Published in AJSE, Vol:20, Issue: 2
Received on 23th February 2021
Revised on 1st April 2021
Accepted on 15th May 2021

AIUB JOURNAL OF SCIENCE AND ENGINEERING
ISSN: 1608 – 3679 (print) 2520 – 4890 (Online)

Identification of Risk of Occurring Skin Cancer (Melanoma) Using Convolutional Neural Network (CNN)

Mayinuzaman Shawon, Anik Majumder, Abir Mahmud, Kazi F. Abedin and Mahbub C. Mishu

***Abstract*— Skin cancer is one of the most common malignancy in human, has drawn attention from researchers around the world. As skin cancer can turn into fatal if not treated in its earliest stages, the necessity of devising automated skin cancer diagnosis system that can automatically detect skin cancer efficiently in its earliest stage in a faster process than traditional one is of crucial importance. In this paper, a computer aided skin cancer diagnosis system based Convolutional Neural Network method has been shown. Our proposed system consists of five stages namely image acquisition, image preprocessing, image segmentation, feature extraction and classification We remove hair any noise from the images using dull then use median filter to smoothen the images. Next, k-means algorithm was applied for image segmentation on the preprocessed images. Finally, the segmented images were fed into CNN model for feature extraction and classification. The developed system can classify benign and melanoma type skin cancers from Dermoscopic images as accurate as 80.47%. While developing the skin cancer detection system, we compare accuracy score of other models such as Artificial Neural Network (ANN), K-Nearest Neighbor (KNN) and Random Forest with our proposed system. The proposed method has been tested on 'ISIC Challenge 2016' test dataset and an accuracy rate of 80.47% was obtained for accurately classifying benign and malignant skin lesions by our proposed model.**

***Index Terms*— Convolutional Neural Network (CNN), Dermoscopy, K-means, Skin cancer, Accuracy**

I. INTRODUCTION

Skin cancer is one of the most common types of cancer, especially among white people and the growth of skin cancer cases has reached epidemic proportions [1, 2]. Skin cancer foundation suggested that one in six people in the U.S. will develop skin cancer during their lifetime [2]. The main types of skin cancers are basal cell carcinoma (BCC), squamous cell carcinoma (SCC) and cutaneous malignant melanoma (CMM) (also simply referred to as 'melanoma') [3].

Mayinuzaman Shawon is a student at American International University Bangladesh (AIUB). (e-mail: mayinuzzaman@gmail.com)
Anik Majumder is a student at American International University Bangladesh (AIUB). (e-mail: anikmj162@gmail.com)
Abir Mahmud is a student at American International University Bangladesh (AIUB). (e-mail: mabirmahmud@gmail.com)
Kazi F. Abedin is a student at American International University Bangladesh (AIUB). (e-mail: fakhrul.shovon@gmail.com)
Mahbub C. Mishu is with American International University-Bangladesh (AIUB). (e-mail: mahbub@aiub.edu).

Among these types of skin cancers, melanoma is the deadliest one and accounts for 75% of all skin cancer related deaths [7]. Basal cell carcinoma and squamous cell carcinoma is also known as nonmelanocytic skin cancers (NMSC) or nonmelanoma skin cancer [2]. Deaths due to both melanoma and nonmelanoma skin cancers can be prevented significantly if both types of skin cancers can be diagnosed and treated in its earliest stage [2]. Thus, it is a matter of utmost importance for diagnosing skin cancer as early as possible.

Dermoscopy, a non-invasive technique for the microscopic examination of skin lesion, is a widely used method for diagnosing skin cancer [4]. Research have shown that the accuracy of detecting melanoma using dermoscopy is considerably higher than detecting melanoma by unaided observation [5]. But still trained dermatologists are required for correctly detecting skin cancer using dermoscopy. Even when experienced skin doctors use dermoscopy for diagnosing melanoma, the accuracy rate of detecting melanoma is about 75%-84% [6, 7]. Because of this, an alternative and more efficient solution for skin cancer diagnosis with higher accuracy is of vital importance.

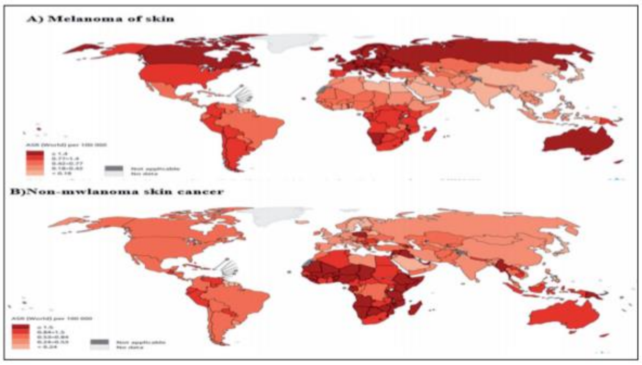

Figure 1. Mortality rate due to melanoma and non-melanoma skin cancer in different parts of the world in 2018 [18]

Many researchers have successfully developed computer aided skin cancer diagnosis systems using image processing and machine learning techniques that have exceeded the accuracy rate obtained by traditional diagnosis system by dermatologists as seen in the result of [8, 9, 10]. Automated skin cancer detection system typically consists of five stages: image acquisition, image preprocessing, image segmentation, feature extraction and finally classification.

In [11], a skin cancer detection system has been developed by H. T. Lau et al. to automatically classify the cancer images into benign or malignant melanoma. The authors used two classifiers namely Back-propagation neural network (BNN) and Auto-associative neural network (AANN) to classify skin cancers and achieved an overall result of 89.9% accuracy for back-propagation neural network and 80.8% accuracy for auto-associative neural network.

In [12] Andre Esteva et al. successfully classified melanoma skin lesions from dermoscopic images using a single convolutional neural network (CNN) that has been trained end-to-end from images directly using only pixels and disease labels as inputs. The authors have applied a GoogleNet Inceptionv3 CNN architecture that was pretrained on approximately 1.28 million images (1,000 object categories) and trained it on their dataset via transfer learning. The authors achieved 72.1±0.9% accuracy from their research output.

M. A. Taufiq et al. has developed a mobile enabled system for early detection of melanoma using support vector machine (SVM) and has obtained 80% accuracy for correctly classifying melanoma [13] whereas S. Alzahrani et al. used seven-point checklist with CNN and achieved an accuracy of 64.3% [14]. In [15], B. Chakradhar et al. detected malignancy on dermis using two decision tree-based algorithms: J48 and random forest classifier. The authors used watershed algorithm for segmentation and ABCD rule for features extraction. The authors achieved 73.66% accuracy for J48 and 85.35% accuracy for tree random forest classifier algorithm.

In this research, a computer aided skin cancer diagnosis system has been developed to automatically detect skin cancers from dermoscopic images and classifying them into either benign or malignant melanoma based on our proposed method.

II. METHODOLOGY

The proposed overall methodology for developing the computer aided skin cancer diagnosis system is given in figure 2.

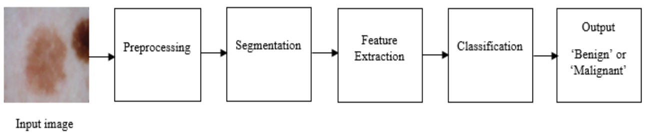

Figure 2. Block diagram of proposed methodology

In our proposed methodology the dermoscopic images are first preprocessed for noise and hair removal and then segmented using k-means algorithm. In next step, CNN is used for both feature extraction and classification. In output, the images are detected either as benign or malignant types. All of these steps are discussed in detail in the subsequent sections.

A. Image Preprocessing

Image preprocessing is an important step before classification that enhances the quality of an image by removing the noises and distortions from the background of the image which are necessary for correctly classifying images.

Figure 3. Dermoscopic image before (left) and after (right) preprocessing.

We have resized all our images into 200×200 pixels. Then remove hairs from the images, Dull Razor with three basic concepts has been used as discussed in [24]. The images were further smoothened using median filter. Histogram was plotted in the proposed skin cancer diagnosis system for getting a visual interpretation of the images (shown in figure 6). An image before and after preprocessing is shown in figure 3.

B. Image Segmentation

Segmentation of images before classification is an important step that can exceedingly affect the results of classification [19]. The main aim of segmentation is to partition of an input image into region. Image segmentation is a method that determines the shape and size of the border and isolate the object from its background based on various features extracted from the image [20, 21]. k-means algorithm has been used to extract the region of interest from the background and to define the boundary of skin lesion [20]. For segmenting the skin lesion from the non-lesion region of the image, two clusters have been used in the k-means algorithm. The output of k-means clustering segmentation is shown in figure 4.

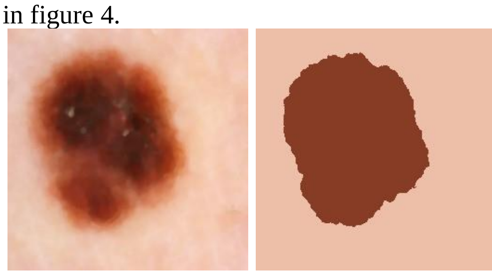

Figure 4. Dermoscopic image before (left) and after (right) image segmentation.

C. Feature Extraction & Classification

Convolutional Neural Network (CNN) is a specialized kind of neural network that has been developed for processing data which possess a grid-like topology, for example, image data that is thought to have 2D grid of pixels [22]. Because of this, CNN is considered a very efficient algorithm for image classification. Besides, one of the biggest advantages of CNN is that we do not need to manually perform feature extraction for classification because CNN can perform feature extraction by its convolution and pooling layers whilst the fully connected layer of CNN can perform the classification. Because of these, we have built a deep learning (DL) model by constructing a convolutional neural network for classifying 'benign' and 'malignant' skin lesions from a given input image for our computer aided skin cancer diagnosis system.

i. CNN Architecture

We have used three convolutional layers with filters 16, 32 and 64 for each layer respectively followed by three pooling layers where each pool size = (2, 2). Two dense layers have also been used for constructing the model where first dense layer consists of 128 neurons and the later one 64 neurons. Finally, the sigmoid activation function was used in the fully connected layer to classify the images into two categories: benign (0) and malignant (1). The proposed CNN model was trained through 100 epochs and binary cross-entropy was set as loss function. For training, ISIC Challenge 2016 training dataset [16] used. The dataset was split into 80:20 ratio where 80% images were used for training and remaining 20% images were used for validation test.

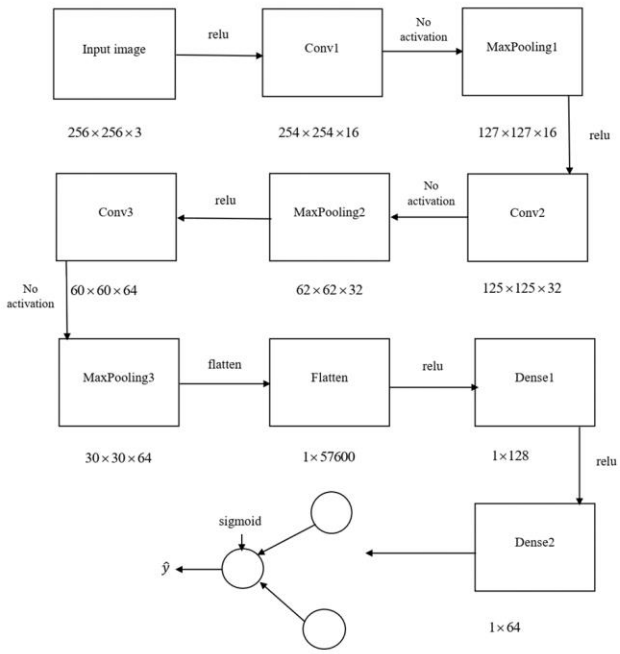

Figure 5. CNN architecture

ii. Mathematical Explanation of Proposed CNN Model

The architecture of our proposed CNN model of figure-5 is explained below with mathematical explanation that shows the different output sizes for different layers of our constructed CNN model. For determining the output sizes, the following formula has been used as discussed in [23],

$$n_{out} = \frac{(n_{in} + (2 \times p) - k)}{s} + 1 \tag{1}$$

were,

n_{in} = input, p = padding, k = kernel size, s = stride and

n_{out} = output

For conv2d_1,

Number of filters = 16, $s = 1$, $p = 0$, $k = 3$,

$$n_{out} = \frac{(256 + (2 \times 0) - 3)}{1} + 1 = 254$$

So, output size = $(254{,}254{,}16)$

For max_pooling2d_1,

$s = 2$, $p = 0$, $k = 2$,

$$n_{out} = \frac{(254 + (2 \times 0) - 2)}{2} + 1 = 127$$

So, output size = $(127{,}127{,}16)$

For conv2d_2,

Number of filters = 32, $s = 1$, $p = 0$, $k = 3$,

$$n_{out} = \frac{(127 + (2 \times 0) - 3)}{1} + 1 = 125$$

So, output size = $(125{,}125{,}32)$

For max_pooling2d_2,

$s = 2$, $p = 0$, $k = 2$,

$$n_{out} = \frac{(125 + (2 \times 0) - 2)}{2} + 1 = 62.5 \approx 62$$

So, output size = $(62{,}62{,}32)$

For conv2d_3,

Number of filters = 64, $s = 1$, $p = 0$, $k = 3$,

$$n_{out} = \frac{(62 + (2 \times 0) - 3)}{1} + 1 = 60$$

So, output size = $(60{,}60{,}64)$

For max_pooling2d_3,

$s = 2$, $p = 0$, $k = 2$,

$$n_{out} = \frac{(60 + (2 \times 0) - 2)}{2} + 1 = 30$$

So, output size = $(30{,}30{,}64)$

For flatten,

Output = $(30 \times 30 \times 64) = 57600$

III. RESULTS

The performance of our proposed model has been evaluated based on accuracy, sensitivity, and specificity. For this purpose, ISIC Challenge 2016 test dataset [16] was used and accuracy, sensitivity and specificity were determined based on confusion matrices as follows:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \tag{2}$$

$$Sensitivity = \frac{TP}{(TP + FN)} \tag{3}$$

$$Specificity = \frac{TN}{(TN + FP)} \tag{4}$$

where TP represents True Positive, TN represents True Negative, FP represents False Positive and FN represents False Negative.

In table 1, the performance of our proposed method has been compared with other existing state-of-the-art methods that also attempted to automatically detect skin cancer from skin lesion images.

TABLE I
UNITS FOR MAGNETIC PROPERTIES

Methods	Sensitivity	Specificity	Accuracy
MED-NODE Texture Descriptor [25]	0.62	0.85	0.76
ANN [17]	0.751	.831	0.791
KNN [26]	.6827	.6251	.6539
Random Forest [26]	.7685	.7179	.7432
Proposed	0.8194	0.5263	0.8047

From the above table, it is clearly seen that our proposed method gives highest accuracy result comparing to the existing methods to detect skin cancers from skin lesion images.

To deploy our trained deep learning model, we have developed a graphical user interface (GUI) that will automatically detect skin cancer. The system first asks the user to give a dermoscopic image as input and later preprocesses and segmented the image and prepare it for classification. Next, with the help of our trained model, it can successfully classify the images into either 'benign' or 'malignant' category as shown in figure 6.

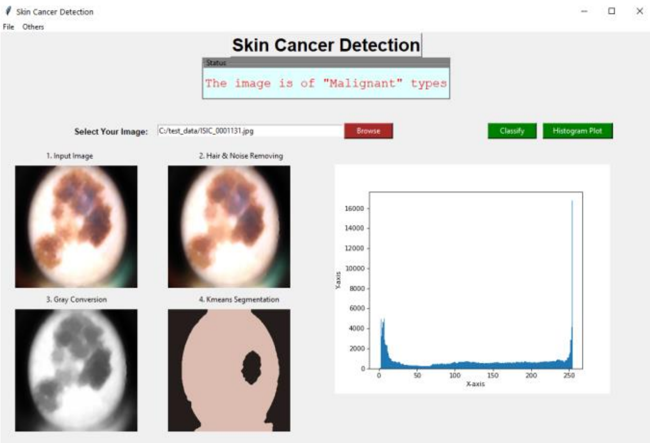

Figure 6. Detecting 'malignant' type skin cancer in our computer aided skin cancer diagnosis system

IV. CONCLUSIONS

In this paper, development of a skin cancer risk diagnosis system is shown based on CNN. The model is used for detecting the risk of skin cancer occurrence. In our developed system, an input of Dermoscopic image is provided after preprocessing and then k-means algorithm is applied for segmentation purposes. Next, the segmented image is fed to our proposed CNN model to classify the dermoscopic image either into benign or malignant type. Overall, we have got 52.63% specificity, 81.94% sensitivity, and 80.47% accuracy from our proposed model which is an improved result comparing with the ones yielded by previous methodologies demonstrated in table 1. Implementing machine learning method in segmentation and deep learning for feature extraction and classification are to be credited for the improved result comparing to the existing systems since most of the existing systems applied machine learning algorithms only for classification.

Our research is limited to only detecting 'benign' and 'malignant' types of skin cancer. In future, our aim will be to adding additional features of detecting more specific types of skin cancers than limiting it to only detecting benign and malignant types of skin cancers in our system in addition to improving the current accuracy rate of our proposed model.

REFERENCES

[1] M. C. F. Simões, J. J. S. Sousa, and A. A. C. C. Pais, "Skin cancer and new treatment perspectives: a review," Cancer Letters, vol. 357, no. 1, pp. 8–42, 2015.

[2] T.L. Diepgen, and V. Mahler, "The epidemiology of skin cancer." British Journal of Dermatology, vol. 146, no. 1, pp. 1-6, 2002.

[3] B.K. Armstrong, and A. Kricker, "The epidemiology of UV induced skin cancer", Journal of photochemistry and photobiology B, Biology, vol. 63, no. 1-3, pp. 8-18, October 2001.

[4] H. Kittler, H. Pehamberger, K. Wolff, and M. Binder, "Diagnostic accuracy of dermoscopy", The lancet oncology, vol. 3, pp. 159-165, 2002.

[5] Bafounta ML, Beauchet A, Aegerter P, Saiag P., "Is dermoscopy (epiluminescence microscopy) useful for the diagnosis of melanoma? Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests", Arch Dermatol, vol. 137, no. 10, pp. 1343-1350, 2001.

[6] G. Argenziano, H. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, and M. Binder, "Dermoscopy of pigmented skin lesions: Results of consensus meeting via the Internet", Journal of the American Academy of Dermatology, vol. 48, no. 5, pp. 679–693, May 2003.

[7] S. Jaina, V. jagtapb, N. Pise, "Computer aided Melanoma skin cancer detection using Image Processing", Procedia Computer Science, Volume 48, pp. 735-740, 2015.

[8] Mahmoud Elgamal, "Automatic Skin Cancer Images Classification", International Journal of Advanced Computer Science and Applications,Vol. 4, No. 3, 2013.

[9] Z. Waheed, A. Waheed, M. Zafar and F. Riaz, "An efficient machine learning approach for the detection of melanoma using dermoscopic images," 2017 International Conference on Communication, Computing and Digital Systems (C-CODE), Islamabad, 2017, pp. 316-319.

[10] D. Choudhury, A. Naug and S. Ghosh, "Texture and color feature based WLS framework aided skin cancer classification using MSVM and ELM," 2015 Annual IEEE India Conference (INDICON), New Delhi, 2015, pp. 1-6.

[11] H. T. Lau and A. Al-Jumaily, "Automatically Early Detection of Skin Cancer: Study Based on Nueral Netwok Classification," 2009 International Conference of Soft Computing and Pattern Recognition, Malacca, 2009, pp. 375-380.

[12] Esteva, A., Kuprel, B., Novoa, R. et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, pp. 115–118, January 2017.

[13] M.A. Taufiq, N. Hameed, A. Anjum, F. Hameed, "m-Skin Doctor: A Mobile Enabled System for Early Melanoma Skin Cancer Detection Using Support Vector Machine," In: Giokas K., Bokor L., Hopfgartner F. (eds) eHealth 360°, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Springer, Cham, vol. 181, pp. 468-475, 2017.

[14] S. Alzahrani, W. Al-Nuaimy and B. Al-Bander, "Seven-Point Checklist with Convolutional Neural Networks for Melanoma Diagnosis," 2019 8th European Workshop on Visual Information Processing (EUVIP), Roma, Italy, 2019, pp. 211-216.

[15] B. Chakradhar, I. S. Siva Rao, V. Jhansy Archana and C. V. M. K. Hari, "Detection of Malignancy On Dermis Using J48 and Random Forest Classifiers," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), Gunupur, India, 2020, pp. 1-6.

[16] https://challenge.isic-archive.com/data

[17] E.Zagrouba andW.Barhoumi, "A preliminary approach for the automated recognition of malignant melanoma," Image Analysis & Stereology, vol. 23, no. 2, pp. 121-135, 2004.

[18] Khazaei Z., Ghorat F., Jarrahi A. M., Adineh H. A., Sohrabivafa M., Goodarzi E., "Global incidence and mortality of skin cancer by histological subtype and its relationship with the human development index (HDI); an ecology study in 2018," WCRJ 2019; 6: e1265

[19] A. N. Hoshyar, A. Al-Jumaily, and R. Sulaiman, "Review on automatic early skin cancer detection," in 2011 International Conference on Computer Science and Service System (CSSS). IEEE, 2011, pp. 4036–4039.

[20] Neoh, S.C., Srisukkham, W., Zhang, L, Todryk, S., Greystoke, B., Lim, C.P., Hossain, A. And Aslam, N., "An Intelligent Decision Support System for Leukaemia Diagnosis using Microscopic Blood Images," Scientific Reports 5 (14938), Nature, October 2015

[21] Zhang, L., Jiang, M., Farid, D. And Hossain, A.M., "Intelligent Facial Emotion Recognition and Semantic-based Topic Detection for a Humanoid Robot," Expert Systems with Applications, Elsevier, vol. 40, no. 13, pp. 5160- 5168, October 2013.

[22] I. Goodfellow and Y. Bengio and A. Courville, *Deep Learning*. Cambridge, MA: The MIT Press, 2016, ch. 9.
[23] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), Antalya, 2017, pp. 1-6.
[24] Lee T, Ng V, Gallagher R, Coldman A, McLean D. DullRazor: a software approach to hair removal from images. Computers in Biology and Medicine, Elsivier, vol. 27, no. 6, pp. 533-43, November 1997.
[25] I. Giotis, N. Molders, S. Land, M. Biehl, M. F. Jonkman and N. Petkov, "MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images," Expert Systems with Applications, Elsevier, vol. 42, no. 19, pp. 6578-6585, November 2015.
[26] Murugan, A., Nair, S.H. & Kumar, K.P.S., "Detection of Skin Cancer Using SVM, Random Forest and kNN Classifiers," J Med Syst, vol. 43, no. 7:269, July 2019.

Mayinuzaman Shawon was born on November 17, 1999 in Gopalganj, Bangladesh. He has completed his B.Sc. in Computer Science & Engineering in 2021 from American International University- Bangladesh (AIUB). Currently, he is working as a software developer at Periscope Labs Limited, Dhaka. His research interests include Artificial Intelligence, Machine Learning, Computer Vision, Image Processing, and Data Mining.

Anik Majumder was born on January 1, 1999 in Chandpur, Chattogram, Bangladesh. He has completed his B.Sc. in Computer Science and Engineering in 2021 from American International University-Bangladesh (AIUB). His research interests and passion include Computer Vision, Machine Learning, Data Analysis, Artificial Intelligence and Algorithms.

Abir Mahmud was born in Mymensingh, Bangladesh on 10 October 1997. He has completed his B.Sc. in Computer Science and Engineering in 2021 from American International University-Bangladesh. He received awards and honors which includes "Dean's List Honors" from American International University-Bangladesh. His research interests are Data science, Machine Learning and algorithms.

Kazi Fakhrul abedin was born on May 23, 1998 in Munshiganj, Bangladesh. He has completed his B.Sc. in Computer Science and Engineering in 2021 from American International University-Bangladesh. His research interests and passion are mostly based on Computer Vision, Image Processing, Data Mining, Artificial Intelligence, and Machine Learning

Mahbub C. Mishu received his PhD from Bournemouth University, UK in 2015 and MSc from the University of Nottingham, UK in 2010. His research interests are mainly focused on Intelligent System Design, Machine Learning and Big Data Analytics. After his PhD, he joined Northern Regional College, UK as a Lecturer. He also obtained Higher Education Teaching and Learning Qualification from Ulster University, UK in 2017. He has started working as an Assistant Professor at American International University-Bangladesh since January 2018. Dr Mishu has published several articles in local and international peer reviewed journals and conferences,