Arabic Braille Numeral Recognition Using Convolutional Neural Networks

Shurouq Alufaisan, Wafa Albur, Shaikha Alsedrah, and Ghazanfar Latif

Abstract Braille is a system that is designed to assist visually impaired individuals to acquire information. It consists of raised dots arranged in a cell of three rows and two columns. Visually impaired individuals rely on the sense of touch to read and write. However, it is difficult to memorize the arrangement of dots that compose a character. This research aims to design an application that recognizes and detects Arabic braille numerals and convert it to plain text and speech by implementing convolutional neural network variation Residual Network (ResNet). A new dataset was collected by capturing Arabic braille numerals using smartphone cameras. The recognition accuracy for Arabic braille numerals achieved 98%, taking into accountability different light and distance conditions.

Keywords Braille recognition · Deep learning · Arabic braille numerals classification · Convolutional neural network · Residual network

1 Introduction

Currently, everywhere in the world operate with the data as it is the most valuable part of our society, simply known as information. Collecting data successfully is the first step to operate effectively and make decisions efficiently. Sharing knowledge can be interpreted as communication, and it is one way of acquiring the needed information. In addition, reading is an important factor to learn and to obtain the information needed to prosper in our society. However, for visually impaired and deaf individuals, it is impossible to acquire information with plain texts, and there is no means of sharing information through communication with sighted individuals [1]. Therefore, a system, known as the braille system, was designed for visually impaired groups of individuals to access and receive information. With such a system, it is now possible to make decisions, operate upon events effectively, and communicate with sighted individuals.

S. Alufaisan · W. Albur · S. Alsedrah · G. Latif (✉)
College of Computer Engineering and Sciences, Prince Mohammad Bin Fahd University, Al Khobar, Saudi Arabia
e-mail: glatif@pmu.edu.sa

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_7

Braille system can be used as a mean of communication to share information. It is used for sighted individuals who wish to communicate with visually impaired individuals through written informational communication. Visually impaired individuals may prosper in our society and can play a significant role in our world. Thus, it is important to have a mean of communication between sight and visually impaired individuals to share information and to learn from each other.

Braille is a system that depends on the sense of touch. It is a system that enables visually impaired individuals to write and read with the help of the touch sense. It uses a series of raised dots that are used to read with the personal sense of touch. The language consists of six dots arranged in a rectangular shape. The dots can be arranged in any of the six position to compose a word, letter, number, or a special character. Braille system can be used to write different languages such as Arabic or English. In addition, the system can be used to write musical compositions and mathematical notations. Reading braille texts, for both Arabic and English, are read by moving the index finger from left to right.

Reading braille language with no previous experience can be difficult at first. To be able to read it effectively, the sense of touch must be trained first while making all the other senses unused. Also, users must memorize the positions of the dots and make sure what each composition of dots mean. Of course, not all individuals were born visually impaired. Some might go through events in life that make a person loses his/her sight, such as a chemical accident. In addition, there are situations, where a person might want to learn the language to teach it or to communicate with a visually impaired individual. Being a beginner in learning, this system can be hard at first. Ideally, if there exists a technology that can make reading and writing braille language easier, then it would be beneficial for the users to learn faster and to acquire the information that is needed efficiently. Such a system would have to translate the braille language into text and speech to be familiar with the system.

Therefore, the objective of this research is to develop an Arabic braille numeral system detection and recognition based on a deep learning algorithm. The proposed system will have the ability to recognize Arabic braille numeral images captured by cameras and to be processed by CNN variation, residual network [2, 3]. The goal of this work is to assist a person in reading and learning Arabic braille numeral. The ultimate goal is to have a pi camera that is integrated with the application, to scan and recognize printed Arabic braille numeral to translate it and display it to the user in plain text and speech [4–6]. This aims to improve the learning process of the Arabic braille system to acquire the required information for visually impaired users which lead to improve their daily life activities [7]. In addition, to help those who wish to communicate with visually impaired individuals.

2 Literature Review

Classifying and recognizing Arabic braille are significant to assist visually impaired individuals to learn and to obtain the necessary knowledge. In addition, to assist who wishes to communicate with a visually impaired individual. Extensive research has been conducted in the area of classifying and recognizing Arabic braille scripts. In [8], the authors proposed the use of find contour and artificial neural network for braille character recognition. Their method consists of preprocessing the image to prepare it for the process of the finding contour to get the black dots on the image for different datasets consisting of tilted images. The authors used segmentation to read the area of the braille cell. Later on, the artificial neural network was conducted as the final step to have the system learn by feeding it data input to obtain the desired data output value. The method achieved 99% of titled images of $-1°$ to $1°$. The authors showed that the level of accuracy decreased when the image is tilted more than $1°$.

In [9], the author's dataset consisted of braille documents with the color of green and yellow that included dots in one side of the document. The author's method was to use image processing techniques to recognize Arabic braille characters. The authors preprocessing steps include: converting the image to grayscale, filtering the image, applying local thresholding for green braille documents, applying adaptive thresholding for yellow braille documents, segmentation, and extracting features. The authors described that Arabic braille characters were successfully recognized. Afterward, the authors proposed to convert Arabic braille character to binary strings which are converted into ASCII code to obtain the correct Arabic translation. The method achieved 98.04–100% for green braille document and 97.08–99.65% for yellow braille document. In [10], the authors proposed performing image preprocessing techniques to prepare the image for feature extraction. In the feature extraction step, the authors computed centroids of dots in the image to extract the relevant information from the image. Afterward, the authors aligned the coordination by applying many operations to rotate the centroids to align the page and the braille dots. Finally, the authors were able to recognize braille cells by grouping the dots to acquire a combination of letters and words. The method achieved between 94 and 99% braille cell recognition accuracy. In [11], the authors proposed converting the scanned document of braille page to gray color. After converting the image, the authors used the threshold technique to obtain three classed of regions. The authors used the three classes to initially identify braille dots. The possibility of valid dot identification was used in braille cell recognition. The method achieved 99% accuracy for skewed, reversed, or workout braille document.

In [12], the authors proposed classifying Arabic braille character using fuzzy classification, character feature extraction, and character recognition. The authors proposed system was developed with the use of segmenting Arabic braille characters. The authors use of fuzzy classification was inspired by the Fuzzy C-Mean (FCM) and fuzzy KNN classification algorithms. The method achieved up to 83% accuracy of classification and recognition of braille character. In [13], the authors suggest using a text-matching method as a way to recognize images. This system implies that starting with observing the interaction between words, then the use of several matching patterns between the phrases and ending up with matching entire sentences. This paper tried two methods to see what gives the highest accuracy, one of which was using paper citation matching, where the authors used a large academic dataset along with their citations and abstracts. The dataset sized 838,908 instances in total, containing 279,636 positive instance pairs and 559,272 negative instance pairs. The negative pairs were selected randomly, where they do not have citations along with them. Moreover, one out of three models that the authors trained outed the other two models. The three models being MP-IND, MP-COS, and MP-DOT, where MP-DOT was the model that gave the highest accuracy with a percentage of 88.73. In [14] seen that camera-enabled smartphone was used as the main method to capture braille characters. The paper suggested an algorithm that manipulates the images of braille documents that interpret the document's highlights and convert them to their equivalent English characters. Taking into account the lighting conditions while capturing the images, the authors obtained an accuracy of over 80% by developing an application that works under Android platform. The authors also stated that under the right conditions while capturing braille texts, the accuracy can go up to 84%.

In [15], the authors proposed a system for recognition double-sided Amharic braille documents that use the identification of three methods. Those methods are recto, verso, and overlapping dots. The system that the authors suggested in this paper works by simply converting braille texts into codes, and those codes are later on being translated into texts again. On top of that, adding the concept of reflection to reverse wrongly scanned braille documents automatically. While the dataset was collected from Addis Ababa University's Kennedy Library that contains good and bad scanned braille documents, the system was evaluated to give a high accuracy of 99.3% for identification and accuracy of 95.6% for translation. The authors in [16] used optical braille recognition based on semantic segmentation network along with auxiliary learning strategies. Using the OBR framework along with BraUNet and morphological post-processing procedures, the authors also used corresponding pixel-level annotations of braille characters with 64 other classes as an input in the system for both training and testing. The results of the methods they applied on DSBI dataset, type recto braille character along with BraUNet gave the best results with an accuracy of 0.9966%. While the regular number of classes for braille classes is 64 the authors in [17] took another approach and added 7 more classes to add up to 71 classes of characters to corresponding to the braille dataset that consists of 37 characters. Moreover, a collective dataset of 26,724 labeled braille images now has 37 braille symbols that correspond to the 71 classes. Using a novel method that pairs ratio character segmentation, RCSA's algorithm was used aside with CNN to translate a line of braille into its English counterpart. By linking the CNN model to two recognition techniques: character and word recognition the system proposed in this paper were able to reach an accuracy of 98.73% on the test set. In [18], the authors used Convolutional Neural Network (CNN) techniques to develop a system that can identify Cyrillic braille characters. After scanning the braille documents, image preprocessing techniques were used to make the recognition process easier. Then, character segmentation was performed to improve recognition accuracy. Subsequently, a modified backpropagation algorithm was used to train neural networks. The system has achieved 95.7% training accuracy and 95% testing accuracy. The authors also concluded that the use of the artificial neural network is very helpful in identifying characters due to the ease of programming the network architecture to train and test with any image sizes as an input. In [19], the authors proposed a module that uses associative memories to recognize single-sided braille documents and then convert it to audio. Their module consisted of two stages, preprocessing and recognition. In the preprocessing stage, different operations are performed on the scanned braille papers such as grayscale conversion and dilation to prepare the images for the next stage. Then, Modify Multi-Connect Architecture (MMCA) and Modify Bidirectional Associative Memory (MBAM) algorithms are used to recognize the characters. The authors compared results of MMCA and MBAM algorithms, where (MMCA) achieved an average accuracy of 98.26% for characters recognition while (MBAM) achieved 91.87%. Afterward, the proposed module converts the recognized text into audio.

In [20], the authors proposed developed an Optical Braille Translator (OBT) system that identifies single-sided Sinhala braille characters and translates it into texts. The systems features were developed based on image processing techniques in MATLAB environment. Their methodology was to apply preprocessing techniques like grayscale conversion, image rescaling, and angle correction functions on both handwritten and computer-generated braille documents. Then, segmentations are used to recognize the braille character cells. Afterward, extracted characters are resized into a 21×16 matrix binary images. Braille characters are regenerated using an algorithm to improve accuracy. The developed system was able to reach an accuracy of over 99%. In [21], the authors proposed a system equipped with a scanner and a webcam. The braille characters are captured by the webcam or the scanner. Then, grayscale conversion, filtering, segmentation, and normalization are done on the captured images using MATLAB IDE. After the image processing stage, Artificial Neural Networks (ANN) and feature extraction are used to recognize the patterns of braille characters as well as obtaining a training model as a dataset. Furthermore, the recognized characters are then converted to audio. The system results showed that the characters captured with webcam resulted in 57.69% accuracy while the characters captured using the scanner resulted in an average of 93.26% accuracy. In [14], the authors proposed a method to process braille documents and convert them into English text and speech. Their method performs Hough's circle detection algorithm on the phone captured braille images to identify the dots. Then, maximum length and width are calculated for all dots to generate a new image of equaled size dots. After that, the authors used image segmentation to divide the rows and columns into cells. Later on, each cell in a row is read to recognize its pattern by the position of its dots. Finally, each cell pattern is mapped to its matched English character. The output text is converted into speech using a text-to-speech engine. In ideal lighting and alignment conditions, the method achieved more than 80% accuracy.

3 Methodology

The overall aim of this research is to design an application that uses deep learning techniques that recognize Arabic braille numerals with high accuracy. Captured images of Arabic braille numerals are used to train the CNN variation Residual Network (ResNet). The model was modified to achieve high accuracy of recognizing Arabic braille numerals.

Training a deep network is a challenge, and it has been proven that the depth of a network degrades the network performance [22, 23]. To address this problem, ResNet is chosen as the building block of our network due to its methods of training a deep network efficiently. Adding more layers to the network leads to the vanishing gradient issue which shows high training error. Thus, the authors in [24] suggested adding skip connections that skip one or more layers. The authors in [24] proved that when the model uses skip connection, the training efficiency improved since the gradients can travel to many layers with the help of skip connections. The proposed model is implemented using python libraries that will accept images and classify them to its respective classes. The application will be able to detect and recognize Arabic braille numerals using a pi camera that will be integrated with raspberry pi 4. Once the model is trained, captured images of Arabic braille numerals will be classified to its respective classes, provide a correct translation of it and then convert it to speech audio representing the numeral that was processed.

3.1 Data Description

Arabic braille numerals were printed on a single side A4 embossed paper with color blue and white documents. Braille dots are arranged in a cell with three rows and two columns. Dot height is approximately 0.02 inches, and the horizontal and vertical spacing between the dots are 0.1 inches. The blank space between the cells is 1.3 inches.

As Table 1 shows, 5000 images of Arabic braille numerals were collected. Smartphone cameras are used to generate the dataset. The images are captured under different lights, such as natural and industrial lighting. The images range from several colors, white, yellow, and gray. Images were captured on all possible angles and heights.

3.2 Preprocessing

In this stage, images are being prepared to make it easier for the model to train and to recognize braille numerals.

Table 1 Arabic braille numerals sample dataset

#	Name in Arabic	Name in English	Sample image	Number of images
0	صفر	Sifer		500
1	واحد	Wahed		500
2	اثنين	Ethnein		500

(continued)

Table 1 (continued)

#	Name in Arabic	Name in English	Sample image	Number of images
3	ثلاثة	Thalatha		500
4	اربعة	Arba-a		500
5	خمسة	Khamsa		500

(continued)

Table 1 (continued)

#	Name in Arabic	Name in English	Sample image	Number of images
6	ستة	Sitta		500
7	سبعة	Sab-a		500
8	ثمانية	Thamanya		500

(continued)

Table 1 (continued)

#	Name in Arabic	Name in English	Sample image	Number of images
9	تسعة	Tis-a		500

3.2.1 Grayscaling

Computers identify RGB images as 3D arrays. Meanwhile, grayscale images are identified as 2D arrays, which make the preprocessing stage more efficient. Thus, captured RGB images of braille numerals are converted to grayscale.

3.2.2 Resizing

Training images of larger size will add computational power to the network as well as time complexity. Therefore, images are resized to 256, along the y and x-axis, before inputting them into ResNet. With this approach, the model will be able to train faster with less pixels, while preserving important features.

3.2.3 Converting to Array

In deep learning, the model is trained with images in a NumPy array form. Thus, grayscaled images will be converted to an array using the NumPy library. Later on, the arrays will be fed to the network to start the training.

3.3 Deep Residual Learning

The depth of a network is significant in any neural network model [25]. However, the deeper the network the harder it is to train the model. It is noticed that as the network gets deeper, there will be a degradation problem which will lead to a decrease in the validation set. In residual block, the layers would be fed into the next layer and the layer after it, while having the advantage to skip the number of layers known as skip connection. The diagram below (Fig. 1) illustrates the skip connection.

The diagram on the right describes a deep network that shows stacked layers one after the other. However, the diagram on the left describes a deep network with stacked layers as before but the original input is now added to the output of the block. This operation is called a skip connection. Skipping the training of a few layers can be beneficial in solving the degradation problem that affects the accuracy negatively due to having more layers than needed for training.

In [17], the authors suggested using the building block, which is defined as:

$$y = F(x, \{W_i\} + x \tag{1}$$

The above equation is the shortcut connection which does not hold extra parameters and computation complexity. The identity mapping is responsible for adding the output from the previous layer to the next layer. Input and output vectors are represented as x and y of the considered layers. $F(x, \{W_i\}$ is the function that is used

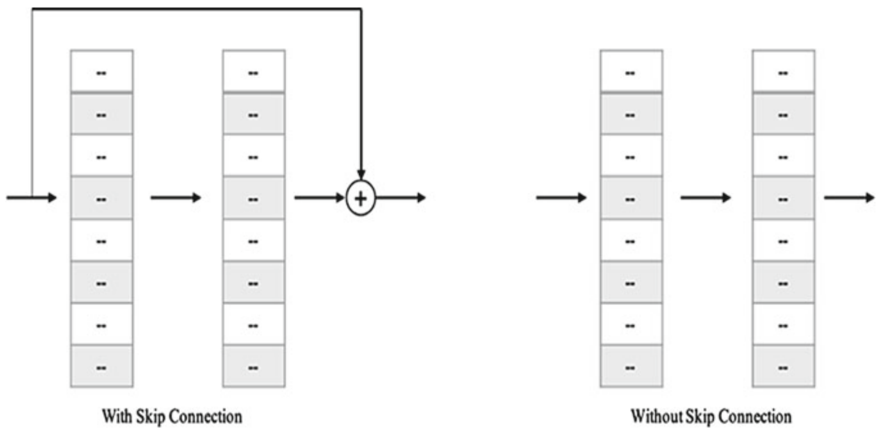

Fig. 1 Skip connection concept

for the residual mapping that will be learned. It is a requirement for the input and output to have equal dimensions to use the shortcut connection equation. Also, in [17], the authors discussed a solution in case when the dimensions are not equal that is linear projection. If the model found that the dimensions are not equal, then it will have an operation which will multiply a linear projection W to the identity mapping to match the dimensions. The equation is defined as:

$$y = F(x, \{W_i\} + W_s x \tag{2}$$

while identity mapping is used to solve the degradation problem multiplying is W_s used when there is a need to match the dimensions [17].

4 Results and Discussion

In this section, the performance of CNN was examined for training and recognizing Arabic braille numerals. A modified ResNet architecture model was used to implement our system. The dataset has been divided based on their corresponding labels to guarantee that each image will be classified based on their labels. The model was trained for 120 epochs on the braille numerals dataset, which means training for 10 classes for 120 cycles. 20% of the dataset images were used for validation and testing. This means that 3000 images were used for training, 1000 for validation, and 1000 for testing. The results using the ResNet model were extremely high, the model achieved a validation accuracy of 98%, as long with high model performance. Figure 2 shows the validation accuracy and test accuracy.

Table 2 shows a comparison between the results of the first and last epoch. The validation has increased from 0.40 to 0.98 as the training has progressed.

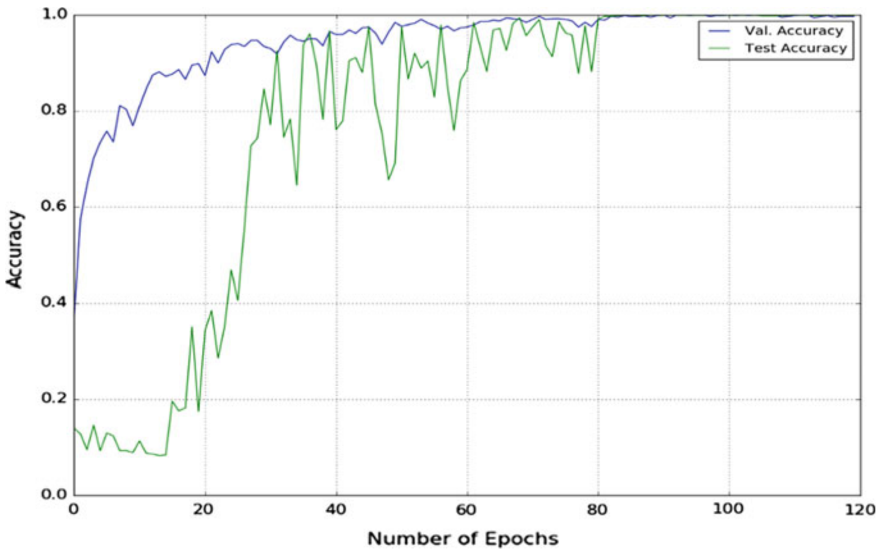

Fig. 2 Model validation and testing accuracy curve

Table 2 First and last epoch results

Epoch no.	Accuracy	
	Test	Val
First epoch	0.14	0.40
Last epoch	0.99	0.98

From Fig. 2 can be observed that the test accuracy has dramatically increased after epoch no. 25.

5 Conclusion

In this article, Arabic braille recognition using deep learning techniques was proposed. Braille is a system that is designed to assist visually impaired individuals to acquire information. Visually impaired individuals rely on the sense of touch to read and write. The proposed system can help teachers, parents, and people who have lost their vision recently. The system uses a modified ResNet model, which overcomes the degradation problem. A validation accuracy of 98% and a test accuracy of 99% were able to achieve. Deep learning techniques have proved their efficiency in recognizing braille characters.

References

1. Latif G, Mohammad N, AlKhalaf R, AlKhalaf R, Alghazo J, Khan M (2020) An automatic arabic sign language recognition system based on deep CNN: an assistive system for the deaf and hard of hearing. Int J Comput Dig Syst 9(4):715–724
2. Alghmgham DA, Latif G, Alghazo J, Alzubaidi L (2019) Autonomous traffic sign (ATSR) detection and recognition using deep CNN. Procedia Comput Sci 163:266–274
3. Shaikh E, Mohiuddin I, Manzoor A, Latif G, Mohammad N (2019) Automated grading for handwritten answer sheets using convolutional neural networks. In: 2019 2nd International conference on new trends in computing sciences (ICTCS), pp 1–6. IEEE, Oct 2019
4. Mahmoud AA, Alawadh INA, Latif G, Alghazo J (2020) Smart nursery for smart cities: ınfant sound classification based on novel features and support vector classifier. In: 2020 7th ınternational conference on electrical and electronics engineering (ICEEE), pp 47–52. IEEE, Apr 2020
5. Al-Hmouz A, Latif G, Alghazo J, Al-Hmouz R (2020) Enhanced numeral recognition for handwritten multi-language numerals using fuzzy set-based decision mechanism. Int J Mach Learn Comput 10(1)
6. Alzubaidi L, Latif G, Alghazo JM, Zikria M (2019) Cloud-based ınteractive hands free e-learning environment for students with disabilities
7. Latif G, Alghazo JM, Maheswar R, Jayarajan P, Sampathkumar A (2020) Impact of IoT-based smart cities on human daily life. In: Integration of WSN and IoT for smart cities. Springer, Cham, pp 103–114
8. Subur J, Sardjono TA, Mardiyanto R (2016) Braille character recognition using find contour and artificial neural network. J Electr Electron Eng 14(1)
9. Authman ZI, Jebr ZF (2012) Arabic Braille scripts recognition and translation using image processing techniques. J Educ Pure Sci 2(3):18–26
10. Mousa A, Hiary H, Alomari R, Alnemer L (2013) Download limit exceeded, citeseerx.ist.psu.edu, Nov 2013. Available: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.403.6856&rep=rep1&type=pdf.
11. Al-Salman A, AlOhali Y, AlKanhal M, AlRajih A (2007) An Arabic optical braille recognition system
12. Al Nassiri A, Abdulla S (2018) A fuzzy classification and recognition system for Arabic Braille segmented characters. Int J Appl Eng Res 13(6):3662–3669
13. Pang L, Lan Y, Guo J, Xu J, Wan S, Cheng X (2016) Text matching as image recognition. arXiv preprint arXiv:1602.06359
14. Venugopal-Wairagade GA (2016) Braille recognition using a camera-enabled smartphone. Int J Eng Manuf 6(4):32–39
15. Seid H, Assabie Y (2017) Recognition of double sided Amharic Braille documents recognition of double sided Amharic Braille documents view project NLP tools for Ethiopian languages view project recognition of double sided Amharic Braille documents. Int J Image Graph Sig Process 4:1–9
16. Li R, Liu H, Wang X, Xu J, Qian Y (2020) Optical Braille recognition based on semantic segmentation network with auxiliary learning strategy. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp 554–555
17. Hsu B-M (2020) Braille Recognition for reducing asymmetric communication between the blind and non-blind. Symmetry 12(7):1069
18. Smelyakov K, Chupryna A, Yeremenko D, Sakhon A, Polezhai V (2018) Braille character recognition based on neural networks. In: 2018 IEEE second ınternational conference on data stream mining & processing (DSMP), pp 509–513. IEEE, Aug 2018
19. Khaled S, Safana H, Abbas (2017) Braille character recognition using associative memories. Int J Eng Res Adv Technol (IJERAT) 1:2454–6135
20. Perera TDSH, Wanniarachchi WKIL (2018) Optical Braille translator for Sinhala Braille system: paper communication tool between vision ımpaired and sighted persons. Int J Multimed Its Appl (IJMA) 10

21. Ramiati, Aulia S, Lifwarda, Nindya Satriani SN (2020) Recognition of ımage pattern to ıdentification of braille characters to be audio signals for blind communication tools. İn: IOP conference series: materials science and engineering, vol 846, p 012008, May 2020
22. Chen Z, Xie Z, Zhang W, Xu X (2017) ResNet and model fusion for automatic spoofing detection. In: INTERSPEECH, Aug 2017, pp 102–106
23. Latif G, Iskandar DNFA, Alghazo J, Butt MM (2020) Brain MR ımage classification for glioma tumor detection using deep convolutional neural network features. Curr Med Imag
24. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition, 2016, pp 770–778
25. Latif G, Alghazo J, Maheswar R, Vijayakumar V, Butt M (2020) Deep learning based intelligence cognitive vision drone for automatic plant diseases identification and spraying. J Intell Fuzzy Syst (Preprint) 1–12