proposed work is to build a MLTL framework that consists of:

(i) The generalized feature extraction (GFE) network to train the source domain, consisting of the non-medical images that are easily available.
(ii) The Intertune MedCNN to train the intermediate domain, consisting of CT abdomen images.
(iii) The target Finetune MedCNN to classify rare liver caroli syndrome and complex renal cell carcinoma diseases.

The existing work available in literature, related to the proposed work is presented below.

2 Related work

In recent years, handling of data in Biomedicine research is eternally increasing. Artificial intelligence (AI) has been successfully applied to the fields of pharma and medicine to handle rare diseases. (Garg et al. 2016) automated the process of identifying a kind of rare disease, Cardiac Amyloidosis, from EHRs by using an ensemble machine learning classifier and experimented on 73 positive (cardiac amyloidosis) and 197 negative instances and achieved F1 score of 0.98.

Li et al. (2019) have improved rare disease classification using imperfect knowledge graph technique. The proposed text classification algorithm delivers robust performance gain when evaluated on two datasets, containing more than one lakh patient records, categorized into 849 diseases. (Singh and Kisku 2018) have proposed an approach for detecting rare genetic diseases based on 2D facial images. The proposed model is trained on the dataset consisting of 1772 facial images from 12 classes of genetic syndrome and provides performance comparable to state of the art methods. Iadanza et al. (2020) have developed a novel application using Chromatic Pupillometry device to automatically detect genetic diseases in pediatric age, in order to support diagnosis of inherited retinal diseases in pediatric subjects.

Fernandes et al. (2019) have proposed gait classification of patients with Fabry's disease (FD) using multiple regression models, achieving an FD classification accuracy of 78.21%. Sun and Wang (2019) adapt four machine learning algorithms to identify people with rare diseases and according to the summary of those misclassification rates, neural network has the least error rate to classify rare diseases.

However, for several rare diseases, fundamental information such as the cause of the disease and epidemiological data is limited or not available. This lack of chronological data for rare diseases poses a great challenge to statistical machine learning based approaches. But, the advances in machine learning techniques (Ganguly et al. 2019; Chang et al. 2017) including multitask learning (Todoroki et al. 2019) and various deep neural networks (Romero et al. 2019; Chen et al. 2019a, b; Chen et al. 2019a, b; Yusuke et al. 2020), have demonstrated their effectiveness in many applications involving CT medical images.

Commandeur et al. (2018) have proposed two multi-task deep learning networks to detect cardiovascular risk in coronary calcium CT datasets. The automated multi-task CNN adopts shape regularization and median filtering technique to segment and classify epicardial and thoracic adipose tissues. Chen et al. (2017) have proposed three multitask learning schemes based on linear regression and random forest algorithm to control the heterogeneous features derived from deep learning models. These models were trained on 2400 CT lung nodules and gives good regression results for better diagnosis and decision support.

Tan et al. (2018) have proposed a multi-task CNN model that adopts distance regression and fusion scheme method to learn task-specific features and has been tested across 6000 2D CT kidney and MRI datasets. Bienias et al. (2019) have built a two-stage CNN model to segment Adenocarcinoma Gland instances, based on a multi-task learning method. The network has been trained on 2015 MICCA1 CT gland challenge dataset and gives good F1 score and dice index measure.

Vlachostergiou et al. (2018) have exploited the use of multitask learning for classifying long term Parkinson's disease. Yu et al. (2020) have developed adversarial reverse mapping multi-task learning framework for exploring task dependencies. The model has been trained on 2900 MRI images and later has been finetuned to run on 2360 CT cardiac images to map the indices to cardiac images.

Also, the effectiveness of multi-task learning (Farhadi and Foruzan 2019) and deep neural networks (Brunetti et al. 2019; Lakshmi Priya et al. 2019) has been demonstrated in several applications involving liver. For example, segmenting liver tumors (Alahmera and Ahmed 2016; Ben-Cohen et al. 2018; Nasiri et al. 2019; Bai et al. 2019) and classifying liver cancer (Al Sadeque et al. 2019) etc. Recently, liver segmentation (Hiraman et al. 2019; Hoang et al. 2019; Cheema et al. 2019; Czipczer and Manno-Kovacs 2019; Song et al. 2019) has attracted a lot of attention in the medical community. Accurate tumor segmentation is a challenging task because of the tumor shape, size, and location.

In order to address the limited availability of training samples, which is a common problem in medical imaging, (Hussein et al. 2019) have proposed an unsupervised transfer learning based 3D CNN model trained on 1018 CT and 171 MRI images to characterize tumors. This model incorporates GIST descriptors to identify the difference between lung and pancreas cyst. The techniques used in Laouid et al. (2019)

and Yousaf et al. (2019) can be adopted to extract most relevant features when collecting clinical data from multiple IoT devices to facilitate safe online transfer learning.

However, there is not much work reported in the literature with regard to classifying rare diseases with limited datasets. This work addresses this problem and proposes a MLTL framework, with three deep CNN models for classifying rare diseases like liver caroli syndrome and renal cell carcinoma.

The rest of the paper is organized as follows. Section 2 presents the proposed MLTL framework and its details. The experimental results and discussions are presented in Sects. 3 and 4.

3 Materials and methods

The architecture of the proposed basic MLTL framework is shown in Fig. 2. The MLTL framework consists of three deep CNN models. The first model, the generalized feature extraction (GFE) network, corresponds to the source domain, which consists of non-medical images that are easily available. The Fine-tune MedCNN model corresponds to the target domain, trained individually on two rare datasets, namely, liver caroli syndrome and renal cell carcinoma. Since the source and target are unrelated domains, an intermediate Intertune MedCNN model is introduced as proposed in Samala et al. (2019). The generalized features learned from the GFE network are transferred to the intermediate domain and the specific features of the medical images learned from the intermediate CT abdomen are transferred to the target domain to give good classification results.

3.1 GFE network model

The generalized feature extraction neural network is constructed in two different structures. The GFE CNN is trained with non-medical images. The GFE Structure-I has two convolutional layers, with each layer followed by a max-pooling layer, one dense layer and then a classification layer. The GFE network Structure-II consists of five convolutional layers, with each convolutional layer followed by a max-pooling layer, and two dense layers followed by a softmax classifier layer. The input images are resized and the convolutional layers extract the features and the feature map is produced as the output. Dimensionality reduction is done using the max-pooling layer and the sub-feature map is generated as the output. The convolutional layer function is described with the help of the following equation:

$$y_k^n = h(b_k^n + \sum w_{x,k}^{n-1} * m_x^n) \tag{1}$$

where n is the number of layers in the coding network. h is the activation function. m_x is the input feature map. y_k is the output feature map. $w_{n,k}$ is the weight kernel for convolutional layers. * denotes convolution operation and b_k is the bias of the coding network.

Max pooling operation is carried out in a window size of 5×5 dimensions with 2 strides. The pooling operation is described by

$$y_{k,r}^x = \max_{o \le a,b < 5} (m_{k.5+a,k.5+b}^x) \tag{2}$$

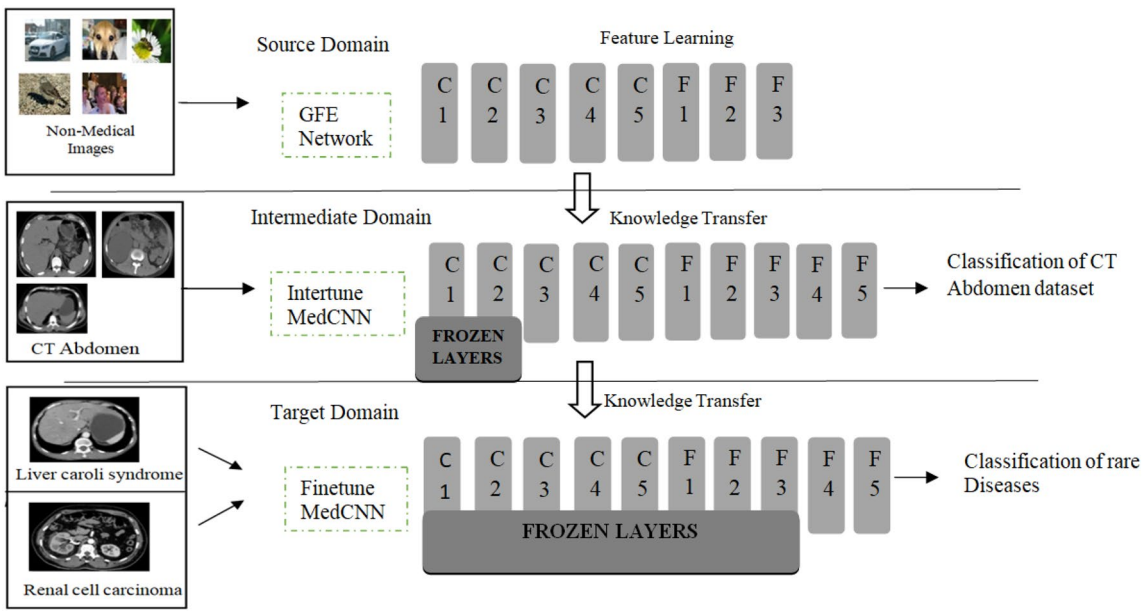


Fig. 2 MLTL framework

where y^x is the output feature map obtained by the 5×5 dimension of pooling layers. There are two activation functions widely used, the sigmoid function and the tanh activation function. Both the functions suffer from gradient diffusion problem and low convergence rate. Therefore, to make the intermediate network efficient, this work uses Rectified Linear Unit, ReLU. The GFE network specification of the two structures is shown in Table 1a and b. The features that are learned from the non-medical images are transferred to the intermediate domain for classification.

Among the two different GFE structures, the one which gives lower loss and high accuracy is chosen for performing further processing.

3.2 Intertune MedCNN model

In the proposed work, the available dataset in the target domain is limited and the source and target distributions are different. Therefore, the learned features from the source domain are not sufficient to fine-tune according to the target task. Therefore, an intermediate domain with CT abdomen dataset, which is related to the target domain, is used to improve the feature learning and the classification of the target task. The intertune MedCNN acts as an auxiliary intermediate domain.

Table 1 (a) GFE CNN-1 Network Specification, (b) GFE CNN-2 Network Specification

Layer type	Size/stride	Output dimension
(a)		
GFE CNN structure-I		
Convolution	3×3/1	148×148×16
Max pool	2×2/1	74×74×16
Convolution	3×3/1	72×72×64
Max pool	2×2/1	36×36×64
Flatten		82,944
Dense		512
(b)		
GFE CNN structure-II		
Convolution	11×11/1	140×140×32
Max pool	2×2/1	70×70×32
Convolution	5×5/1	66×66×64
Max pool	2×2/1	33×33×64
Convolution	3×3/1	31×31×128
Convolution	3×3/1	29×29×192
Convolution	3×3/1	27×27×256
Max pool	2×2/1	13×13×64
Flatten		43,294
Dense		384
Dense		512

The medical images of CT abdomen consisting of several organs like spleen, ISOP food pipe, artery vessel, intestine, gallbladder, uterus, left kidney, right kidney, liver, and pancreas are used as the intermediate domain data. The CT abdomen images consisting of all these organs are classified into three classes, namely, normal, cyst, and calculi. The CT abdomen images dataset is classified with the various parameter transfer techniques using pre-trained GFE CNN. This is designed by finetuning the GFE network and adding new layers according to the CT abdomen dataset. In the CNN, the layers near the input are generic and the layers towards the output are specific to the input. Hence, the layers near the input are frozen and the other layers are fine-tuned according to the CT abdomen dataset. Two dense layers, which correlate the high-level features more strongly to a particular class, are added. The regularization is done using drop out strategy, which introduces randomness in the network. This is done by randomly selecting some nodes and removing some nodes for the incoming and outgoing connections, in order to avoid over fitting. The classifier layer is modified, as the number of output classes is different in the source and intermediate domains.

Four different types of intermediate structures are designed based on parameter transfer techniques by—freezing all the layers (C1-F3), freezing the first convolutional layer (C1), freezing the first two convolutional layers (C1–C2) and freezing the first three convolutional layers (C1–C3), respectively. This work is compared to the model proposed by Samala et al. (2019) which used ImageNet trained Alexnet CNN structure as the source network. The one that gives the maximum area under receiver operating characteristic (RCC) curve is chosen for classification of the target domain. It consists of five convolutional layers and three fully connected layers, connected with max pooling and normalization layers. The specific features of the CT abdomen dataset are learned and the Optimal Intertune MedCNN transfer network is chosen for the next stage of transfer learning.

3.3 Finetune MedCNN model

The target fine-tune MedCNN is designed from the intertune MedCNN by restoring and fine-tuning, according to the rare liver caroli syndrome and renal cell carcinoma dataset. The learned knowledge from the intertune MedCNN with the CT abdomen dataset is transferred to train the target fine-tune MedCNN for classifying the two different rare diseases. The related auxiliary information from the intermediate domain, consisting of the CT abdomen dataset, is expected to improve the performance of the target domain significantly.

4 Results

4.1 Dataset

The source domain dataset consists of 25,000 images from five different classes such as birds, cars, dogs, flowers, and humans. These images are a part of the ImageNet dataset. The database is publicly available at http://www.image-net.org/ and the intermediate domain dataset consists of 1303 full CT abdomen images with three different classes. The target domain dataset consists of 120 images of the rare CT liver disorder and another 250 images of renal cell carcinoma disease. An example of rare liver caroli syndrome and renal cell carcinoma, which is used for training the proposed target MedCNN, is shown in Fig. 3a, b, respectively.

These are real-time images collected from Saveetha Institute of Medical and Technical Sciences of India. The details of the datasets used are summarized in Table 2.

Table 2 Summary of Dataset

	Dataset name	No. of. instances
Source domain	Imagenet	25,000
Intermediate domain	Full CT abdomen	1303
Target domain 1	CT liver caroli Syndrome	120
Target domain 2	CT renal cell carcinoma	250

4.2 Evaluation metrics

The metrics chosen for evaluating the proposed MLTL framework using real-time CT images is the 'area under receiver operating characteristic (ROC) curve' and the 'Accuracy (ACC)'. This paper deals with a medical dataset, which is extremely imbalanced. Hence, the area under ROC curve and ACC is best applicable, compared to other metrics such as precision, recall, and f-1 score. Since it is the medical dataset, the main goal is to differentiate the people identified with caroli syndrome and positive carcinoma cases from normal cases. The output from the softmax layer of the fine-tune MedCNN is used as the input for ROC curve analysis. Higher the area under ROC curve, the model is better in distinguishing between patients with and without the disease. ROC curve depicts true positive rate on the Y-axis, and false positive rate on the X-axis.

$$\text{True Positive Rate} = \text{TP}/((\text{TP} + \text{FN})) \tag{3}$$

$$\text{FalsePositive Rate} = \text{FP}/((\text{TN} + \text{FP})) \tag{4}$$

Using Eqs. 3 and 4, True Positive Rate gives the measure of the ratio of the images correctly classified as positive with that of the total predicted positive images. False Positive

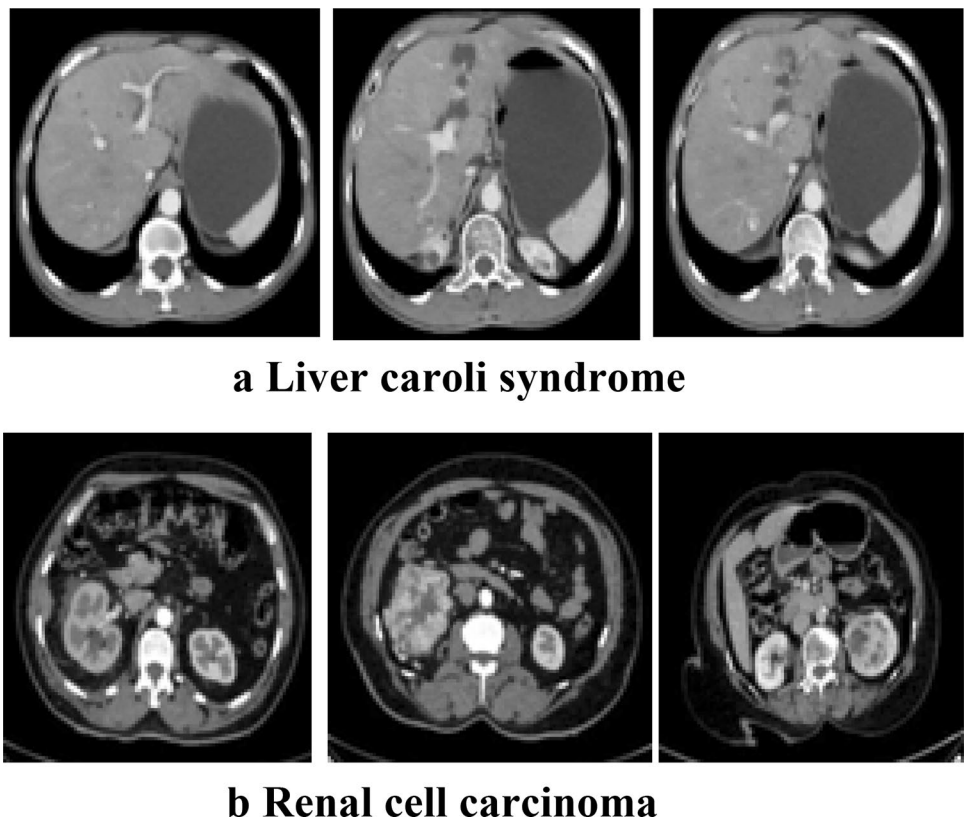


Fig. 3 **a** Liver caroli syndrome. **b** Renal cell carcinoma

Rate gives the measure of the ratio of the images wrongly classified as positive with that of the total actual negative images. All three networks discussed in the previous section are trained with mini-batch stochastic gradient descent optimization using a batch size of 64 on a GeForce GTX TITAN X GPU, with an initial learning rate of 0.01 and 1500 epochs.

4.3 GFE network

The 25,000 non-medical images are used for feature learning in the source domain. The source domain GFE network is designed with two different network structures and their performances are evaluated. The first network structure consists of two convolutional layers each, followed by max-pooling layer and two dense layers. This network yields a training accuracy of 0.75 and is demonstrated in Fig. 4a.

The second network structure consists of five convolutional layers each, followed by max-pooling layer and three dense layers. This network, after adequate training gives 0.90 training accuracy as shown in Fig. 4b.

As shown in Table 3, the second network structure provides higher training accuracy, and hence this network structure is selected to perform multi-level transfer learning.

4.4 Intertune MedCNN

The full CT abdomen dataset with 1303 images is used for intermediate domain feature learning and classification. The pre-trained GFE network using non-medical images is used as the basis for the development of the intertune MedCNN. The area under ROC curve of the intertune MedCNN is depicted in Fig. 5.

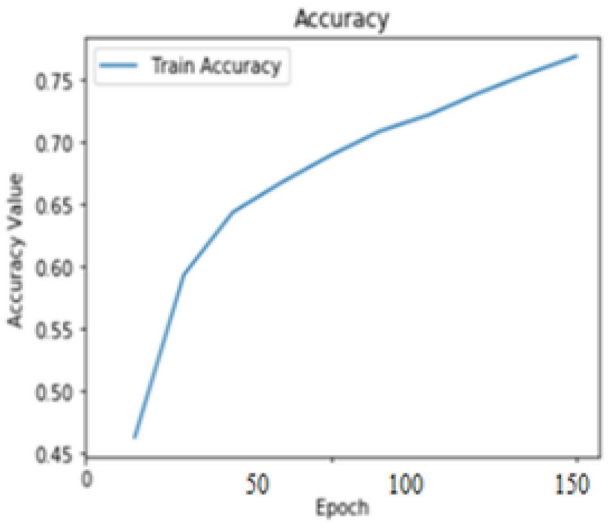
Four different network structures are evolved by freezing the first two convolutional layers (C1–C2) as shown in Fig. 5a, freezing the first convolutional layer (C1) as given in Fig. 5b, freezing all the layers (C1-F3) as in Fig. 5c, and freezing the first three convolutional layers (C1–C3) as shown in Fig. 5d, respectively.

The area under ROC curve obtained from all four network structures are 0.74, 0.77, 0.83 and 0.86, respectively.

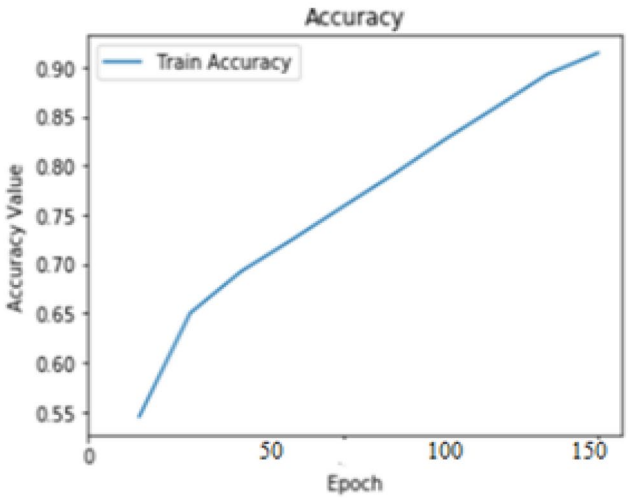
The network structure having a higher area under ROC curve of 0.86, obtained by freezing the first two convolutional layers, is selected to perform transfer learning and to design the target fine-tune MedCNN.

4.5 Finetune MedCNN

The target domains fine-tune MedCNN trained on rare liver caroli syndrome gives 0.90 area under ROC curve, as shown in Fig. 6a.



a GFE CNN Structure-1



b GFE CNN Structure-2

Fig. 4 **a** GFE CNN structure-1. **b** GFE CNN structure-2

Table 3 Source domain network performance

Network	Accuracy
GFE CNN structure-I	0.75
GFE CNN structure-II	**0.90**

The fine-tune MedCNN trained on renal cell carcinoma Syndrome gives 0.89 area under ROC curve, as shown in Fig. 6b. The experimental results of the proposed MLTL framework is shown in Table 4.

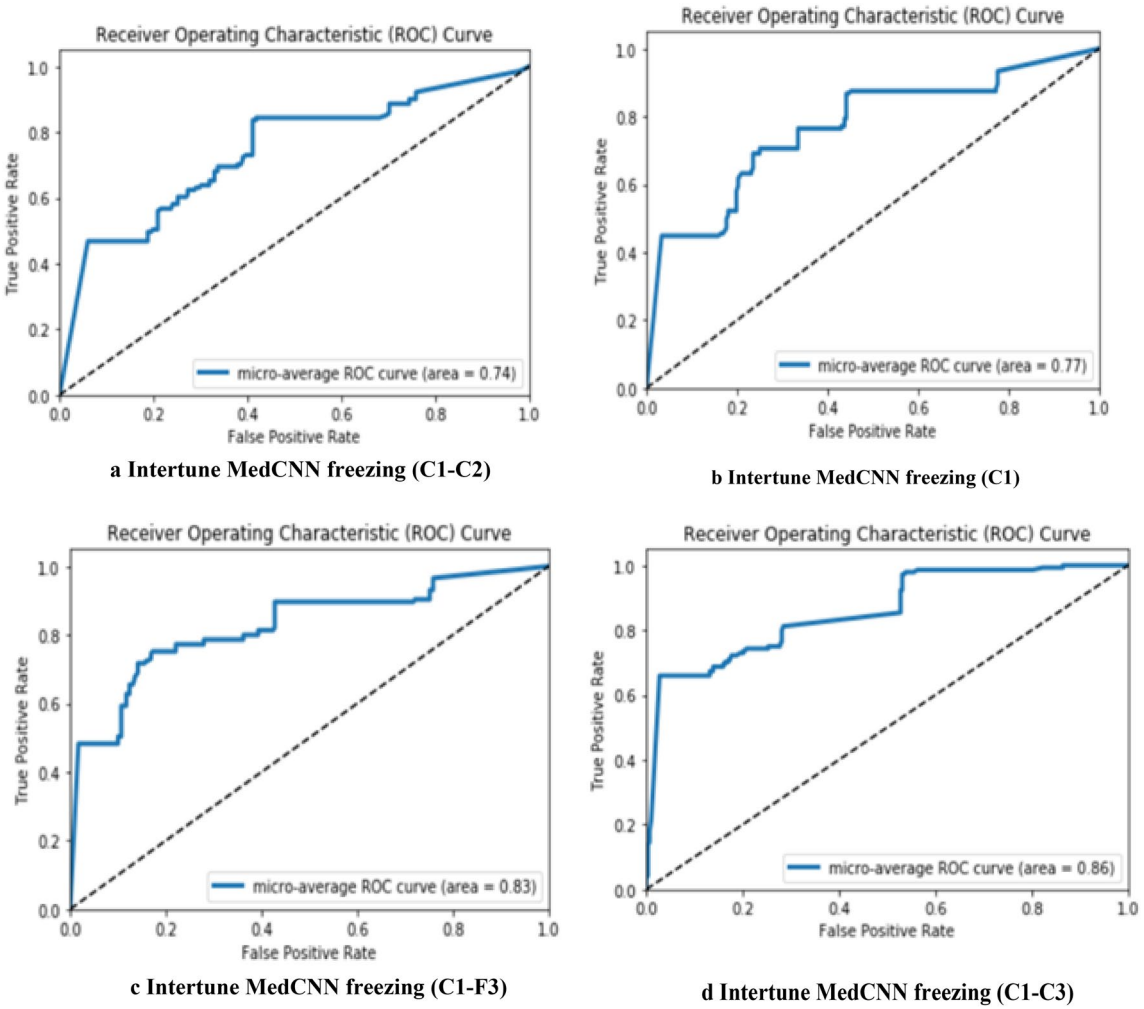


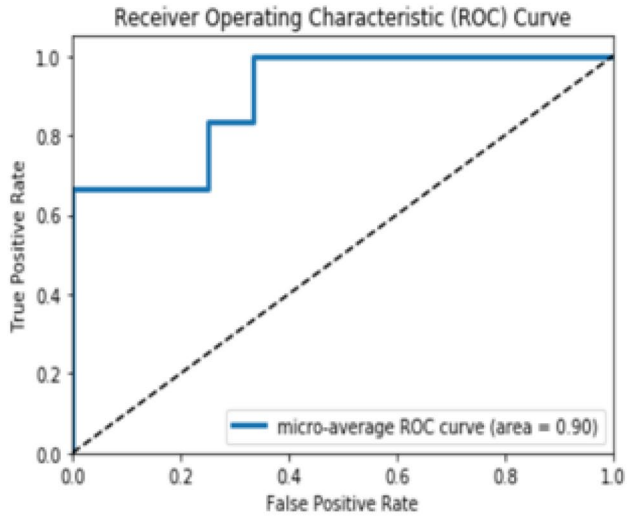
Fig. 5 **a** Intertune MedCNN freezing (C1–C2). **b** Intertune MedCNN freezing (C1). **c** Intertune MedCNN freezing (C1-F3). d Intertune MedCNN freezing (C1–C3)

5 Discussion

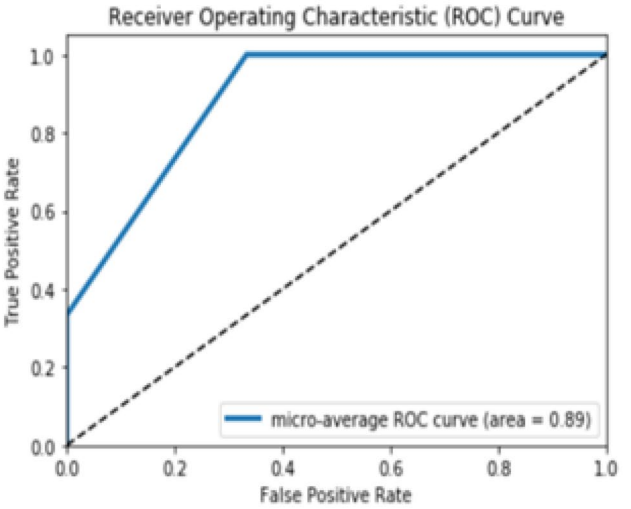
As seen from Table 4, the proposed model classifies well, even when the target domain dataset has only fewer images and is different from the source domain distribution. Table 4 provides a comparison between the existing work (Samala et al. 2019) and the proposed work. The existing work (Samala et al. 2019) reports an area under ROC curve of 0.91 with the source domain Alexnet, which is trained with ImageNet dataset consisting of 1.2 million non-medical images from 1000 classes and the intermediate domain consisting of 19,632 mammography images and the target domain consisting of 12,680 DBT images.

The proposed MLTL based framework gives better performance, with the area under ROC curve of 0.90 and 0.89 for the classification of the rare liver and kidney dataset, respectively, with just 2.08% of the source domain dataset, 6.6% of the intermediate domain dataset, and less than 10% of the rare target domain dataset.

To bring in the effectiveness of the proposed model even more clearly, the accuracy of the proposed method is compared with some of the existing classifier models, reported in literature. Table 5 provides a comparison between the proposed work and other machine learning techniques classifying different rare diseases. Singh and Kisku (2018) have reported an accuracy of 97.45% using data augmentation technique. However, it is to be noted that the proposed framework provides an accuracy of 96.82% with very limited dataset, which is better than the reported works. Thus, the proposed MLTL framework classifies rare diseases accurately, throwing open the avenues for multi-transfer learning, even in cases where there is very limited dataset available. Comparison of the proposed work with some of the work reported in literature, with reference to the direct handling of CT images, for identifying different types of abnormalities



a Target Finetune MedCNN on Liver Caroli Syndrome



b Target Finetune MedCNN on Renal Cell Carcinoma Syndrome

Fig. 6 **a** Target finetune MedCNN on liver Caroli syndrome. **b** Target finetune MedCNN on renal cell carcinoma syndrome

in the liver, is provided in Table 6. It can be seen that the proposed MLTL framework provides better performance, an accuracy of 96.82%, compared to the reported works.

6 Conclusion

This work has thus successfully implemented the MLTL framework for classifying rare liver caroli syndrome and renal cell carcinoma, with the information obtained from the general CT abdomen images and non-medical images.

Table 4 Comparison with the earlier reported work

	Source domain		Intermediate domain		Target domain 1		Performance (area under ROC curve)	Target domain 2		Performance (area under ROC curve)
	No. of instances	No. of classes	No. of instances	No. of classes	No. of instances	No. of classes		No. of instances	No. of classes	
Samala et al. (2019)	1.2 Million	1000	19,632	2	12,680	2	0.91	–	–	–
Proposed work	25,000 (2.08%)	5	1303 (6.6%)	3	120 (9.4%)	3	0.90	250	3	0.89

Table 5 Comparison with machine learning techniques classifying other rare diseases

References	Rare disease	Classification method	Accuracy
Singh and Kisku (2018)	12 rare diseases	Transfer learning	97.45 After data augmentation
Ganguly et al. (2019)	Eye Melanoma	CNN	91.76
Iadanza et al. (2020)	Genetic Diseases	SVM	84.6
Zou et al. (2019)	Wilson's disease	SVM, LDA, LR	96.1
Proposed	Liver Caroli Syndrome and Renal Cell Carcinoma	Multistage transfer learning approach	96.82 Without data augmentation

Table 6 Comparison table involving CT images

References	Lesion type	Machine learning technique	ACC
Ben-Cohen et al. (2018)	Cyst, Metastasis	CNN	75.0 ± 5.8
Chin-Chen et al. (2017)	Benign, Malignant	Computer-aided diagnosis (CAD) system	81.69
Al Sadeque et al. (2019)	HCC, Malignant	Hog-SVM model	94
Proposed	Rare diseases	MLTL Framework Without data augmentation	96.82

Experimental results on real-time medical data sets indicate that the proposed algorithms are able to achieve good performance against the compared baselines. The significant contribution of this work is the use of the intermediate domain in the transfer learning technique, which has clearly proved that it is possible to apply this framework to any rare disease domains, where data are sparse.

Future work will focus on exploring the possibility of extending this framework with new feature extraction techniques for the target domain dataset. Further, the model can be tailored to accept multi-input data of diverse distributions. The proposed method might also be extended to medical images acquired from other imaging modalities such as MRI, PET, or ultrasound.

Acknowledgements This work was supported by the Department of Radiology, Saveetha Medical College & Hospital (SMCH), and Saveetha Institute of Medical and Technical Sciences of India. We thank Ms. R. Aishwarya, PG Student, College of Engineering, Guindy, for helping us in doing the simulations.

References

Al Sadeque Z, Khan TI, Hossain QD, Turaba MY (2019) Automated detection and classification of liver cancer from CT images using HOG-SVM model. In: Conference on proceeding of the 5th international conference on advances in electrical engineering (ICAEE), Dhaka, Bangladesh, pp 21–26

Alahmera H, Ahmed A (2016) Computer-aided classification of liver lesions from CT images based on multiple ROI. Proc Comput Sci 90:80–86

Bai Z, Jiang H, Li S, Yao Y-D (2019) Liver tumor segmentation based on multi-scale candidate generation and fractal residual network. IEEE Access 7:82122–82133

Ben-Cohen A, Mechrez R, Yedidia N, Greenspan H (2018) Improving CNN training using disentanglement for liver lesion classification in CT. Computer Vision and Pattern Recognition, arXiv:1811.00501:886-889

Bienias LT, Guillamon JR, Nielsen LH, Alstrom TS (2019) Insights into the behavior of multi-task deep neural networks for medical image segmentation. In: Conference on proceedings of the 29th conference MLSP 2019, Pittsburgh, PA, USA

Brunetti A, Carnimeo L, Trotta GF, Bevilacqua V (2019) Computer-assisted frameworks for classification of liver, breast and blood neoplasias via neural networks: a survey based on medical images. Neurocomputing 335:274–298

Chang C-C, Chen H-H, Chang Y-C, Yang M-Y, Lo C-M, Ko W-C, Lee Y-F, Liu K-L, Chang R-F (2017) Computer-aided diagnosis of liver tumors on computed tomography images. Comput Methods Programs Biomed 145:45–51

Cheema MN, Nazir A, Sheng B, Li P, Qin J, Feng DD (2019) Liver extraction using residual convolution neural networks from low-dose CT images. IEEE Trans Biomed Eng 66(9):2641–2650

Chen S, Qin J, Ji X, Lei B, Wang T, Ni D, Cheng J-Z (2017) Automatic scoring of multiple semantic attributes with multi-task feature leverage: a study on pulmonary nodules in CT images. IEEE Trans Med Imaging 36(3):802–814

Chen L, Song H, Li Q, Cui Y, Yang J, Hu XT (2019) Liver segmentation in CT images using a nonlocal fully convolutional neural network. In: Conference on proceedings of the 2019 IEEE international conference on bioinformatics and biomedicine (BIBM), pp 639–642

Chen X, Zhang R, Yan P (2019) Feature fusion encoder decoder network for automatic liver lesion segmentation. In: Conference on proceedings of the IEEE 16th international symposium on biomedical imaging (ISBI 2019) Venice, Italy, pp 430–433, April 8–11, 2019

Commandeur F, Goeller M, Betancur J, Cadet S, Doris M, Chen X, Berman DS, Slomka PJ, Tamarappoo BK, Dey D (2018) Deep learning for quantification of epicardial and thoracic adipose tissue from non-contrast CT. IEEE Trans Med Imaging 37(8):1835–1846

Czipczer V, Manno-Kovacs A (2019) Automatic liver segmentation on CT images combining region-based techniques and convolutional

features. In: Conference on proceedings of the 2019 international conference on content-based multimedia indexing (CBMI) IEEE, Piscataway (NJ), pp 1–6

Farhadi M, Foruzan AM. Data augmentation of CT images of liver tumors to reconstruct super-resolution slices based on a MultiFrameApproach. In: Conference on proceedings of the 27th Iranian conference on electrical engineering, (ICEE2019), pp 1783–1786

Fernandes C, Ferreira F, Gago M (2019) Gait classification of patients with Fabry's disease based on normalized gait features obtained using multiple regression models. In: Conference on proceedings of IEEE international conference on bioinformatics and biomedicine (BIBM), ISBN 978-1-7281-1867-3, pp 2288–2295

Ganguly B, Biswas S, Ghosh S, Maiti S, Bodhak S (2020) A deep learning framework for eye melanoma detection employing convolutional neural network. In: 2019 international conference on computer, electrical & communication engineering (ICCECE), 20 February 2020 Kolkata, India, https://doi.org/10.1109/ICCECE44727.2019.9001858

Garg R, Dong S, Shah S, Jonnalagadda SR (2016) Bootstrap machine learning approach to identify rare disease patients from electronic health records. Cornell University. arXiv:1609.01586v1.2016

Hiraman A, Viriri S, Gwetu M (2019) Efficient region of interest detection for liver segmentation using 3D CT scans. In: Conference on proceedings of conference on information communications technology and society (ICTAS), ISBN 978-1-5386-7365-2

Hoang HS, Pham CP, Franklin D, van Walsum T, Luu MH (2019) An evaluation of CNN-based liver segmentation methods using multitypes of CT abdominal images from multiple medical centers. In: Conference on proceedings of the 19th international symposium on communications and information technologies (ISCIT) 2019, pp 20–15

Hussein S, Kandel P, Bolan CW, Wallace MB, Bagci U (2019) Lung and pancreatic tumor characterization in the deep learning era: novel supervised and unsupervised learning approaches. IEEE Trans Med Imaging 38(8):1777–1787

Iadanza E, Goretti F, Sorelli M, Melillo P, Pecchia L, Simonelli F, Gherardelli M (2020) Automatic detection of genetic diseases in pediatric age using pupillometry. IEEE Access 8:34949–34961

Lakshmi Priya B, Jayanthi K, Pottakkat B, Ramkumar G (2019) GoogLeNet based ensemble FCNet classifier for focal liver lesion diagnosis. IEEE J Biomed Health Inform. https://doi.org/10.1109/JBHI.2019.2942774

Laouid A, Dahmani A, Hassen HR et al (2019) A self-managing volatile key scheme for wireless sensor networks. J Ambient Intell Human Comput 10:3349–3364

Li X, Wang Y, Wang D, Yuan W, Peng D, Mei Q (2019) Improving rare disease classification using imperfect knowledge graph. BMC Med Inform Decision Making. 19-S(5):1–10

Nasiri N, Foruzan AH, Chen YW (2019) A controlled generative model for segmentation of liver tumors. In: Conference on proceedings of the 27th Iranian conference on electrical engineering, (ICEE2019), pp 1742–1745

Romero F, Diler A, Bisson-Gregoire G (2019) End-To- End discriminative deep network for liver lesion classification. In: Conference on proceedings of IEEE 16th international symposium on biomedical imaging, Venice, Italy, (ISBI 2019), April 8–11, pp 1243–1246

Samala RK, Chan H-P, Hadjiiski L, Helvie MA, Richter CD, Cha KH (2019) Breast cancer diagnosis in digital breast tomosynthesis-effects of training sample size on multi-stage transfer learning using deep neural nets. IEEE Trans Med Imaging 38(3):686–696

Singh A, Kisku DR (2018) Detection of rare genetic diseases using facial 2D images with transfer learning. In: Conference on Proceedings of the 2018 8th international symposium on embedded computing and system design (ISED), https://doi.org/10.1109/ISED.2018.8703997

Song X, Deng G, Zhuang Y, Zeng N (2019) An improved confidence connected liver segmentation method based on three views of CT images. IEEE Access 7:58429–58434 (**special section on data-enabled intelligence for digital health**)

Sun X, Wang R (2019) Machine learning approaches to identify rare diseases. Pharma SUG 2019—Paper ST-325. pp 1–5

Tan C, Zhao L, Yan Z, Li K, Metaxas D, Zhan Y (2018) Deep multitask and task-Specific feature learning network for Robust shape preserved organ segmentation. In: Conference on Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, D.C., USA, April 4–7, pp 1221–1224

Todoroki Y, Iwamoto Y, Lin L, Hu H, Chen Y-W (2019) Automatic detection of focal liver lesions in multi-phase CT images using a multi-channel & multi-scale CNN. In: Conference on proceedings of the 41st conference, Berlin, Germany EMBC 2019, pp 872–875

Vlachostergiou A, Tagaris A, Stafylopatis A, Kollias S (2018) Multitask learning for predicting Parkinson's disease based on medical imaging information. In: Conference on proceedings of the 2018 IEEE International conference on image processing, ICIP 2018, Athens, Greece, pp 2052–2056

Yoshinobu Y, Iwamoto Y, Xianhua HA, Lanfen LI, Hongjie HU, Zhang Q, Yen-Wei CH (2020) Deep learning method for content-based retrieval of focal liver lesions using multiphase contrast-enhanced computer tomography images. In: Conference on proceedings of the 2020 IEEE international conference on consumer electronics ICCE, ISBN 978-1-7281-5186-1

Yousaf S, Shafi I, Din S, Paul A, Ahmad J (2019) A big data analytical framework for analyzing solar energy receptors using evolutionary computing approach. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-019-01443-7

Yu C, Gao Z, Zhang W, Yang G, Zhao S, Zhang H, Zhang Y, Li S (2020) Multitask learning for estimating multitype cardiac indices in MRI and CT based on adversarial reverse mapping. In: IEEE Transactions on Neural Networks and Learning Systems, April 30, 2020

Zou L, Song Y, Chu J, Tang X (2019) Whole brain volume and cortical thickness based automatic classification of Wilson's disease. In: Conference on proceedings of the 2019 IEEE international conference on systems, man and cybernetics (SMC), Bari, Italy. October 6–9, 2019.