ICIC Express Letters
Volume **17**, Number **3**, March **2023**
ICIC International ©2023 ISSN 1881-803X
pp. **359–366**

OFFLINE SIGNATURE VERIFICATION USING TRANSFER LEARNING AND DATA AUGMENTATION ON IMBALANCED DATASET

Aeron Mitchell[1], Evan Edbert[1,*], Gregorius Natanael Elwirehardja[2] and Bens Pardamean[1,2]

[1]Computer Science Department, BINUS Graduate Program – Master of Computer Science
[2]Bioinformatics and Data Science Research Center
Bina Nusantara University
Jl. K. H. Syahdan No. 9, Kemanggisan, Palmerah, Jakarta 11480, Indonesia
aeron.mitchell@binus.ac.id; { gregorius.e; bpardamean }@binus.edu
*Corresponding author: evan.edbert@binus.ac.id

Received March 2022; accepted June 2022

Abstract. *Handwritten signature verification has attracted a lot of attention in recent years for authentication purposes and yet, it is still prone to mistakes of human error. Thus, with the development of deep learning, many methods of signature forgery classification were proposed in previous studies. However, not many studies have addressed the problem of imbalanced datasets which have been known to affect the models' performance. Therefore, this paper aims to solve the problem of imbalanced dataset by using transfer learning and multiple data augmentations. Transfer learning provides efficiency by using a pre-trained model for enhanced performance. Additionally, data augmentation applies transformations to enlarging datasets. Hence, four approaches were evaluated in this study. The first three approaches used a custom Convolutional Neural Network (CNN) trained on different datasets: 1) an imbalanced dataset, 2) a balanced dataset, and 3) an augmented balanced dataset. Meanwhile, the fourth approach used a pre-trained VGG16 via transfer learning trained on an augmented balanced dataset. The result of our study shows that among our proposed approaches, the transfer learning of the VGG16 model trained on an augmented balanced dataset outperformed the replicated baseline models with 86.7% accuracy.*

Keywords: Deep learning, Data augmentation, Imbalance, Offline signature verification, Transfer learning, Convolutional neural network

1. **Introduction.** In recent years, biometric technology has become popular in information technology due to its ability to recognize individual behavior or physical characteristics such as fingerprints, DNA, voice, and signature [1,2]. For example, handwritten signature verification is generally used due to its non-invasive characteristic and most people have their unique signatures for authentication purposes such as banking, and politics. Nevertheless, manual handwritten signatures verification often causes mistakes due to human error. Thus, many studies have proposed several approaches of identification using Artificial Intelligence (AI) as a viable option to improve signature verification [3-5]. Generally, there are two types of signature verification which are online and offline verification. The online type uses features based on a signature track by using an intelligent pen such as pressure, and coordinate. Meanwhile, the offline type uses bitmap images of finished signatures for further analysis with image processing techniques [6-8].

Deep learning is a subfield of machine learning that replicates the human brain in the form of a neural network for calculating and analyzing data. Although conventional machine learning is still feasible, deep learning still provides better results than the former

DOI: 10.24507/icicel.17.03.359

[9,10]. Due to deep learning influence, a lot of technologies have implemented computer vision applications in daily lives such as moving object detection, image classification, and object counting [11,12]. However, deep learning has the downside of requiring an enormous amount of balanced data to avoid overfitting and achieve great results. Many previous studies in signature verification have encouraged data augmentation or transfer learning to handle the lack of data and avoid overfitting. However, only a few of them have addressed the problem of imbalanced data in their studies [13-15]. Thus, this study aims to create an offline signature verification with deep learning to solve imbalanced dataset problems. For that purpose, we conduct several approaches for comparison using our proposed Convolutional Neural Network (CNN) and transfer learning model with data augmentation. The main contributions of this research are listed below:

1) Utilizing Gaussian noise injection to oversample the imbalanced handwritten-signature dataset;
2) Improving the performance of deep learning models for signature verification by using multiple data augmentation techniques;
3) Using CNN transfer learning method of VGG16 supported by multiple data augmentation techniques to increase accuracy.

This research is composed of four sections. The second section presents the works of previous studies regarding their methods and purposes related to this paper, the third section explains our methodology in this study, the fourth section discusses the final result of our experiment, and finally, the fifth section discusses our contributions of experiment and possible future works.

2. **Related Works.**

2.1. **Data augmentation.** In previous studies, data augmentation had been used as a technique to increase dataset size whether to improve generalizations, avoid overfitting, or create a balanced dataset for image classification. For instance, in a study by Muljo et al. [16], they used multiple data augmentations like random horizontal flip, scaling, and center crop to increase data size and improve accuracy. A 5-crop augmentation was used in a study by Pardamean et al. [17] to overcome the limited mammogram data for breast cancer detection. Multiple data augmentations like rotation, flipping, shifting, resizing and gamma correction was used in a study by Kashyap [18] to multiply data size and avoid overfitting. Meanwhile, Suharjito et al. [19] used data augmentation to avoid overfitting by using a modified version of 5-crop augmentation called 9-angle crop. Other studies such as the ones from Gunawan et al. [20] and Dominic et al. [21] used data augmentation to handle imbalanced data and generate more data by combining random background noise, shearing, brightness, and zoom adjustment. These previous studies of data augmentation had proven to be useful for image classification in deep learning.

2.2. **Transfer learning.** Transfer learning had been used in previous studies by implementing a pre-trained model to improve deep learning model of another classification problem by providing an easy and faster new training segment while avoiding overfitting. For instance, EffecientNetB0 was used in a study by Gunawan et al. [20] to avoid overfitting in their owl sound classification. In another study by Muchtar et al. [22], YOLOv3 was adopted to develop a swift computing process for pedestrian detection systems. Other studies like Pardamean et al. [23] proposed using AlexNet, VGG16, GoogLeNet, ResNets, and DenseNet121 which had been pre-trained on ImageNet to accommodate smaller datasets. Transfer learning was also proposed in Harsono et al.'s [24] research using I3DR-Net to increase performance and reduce training time for lung nodule detection to determine malignancy. Other examples could be found in the work of Marcellino et al. [25] as they implemented the UNET++ model with VGG backbone to improve deep learning model in the crowd counting system. These previous studies of transfer learning had proven to be useful for image classification in deep learning.

2.3. **Handwritten signature verification.** Besides this study, there were other studies regarding handwritten signature verification as well. For instance, both studies by Alajrami et al. [7] and Kiran et al. [26] used CNN for template matching to identify whether the image match input is similar to the given template and both results achieved 99.7% accuracy. Merlin et al. [27] used CNN with data augmentation consisting of rescaling, zoom range, horizontal flip, and resizing with transfer learning of AlexNet and LeNet. Other studies like Pinzón-Arenas et al. [28] proposed the use of DAG-CNN with random shifting augmentation which achieved an accuracy of 99.4%, Rabbi et al. [29] proposed the use of CNN with rotation, shifting, and zooming in and out augmentation which resulted in 98.33% accuracy, and Yapıcı et al. [30] used CNN with cycle GAN augmentation to achieve a better result of CNN models for offline signature.

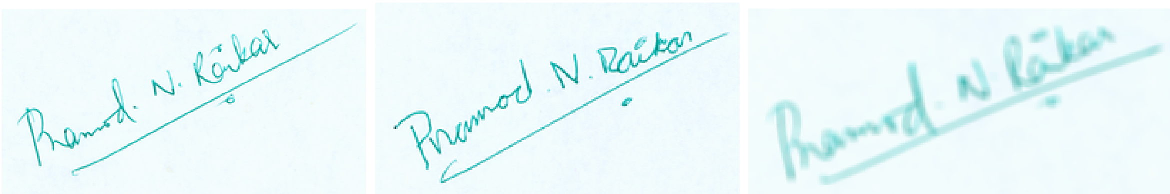
Meanwhile, a study by Longjam and Kisku [31] used AlexNet, VGG16, and VGG19 model with Support Vector Machine (SVM) classifier for comparison and VGG16 achieved the best accuracy of 86.8%, false acceptance rate of 38, and rejection rate of 4.6. Manikantha et al.'s [32] study compared different sets of classifiers such as linear regression, Gaussian Naïve Bayes, K Nearest Neighbor (KNN), and SVM in four models of transfer learning such as VGG16, MobileNetV2, DenseNet121, and Xception tested on different datasets. As a result, the VGG16 model achieved the best Euclidean distance of 100% accuracy on the Cedar and Kaggle datasets. Jahandad et al. [33] compared two models of InceptionV1 and InceptionV3 for the signature verification which resulted in 83% and 75% accuracy, respectively.

Previous studies introduced different data augmentations and CNN architectures with various results. However, there were two points that made this study different. The first one was utilizing data augmentation twice to balance our dataset first before further expanding its size. Thus, the dataset is both balanced and larger which can yield better results for deep learning models. The second one was comparing our approaches using the dataset before and after data augmentation.

3. **Methodology.** In this study, we proposed the four approaches we had taken. The first three approaches used a custom Convolutional Neural Network (CNN) trained on different datasets: 1) an imbalanced dataset, 2) a balanced dataset, and 3) an augmented balanced dataset. Meanwhile, the fourth approach used a pre-trained VGG16 via transfer learning trained on an augmented balanced dataset. To explain those four approaches, this section covered our methodology consisting of the dataset, image processing, model architecture, and training configuration.

3.1. **Dataset.** We used a dataset called "handwritten signatures – Genuine and Forged Signature Examples" from the public Kaggle repository. The dataset had four folders and each folder contained 180 images mixed with the real and forged signatures, adding up to the total of 720 images. After further analysis, there were a total of 144 different types of signatures based on two factors which were whose signature and the signature authenticity. For instance, the signature belongs to person A but the one who wrote it is person B which means that the signature is forged. Thus, we divided the main four folders into 144 folders based on the previous category statement.

3.2. **Image preprocessing.** After sorting out the folders, we split the dataset by collecting the first and second images from each folder into the test data while leaving the rest for training data. During the process, several folders had a smaller number of images than the others so we decided to take only one image out. As a result, we obtained 287 test images and 433 training images. Considering that some training folders only had two to four images, the dataset became imbalanced. Hence, we added more images using data augmentation until each folder contained four images by injecting noise of Gaussian blur with the kernel size of five. Figure 1 shows the example images of the dataset. Thus, our training dataset consisted of 576 images. Moreover, we split 20% of the training data as the validation data during the model training process.



(a) (b) (c)

FIGURE 1. The example of images from the dataset: (a) Real signature, (b) forged signature, and (c) blurred real signature

Furthermore, to expand our dataset for the third and fourth approaches, we used data augmentation again such as Gaussian noise injection with the kernel size of three, increasing and decreasing brightness by 20 pixels, and zooming in the images to 80%. The examples of data augmentation can be seen in Table 1. Therefore, the dataset expanded from 576 images to 2880 images and each folder contained 20 images.

TABLE 1. Examples of data augmentation for image

Image with Gaussian blur (3, 3) kernel	**Image with 20-pixel brightness**	**Image with 20-pixel dark contrast**	**Image with 80% zoom-in**
			

3.3. **Model architecture.** In this experiment, we used two models which were our proposed CNN model for the first three approaches and VGG16 for the fourth approach. Both models had the same input size of 224×224 and the output was classified using Support Vector Machine (SVM) to determine whether the signature was real or forged. First, our proposed CNN architecture consisted of five convolution layers and five max-pooling layers and finally, a Multi-Layer Perceptron (MLP) block consisting of a flatten layer, a dropout layer, and two dense or fully connected layers. The detail of our proposed CNN model could be seen in Figure 2.

Second, we used a pre-trained model of VGG16 via transfer learning for our fourth approach. Transfer learning had proven to be efficient and useful for achieving high accuracy in previous studies [17,20]. Thus, we chose a pre-trained VGG16 model due to its implementation of a kernel size of three instead of five or more for less parameters and more efficiency plus stacking two or three layers provided deeper results for a more discriminative decision making in the model [34]. The VGG16 model contained pre-trained weights from ImageNet and consisted of 19 layers which were all frozen so that the model could be implemented immediately. Moreover, we added the same MLP block from our proposed CNN by removing the top layers from VGG16 to match our needs of classification. The detail of the VGG16 model could be seen in Figure 3.

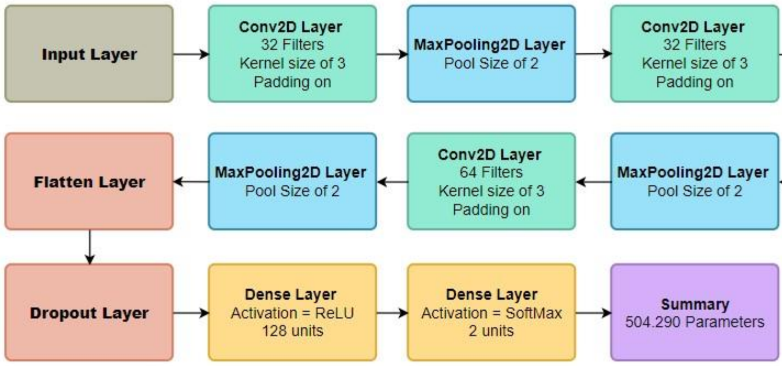


Figure 2. Proposed CNN model architecture

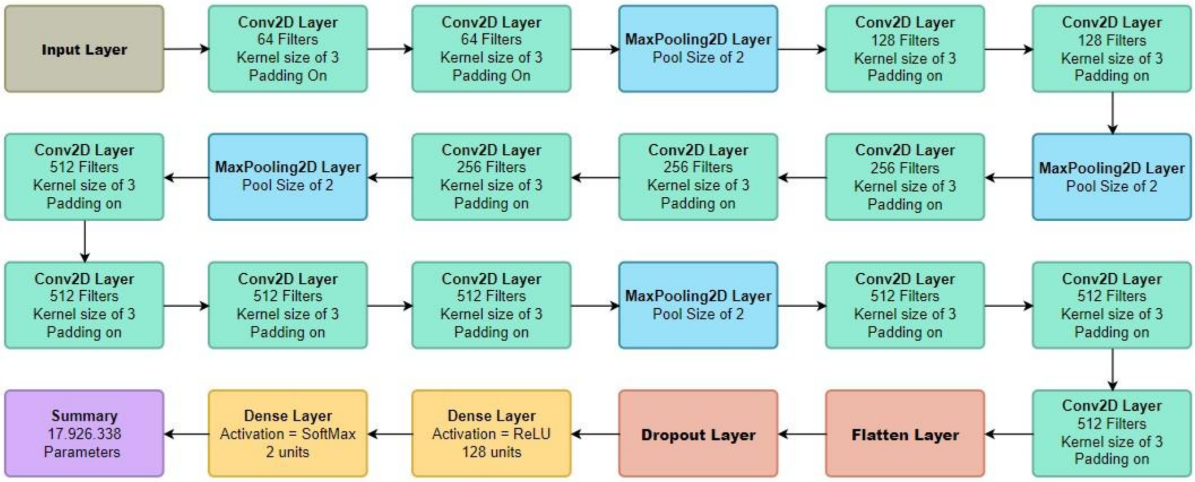


Figure 3. VGG16 model architecture

3.4. **Training configuration.** For our model training segment, several hyper-parameters such as a dropout layer with a factor of 0.5 and an L2 kernel regularizer of 0.1 in both dense layers were added for better generalization and SVM classifier. Then, we used Adam optimizer with the learning rate of 0.005 and squared hinge as the loss function for further optimization. Finally, all models were trained for 20 epochs equipped with an early stopping function in case the result was stagnant. Once the training was finished, the models were used to predict accuracy, precision, and F1-score results from the test dataset.

4. **Result and Discussion.** Table 2 shows the comparison results of accuracy, precision, and F1-score for our four approaches which were taken from the models' prediction on the test dataset.

Table 2. Comparison of model performance

Approaches	Accuracy	Precision	F1-score
Custom CNN trained on imbalanced dataset	52.9%	0.0559	0.1059
Custom CNN trained on balanced dataset	72.8%	0.7552	0.7346
Custom CNN trained on balanced dataset with data augmentation	83.2%	0.8181	0.8297
VGG16 model on balanced dataset with data augmentation	86.7%	0.9510	0.8774

Based on the first and second approaches, the results from the latter model are better than the former because the second one used a balanced dataset. This proves that using a balanced dataset allows the model to generalize better than the ones trained on the imbalanced dataset. Meanwhile, by comparing the second and third approaches, it appears that the third approach has the better results because the third model had a bigger size of training dataset than the second approach. Thus, it is proved that using data augmentation increases the model performance for signature verification. On the other hand, by comparing the fourth and the third approaches, the fourth model has better results which shows that transfer learning supported by data augmentation improves performance. This could happen because the pre-trained weights were obtained from ImageNet, allowing the VGG16 model to capture more relevant features from our augmented balanced dataset for better performance.

Table 3 shows the comparison results using the proposed VGG16 model as our best approach with other CNN models from other studies by replicating and testing them on the same augmented balanced dataset that was used for our third and fourth approaches. However, we changed their top layers using an MLP block from our proposed CNN and used the same hyper-parameters in the training configuration to match our needs. As a result, our approach achieves better results compared to the replicated baseline models.

Table 3. Comparison to other methods

Method	Accuracy	Precision	F1-score
The proposed VGG16 model	86.7%	0.9510	0.8744
Alajrami et al.'s CNN model [7]	49.8%	1.0	0.6651
Rabbi et al.'s CNN model [29]	86.0%	0.8181	0.8540
InceptionV3 model [33]	83.6%	0.7062	0.8112

5. **Conclusions.** This paper explains how the dataset can affect the results of the signature verification on imbalanced data, balanced data, augmented balanced data, and augmented balanced data with transfer learning. We proposed using Gaussian noise injection to oversample the imbalanced dataset problem. Thus, by comparing the result of the first and second approaches, it was proven that using Gaussian noise injection could oversample the imbalanced dataset problem. Next, we also proposed using multiple data augmentation techniques to improve the model performance by adding more augmented images. It had been proven from the results of comparing the second and third approaches that our proposed data augmentation techniques could improve the performance of our proposed CNN model. Finally, we proposed the use of transfer learning of VGG16, supported by multiple data augmentation to improve its performance. By comparing the third and fourth approaches, transfer learning of VGG16 has better results than the previous approaches. Moreover, our proposed VGG16 model managed to outperform the replicated models from previous works.

For future works, the performance of signature verification models can be further improved by expanding the dataset size for both training and testing using other types of data augmentation to improve data generalization and avoid overfitting. Another idea that could be done is unfreezing some of the layers of VGG16 to be fine-tuned for signature verification. Thus, a comparative study of how many blocks should be left unfrozen can also be conducted in the future.

Acknowledgment. This work is supported by Bina Nusantara University. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] Vandana and N. Kaur, A study of biometric identification and verification system, *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE2021)*, pp.60-64, DOI: 10.1109/ICACITE51222.2021.9404735, 2021.

[2] E. A. M. Alrahawe, V. T. Humbe and G. N. Shinde, A biometric technology-based framework for tackling and preventing crimes, *Intelligent Data Analytics for Terror Threat Prediction*, pp.133-160, DOI: 10.1002/9781119711629.ch7, 2021.

[3] N. Neelima, Social ski driver-jaya optimization-enabled deep convolution neural network for signature verification, *Turkish Journal of Computer and Mathematics Education*, vol.12, no.12, pp.62-77, https://turcomat.org/index.php/turkbilmat/article/view/7260, 2021.

[4] N. M. T. Tahir, A. N. Ausat, U. I. Bature, K. A. Abubakar and I. Gambo, Off-line handwritten signature verification system: Artificial neural network approach, *International Journal of Intelligent Systems and Applications*, vol.13, no.1, pp.45-57, DOI: 10.5815/ijisa.2021.01.04, 2021.

[5] A. Jain, S. K. Singh and K. P. Singh, Handwritten signature verification using shallow convolutional neural network, *Multimedia Tools and Applications*, vol.79, nos.27-28, DOI: 10.1007/s11042-020-08728-6, 2020.

[6] Y. Zhou, J. Zheng, H. Hu and Y. Wang, Handwritten signature verification method based on improved combined features, *Applied Sciences*, vol.11, no.13, DOI: 10.3390/app11135867, 2021.

[7] E. Alajrami et al., Handwritten signature verification using deep learning, *International Journal of Academic Multidisciplinary Research*, vol.3, no.12, pp.39-44, www.ijeais.org/ijamr, 2019.

[8] A. Sharma and S. Sundaram, A novel online signature verification system based on GMM features in a DTW framework, *IEEE Trans. Information Forensics and Security*, vol.12, no.3, DOI: 10.1109/TIFS.2016.2632063, 2017.

[9] N. K. Chauhan and K. Singh, A review on conventional machine learning vs deep learning, *2018 International Conference on Computing, Power and Communication Technologies (GUCON2018)*, pp.347-352, DOI: 10.1109/GUCON.2018.8675097, 2019.

[10] R. Y. Choi, A. S. Coyner, J. Kalpathy-Cramer, M. F. Chiang and J. P. Campbell, Introduction to machine learning, neural networks, and deep learning, *Translational Vision Science and Technology*, vol.9, no.2, DOI: 10.1167/tvst.9.2.14, 2020.

[11] K. Muchtar, F. Rahman, T. W. Cenggoro, A. Budiarto and B. Pardamean, An improved version of texture-based foreground segmentation: Block-based adaptive segmenter, *Procedia Computer Science*, vol.135, DOI: 10.1016/j.procs.2018.08.228, 2018.

[12] T. W. Cenggoro, F. Tanzil, A. H. Aslamiah, E. K. Karuppiah and B. Pardamean, Crowdsourcing annotation system of object counting dataset for deep learning algorithm, *IOP Conference Series: Earth and Environmental Science*, vol.195, no.1, DOI: 10.1088/1755-1315/195/1/012063, 2018.

[13] C. Shorten and T. M. Khoshgoftaar, A survey on image data augmentation for deep learning, *Journal of Big Data*, vol.6, no.1, DOI: 10.1186/s40537-019-0197-0, 2019.

[14] L. Alzubaidi, O. Al-Shamma, M. A. Fadhel, L. Farhan, J. Zhang and Y. Duan, Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model, *Electronics (Switzerland)*, vol.9, no.3, DOI: 10.3390/electronics9030445, 2020.

[15] J. M. Johnson and T. M. Khoshgoftaar, Survey on deep learning with class imbalance, *Journal of Big Data*, vol.6, no.1, DOI: 10.1186/s40537-019-0192-5, 2019.

[16] H. H. Muljo, B. Pardamean, K. Purwandari and T. W. Cenggoro, Improving lung disease detection by joint learning with COVID-19 radiography database, *Communications in Mathematical Biology and Neuroscience*, vol.2022, DOI: 10.28919/cmbn/6838, 2022.

[17] B. Pardamean, T. W. Cenggoro, R. Rahutomo, A. Budiarto and E. K. Karuppiah, Transfer learning from Chest X-Ray pre-trained convolutional neural network for learning mammogram data, *Procedia Computer Science*, vol.135, DOI: 10.1016/j.procs.2018.08.190, 2018.

[18] R. Kashyap, Breast cancer histopathological image classification using stochastic dilated residual ghost model, *International Journal of Information Retrieval Research*, vol.12, no.1, DOI: 10.4018/ijirr.289655, 2022.

[19] Suharjito, G. N. Elwirehardja and J. S. Prayoga, Oil palm fresh fruit bunch ripeness classification on mobile devices using deep learning approaches, *Computers and Electronics in Agriculture*, vol.188, DOI: 10.1016/j.compag.2021.106359, 2021.

[20] K. W. Gunawan, A. A. Hidayat, T. W. Cenggoro and B. Pardamean, A transfer learning strategy for owl sound classification by using image classification model with audio spectrogram, *International Journal on Electrical Engineering and Informatics*, vol.13, no.3, DOI: 10.15676/IJEEI.2021.13.3.3, 2021.

[21] N. Dominic, Daniel, T. W. Cenggoro, A. Budiarto and B. Pardamean, Transfer learning using inception-resnet-v2 model to the augmented neuroimages data for autism spectrum disorder classification, *Communications in Mathematical Biology and Neuroscience*, vol.2021, DOI: 10.28919/cmbn/5565, 2021.

[22] K. Muchtar et al., Lightweight pedestrian detection through guided filtering and deep learning, *2021 IEEE 10th Global Conference on Consumer Electronics (GCCE2021)*, pp.629-630, DOI: 10.1109/GCCE53005.2021.9622111, 2021.

[23] B. Pardamean, H. H. Muljo, T. W. Cenggoro, B. J. Chandra and R. Rahutomo, Using transfer learning for smart building management system, *Journal of Big Data*, vol.6, no.1, DOI: 10.1186/s40537-019-0272-6, 2019.

[24] I. W. Harsono, S. Liawatimena and T. W. Cenggoro, Lung nodule detection and classification from Thorax CT-scan using RetinaNet with transfer learning, *Journal of King Saud University – Computer and Information Sciences*, DOI: 10.1016/j.jksuci.2020.03.013, 2020.

[25] Marcellino, T. W. Cenggoro and B. Pardamean, UNet++ with scale pyramid for crowd counting, *ICIC Express Letters*, vol.16, no.1, pp.75-82, DOI: 10.24507/icicel.16.01.75, 2022.

[26] L. C. Kiran, G. A. Chowdary, M. S. Raju and K. Gopi, Digital signature forgery detection using CNN, *International Research Journal of Engineering and Technology*, vol.8, no.6, pp.2969-2973, www.irjet.net, 2021.

[27] P. B. Merlin, G. Santhiya and S. Vijayalakshmi, Human signature verification using CNN with Tensorflow deployment using Django framework, *International Advanced Research Journal in Science, Engineering and Technology (IARJSET)*, vol.8, no.8, pp.2506-2516, DOI: 10.17148/iarjset.2021.8856, 2021.

[28] J. O. Pinzón-Arenas, R. Jiménez-Moreno and C. G. Pachón-Suescún, Offline signature verification using DAG-CNN, *International Journal of Electrical and Computer Engineering*, vol.9, no.4, DOI: 10.11591/ijece.v9i4.pp3314-3322, 2019.

[29] M. T. F. Rabbi et al., Handwritten signature verification using CNN with data augmentation, *The Journal of Contents Computing*, vol.1, no.1, DOI: 10.9728/jcc.2019.12.1.1.25, 2019.

[30] M. M. Yapıcı, A. Tekerek and N. Topaloelu, Deep learning-based data augmentation method and signature verification system for offline handwritten signature, *Pattern Analysis and Applications*, vol.24, no.1, pp.165-179, DOI: 10.1007/s10044-020-00912-6, 2021.

[31] T. Longjam and D. R. Kisku, Writer independent manipuri offline signature verification using transfer learning technique of convolutional neural network, *International Journal of Innovative Technology and Exploring Engineering*, vol.9, no.1S, DOI: 10.35940/ijitee.a1018.1191s19, 2019.

[32] K. Manikantha, A. R. Bhat, P. Nerella, P. Baburaj and S. SharvariK, A comparative study of transfer learning models for offline signature verification and forgery detection, *Journal of University of Shanghai for Science and Technology*, vol.23, no.7, pp.1129-1139, DOI: 10.51201/jusst/21/07272, 2021.

[33] Jahandad, S. M. Sam, K. Kamardin, N. N. A. Sjarif and N. Mohamed, Offline signature verification using deep learning convolutional neural network (CNN) architectures GoogLeNet Inception-v1 and Inception-v3, *Procedia Computer Science*, vol.161, pp.475-483, DOI: 10.1016/j.procs.2019.11.147, 2019.

[34] J. Pardede, B. Sitohang, S. Akbar and M. L. Khodra, Implementation of transfer learning using VGG16 on fruit ripeness detection, *International Journal of Intelligent Systems and Applications*, vol.13, no.2, pp.52-61, DOI: 10.5815/ijisa.2021.02.04, 2021.