Medical Image Classification using Pre-trained Convolutional Neural Networks and Support Vector Machine

Ali Ahmed

King Abdulaziz University-Rabigh, 21589, Rabigh, Saudi Arabia

Abstract
Recently, pre-trained convolutional neural network CNNs have been widely used and applied for medical image classification. These models can utilised in three different ways, for feature extraction, to use the architecture of the pre-trained model and to train some layers while freezing others. In this study, the ResNet-18 pre-trained CNNs model is used for feature extraction, followed by the support vector machine for multiple classes to classify medical images from multi-classes, which is used as the main classifier. Our proposed classification method was implemented on Kvasir and PH2 medical image datasets. The overall accuracy was 93.38% and 91.67% for Kvasir and PH2 datasets, respectively. The classification results and performance of our proposed method outperformed some of the related similar methods in this area of study.

Key words:
Pre-trained convolution neural networks, medical image classification, support vector machine

1. Introduction

Pre-trained convolution neural networks (CNNs) models have become popular for use in classification processes in general and in medical images in particular, due to their accuracy and good performance in classification and prediction tasks. These pre-trained CNNs have been trained with thousands and even millions of types of images from different categories and can be used for other classification processes in various problem domains [1, 2]. The main feature found in these pre-trained CNNs models which was not found in other neural networks is the initial values of the weights that were reached after repeated training processes. This feature enables these models to take advantage of their great capabilities and high accuracy and make them suitable for use in different classification processes, after the necessary adjustments to the input images and careful configuration of their parameters [3]. If some pre-trained CNNs can provide an equivalent or better classification performance than other sophisticated CNNs, then the detection and prediction of many diseases or cancers in the early stages can be more rapid and cost-effective [4]. There are three ways to utilise the pre-trained CNNs models and transfer learning concepts, feature extraction, use the pre-trained model architecture and train some layers of the model while freezing others. This study benefits from the first way and extracts the vector features from each input image. It then proposes the use of a support vector machine for classification purposes. The contributions of this study can be summarised as follows:

(a) To utilise the ResNet-18 pre-trained CNNs model for feature extraction from medical images.
(b) To apply a multi-class support vector machine for medical image classification for gastrointestinal and skin lesion cancer detection.
(c) To enhance the accuracy of classification and compare the proposed method with some other state-of-the-art methods.

2. Related Works

Transfer learning and pre-trained CNNs models have been widely used in various areas of applications for various classification problems. Four types of pre-trained CNNs models were tried and used by [5] for paddy leaf .disease detection They found that ResNet-V2 achieved better results than VGG-19, ResNet-101 and XceptionNet. The authors of study [6] conclude that ResNet-101 and XceptionNet CNNs give better results in chicken disease detection. Recently, the implementation and usage of ResNet-18 pre-trained CNNs models was found to be successful in similar studies, as in [7-9]. In their studies, the concept of transfer learning was applied to classify various types of diseases, such as Alzheimer's disease detection, on MRI Images and Covid-19 computed tomography (CT) images. Utilising of ResNet-18 pre-trained CNNs model for feature extraction, as in these studies, was widely used, as in [10, 11]. Samples of studies that have good performance results for the ResNet-18 pre-trained CNNs model, and support for the vector machine SVM, can be found in [12-14]. A successful combination of support vector machine SVM for multiple classes' and the ResNet-18 pre-trained CNNs model was achieved by [15, 16]. The authors in study [17] successfully implemented SVM for multi-classes classification and separation between normal and abnormal tumors.

Manuscript received June 5, 2021
Manuscript revised June 20, 2021
https://doi.org/10.22937/IJCSNS.2021.21.6.1

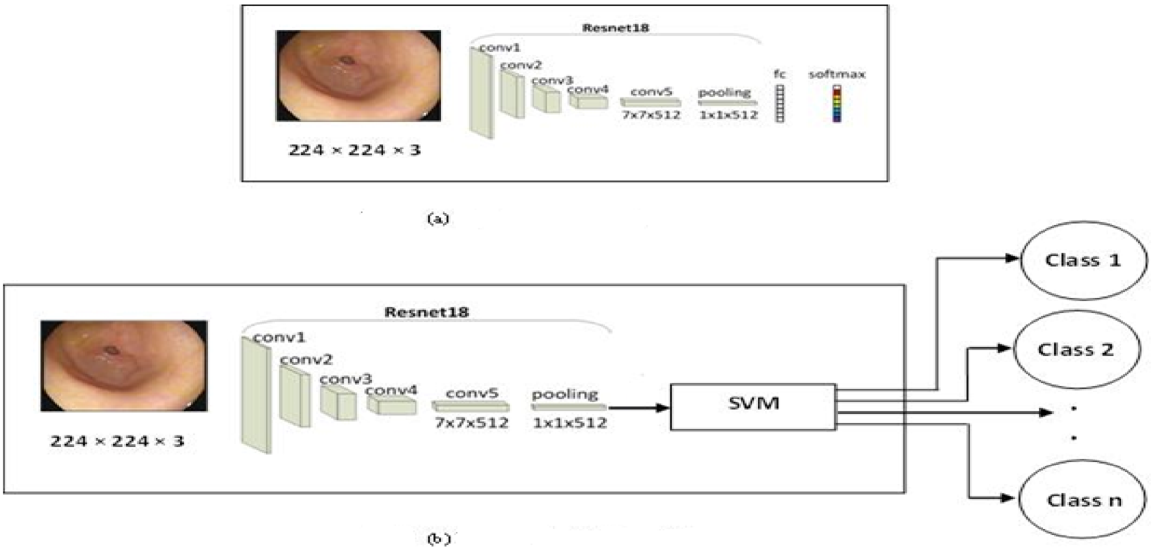

Fig. 1 (a) ResNet18 pre-trained CNN **(b)** Customized proposed classification model

Successful using of SVM in mammogram medical images found in [18-21] .Comparisons of many models for the support vector machine SVM are to be found in [22]. In their research, they compared and evaluated different models of SVM for use in proper applications. They highlighted the usage of SVM for large datasets, unbalanced and multi-classes. This study proposes a medical image classification model which combines ResNet-18 pre-trained CNNs at the features extraction stage and multiple classes support vector machineSVM as the main classifier. The proposed model will be used to classify medical images from two types of diseases, gastrointestinal and skin lesion.

3. Proposed Methodology

ResNet-18, a convolutional neural network that is 18 layers deep, was developed by [23]. Like other types of pre-trained convolution neural network, it was trained on more than a million images from the ImageNet database [24]. This pre-trained neural network can be used for both classifications based on the transfer learning concept or for feature extraction, as in this study. In this study, the ResNet-18 pre-trained neural networks are used for feature extraction. Then, the features extracted from the training images are used as predictor variables and fit a multiclass support vector machine. Finally, during the classification or test stage, the trained model of the multiclass support vector machine is used to classify the test images based on features extracted from the test images subset. The framework of our proposed method is shown in Fig.1. The original architecture of ResNet-18, as shown in Fig.1(a) above, has been used as a features extraction method after some tuning adjustment processes. The first layer, which represents the image input layer in the ResNet-18 architecture, requires input images of size 224-by-224-by-3, where 3 are the RGB colour channels. The layer name, output size and learnable of ResNet-18 are shown in Table 1. Here, all the input images are resized to match the network architecture input size, as shown in Fig.1 (b). Only the pools of 512 features are required to fit and train the multiclass support vector machine. Lastly, the output of our classifier depends on the number of classes in each of our datasets used here, which are three or eight, as we will explain in dataset description section. The following algorithm gives more detail for the two main steps of our proposed methodology. The total number of extracted features from each image is 512, while the number of training and testing rows for each dataset depends on splitting the percentages of data, which are 80% for training and 20% for testing. The final output of the algorithm is the predicted class image, which is either one out of eight in the Kvasir dataset or one out of three in the PH2 dataset. The main advantages of this type of pre-trained CNN and similar other types is their ability to extract highly accurate numeric features which then use for many purposes such as object recognition, content based image retrieval and image classification such as in this study. Here, all extracted features represent the image .descriptors for both query and databases images

```
Proposed Algorithm: Medical Image Classification

Input: Training set and testing set of images
Output: Prediction class for testing images
```

```
1: Start
2: Load the ResNet-18 pre-trained neural networks
3: Resize input images to match the input layer of ResNet-18
4: Adjust the output layer of ResNet-18 to match number of the dataset classes
5: Use generated features for training set to train the SVM
6: Use generated features for testing set to predict the image class
7: Calculate values of the classification measures
8: End   // Start
```

Table 1: Basic properties of ResNet-18 pre-trained convolutional CNN

Layer Name	*Output Size*	*Learnables*
conv1	112 × 112 × 64	7 × 7, 64, stride 2
conv2	56 × 56 × 64	3 × 3 max pool, stride 2 $\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
conv3	28 × 28 × 128	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
conv4	14 × 14 × 256	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
conv5	7 × 7 × 512	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$
average pool	1 × 1 × 512	7 × 7 average pool

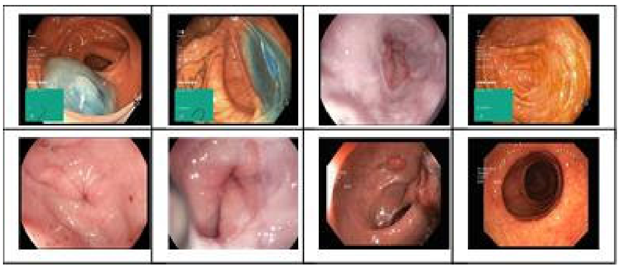

Fig. 2 Image sample from Kvasir medical images

Fig. 3 Image sample from PH2 medical images

4. Experimental Design

The classification method proposed here consists of two main stages, features extraction based on ResNet18 and the main classifier method using a support vector machine for multiple classes. The classification method is implemented on two medical image datasets, Kvasir medical images [25] and the PH2 dataset [26]. ResNet-18 is a pre-trained convolutional neural network model which is used as the base method for features extraction, as we explained in the above section, while the SVM is used as multiclass classification.

4.1 Medical Images Datasets

Two of most common medical datasets are used in this study. The first dataset used here is Kvasir medical images, which consists of 4,000 coloured endoscopic images annotated by medical experts. These images are arranged into eight groups, with 500 images in each class. The second dataset is a dermoscopic image database for research and benchmarking and has a total of 200 images (here only 120 images are taken and grouped into three classes). These two datasets have been used in many recent studies [27, 28]. They have been used for content-based medical image retrieval, as in [29, 30], and also in other studies for medical image classification and clustering purposes [31]. Samples of images for each class of the two datasets are shown in Fig2 and Fig3 respectively.

4.2 Evaluation Measures

For performance evaluation of the classification method proposed here, five evaluation measures are used: accuracy, sensitivity, specificity, precision and F1 score. The following equations describe each of these measures. For each of the following equations (TP is a true positive, TN is a true negative, FP is a false positive and FN is a false negative) and represents the pillar for each matric, as shown below :

$$SEN = TP/(TP+FN) \times 100 \quad (1)$$

$$SPE = TN/(TN+FP) \times 100 \quad (2)$$

$$ACC = (TP+TN)/(TP+TN+FP+FN) \times 100 \quad (3)$$

$$PRE = TP/(TP+FP) \times 100 \quad (4)$$

$$F1 = 2TP/(2TP+FP+FN) \quad (5)$$

Table 2: Confusion matrix for Kvasir dataset

	C1	C2	C3	C4	C5	C6	C7	C8
C1	188	10	0	0	0	0	1	1
C2	10	189	0	0	0	0	1	0
C3	0	0	186	0	1	13	0	0
C4	0	0	0	189	0	0	6	5
C5	0	0	2	0	195	3	0	0
C6	0	0	10	0	0	182	5	3
C7	0	1	1	6	0	0	183	9
C8	2	0	1	5	4	2	4	182

Table 3. Confusion matrix for PH2 dataset

	C1	C2	C3
C1	15	1	0
C2	1	14	1
C3	0	1	15

Table 4. Performance measures of Kvasir dataset

	ACC (%)	SEN (%)	SPE (%)	PRE (%)	F1-Score (0 – 1)
Mehtod1 [32]	90.86	71.6	97.0	88.6	0.79
Mehtod1 [33]	88.94	71.4	94.4	80.0	0.76
Mehtod1 [34]	91.0	71.0	N/A	88.1	0.74
Ours	**93.38**	**94.95**	**97.4**	**94.93**	**0.94**

Table 5. Performance measures of PH2 dataset

	ACC (%)	SEN (%)	SPE (%)	PRE (%)	F1-Score (0 – 1)
Mehtod1 [35]	90.2	90.5	**99.2**	92.1	0.89
Mehtod1 [36]	80.5	71.0	89,0	N/A	N/A
Mehtod1 [37]	87.2	54.7	95.0	75.1	0.87
Ours	**91.67**	**93.75**	90.63	**93.75**	**0.94**

4.3 Training and Testing Phases

Each collection of medical images is divided randomly into 60% for training and 40% for testing purposes. All images are resized to 224-by-224-by-3 to match the first layer of ResNet-18. The total of 512 features are extracted from each image represent image attributes input ResNet-18. For our two experiments, the batch size was set to 10, the number of epochs was 20 and the learning rate was set to 10e-4 for both datasets. The total numbers of iterations were 160 and 5,600 for the Kvasir and the PH2 dataset, respectively.

5. Results and Discussion

The main pillar of any classification measure is derived from the confusion matrix components. Tables 2 and 3 illustrate the confusion matrix for testing a portion of each dataset. The percentage of our dataset splitting is set to 60% for training and 40% for testing. The totals of test images are 200 for the Kvasir dataset and 16 for the PH2 dataset. As shown in Tables 2 and 3, of confusion matrices the diagonal element values represent the summation of true predicted classes. As seen, our proposed method performed better for both datasets. Overall accuracy is the most important metric, since it combines all four components of the confusion matrix. The overall accuracy as well as other performance evaluation metrics for the two datasets are shown in Tables 4 and 4. Observation from these metrics reflect the good performance of all these five measures, which were higher values. The results of this proposed method are compared with some state-of-the-art studies, as shown in the same tables. The proposed method outperformed some of these similar studies.

The performance measure values of our proposed method for two datasets as shown in Table 4 and Table 5 for the two datasets show that the proposed classification method performed better compared with three state-of-the-art methods for both datasets. For the Kvasir dataset, our classification method has best accuracy compared with the three methods and also most of the measures. Also, for the PH2 skin dataset our proposed classification method has acceptable performance measures compared with the three similar studies, as shown in Table 5.

6. Conclusion

This study successfully implemented ResNet18 pre-trained CNNs for the classification of two types of medical image to classify gastrointestinal and skin lesion images abnormality. Our proposed method performed best compared to some of the state-of-the-art studies, as shown in our findings and results. Pre-trained CNNs have accurate classification results because of their training process on a wide range and various types and the huge number of images. Their main disadvantages can be summarised as the memory needed and quite running time required. Pre-trained CNNs could be used for and applied to classification in three different ways, firstly for feature extraction, secondly to use the architecture of the pre-trained model and thirdly to train some layers while freezing others. In this study, we used the first scenario, so further studies could utilise the other two methods. For further studies researchers could try using more than one pre-trained convolution neural network models then a suitable method of merging such as mapping similar features from each model could proposed.

Acknowledgments

This research was funded by the Deanship of Scientific Research (DSR) at King Ab-dulaziz University, Jeddah, Saudi Arabia. The authors, therefore, gratefully acknowledge the DSR for their technical and financial support.

References

[1] Zhuang, F., Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong and Q. He., A comprehensive survey on transfer learning. Proceedings of the IEEE, 2020. 109(1): p. 43-76.

[2] Kandel, I. and M. Castelli, Transfer learning with convolutional neural networks for diabetic retinopathy image classification. A review. Applied Sciences, 2020. 10(6): p. 2021.

[3] Gopalakrishnan, K., S. K. Khaitan, A. Choudhary and A. Agrawal., Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. Construction and Building Materials, 2017. 157: p. 322-330.

[4] Pham, T.D., Classification of COVID-19 chest X-rays with deep learning: new models or fine tuning? Health Information Science and Systems, 2021. 9(1): p. 1-11.

[5] Islam, M. A., M. N. R. Shuvo, M. Shamsojjaman, S. Hasan, M. S. Hossain and T. Khatun., An Automated Convolutional Neural Network Based Approach for Paddy Leaf Disease Detection. International Journal of Advanced Computer Science and Applications,, 2021. 12(1): p. 280-288.

[6] Cuan, K., T. Zhang, J. Huang, C. Fang and Y. Guan., Detection of avian influenza-infected chickens based on a chicken sound convolutional neural network. Computers and Electronics in Agriculture, 2020. 178: p. 105688.

[7] Ayyachamy, S., V. Alex, M. Khened and G. Krishnamurthi. Medical image retrieval using Resnet-18. in Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications. 2019. International Society for Optics and Photonics.

[8] Ebrahimi, A., S. Luo, and R. Chiong. Introducing Transfer Leaming to 3D ResNet-18 for Alzheimer's Disease Detection on MRI Images. in 2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ). 2020. IEEE.

[9] Pham, T.D., A comprehensive study on classification of COVID-19 on computed tomography with pretrained convolutional neural networks. Scientific Reports, 2020. 10(1): p. 1-8.

[10] Ou, X., P. Yan, Y. Zhang, B. Tu, G. Zhang, J. Wu and W. Li., Moving object detection method via ResNet-18 with encoder–decoder structure in complex scenes. IEEE Access, 2019. 7: p. 108152-108160.

[11] Zhou, Y., F. Ren, S. Nishide and X. Kang. Facial Sentiment Classification Based on Resnet-18 Model. in 2019 International Conference on Electronic Engineering and Informatics (EEI). 2019. IEEE.

[12] Novitasari, D. C. R., R. Hendradi, R. E. Caraka, Y. Rachmawati, N. Z. Fanani, A. Syarifudin, T. Toharudin and R. C. Chen., Detection of covid-19 chest x-ray using support vector machine and convolutional neural network. Commun. Math. Biol. Neurosci., 2020. 2020: p. Article ID 42.

[13] Sharifrazi, D., R. Alizadehsani, M. Roshanzamir, J. H. Joloudari, A. Shoeibi, M. Jafari, S. Hussain, Z. A. Sani, F. Hasanzadeh and F. Khozeimeh., Fusion of convolution neural network, support vector machine and Sobel filter for accurate detection of COVID-19 patients using X-ray images. Biomedical Signal Processing and Control, 2021: p. 102622.

[14] Win, K. Y., N. Maneerat, K. Hamamoto and S. Sreng., Hybrid Learning of Hand-Crafted and Deep-Activated Features Using Particle Swarm Optimization and Optimized Support Vector Machine for Tuberculosis Screening. Applied Sciences, 2020. 10(17): p. 5749.

[15] Talo, M., O. Yildirim, U. B. Baloglu, G. Aydin and U. R. Acharya., Convolutional neural networks for multi-class brain disease detection using MRI images. Computerized Medical Imaging and Graphics, 2019. 78: p. 101673.

[16] Mahbod, A., G. Schaefer, C. Wang, R. Ecker and I. Ellinge. Skin lesion classification using hybrid deep neural networks. in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2019. IEEE.

[17] Kang, C., Y. Huo, L. Xin, B. Tian and B. Yu., Feature selection and tumor classification for microarray data using relaxed Lasso and generalized multi-class support vector machine. Journal of theoretical biology, 2019. 463: p. 77-91.

[18] Mohammed, A., A. Ahmed, W. Mohammed, G. Viju and M. Taha, Mammogram Images Classification using Linear Discriminant Analysis. 2020. International Research Journal of Engineering and Technology (IRJET), 2020. 7(6):p.6656-6662.

[19] Ahmed, A. and S. Malebary, Feature selection and the fusion-based method for enhancing the classification accuracy of SVM for breast cancer detection. Int. J. Comput. Sci. Netw. Secur., 2019. 19(11): p. 55.

[20] Ahmed, A. and O.B. El Sadig, Heterogeneous multi-classifier method based on weighted voting for Breast Cancer Detection. Int. J. Adv. Sci. Eng. Technol., 2019. 7(4): p. 36-41.

[21] Ibrahim, A. O., A. Ahmed, A. Abdu, R. Abd-alaziz, M. A. Alobeed, A. Y. Saleh and A. Elsafi , Classification of Mammogram Images Using Radial Basis Function Neural Network. in International Conference of Reliable Information and Communication Technology. 2019. Springer..p.311-320:

[22] Cervantes, J., F. Garcia-Lamont, L. Rodríguez-Mazahua and A. Lopez., A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing, 2020. 408: p. 189-215.

[23] He, K., X. Zhang, S. Ren and J. Sun. Deep residual learning for image recognition. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[24] ImageNet. http://www.image-net.org

[25] Dang-Nguyen, M. Lux and P. T. Schmidt. Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. in Proceedings of the 8th ACM on Multimedia Systems Conference. 2017.

[26] Mendoncÿa, T., P. Ferreira, J. Marques, A. Marcÿal and J. Rozeira., A dermoscopic image database for research and benchmarking. Presentation in proceedings of PH2 IEEE EMBC, 2013.

[27] Jha, D., P. H. Smedsrud, M. A. Riegler, P. Halvorsen, T. de Lange, D. Johansen and H. D. Johansen. Kvasir-seg: A segmented polyp dataset. in International Conference on Multimedia Modeling. 2020. Springer.

[28] Öztürk, Ş. and U. Özkaya, Residual LSTM layered CNN for classification of gastrointestinal tract diseases. Journal of Biomedical Informatics, 2021. 113: p. 103638.

[29] Ahmed, A. and S.J. Malebary, Query Expansion Based on Top-Ranked Images for Content-Based Medical Image Retrieval. IEEE Access, 2020. 8: p. 194541-194550.
[30] Ahmed, A., Implementing Relevance Feedback for Content-Based Medical Image Retrieval. IEEE Access, 2020. 8: p. 79969-79976.
[31] Zhong, J., W. Wang, H. Wu, Z. Wen and J. Qin. PolypSeg: An Efficient Context-Aware Network for Polyp Segmentation from Colonoscopy Videos. in International Conference on Medical Image Computing and Computer-Assisted Intervention. 2020. Springer.
[32] Ma, Y., X. Chen, and B. Sun. Polyp detection in colonoscopy videos by bootstrapping via temporal consistency. in 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). 2020. IEEE.
[33] Bernal, J., N. Tajkbaksh, F. J. Sánchez, B. J. Matuszewski, H. Chen, L. Yu, Q. Angermann, O. Romain, B. Rustad and I. Balasingham., Comparative validation of polyp detection methods in video colonoscopy: results from the MICCAI 2015 endoscopic vision challenge. IEEE transactions on medical imaging, 2017. 36(6): p. 1231-1249.
[34] Yu, L., H. Chen, Q. Dou, J. Qin and P. A. Heng., Integrating online and offline three-dimensional deep learning for automated polyp detection in colonoscopy videos. IEEE journal of biomedical and health informatics, 2016. 21(1): p. 65-75.
[35] Codella, N. C., D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra and H. Kittler. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). in 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). 2018. IEEE.
[36] Jianu, S.R.S., L. Ichim, and D. Popescu. Automatic diagnosis of skin cancer using neural networks. in 2019 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE). 2019. IEEE.
[37] Menegola, A., J. Tavares, M. Fornaciali, L. T. Li, S. Avila and E. Valle., RECOD titans at ISIC challenge 2017. arXiv preprint arXiv:1703.04819, 2017.

Ali Ahmed received his B.Sc. from Karary University (Sudan) in computer engineering and his M.Sc. degree in computer science, from Khartoum University (Sudan). He received his PhD and post-doctoral in computer science from UTM University (Malaysia). He has been worked as an assistant professor for more than four years (teaching and supervision master students). Now he is an associate professor in Faculty of Computing and Information Technology at King Abdulaziz University in KSA, Rabigh Branch. His research interests include data science, machine learning, data mining, information retrieval, image processing and retrieval. He is a member of IEEE, ACM and SEA Sudan".