Deep Transfer Learning for Human Identification Based on Footprint: A Comparative Study

Mabroukah M. A. Abuqadumah[1], Musab A. M. Ali[1], Ali Abd Almisreb[2], Benjamin Durakovic[2]

[1] Faculty of Information Science and Engineering, Management and Science University, Malaysia
[2] Faculty of Engineering and Natural Sciences, International University of Sarajevo
Hrasnicka cesta 15, Ilidža 71210 Sarajevo, Bosnia and Herzegovina

Article Info	ABSTRACT
  	Identifying people based on their footprint has not yet gained enough attention from the researchers. Therefore, in this paper, an investigation of human identification conducted based on the footprint. Transfer Learning used as the main concept of this investigation. The aim of using Transfer Learning is to overcome the need for a large-scale dataset and achieve high accuracy with a low-scale dataset. Five well-known models used, namely, Alexnet, Vgg16, Vgg19, Googlenet, and Inception v3. Each of these models fine-tuned to fit-in the paper's topic. A dataset of 30 individuals constructed in order to train the models. The right and left footprint of each individual captured with iPhone camera. The models trained and evaluated based on the same settings. The evaluation shows that Inception v3 model achieved the highest accuracy compared to all other four models.

Corresponding Author:

Mabroukah M. A. Abuqadumah,
Faculty of Information Science and Engineering,
Management and Science University,
Malaysia.
Email: omomran8884@gmail.com

1. Introduction

Many biometric technologies based recognition were deployed in several applications to recognize adults and teenagers such as the ear [1]–[3], face, iris, and fingerprint. In addition to other kind of classification such as lung cancer [4], tumer cancer [5], and brain MR images classification [6].

In the past few years, footprint recognition for infants has received increasing attention. Footprint recognition for infants and newborns was deployed in several applications. For instance, tracking child vaccination and identifying missing children are the main applications. In contrast, the existing methods show that the infant recognition accomplished by identifying the parents or the certificates of identity because of the lack of efficient identification methods [7]. The human footprint carries many characteristics that play important roles in forensic investigation. For an example of these characteristics the walking habits or standing, skin texture of the foot sole and anatomical structures of the foot [8]. Therefore, these characteristics help to overcome an increasing happening issue in hospitals, birthing centers, health centers where multiple births occur simultaneously. Examples of such issues there are; infant missing, swapping, abduction, kidnapping and illegal adoptions [9]. In addition, in order to prevent the occurrence of mix-up among newborns and infants in hospitals, capturing the footprint for the infants assists the medical staff to ensure such issues will not happen [10].

However, the footprint shows some advantages of animal recognition such as stated in [11] and [12]. One of the main challenges in recognizing the animals through the footprint is the features extraction from the images, which are usually located in the boundary curve of the footprints [8]. In general, there are several

features can be used to identify the animal through the footprints such as the number and the size of blobs which are usually used by the humans to identify the animals in reality. Thus, this is confirmed the ability to use the footprint features for animal recognition [12].

The remaining of this paper is structured as follows: the next right section will highlight the most significant research work in this area, followed by the proposed transfer learning models and the dataset used for training and testing. The section covers the result and discussions. The goal of this paper is to investigate the application of five different up-to-date transfer-learning models on human footprint identification and compare the result with the aim of asseing the accuracy of each model.

2. Related work

As research in biometric sector is attracting more attention in the past few years, people identification is becoming an essential component to numerous applications within our daily life. Therefore, There are certain types of biometrics based systems such as face fingerprint, hand geometry and iris gained more efforts of the researchers such as in [13] compare to less efforts to explore the utilization of footprint features in biometrics applications. The work reported in [14] presents a footprint personal recognition using scanning technology to recognize a person's identity based on computing details of features related to the height, weight and body mass index, and in addition to that, the researchers consider additional features such as foot length and foot category. The researchers used Matlab platform to collect their database. Then, the collected footprint mages enhanced and enhanced and feature extraction applied to collect the uniqueness of each footprint. In order to discover the relationship between the footprint parameters, a correlation analysis is conducted. The researchers found a potential correlation between height and weight, actual height and foot length, and actual height and toes.

Work by [15] explored the application of Modified Sequential Haar Energy Transform (MSHET) approach for footprint recognition. In order to collect the Modified Haar Energy (MHE) features, the researchers resized the footprint images and then applied the MSHET. For the purpose of comparison, Euclidean Distance is used to compare the MHE features and the feature vectors that stored in a database. The researchers reported accuracy of 92.375% based on the proposed MHE feature. Another footprint recognition, which uses Minutia descriptor, has also been designed in [16]. In order to measure the minutia similarity for newborns, the researchers used deep convolutional neural network to propose a novel Minutia descriptor. The work investigated the potential use of footprint recognize infants by relying on features collected by a 500 ppi commodity friction ridge sensor. lastly, the researchers conducted validation experiments to show the impact of age and time gap on matching performance and the impact of both single enrolled template and fusion of multiple enrolled templates.

gender recognition based on footwear is reported in [17], and a reprehensive footwear database is created for the same purpose. This conducted research founds that footwear shape is a possible way to recognize humans based on gender. In addition, it is recommended that footwear can be used jointly with other biometrics features to enhance the overall performance. Another person identification systems proposed based on footprint using a single bare or socked footprint is proposed [8]. The researchers used both of Pressure Radial Gradient Map and Geometrical Shape Spectrum Representation to simulate a footprint. The proposed methodology showed outperforms compare to the state-of-the-art algorithms, and in terms of the recognition rate, it achieved 98.75%.

Another paper reported in [18] used Image Parameters to classify the human footprints. The researchers considered parameters such as Footprint Geometry Index (FGI), Footprint Index (FPI) footprint image parameters, A intercept and B intercept. Based on these parameters, the Human foot can be categorized into several classes, for instance, High Arch Foot, Normal Foot and Flat Foot. Therefore, there is a promising potential of using the achieved outcome for additional diagnosis and treatment, and ensure it will be delivered to the corresponding patient in a proper manner. An additional work by [19] presented the usage of small footprint keyword spotting with deep neural networks. The scholars used the deep neural network with a small memory footprint, low computational cost, and high precision. However, the proposed method achieved 45% relative improvement compare to Hidden Markov Model-based system. On the other hand, in the presence of babble noise, the performance shows 39% relative improvement.

In addition to the ability to use this study in human identification, it is also hold a great potential to be used for several other applications such as; forensic and non-forensic purposes.

3. Methodology

The proposed methodology includes a comparison between Alexnet, Vgg16, Vgg19, Googlenet and Inception v3 Convolutional Neural Networks (CNNs). Each of the CNNs fine-tuned to meet with the dataset used in this paper.

3.1. Dataset

In this paper, we collected footprint data from 30 individuals. iPhone is used to take the picture of the subjects' footprints. Firstly, we recorded a video for the left and the right foot for each person. Then, each video is framed into 198 images for each foot (198 images for the left foot, 198 for right foot). Therefore, the total images for each person are 396 images. The total number of images is 11880. To train the model, we allocated 190 images for training and the remaining (8 images) are for the testing. Thus, the total training and testing images are 11400 (190*60) and 480 respectively. Figure 1 shows samples of the footprints.

Figure 1: Samples of collected footprints

3.2. Alexnet

Alexnet is a well-known Convolutional Neural Network created and explained in [20]. Alexnet model is preferred due to the reason that it is the most used model [21]. Furthermore, Alexnet model has the ability to combine between two important factors; namely, speed and accuracy [21]. It consists of 8 pre-trained layers, 5 of these layers are of the type of convolutional layers and other 3 layers are so called fully-connected layers. The last fully-connected layer is designed to classify 1000 object and the remaining layers work to extract features from the image. Alexnet generates feature vector of size 4096-dimensional for each image. The feature vector includes details about the activations of all the layers immediately before the output layer. Alexnet model receives image of size 227x227x3, which is passed to the input layer.

3.3. Vggnet

The Vgg CNN model designed and created after the release of Alexnet, therefore, it carries improvement in terms of its architecture over Alexnet. An example of these enhancements include; using multiple 3X3 kernel-sized filters rather than the 11 and 5 kernel-sized filters engaged by Alexnet in the first and second convolutional layers. The advantage of adding smaller filters, it helps to increase the depth of Vgg model which in turn lead the ability of learning more complicated features. The width of the filter in each convolutional layer in Vgg models is relatively small. The filters size increased by factor of 2 after each max-pooling layer and it begins with 64 in the first layer until the size 512 at the last convolutional layer. In order to identify individuals according to their footprints, both of Vgg-16 and Vgg-19 are chosen.

3.3.1 Vgg16

Vgg16 model consists of 41 layers. Layers with learnable weights are 16 layers and 13 of them are convolutional layers. The rest are fully connected layers [22]. This model receives images with a size of 224x224x3 on its input layer.

3.3.2 Vgg19

Vgg19 is slightly deeper than Vgg16, it has 47 layers. Out of these layers, 19 layers with learnable weights. Particularly, Vgg19 contains 16 convolutional layers, and the remaining are fully connected layers [22]. Vgg19 and Vgg16 share the same input image size, which is 224x224x3.

3.4. Googlenet

As for Googlenet, it is a convolutional neural network implementing a deep module called the inception as described in [23] with 22 layers. Googlenet accepts input images with size 224x224x3.

3.5. Inception-v3

The inception-v3 network has 316 layers. Inception-v3 represented as Directed Acyclic Graph (DAG) object [24]. Inception-v3 model accept input image with size of 229x299x3.

3.6. Fine-tuning CNN models

Fine-tuning rely on retaining the layers of the pre-trained CNN model which is responsible for feature extraction. In transfer learning, this first step is to place a set of new layers able to classify 30 classes (based on our dataset) instead of the last three layers in each of the pre-trained CNN.
 The fine-tuning is achieved by adding one fully connected layer with filter size 64x64, in order to fit in with our new dataset (30 subjects). Another layer is added, namely, Rectified Linear Unit (ReLU) layer, or as often referred as Softmax layer. The main purpose of adding this layer as suggested by [25] is to improve the non-linear problem-solving ability. In addition, this layer is not only able to improve the performance of model, but it is also, due to activation non-linearities of Sigmoid or Tanh, does not produce any gradient vanishing effect units [26]. Another fully connected layer is added and equipped with 30 output neurons in order to facilitate the classification of our 30 subjects. To boost up the learning rate of the newly added layers than in the transferred layer, the weights of the last fully connected layer are initialized with 10. Furthermore, the neuron biases in these layers are also initialized with the constant 20. This is help to accelerate the early stages of learning by providing the ReLUs with positive inputs.

4.   Results and discussion

This section includes a comparison between the transfer learning models in terms of accuracy and loss.

4.1. Alexnet

Alexnet is one of the earliest successful deep learning models. Applying Alexnet to footprint dataset shows that Alexnet Outperforming in two asscpects. Firstly, Alexnet just performed faster than Inceptionv3 model but less accuracy. The performance of Alexnet is shown in Figure 2 (a). The figure shows that Alexnet reached its highest training point after 1200 iteration. Figure 2 (b) shows the corresponding loss.
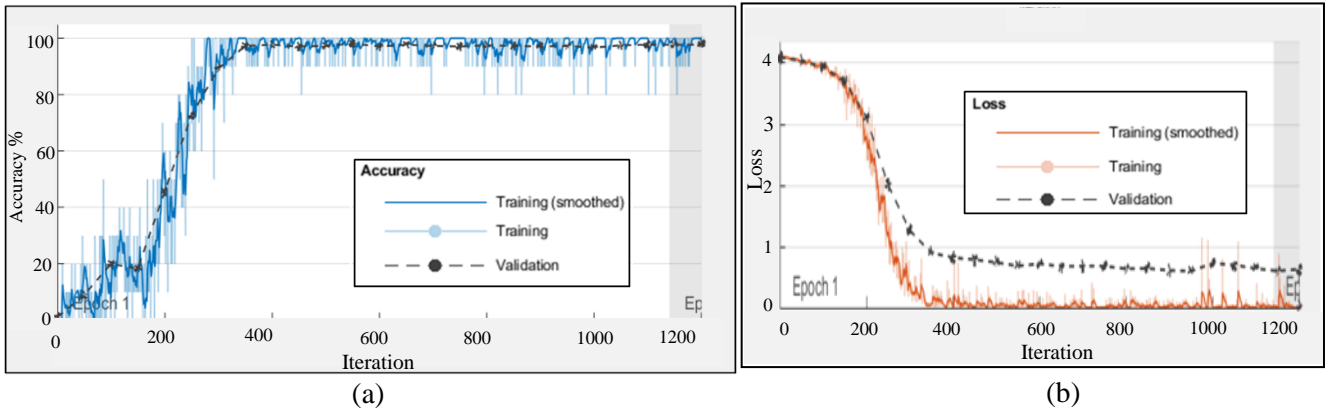
Figure 2: (a) Shows Alexnet training. (b) Shows the loss

4.2. Vgg16

Vgg16 is the second fastest model after Googlenet and the third best training accuracy after Inception-v3 and Alexnet respectively. Figure 3 (a) shows the training performance of Vgg16 and Figure 3 (b) shows the loss.
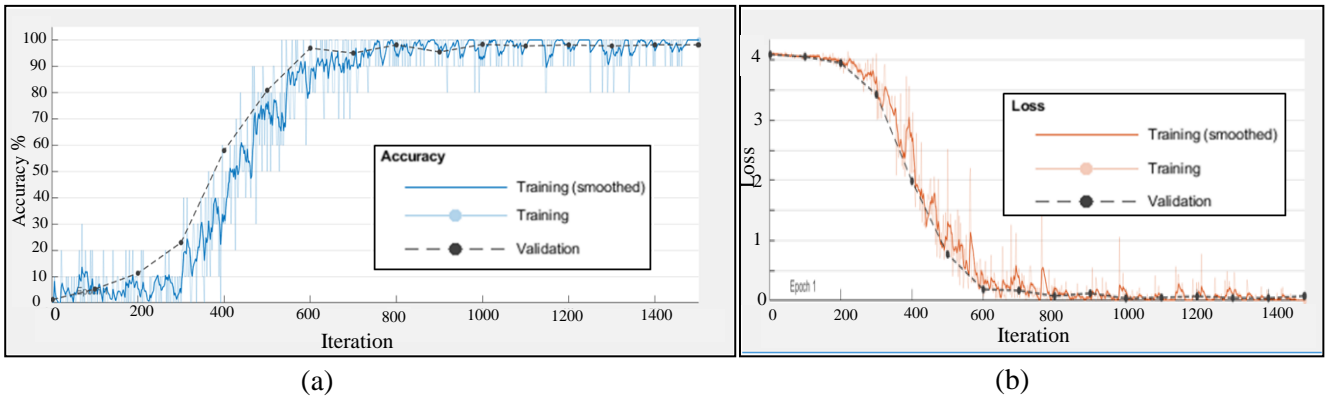

Figure 3: (a) Shows Vgg16 training. (b) Shows the loss

4.3. Vgg19

Vgg19 scored less than Vgg16 eventhough it is deeper. In contrast it is also slower than Vgg16. Therefore, Vgg19 is not recommended for footprint applications. On the other hand, Vgg19 scored only higher than Googlenet. Figure 4 (a) and (b) shows the training and loss, respectively.
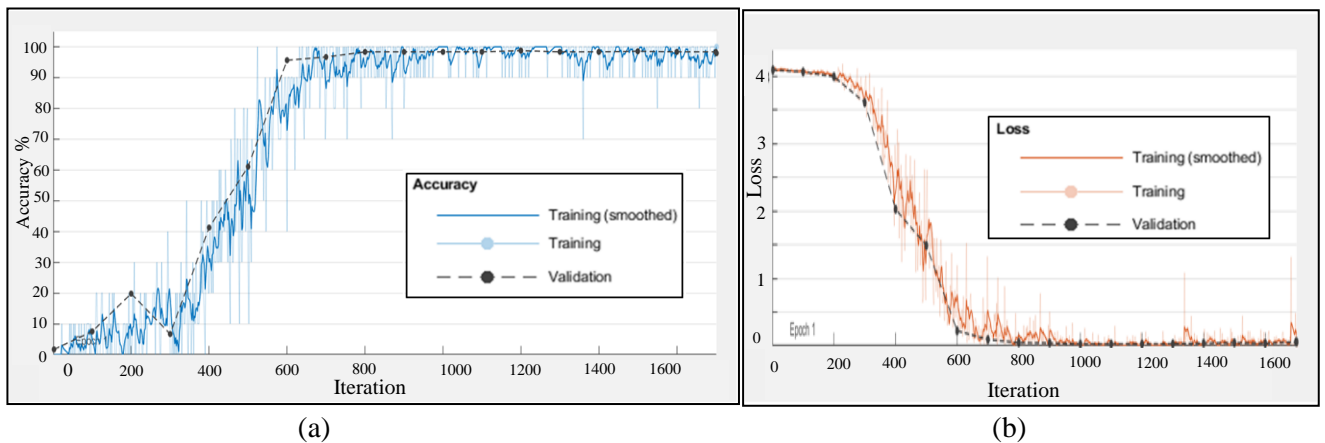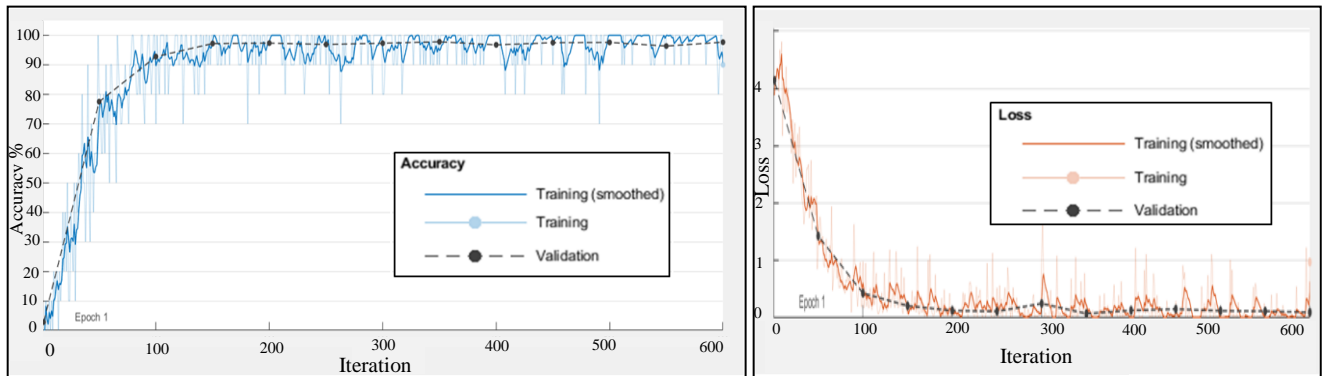

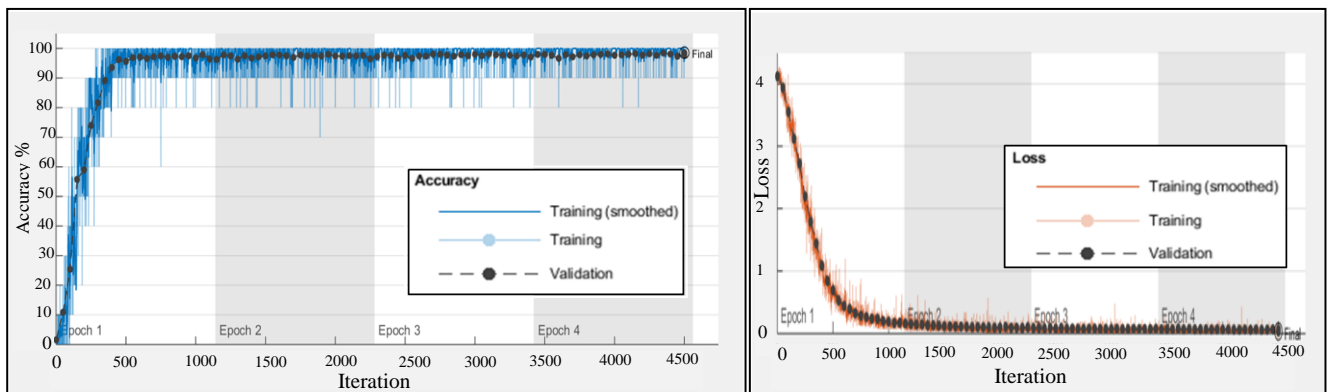Figure 4: (a) Shows Vgg19 training. (b) Shows the loss

4.4. Googlenet

Googlenet shows less effeciency in terms of the training accuracy compare to all other models. In contrast, it shows high speed in terms of the computational time as shown in Figure 5 (a). Furthermore, this figure shows that Googlenet model has better learining curve compare to Alexnet, Vgg16 and Vgg19.

Figure 5: (a) Shows Googlenet training. (b) Shows the loss

4.5. Inception–v3

Inception v3 gained the outmost performance compare to all other models in this this paper. Its performace as shown in Figure 6 (a) indecate that the learining process acheived the highest score but it required longer time. It also shows that the learning process is more stable and smooth. Therefore, in this paper, inception v3 model is recommended for foorprint applications over other mentioned models.

Figure 6: (a) Shows Googlenet training. (b) Shows the loss

The overall perforamce of each model presented in Figure 7. As asummary, this figure shows that Inception v3 acheived the highest accuracy followd by Alexnet, Vgg16, Vgg19 and finally Googlenet.
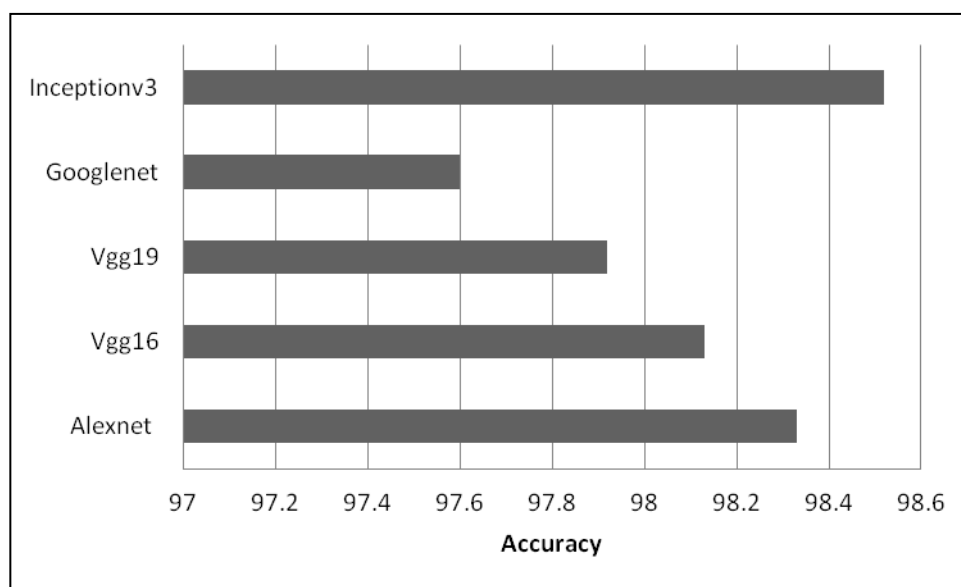
Figure 7: The accuracy of each model

5. Conclusion

In this paper, we utilized five deep learning models to inviestigate the applicability of transfer learning in hurman recognition based on footprint. The performance comparison between the models conducted based on the acheived accuracy and the computational time. It showed that Inception v3 model better than all other models in this paper with over all accuracy 98.52%. Followed by Alexnet, Vgg16, Vgg19, and Googlenet with accuracy of 98.33%, 98.13%, 97.92%, and 97.60% respectively. On the other hand, Inception v3 model consumed more time than any of the other four models. More work is going to be done in the future by using the object detection algorithms and process the area of footprint only. In addition, the sample size for the future work is planned to be larger. This can lead to more universal human identification system.

References

[1]   A. A. Almisreb and N. Jamil, "Automated ear segmentation in various illumination conditions," in Proceedings - 2012 IEEE 8th International Colloquium on Signal Processing and Its Applications, CSPA 2012, 2012.

[2]   A. A. Almisreb, N. Jamil, and N. M. Din, "Utilizing AlexNet Deep Transfer Learning for Ear Recognition," in Proceedings - 2018 4th International Conference on Information Retrieval and Knowledge Management: Diving into Data Sciences, CAMP 2018, 2018.

[3]   N. Jamil, A. AlMisreb, and A. A. Halin, "Illumination-invariant ear authentication," Procedia Comput. Sci., vol. 42, no. C, pp. 271–278, 2014.

[4]   S. R. A. Ahmed, I. Al-Barazanchi, H. R. Abdulshaheed, and A. Mhana, "Lung cancer classification using data mining and supervised learning algorithms on multi-dimensional data set," Period. Eng. Nat. Sci., vol. 7, no. 2, pp. 438–447, 2019.

[5]   S. F. Jabbar, "A classification model on tumor cancer disease based mutual information and firefly algorithm," Period. Eng. Nat. Sci., vol. 7, no. 3, pp. 1152–1162, 2019.

[6]   S. Kumar, D. Jayadevappa, and M. V. Shetty, "A novel approach for segmentation and classification of brain MR images using cluster deformable based fusion approach," Period. Eng. Nat. Sci., vol. 6, no. 2, pp. 237–242, 2018.

[7]   E. Liu, "Infant Footprint Recognition," Proc. IEEE Int. Conf. Comput. Vis., pp. 1662–1669, 2017.

[8]   X. Wang, H. Wang, Q. Cheng, N. L. Nankabirwa, and T. Zhang, "Single 2D pressure footprint based person identification," *IEEE Int. Jt. Conf. Biometrics, IJCB 2017*, pp. 413–419, 2018.

[9]   S. Sumathi and S. Balameenakshi, "Biometric Recognition of Newborns : Identification using Palmprints," *in IEEE Conference on Information and Communication Technologies (ICT 2013)*, 2013, no. Ict, pp. 1–6.

[10]   J. Kotzerke, S. Davis, K. Horadam, and J. McVernon, "Newborn and infant footprint crease pattern extraction," *2013 IEEE Int. Conf. Image Process. ICIP 2013* - Proc., pp. 4181–4185, 2013.

[11]   N. Lokare, Q. Ge, W. Snyder, Z. Jewell, S. Allibhai, and E. Lobaton, "Manifold learning approach to curve identification with applications to footprint segmentation," *IEEE SSCI 2014 - 2014 IEEE Symp. Ser. Comput. Intell. - CIMSIVP 2014 2014 IEEE Symp. Comput. Intell. Multimedia, Signal Vis. Process. Proc.*, 2015.

[12]   M. N. Alli and S. Viriri, "Animal identification based on footprint recognition," *IEEE Int. Conf. Adapt. Sci. Technol. ICAST*, 2013.

[13]   A. Uhl and P. Wild, "Personal identification using Eigenfeet, Ballprint and Foot geometry biometrics," *IEEE Conf. Biometrics Theory, Appl. Syst. BTAS'07*, pp. 1–6, 2007.

[14]   R. Khokher and R. C. Singh, "Footprint-based personal recognition using scanning technique," *Indian J. Sci. Technol.*, vol. 9, no. 44, 2016.

[15]   V. Kumar and M. Ramakrishnan, "Footprint Recognition using Modified Sequential Haar Energy Transform (MSHET)," *Int. J. Comput. Sci.* Issues, vol. 7, no. 3, pp. 47–51, 2010.

[16]   E. Liu, "Infant Footprint Recognition," *Proc. IEEE Int. Conf. Comput*. Vis., vol. 2017-Octob, pp. 1662–1669, 2017.

[17]   Y. Yuan, Y. Pang, and X. Li, "Footwear for gender recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 1, pp. 131–135, 2010.

[18]   P. S. Kulkarni and V. B. Kulkarni, "Human footprint classification using image parameters," in 2015 *International Conference on Pervasive Computing: Advance Communication Technology and Application for Society, ICPC 2015*, 2015, vol. 00, no. c.

[19]   G. Chen, C. Parada, and G. Heigold, "Small-footprint keyword spotting using deep neural networks," *Icassp*, no. i, pp. 4115–4119, 2014.

[20]   A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Adv. Neural Inf. Process. Syst.,* pp. 1–9, 2012.

[21]   H. C. Shin et al., "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," *IEEE Trans. Med. Imaging*, 2016.

[22]   K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *Int. Conf. Learn. Represent.*, 2015.

[23]   C. Szegedy et al., "Going deeper with convolutions," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, pp. 1–9, 2015.

[24]   C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *in Proceedings of the IEEE conference on computer vision and pattern recognition,* 2015.

[25]   E. Azarov, M. Vashkevich, D. Likhachov, A. Petrovsky, and P. Brovky, "Real-time Voice Conversion Using Artificial Neural Networks with Rectified Linear Units," *INTERSPEECH 2013*, no. August, pp. 1032–1036, 2013.

[26]   X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," *AISTATS '11 Proc.* 14th Int. Conf. Artif. Intell. Stat., 2011.