

An Optimized Transfer Learning Model Based Kidney Stone Classification

S. Devi Mahalakshmi*

Department of Computer Science and Engineering, Mepco Schlenk Engineering College, Sivakasi, 626005, Tamilnadu, India

*Corresponding Author: S. Devi Mahalakshmi. Email: drdevimahalakshmi123@gmail.com

Received: 21 January 2022; Accepted: 23 February 2022

Abstract: The kidney is an important organ of humans to purify the blood. The healthy function of the kidney is always essential to balance the salt, potassium and pH levels in the blood. Recently, the failure of kidneys happens easily to human beings due to their lifestyle, eating habits and diabetes diseases. Early prediction of kidney stones is compulsory for timely treatment. Image processing-based diagnosis approaches provide a greater success rate than other detection approaches. In this work, proposed a kidney stone classification method based on optimized Transfer Learning(TL). The Deep Convolutional Neural Network (DCNN) models of DenseNet169, MobileNetv2 and GoogleNet applied for classification. The combined classification results are processed by ensemble learning to increase classification performance. The hyperparameters of the DCNN model are adjusted by the metaheuristic algorithm of Gorilla Troops Optimizer (GTO). The proposed TL model outperforms in terms of all the parameters compared to other DCNN models.

Keywords: DCNN; GTO; kidney stone; transfer learning

1 Introduction

A kidney stone is the growth of solid tiny crystals in the inner portions of a kidney. It can stick or leave the ureters based on the size of the stone [1]. Diet, excess obesity, some medical treatment, and dehydration are the major causes of kidney stones. Early prediction of kidney stones can support avoiding further difficulties of urinary tract damages and kidney infections. Postponing treatments may cause permanent kidney failure or may even be an origin of cancer [2].

In the medical field, experts learn the inner structures and tissues of organisms to detect and treat anomalies faster. Medical image processing techniques help doctors ultimately to diagnose diseases speedily and accurately [3]. Due to the development of computer-aided automation, the application of artificial intelligence techniques reaches superior medical image processing effects to a particular level like different machine learning (ML) and deep learning (DL) algorithms in image processing [4].

The proposed work aims to introduce a new optimized TL model for effective kidney stone classification. The hyperparameters present in the learning model are adjusted with the guidance of the GTO algorithm.



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The remaining sections of this work structure are as follows: Section 2, describes the existing DL and TL architectures for kidney stone classifications. In Section 3, the proposed TL model with IP GTO is explained. In Section 4, Date set details and experimental results are discussed. Finally, conclude the work in Section 5.

2 Related Work

This section explains various DL and TL algorithms used in kidney stone segmentation and classification algorithms. Soni et al. 2020 [5] proposed an automated kidney stone classification using machine learning. An efficient ML algorithm of Support Vector Machine is used to classify the images into normal and stone affected kidney images. The proposed ML approach achieved an accuracy of 95.8% in the CT image database.

Viswanath et al. 2015 [6] proposed a level set segmentation-based kidney stone detection and classification method. Compared to other imaging systems, ultrasound images are easily vulnerable to speckle noises. Initially, the images are preprocessed using a Gabor filter for noise removal. To get a stone portion, level set segmentation is applied. The technique of Daubechies lifting scheme wavelet is used to extract the energy levels of extracted stone portions. Finally, a feed-forward neural network model is applied to classify the stone levels.

Yuan et al. 2020 [7] analyzed the performance of various AI algorithms in the field of kidney stone detection. The effect of applying artificial neural networks in the segmentation process is analyzed for different kidney stone image databases. Martinez et al. 2020 [8] introduced a classification algorithm for kidney stone images captured with ureteroscopes. The proposed classification model includes two steps: encoding and classification. In the encoding step, the extracted features from kidney stones are encoded as a vector. Then, the combined classifiers of Random Forest and ensemble K Nearest Neighbor are used to classify the normal and abnormal stages. Experimental results show that the proposed combined model improves the overall accuracy by about 10% when compared to other classification algorithms.

The concept of DL is applied to overcome the problem of the time-consuming and tedious process of accurate stone spotting and localization. The five different DL models for kidney stone processing is proposed by Francisco Lopez et al. 2021 [9]. Among the five models, the DL model with XBOOST classifier produces better results in terms of accuracy, precision and recall rates. Cui et al. 2019 [10] proposed a radiomics algorithm for the classification of kidney stones in non-contrast CT images. The proposed radiomics algorithm uses multivariable logistic regression analysis for stone detection by extracting the morphological and textural features. Finally, ensemble learning with bagged trees is used for classification. Decision curve results proved that the proposed algorithm has a higher potential for stone classification. Kazemi et al. 2018 [11] analyzed the performance of the types of classifiers: Decision Trees, Artificial Neural Networks, and Rule-based Classifiers for stone classification. The proposed model is integrated with ensemble learning to increase classification accuracy. The weight assignment and adjustment for each classifier is carried out by a genetic algorithm. The overall accuracy of the classifier is increased to 97.2% through the proper adjustments of weight by the genetic algorithm.

Yin et al. 2019 [12] proposed a threefold boundary distance mapping model for kidney stone segmentation in ultrasound images. Initially, the high-level feature is extracted from ultrasound images using pre-trained learning model. Then, a boundary distance regression network is used for segmentation. Finally, the pixel-wise classification network is applied to classify the images for the presence of stones. Yildirim et al. 2021 [13] presented a DL model of cross-residual network (XResNet-50) for kidney stone classification. The proposed XResNet-50 includes four stages of layer processing for accurate diagnosis. The inclusion of ResNet layers in every stage improves the feature extraction capability of the model. Experimental results show that the proposed cross-layered model achieves an accuracy of 96.23% which is higher than previously proposed models.

Sudharson et al. 2020 [14] proposed TL based model for stone classification in B-mode kidney ultrasound images. The pre-trained model of ResNet-101, ShuffleNet, and MobileNet-v2 are integrated to produce a classification result. The concept of data augmentation is further utilized to increase the number of training images. Rui et al. 2021 [15] proposed TL based U-Net for kidney stone classification. The layers of the tractional U-Net model are modified to extract more features from the images to deeply learn the features for classification. The accuracy of extracting semantic features from the image is higher than the other models like VGG and ResNet.

The selection and tuning of layers in TL are complex issues. the accuracy of TL is depending on layer performance. These issues are considered and solved by Vrbancic et al. 2020 [16]. the Differential Evolution based Fine-Tuning (DEFT) techniques are proposed to tune the layers in the learning model in order to increase efficiency. Lee et al. 2020 [17] proposed cross image modality-based TL for accurate classification of medical images. The different modality includes brain MRI and mammogram tumour images. Due to cross-modality, the learning capability of the model is increased greatly.

3 Proposed Optimized TL for Kidney Stone Classification

TL is knowledge transfer from one block to another block in order to improve the efficiency of the learning model. In kidney stone classification, there is a huge need for an effective classifier in order to differentiate structural and intensity similarities between pixels. These difficulties can be solved by using TL models. Further, the performance of DCNN models is mainly based on the hyperparameters in convolutional layers. These parameters values are adjusted correctly to get improved performance.

In this work, the pre-trained DCNN models of DenseNet169, MobileNetv2 and GoogleNet are combined by ensemble method to classify the kidney stones. The proposed approach is divided into four stages: Preprocessing stage, training and optimization stage, classification stage and performance analysis stage as shown in Fig. 1. The preprocessing stage involves data collection and argumentation processes like flipping around the x-axis, right/left mirroring, salt noise creation and image rotation etc. The TL model hyperparameters are tuned for optimization. By performing the ensemble method, the classification results are combined to classify the stone types. Finally, the performance of the proposed model is analyzed for various parameters.

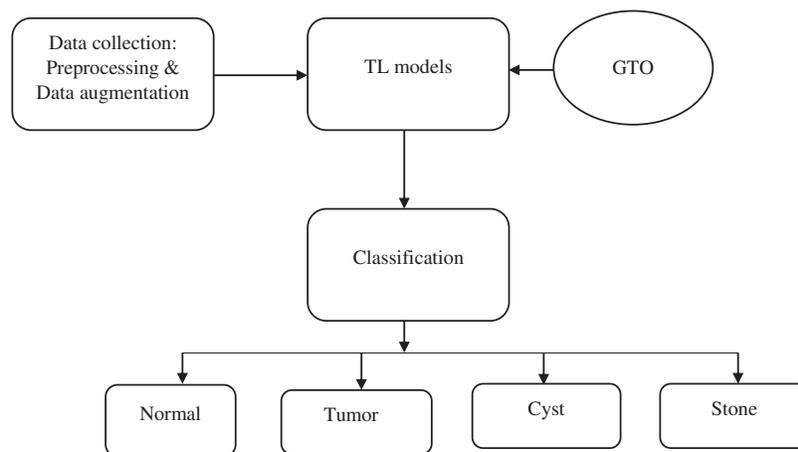


Figure 1: The overall workflow of the proposed model

3.1 DCNN Models

3.1.1 DenseNet

In Resnet, the output of the previous layer is merged with the future layer. But, in DenseNet, the concatenation is carried out between the previous layer to future layer. it has three different variants based on a number of layers in the network like DenseNet-121, DenseNet-160, DenseNet-201, etc. The entire model is constructed by Denseblock. The number of filters used in every block is different but the dimensions are the same.

3.1.2 Google Net

Google Net is proposed to reduce an error rate in image classification when compared to AlexNet and VGG Nets. It includes 1*1 convolution operation and average pooling operations which enables deeper architectures for further performance improvements. By the 1*1 convolution operation, the number of parameters for tuning is greatly reduced.

In other models, the fully connected layers are used at the end of the model. This will lead to an increased number of parameters and increased computation cost. In GoogleNet, the average pooling layers converts 7*7 into 1*1 to reduce a computation cost. In the Inception module, the convolution and max pooling operations are performed parallelly to handle objects at the multi-scale matter. Further, the intermediate Auxiliary Classifier in the middle layers supports combating gradient vanishing problem and also offer regularization.

3.2 GTO

Metaheuristics algorithms are used to solve optimization problems in various fields like medical, signal processing, image processing and cloud computing. These algorithms are inspired by the behavior of nature or animals to find the solution for complex problems. This work uses Gorilla Troops Optimizer (GTO) for TL parameter tuning. GTO is encouraged by the strategy followed by gorillas for their survival. The behaviour of the gorilla is modelled mathematically to execute exploration and exploitation stages in optimization.

The group of gorillas called droops consist of adult males and female gorillas of different ages. The male gorilla with silver-coloured hair is named silverbacks. The members in droops obey the commands of silverbacks for a location change and food searching. GTO includes two stages for optimization: exploration and exploitation.

Exploration stage:

The movement of gorillas to a known or unknown location and migrating to another droop is mathematically modelled in this stage.

$$SX(t+1) = \begin{cases} (UL - LL) \times rand_1 + LL, & rand < p, \\ (rand_2 - c1) \times X_r(t) + c2 \times c3, & rand \geq 0.5, \\ X(i) - c2 \times (c2 \times (X(t) - SX_r(t)) + rand_3 \times (X(t) - SX_r(t))), & rand < 0.5, \end{cases} \quad (1)$$

In the above equation, $SX(t+1)$ is the future position of a gorilla with respect to current position $X(t)$. P denotes the probability of migration varies from zero to one. X_r and $SX_r(t)$ is one member of the gorilla selected from the group. $rand_1$, $rand_2$ and $rand_3$ are the random values varying between zero to one. The coefficient values $c1$, $c2$ and $c3$ is computed as follows:

$$c1 = F * (1 - (1/(MaxIt))) \quad (2)$$

$$F = \cos(2 * rand4) + 1 \quad (3)$$

$$c2 = c1 * l \tag{4}$$

$$c3 = z * x(t) \tag{5}$$

where l is the random values varies between -1 to 1 . $MaxIt$ is the maximum number of iterations performed to get an optimal solution. Z is the random dimension that varies between $-c$ to $+c$. The end of this stage is group formation with silverback. The identification of silverback is considered as the best solution in this stage.

Exploitation stage:

In this stage, obey the silverback and competition for mature females are carried out. All the gorillas in a group should obey silverback commands to find a portion of food and to survive their life.

Obey the silverback:

Compared to other gorillas, the silverback is so young and healthy. All members follow the silverback instruction to search a food effectively. The number of members in droop affects the level of the movement

$$SX(t + 1) = c1 \times T \times (X(t) - X_{silverback}) + X(t) \tag{6}$$

$$T = \left(\left| \frac{1}{N} \sum_{i=1}^N SX_i(t) \right|^g \right)^{\frac{1}{g}} \tag{7}$$

$$g = 2^L \tag{8}$$

Competition for adult females:

The young male gorillas after getting maturity, fight with other males to select females and to increase the size of the group is methodically modelled as follows:

$$SX(i) = X_{silverback} - (X_{silverback} \times U - X(t) \times U) \times K \tag{9}$$

$$U = 2 \times r_5 - 1 \tag{10}$$

$$K = \beta \times D \tag{11}$$

$$D = \begin{cases} N_1, & rand \geq 0.5 \\ N_2, & rand < 0.5 \end{cases} \tag{12}$$

where, $X_{silverback}$ is a position of the silverback and U is simulating force. β is the initial parameter value. The proposed optimized classification algorithm is given in algorithm 1.

Algorithm 1: An optimized TL model pseudocode

Input: Number of populations and iterations, DCNN ,Data sets and image labels

Output: classification results

Begin

Initialize

Optimizers

Batch size

Drop out ratio

(Continued)

Algorithm 1: (continued)

$P_{sen}, P_{spe}, P_{acc}, F1Score, P_{pre}, P_{rec}$ and P_{dc}
 fitnessScores $\leftarrow []$
 data \leftarrow datasplit(train , validation)
initialize GTO
 solution \leftarrow Get Solution(population, np)
 TL \leftarrow LearnDCNN(Model, solution, Validation, test, $P_{sen}, P_{spe}, P_{acc}, F1Score, P_{pre}, P_{rec}$ and P_{dc})
 metrics \leftarrow TestDCNN(Model, Metrics, test, train)
 population **increment**
end
 Arranging the solutions
 population \leftarrow updateGTO(sorted solutions , tuned parameters)
Increment
End

3.3 Ensemble Method

In this work, the ensemble method is used to combine the classification results of three DCNN models to get a final decision. Majority voting is applied to get boosted or strong classification outputs. It also supports reducing the variance and bias of the classification outputs. The final classification results include four types: normal, stone, cyst and tumor.

4 Experimental Results

The ultrasound kidney images are obtained from different locations [18–20]. These data sets consist of kidney images with the types of normal, stone, cyst and tumor. The images are collected using a Toshiba PowerVision 7000 SSA-380A ultrasound scanner, embedded with 3.5 MHz linear array scanning probes. After deleting the marks, margins and dimensions information's, the dataset comprises 705 different images in total. The proposed model is evaluated using the different parameters of sensitivity (P_{sen}), specificity (P_{spe}), accuracy (P_{acc}), F1Score, precision (P_{pre}), recall (P_{rec}) and dice coefficient (P_{dc}) as specified in Eqs. (13) to (19). FN, FP, and TN represent True Positive, False Negative, False Positive and True Negative, respectively.

$$P_{sen} = \frac{TP}{TP + FN} \quad (13)$$

$$P_{spe} = \frac{TN}{TN + FP} \quad (14)$$

$$P_{acc} = \frac{TP + TN}{TP + FP + TN + FN} \quad (15)$$

$$F1Score = \frac{2TP}{2TP + FP + FN} \quad (16)$$

$$P_{pre} = \frac{TP}{TP + FP} \tag{17}$$

$$Prec = \frac{TP}{TP + FN} \tag{18}$$

$$P_{dc} = \frac{2 * TP}{TP + FP + TN + FN} \tag{19}$$

For every model, the performance is analyzed for an average of P_{sen} , P_{spe} , P_{acc} , F1Score, P_{pre} , P_{rec} and P_{dc} after applying 15 iterations. The average results comprise tuning different hyperparameters.

The performance of the four models are given in [Tab. 1](#) and graphically represented in [Fig. 2](#). The results of DenseNet121 model in all iterations for the P_{acc} , P_{sen} , P_{spe} , F1Score, P_{pre} , P_{rec} and P_{dc} are 97.05, 98.05, 97.93, 98.3, 97.56, 97.1 and 98.33 respectively. The results of MobileNetV3 model in all iterations for the P_{acc} , P_{sen} , P_{spe} , F1Score, P_{pre} , P_{rec} and P_{dc} are 96.62, 97.23, 97.32, 97.83, 97.4, 96.92 and 96.18 respectively. The results of the GoogleNet model in all iterations for the P_{acc} , P_{sen} , P_{spe} , F1Score, P_{pre} , P_{rec} and P_{dc} are 94.89, 96.24, 95.8, 96.73, 95.9, 95.72 and 94.71 respectively. The results of the Ensemble model in all iterations for the P_{acc} , P_{sen} , P_{spe} , F1Score, P_{pre} , P_{rec} and P_{dc} are 98.49, 98.7, 98.21, 98.5, 99.51, 97.69 and 98.82 respectively. Among these, the best result is obtained by the proposed combined ensemble model.

Table 1: Results of proposed models

Method	P_{acc}	P_{sen}	P_{spe}	F1Score	P_{pre}	P_{rec}	P_{dc}
DenseNet121	97.05	98.05	97.93	98.3	97.56	97.1	98.33
MobileNetV3	96.62	97.23	97.32	97.83	97.4	96.92	96.18
GoogleNet	94.89	96.24	95.8	96.73	95.9	95.72	94.71
Ensemble model	98.49	98.7	98.21	98.5	99.51	97.69	98.82

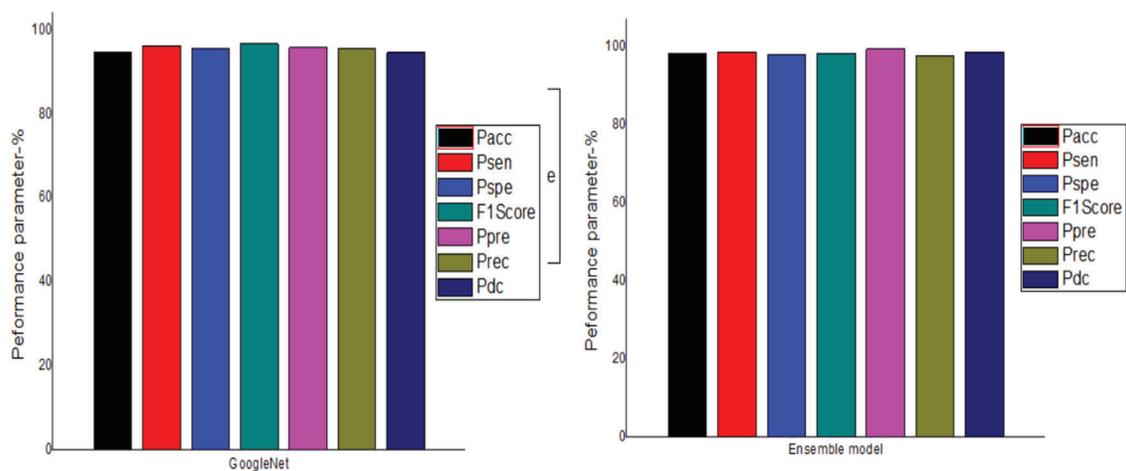


Figure 2: Comparison of different models

[Tab. 2](#). shows the comparison of the proposed TL model to existing TL models in stone classification. The proposed TL model achieved a P_{acc} , P_{sen} , P_{spe} , F1Score, P_{pre} , P_{rec} and P_{dc} values of 98.49, 98.7,

98.21, 98.5, 99.51, 99.51, 97.69 and 98.82 respectively. The proposed optimized TL model achieved the best result in all the parameters as shown in Fig. 3.

Table 2: Performance comparison

Method	P _{acc}	P _{sen}	P _{spe}	F1Score	P _{pre}	P _{rec}	P _{dc}
Proposed	98.49	98.7	98.21	98.5	99.51	97.69	98.82
Sudharson et al	97.5	96.89	98.1	97.52	97.23	96.6	98.4
Yildirim, K et al	95.66	96.52	95.33	95.52	96.27	95.33	96.7
Cui, X. et al.	96.32	96.46	98.15	96.71	96.23	97.18	95.9

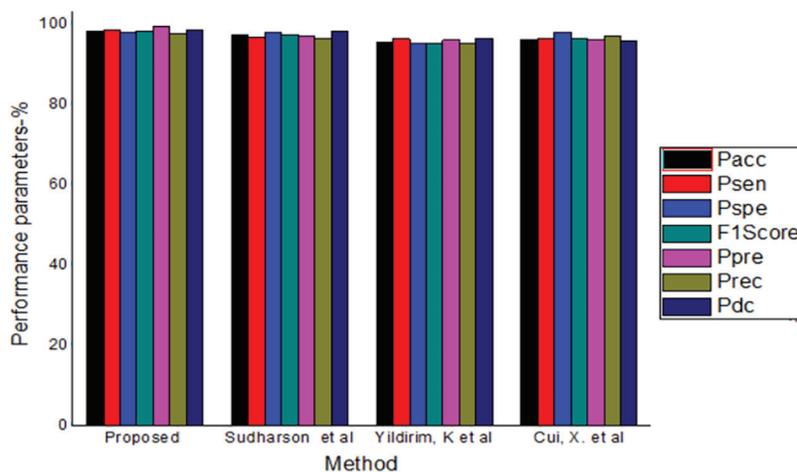


Figure 3: Comparison of different methods

5 Conclusion

Accuracy is the essential parameter in the field of medical diagnosis. This work proposed an optimized TL based kidney stone classification to achieve a less processing instant and to get greater accuracy. The developed model aims to support the doctors in automating the classification of kidney stones and their severity level based on ultrasound images. The concept of ensemble voting in the classification increases the classification accuracy. The adjustments of hyperparameters using GTO yields proper tuning of parameters. Experimental results show the proposed model achieved an overall accuracy of 98.49 which is greater than previously proposed techniques.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] K. K. Shung, "High frequency ultrasonic imaging," *Journal of Medical Ultrasound*, vol. 17, no. 1, pp. 25–30, 2009.
- [2] J. S. Jose, "An efficient diagnosis of kidney images using association rule," *International Journal of Computer Technology Electronic Engineering*, vol. 12, no. 2, pp. 14–20, 2012.
- [3] C. Cortes and V. Vapnik, "Support-vector networks," *Journal of Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

- [4] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *IEEE Conf. on Computer Vision and Pattern Recognition*, Las Vegas, USA, pp. 770–778, 2016.
- [5] A. Soni and A. Rai, "Kidney stone recognition and extraction using directional emboss & SVM from computed tomography images," in *2020 Third Int. Conf. on Multimedia Processing, Communication & Information Technology (MPCIT)*, Shivamogga, India, pp. 172–183, 2020.
- [6] K. Viswanath and R. Gunasundari, "Analysis and implementation of kidney stone detection by reaction diffusion level set segmentation using xilinx system generator on FPGA," *VLSI Design*, vol. 5, no. 3, pp. 573–581, 2159.
- [7] Q. Yuan, H. Zhang and T. Deng, "Role of artificial intelligence in kidney disease," *International Journal of Medical Sciences*, vol. 17, no. 7, pp. 970–984, 2009 .
- [8] A. Martinez, D. Trinh and J. Hubert, "Towards an automated classification method for ureteroscopic kidney stone images using ensemble learning," in *2020 42nd Annual Int. Conf. of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, pp. 1023–1036, 2020.
- [9] L. Fransico and H. Oscar, "Assessing deep learning methods for the identification of kidney stones in endoscopic images," in *2021 43rd Annual Int. Conf. of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Mexico, pp. 890–902, 2021.
- [10] X. Cui, F. Che and N. Wang, "Preoperative prediction of infection stones using radiomics features from computed tomography," *IEEE Access*, vol. 7, no. 1, pp. 122675–122683, 2019.
- [11] Y. Kazemi and S. A. Mirroshandel, "A novel method for predicting kidney stone type using ensemble learning," *Artificial Intelligence in Medicine*, vol. 84, no. 7, pp. 117–126, 2018.
- [12] S. Yin, Q. Peng and Z. Zhang, "Automatic kidney segmentation in ultrasound images using subsequent boundary distance regression and pixelwise classification networks," *Medical Image Analysis*, vol. 5, no. 6, pp. 565–571, 2020.
- [13] K. Yildirim, P. G. Bozdog and M. Talo, "Deep learning model for automated kidney stone detection using coronal CT images," *Computers in Biology and Medicine*, vol. 135, no. 9, pp. 104569–10582, 2020.
- [14] S. Sudharson and P. Kokil, "An ensemble of deep neural networks for kidney ultrasound image classification," *Computer Methods and Programs in Biomedicine*, vol. 3, no. 5, pp. 283–292, 2020.
- [15] L. Rui, Y. Zhao and Y. Dai, "Endoscopic segmentation of kidney stone based on transfer learning," in *2021 40th Chinese Control Conf. (CCC)*, Shanghai, China, pp. 667–672, 2021.
- [16] G. Vrbancic and V. Podgorelec, "Transfer learning with adaptive fine-tuning," *IEEE Access*, vol. 8, no. 4, pp. 196197–196211, 2020.
- [17] J. Lee and R. M. Nishikawa, "Cross-organ, cross-modality transfer learning: Feasibility study for segmentation and classification," *IEEE Access*, vol. 8, no. 1, pp. 210194–210205, 2020.
- [18] Y. Te Liao, C. Hung and K. Chen, "Data augmentation based on generative adversarial networks to improve stage classification of chronic kidney disease," *Recent Advances in Deep Learning for Image Analysis*, vol. 12, no. 1, pp. 1–10, 2022.
- [19] T. Geertsma, *Ultrasoundcases.info* (2011) [Online]. Available: [http://www. ultrasoundcases.info/](http://www.ultrasoundcases.info/), Accessed: March, 2020.
- [20] J. Antony, *Ultrasound-images.com* (2015). [Online]. Available: [https://www. ultrasound-images.com/](https://www.ultrasound-images.com/), Accessed: March, 2020.