IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 2, April 2025, pp. 925~932
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i2.pp925-932

DualVitOA: A dual vision transformer-based model for osteoarthritis grading using x-ray images

Qiu Ruiyun[1], Siti Khatijah Nor Abdul Rahim[1], Nursuriati Jamil[1], Raseeda Hamzah[2]
[1]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Selangor, Malaysia
[2]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Melaka, Malaysia

Article Info

Article history:

Received Mar 18, 2024
Revised Nov 6, 2024
Accepted Nov 14, 2024

Keywords:

Deep learning
Dual-vision transformer
Grading
Knee osteoarthritis
Pre-process

ABSTRACT

Knee osteoarthritis (OA) is a primary factor contributing to reduced activity and physical impairment in older individuals. Early identification and treatment of knee OA can assist patients in delaying the advancement of the condition. Currently, knee OA is detected early using X-ray images and assessed based on the Kellgren-Lawrence (KL) grading system. Doctors' assessments are subjective and can vary among different doctors. The automatic knee OA grading and diagnosis can assist doctors and help doctors reduce their workload. A new novel network called dual-vision transformer (ViT) OA is proposed to automatically diagnose knee OA. The network utilizes pre-processing technologies to process the data before doing classification operations using the Dual-ViT network. The suggested network outperformed neural networks like ResNet, DenseNet, visual geometry group (VGG), inception, and ViT in terms of accuracy and mean absolute error (MAE), and achieved an accuracy of 78.4 and MAE of 0.471, demonstrating its effectiveness.

Corresponding Author:

Siti Khatijah Nor Abdul Rahim
College of Computing, Informatics and Mathematics, Universiti Teknologi MARA
Selangor, Malaysia
Email: sitik781@uitm.edu.my

1. INTRODUCTION

Knee osteoarthritis (OA), characterized by the degeneration of articular cartilage, is the most common form of arthritis and one of the main factors causing limited activity and physical disability in the elderly [1]. Pain and other symptoms caused by knee OA seriously affect the quality of life a lot. Unfortunately, no treatment option can inhibit the degenerative structural changes in the progression of knee OA. However, early detection and treatment can help older adults delay the progression of OA and improve their quality of life. Imaging examination is an important tool to determine the progression of OA, among which X-rays, and magnetic resonance imaging (MRI) are the most used. MRI can reflect the three-dimensional structure of the knee joint. However, MRI is only available in large medical centers, and the high cost of examination makes MRI unsuitable for routine diagnosis of knee OA [2]. On the contrary, X-ray examination has the characteristics of safety, low cost, and wide popularity, and has always been used as the gold standard for knee OA examination. The Kellgren-Lawrence (KL) grading system used by the World Health Organization in 1961 is the most used system for grading the severity of knee OA [3]. The KL system divides the severity of knee OA into 5 grades, represented by grades 0 to 4 respectively. The KL grading system is shown in Figure 1. Doctors usually check the scanned X-ray images of the knee joint and then give the KL grade of the knee joint in a very short time, the accuracy of diagnosis is highly dependent on the doctor's experience and carefulness, the misclassification will lead to the wrong diagnostic method and prognosis, so, if the computer-aid technology can help automatically grade the KL level, it will benefit the diagnosis of OA [4].

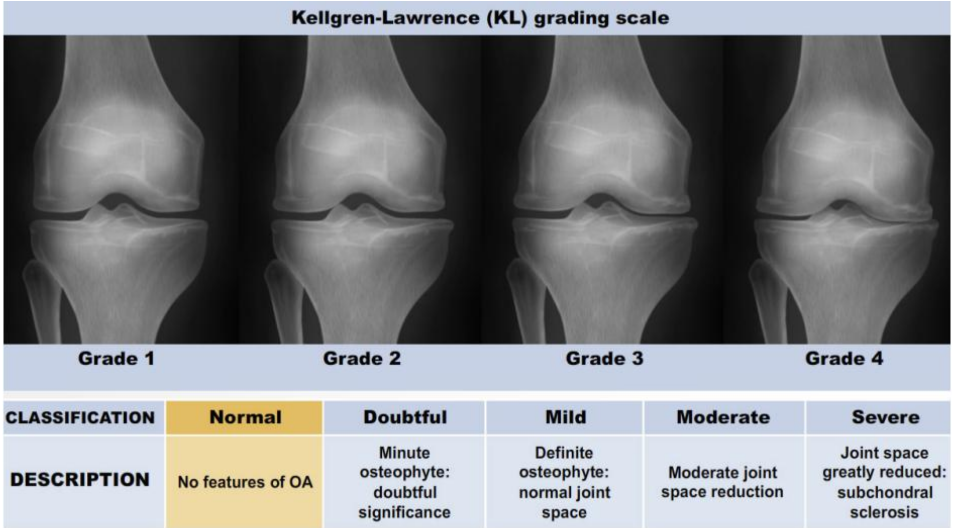

Figure 1. The KL grading system to assess the severity of knee OA

As early as 2009, a lot of excellent work was done for the classification of KL grade of knee OA. Shamir *et al* [5] proposed a weighted nearest neighbor algorithm, but this method requires the manual design of features, including texture features, Chebyshev statistics, and Haralick texture features. In recent years, with the development of deep learning, there has been relatively mature convolutional neural network (CNN) classification network models such as ResNet [6], VGG [7], Inception-V3 [8], and DenseNet [9]. Antony *et al.* [10]–[12] tried methods based on deep learning for the task of KL classification. They designed a new CNN model and optimized the weighted combination of cross-entropy loss and mean square error loss and obtained a 63.6% recognition rate [10]–[12]. Górriz *et al.* [13] proposed a novel end-to-end CNN framework to automatically quantify the severity of knee arthritis using X-ray images, achieving a recognition result of 64.3%. Tiulpin *et al.* [14] used a deep convolutional network to grade the severity of knee OA and achieved a result of 66.7%. Zhang *et al.* [15] proposed a CNN knee OA KL grade classification model under the attention mechanism. ResNet was first used to extract knee joint features from X-rays and then combined with the features extracted by the convolutional attention module and automatically predict KL grades [15].

The main contributions of the proposed method are as follows:

- We proposed dual-vision transformer (ViT) OA, a Dual-ViT network to automatically grade the knee OA severity, by using this method we could achieve a classification accuracy of 78.4% after the probable pre-processing.
- We conducted intensive experiments on the different pre-processing, and the results have shown the combination of eqhist, cutout, and flip pre-processing used on the image data can achieve the best performance.

2. METHOD

In this section, a Dual-ViT network [16] combined with image data pre-processing-DualVitOA is proposed. The Dual-ViT network structure and the data pre-processing will be illustrated respectively. The overall workflow will be better understood.

2.1. Dual-ViT network structure

The flowchart of the Dual-ViT network is shown in Figure 2. The whole process is composed of preparation, training, and output. After that, different pre-processing methods are conducted to get an ideal performance.

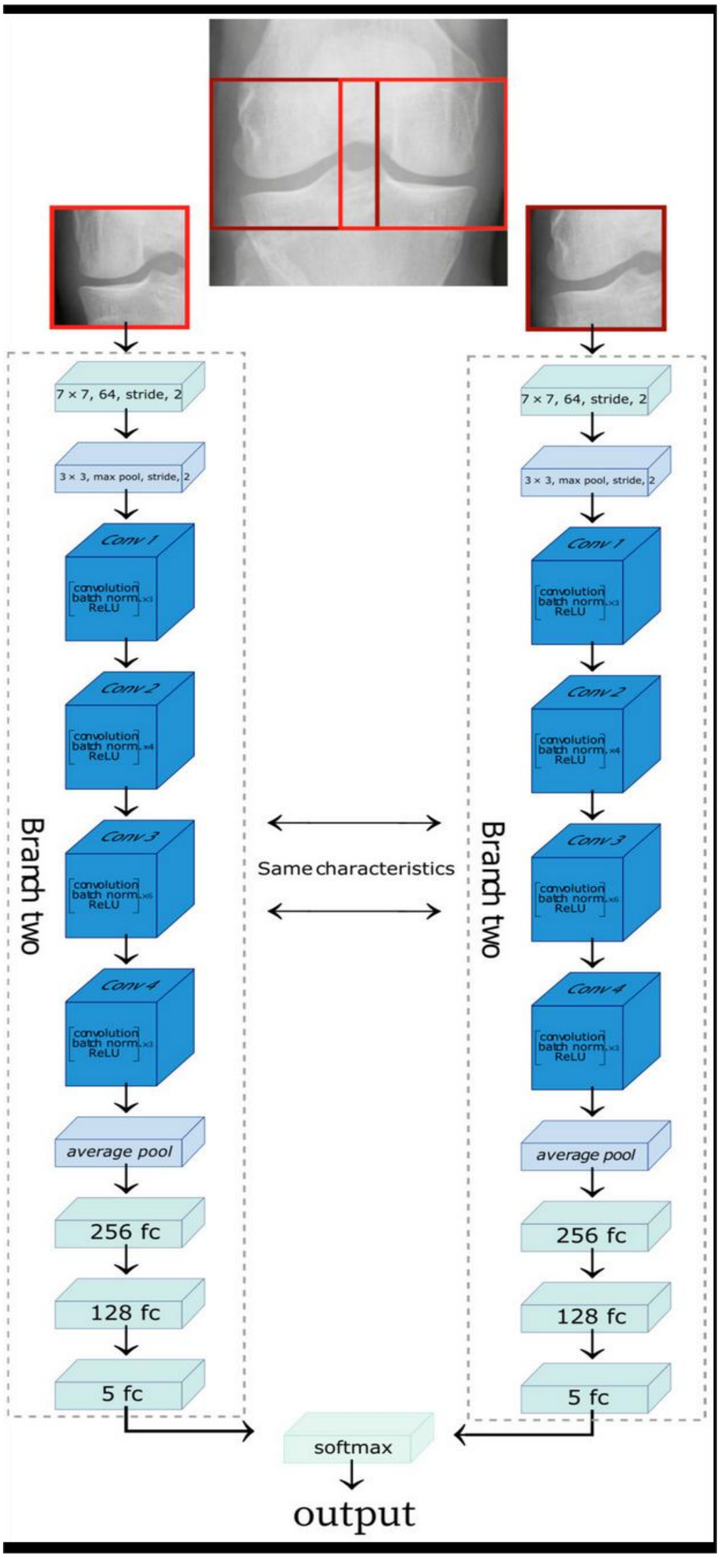

Figure 2. The architecture of the proposed network

2.1.1. Preparation

The dataset is first cut into left and right parts at the gap, the cutting size is 112, which is the size that can include the complete half of the joint gap in the entire data set after testing, this can ensure that no joint gap will be cut out. The row delimitation range is [56:168], and the columns are [0:112] and [112:224], and then the size is enlarged to the same size as the original images are consistent and the default size is 224. To determine the similarity between the left and right sides, the right part is flipped horizontally so that the distribution of the gaps on the right and the left is similar. The objects put into the network are three images, namely the original image, the left side of the gap, and the right side of the gap that has been flipped horizontally.

2.1.2. Training process

The network uses the ViT as the base pre-trained model [17], [18]. The model performs patch embedding on images of size 16 and stretches them into a sequence of 768. The model depth has 12 layers of ViT blocks. In the original structure, the last classification layer is a fully connected layer. To facilitate the acquisition of class activation mapping (CAM), it is replaced with a convolutional layer with a kernel of 1. The last layer of the feature map is composed of [B,196+1,768]. The classification result classifies the class token in the top row, and CAM extracts the remaining [B,196,768]. All the parameters are set according to the original parameters of ViT.

The original image is put into the network to only calculate the classification loss. Since the verification program is expected to input the complete image once for prediction, the original image is added to calculate the classification loss to let the network perceive the complete image. The left and right images of the gap are put into the network respectively to obtain the classification results and CAM. The classification results are calculated separately. The classification loss is calculated. The L1 loss is calculated between the two CAMs to learn the difference between the left and right gaps. Finally, loss feedback is unified.

2.1.3. The network output

The classification accuracy calculation only calculates the classification results obtained by putting the complete image into the network. Image cutting is not performed during the verification and testing stages. The classification accuracy is calculated after the complete image is put into the network. CAM can extract the activation effect map of the target during the verification and testing phases.

2.2. Pre-processing

The data is pre-processed into 7 datasets. The original data set has a total of 8,150 pictures after excluding 10 pictures with reverse colors. The data set is divided according to the training set and verification set ratio of 9:1, the verification set randomly selects 815 pictures from them. The remaining 7,335 pictures are for the training set. The dataset is divided according to random sampling, using numpy.random choice in the name sequence of the total image randomly selected. Data enhancement is performed on this basis.

2.2.1. Histogram equalization

Histogram equalization is a visual depiction of how data is spread out in a dataset. It shows the distribution of data pixels in specific intervals or bins, providing a visual representation of the range of intensities of an image [19]. The tool determines the number of pixels for each intensity level examined and improves the contrast in an image by widening the range of intensities. The process entails converting a provided histogram into a broader and more consistent distribution of intensity values, guaranteeing an equitable distribution of intensity values throughout the whole range.

To accomplish equalization, the remapping process should rely on the cumulative distribution function (CDF). The cumulative distribution $H'(i)$ of the histogram H(i) is determined as follow:

$$H'(i) = \sum_{0 \le j < i} H(j)$$

We use a simple remapping procedure to obtain the intensity values of the equalized image.

$$equalized(x, y) = H'(src(x, y))$$

2.2.2. Cutout

Cutout the nearby parts of the input in CNN using masks can improve the model's robustness and lead to better performance [20]. Implementing cutout involves applying a fixed-size zero mask to a random point of each input image during each training epoch. To optimize efficiency, we should normalize the dataset around zero to minimize the impact of modified images on the predicted batch statistics. Typically, we perform all studies using a square patch as the cutout area. We select a random pixel coordinate within the image as the center point and apply the cutout mask around that area. This method enables the cutout mask to extend beyond the image borders, with certain elements of the patches potentially being beyond the image.

2.2.3. Flip

Horizontal flip or image mirroring is the process of turning an image horizontally to create a mirrored counterpart of the original image. This transformation basically exchanges the left and right sides of the image, resulting in a mirror-image effect [21]. In our study, we flip the right and left part of the knee horizontally to evaluate the grade of the knee OA, because the more different the two parts are, the more likely the severe the knee OA is [5].

2.2.4. The pre-processed dataset

The original dataset has been processed for 7 different datasets after histogram equalization, cutout, flip, and the combination of them, and some datasets included images that only focus on the gap area. The datasets are represented in Table 1. Each of the datasets has been divided into three distinct categories: training, validation, and test sets. The distribution of the data is presented in Tables 2 and 3.

Table 1. The pre-processing of each dataset

Dataset	Region	Histogram Equalization	Flip	Cutout
origin	full image			
origin_fullimg	full image		√	
origin_onlygap	gap		√	
eqhist_fullimg	full image	√	√	
eqhist_onlygap	gap	√	√	
cutout	full image			√
eqhist_cut_flip	full image	√	√	√

Table 2. The training set data distribution

Dataset	Training set	Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
origin	7,335	2,861	1,353	1,929	962	230
origin_fullimg	14,670	5,722	2,706	3,858	1,924	460
origin_onlygap	14,670	5,722	2,706	3,858	1,924	460
eqhist_fullimg	14,670	5,722	2,706	3,858	1,924	460
eqhist_onlygap	14,670	5,722	2,706	3,858	1,924	460
cutout	7,335	2,876	1,335	1,924	970	230
eqhist_cut_flip	29,340	11,504	5,340	7,696	3,880	920

Table 3. The validation set data distribution

Dataset	Validation set	Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
origin	815	344	134	207	110	20
origin_fullimg	815	344	134	207	110	20
origin_onlygap	815	344	134	207	110	20
eqhist_fullimg	815	344	134	207	110	20
eqhist_onlygap	815	344	134	207	110	20
cutout	815	329	152	212	102	20
eqhist_cut_flip	815	329	152	212	102	20

3. RESULTS AND DISCUSSION

3.1. Datasets

The knee X-ray images utilized for evaluation are sourced from the osteoarthritis initiative (OAI), a study focusing on knee OA. The OAI has 4,796 people aged 45 to 79 and aims to identify biomarkers related to the initiation and advancement of OA. We utilized the dataset of Chen which consists of 4,130 X-ray images [22]. The knee X-ray images were divided randomly into training, validation, and test sets with a ratio of 7:1:2. The splitting method maintains a consistent distribution of levels among the training set, validation set, and test set. The knee joint test set has been divided into 1,656 X-ray images, consisting of 639 grade 0 images, 296 grade 1 images, 447 grade 2 images, 223 grade 3 images, and 51 grade 4 images.

3.2. Evaluation metrics

This research primarily utilizes two evaluation metrics, accuracy and mean absolute error (MAE), to evaluate the results of the network. These two metrics are normally used in the evaluation of the OA grading network. The next section provides a detailed introduction to these two measurements.

3.2.1. Accuracy

It is the proportion of accurate predictions on the test data to the total number of predictions made [23]. The bigger the accuracy is, the better the network performs. Mathematically, accuracy can be defined as follows:

$$accuracy = \frac{tp+tn}{tp+tn+fp+fn}$$

Where, *tp* represents true positive, *tn* stands for true negative, *fp* indicates false positive, and *fn* refers to false negative.

3.2.2. Mean absolute error

Accuracy can only be used as a standard measure of model effectiveness, for the KL grading task, in addition to accuracy, doctors are more concerned about the distance between the wrong category and the real category [24]. For example, if the real label is 0, the cost of a misjudgment of 4 is obviously much higher than the cost of a misjudgment of 1, so it is necessary to use another indicator to measure the cost of a misjudgment. The MAE can describe this point very well, and its definition is shown as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|P_i - T_i|$$

Where P_i stands for the prediction, T_i $represents$ the true value, and n refers to the total number of data points. Obviously, the smaller the MAE, the lower the cost of misjudgment of the model, which is more important than the accuracy metric to a certain extent. Therefore, this research takes the MAE index as the most important evaluation metric.

3.3. Results

3.3.1. The comparison of different kinds of networks

Table 4 displays the experimental findings for several networks. CE stands for cross-entropy loss, ordinal represents the sequence penalty weight loss introduced by Chen *et al.* [24], and proposed denotes the approach introduced by Liu *et al.* [25]. As can be shown in the table, the proposed network achieves the best performance in accuracy, but the second best in MAE.

Table 4. The comparison of different network's performance

Model		Accuracy	MAE
ResNet	ResNet-34-CE	66.4	0.585
	ResNet-34-ordinal	65.8	0.550
	ResNet-34-proposed	64.3	0.513
	ResNet-50-CE	65.2	0.618
	ResNet-50-ordinal	62.7	0.591
	ResNet-50-proposed	63.2	0.563
DenseNet	DenseNet-161-CE	66.6	0.519
	DenseNet-161-ordinal	65.2	0.524
	DenseNet-161-proposed	66.9	0.515
VGG	VGG-16-proposed	69.1	0.428
	VGG-19-ordinal	68.1	0.450
Inception	Inception-V3-CE	66.5	0.563
	Inception-V3-proposed	65.8	0.506
ViT	Original dataset	72.9	0.574
	Augment dataset	68.3	0.489
Dual-ViT	ours	78.4	0.471

3.3.2. The performance of different pre-process

Table 5 shows the performance of ViT and the proposed Dual-ViT network when utilizing different pre-processing methods. As we can see from the table, the Dual-ViT performs the best when the data is pre-processed with a combination of histogram equalization, cutout, and flip. This indicates the method is effective.

Table 5. The performance of using different pre-processing methods

Model	Dataset	Accuracy	MAE
ViT	origin_onlygap	69.6	0.574
	Origin_fullfilling	72.9	0.612
	Eqhist_fullfing	70.6	0.594
	Eqhist_onlygap	70.4	0.510
	cutout	66.1	0.524
	Eqhist_cutout_flip	68.3	0.489
Dual-ViT	origin_onlygap	71.4	0.562
	Origin_fullfilling	69.9	0.549
	Eqhist_fullfing	72.3	0.553
	Eqhist_onlygap	71.4	0.502
	cutout	66.4	0.510
	Eqhist_cutout_flip	78.4	0.471

3.3.3. Discussion

This study examined the impact of knee OA grading. Prior research has demonstrated a 69.1% accuracy and a MAE of 0.428. However, there is a strong possibility for the accuracy of OA grading to increase. In order to accomplish this, a DualViTOA network is suggested, which achieves an accuracy grading result of 78.4% and a MAE of 0.471. This performance surpasses that of both traditional and cutting-edge networks. Our study not only proposed a Duai-ViT network, but also explored several pre-processing techniques and their combinations through extensive experimentation. As a result, we identified the most effective approach to pre-processing. Nevertheless, the MAE has not yet achieved optimal performance. Subsequent investigations will focus on addressing this issue more extensively.

4. CONCLUSION

We propose a new end-to-end network, DualViTOA, for knee OA grading. This network first uses pre-processing technology to process the data and then performs classification operations through the Dual-ViT network. Experiments show that compared with neural networks such as ResNet, DenseNet, VGG, Inception, and ViT, the proposed network has achieved better performance in both accuracy and MAE metrics, indicating that the network is effective. Compared with other networks, the MAE of the proposed method still has room for improvement. In future research, we will improve the strategy and further improve the model recognition accuracy by strengthening the adaptive strategy and adding model combination methods. KL automatic grading of OA provides a better auxiliary diagnosis.

ACKNOWLEDGEMENTS

The authors would like to thank Universiti Teknologi MARA for the continuous support and encouragement in conducting this research project.

REFERENCES

[1] J. Wesseling *et al.*, "Cohort hip and cohort knee (CHECK): Similarities and differences with the osteoarthritis initiative," *Annals of the Rheumatic Diseases*, vol. 68, no. 9, pp. 1413–1419, 2009, doi: 10.1136/ard.2008.096164.

[2] D. Li, S. Li, Q. Chen, and X. Xie, "The prevalence of symptomatic knee osteoarthritis in relation to age, sex, area, region, and body mass index in China: a systematic review and meta-analysis," *Frontiers in Medicine*, vol. 7, 2020, doi: 10.3389/fmed.2020.00304.

[3] J. H. Kellgren and J. S. Lawrence, "Radiological assessment of osteo-arthrosis," *Annals of the rheumatic diseases*, vol. 16, no. 4, pp. 494–502, 1957, doi: 10.1136/ard.16.4.494.

[4] D. Saini, A. Khosla, T. Chand, D. K. Chouhan, and M. Prakash, "Automated knee osteoarthritis severity classification using three-stage preprocessing method and VGG16 architecture," *International Journal of Imaging Systems and Technology*, vol. 33, no. 3, pp. 1028–1047, 2023, doi: 10.1002/ima.22845.

[5] L. Shamir *et al.*, "Knee X-ray image analysis method for automated detection of osteoarthritis," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 2, pp. 407–415, 2009, doi: 10.1109/TBME.2008.2006025.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.

[7] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.

[8] C. Szegedy *et al.*, "Going deeper with convolutions," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015, doi: 10.1109/CVPR.2015.7298594.

[9] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," *30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, pp. 2261–2269, 2017, doi: 10.1109/CVPR.2017.243.

[10] J. Antony, K. McGuinness, N. E. O'Connor, and K. Moran, "Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks," *International Conference on Pattern Recognition*, pp. 1195–1200, 2016, doi: 10.1109/ICPR.2016.7899799.

[11] J. Antony, K. McGuinness, K. Moran, and N. E. O'Connor, "Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks," in *Machine Learning and Data Mining in Pattern Recognition*, pp. 376–390, doi: 10.1007/978-3-319-62416-7_27.

[12] J. Antony, K. McGuinness, K. Moran, and N. E. O'Connor, "Feature learning to automatically assess radiographic knee osteoarthritis severity," *Intelligent Systems Reference Library*, vol. 186, pp. 9–93, 2020, doi: 10.1007/978-3-030-42750-4_2.

[13] M. Górriz, J. Antony, K. McGuinness, X. Giró-I-Nieto, and N. E. O'Connor, "Assessing knee OA severity with CNN attention-based end-to-end architectures," *Proceedings of Machine Learning Research*, vol. 102, pp. 197–214, 2019.

[14] A. Tiulpin, J. Thevenot, E. Rahtu, P. Lehenkari, and S. Saarakkala, "Automatic knee osteoarthritis diagnosis from plain radiographs: A deep learning-based approach," *Scientific Reports*, vol. 8, no. 1, 2018, doi: 10.1038/s41598-018-20132-7.

[15] B. Zhang, J. Tan, K. Cho, G. Chang, and C. M. Deniz, "Attention-based CNN for KL grade classification: Data from the Osteoarthritis initiative," *International Symposium on Biomedical Imaging*, pp. 731–735, 2020, doi: 10.1109/ISBI45749.2020.9098456.

[16] A. Dosovitskiy *et al.*, "An image is worth 16X16 words: transformers for image recognition at scale," *ICLR 2021 - 9th International Conference on Learning Representations*, pp. 1-21, 2021.

[17] Z. Li, N. Zhang, H. Gong, R. Qiu, and W. Zhang, "MFA-Net: multiple feature association network for medical image segmentation," *Computers in Biology and Medicine*, vol. 158, 2023, doi: 10.1016/j.compbiomed.2023.106834.

[18] Y. Tay, M. Dehghani, D. Bahri, and D. Metzler, "Efficient transformers: a survey," *ACM Computing Surveys*, vol. 55, no. 6, 2022, doi: 10.1145/3530811.

[19] S. S. Bagade and V. K. Shandilya, "Use of histogram equalization in image processing for image enhancement," *International Journal of Software Engineering Research & Practices*, vol. 1, no. 2, pp. 6–10, 2011.
[20] T. DeVries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," *arXiv-Computer Science*, pp. 1-8, 2017.
[21] J. Shijie, W. Ping, J. Peiyi, and H. Siping, "Research on data augmentation for image classification based on convolution neural networks," *2017 Chinese Automation Congress, CAC 2017*, pp. 4165–4170, 2017, doi: 10.1109/CAC.2017.8243510.
[22] P. Chen, "Knee osteoarthritis severity grading dataset," *Mendeley Data*, V1, 2018, doi: 10.17632/56rmx5bjcr.1.
[23] Z. Li *et al.*, "CDA-Net: A contrastive deep adversarial model for prostate cancer segmentation in MRI images," *Biomedical Signal Processing and Control*, vol. 83, 2023, doi: 10.1016/j.bspc.2023.104622.
[24] P. Chen, L. Gao, X. Shi, K. Allen, and L. Yang, "Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss," *Computerized Medical Imaging and Graphics*, vol. 75, pp. 84–92, 2019, doi: 10.1016/j.compmedimag.2019.06.002.
[25] W. Liu, L. Luo, H. Peng, Q. Zhang, and W. Huang, "Grading scoring of knee osteoarthritis based on adaptive ordinal penalty weighted deep neural networks," *Chinese Journal of Scientific Instrument*, vol. 42, no. 7, pp. 145–154, 2021, doi: 10.19650/j.cnki.cjsi.J2107390.

BIOGRAPHIES OF AUTHORS

Qiu Ruiyun holds a Master of Computer Technology degree from Nanchang University, China in 2010. She also received her B.Sc. from Jiangxi Normal University, China in 2004. She is currently a lecturer at the School of Mathematics and Computer Sciences in Nanchang University, China. Her research includes medical images, machine learning, and deep learning. She has published over 10 papers in journals and conferences, 5 indexed by SCIE. She can be contacted at email: qiuruiyun@ncu.edu.cn or qryun1228@qq.com.

Siti Khatijah Nor Abdul Rahim is a senior lecturer at the College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Shah Alam Selangor. She gained her Ph.D. in Computer Science from the University of Nottingham in 2015. Before that, she received her Bachelor's Degree in Computer Science and M.Sc. of Science (Computer Science) from Universiti Sains Malaysia (USM), Penang, Malaysia, in 2001 and 2003, respectively. Her main research areas lie in the areas of computational intelligence, granular computing, scheduling, optimization, and applied computing. She can be contacted at email: sitik781@uitm.edu.my.

Nursuriati Jamil is a Professor at the College of Computing, Informatics and Mathematics, Universiti Teknologi MARA (UiTM), Malaysia. She is currently heading the Digital Image, Audio and Speech Technology Research Group and is the Director of the National Autism Analytics Centre in UiTM. She has authored 2 books and published over 100 scientific papers on speech synthesis and speech recognition of the Malay language; biometrics; image segmentation and recognition in agriculture and medical domain; gait analysis of autistic children; and image retrieval. She can be contacted at email: lizajamil@computer.org.

Raseeda Hamzah has been a lecturer at Universiti Teknologi MARA since August 12, 2016. She is currently a senior lecturer at the Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam Selangor. Secured her Ph.D. in Information Technology and Quantitative Sciences, Universiti Teknologi Mara, Cawangan Shah Alam. Her area of expertise is machine learning, artificial intelligence, and digital signal processing. She is also actively doing research in urban farming and IoT. She can be contacted at email: raseeda@uitm.edu.my.