Digital handwriting characteristics for dysgraphia detection using artificial neural network

Mohamed Ikermane[1], Abdelkrim El Mouatasim[2]
[1]Laboratory LABSI, Faculty of Sciences, Ibn Zohr University, Agadir, Morocco
[2]Department of Mathematics and Management, Faculty of Polydisciplinary Ouarzazate (FPO), Ibn Zohr University, Ouarzazate, Morocco

Article Info

Article history:

Received Aug 17, 2022
Revised Nov 11, 2022
Accepted Dec 5, 2022

Keywords:

Artificial neural network
BHK test
Digital handwriting
Digitizer tablet
Dropout dysgraphia
Keras

ABSTRACT

Despite all of the technical advancements in writing and text editing with keyboards on numerous devices, writing with a pen remains a fundamental ability in modern human existence. Handwriting disabilities are referred to as dysgraphia. Nonetheless, how well they are taught to write in school, 10-30% of children never attain a respectable level of handwriting. Early identification is critical because it can help children avoid difficulties in their behavioral and academic development. On blank papers attached to digital tablets, 280 individuals were asked to complete the concise evaluation scale for children's handwriting (BHK), with 218 having typical handwriting and 62 having dysgraphia. In addition to their age and BHK quality and speed scores, 12 variables identifying digital handwriting across several domains (static, kinematic, pressure, and tilt) were collected. In this paper, we provided a rapid and automated dysgraphia classification approach using an artificial neural network (ANN) model. Using digital handwriting traits as an input to the ANN approach, the prediction findings were encouraging and very accurate, reaching 96% accuracy, and they could lead to the development of a new self-administered dysgraphia screening tool.

Corresponding Author:

Mohamed Ikermane
Laboratory LABSI, Faculty of Sciences, Ibn Zohr University
Agadir, Morocco
Email: ikermane.mohamed@gmail.com

1. INTRODUCTION

Dysgraphia is a disability of writing competency that can manifest by poor or unreadable writing or writing that necessitates a significant amount of effort and time to accomplish [1], [2]. We should make a distinction between the two forms of dysgraphia, acquired dysgraphia which occurs when established brain connections are interrupted by an event (for example, a brain injury) resulting in the loss of previously acquired writing abilities. Developmental dysgraphia, on the other hand, is the difficulty in gaining writing abilities despite adequate learning opportunities and cognitive capability [3]. Regardless of how much they practice, 10 to 30 percent of children never develop sufficient levels of good handwriting [4]; boys are more likely than girls to have dysgraphia [5]. Such handwriting impairments reduce writing legibility and/or speed, which may negatively affect a child's academic and behavioral development [6]. Additionally, it may have long-term effects since individuals who have trouble writing may continue to have difficulty progressing professionally and performing daily tasks [7].

Different tests can be used to determine the existence of dysgraphia, such as the detailed assessment of speed of handwriting (DASH) [8], the ajuriaguerra scale (E scale) [9], the concise evaluation scale for children's handwriting (BHK) [10], and also the hebrew handwriting evaluation (HHE) [11]. Table 1 lists all characteristics of the most common dysgraphia detection tests, which consist of repeatedly writing a sentence or copying a lengthy passage on paper with a pen or pencil. These tests only evaluate the subject's handwriting quality and speed; they do not consider the dynamics of the hand movements during the writing process. Using digital handwriting, the implementation of a graphics tablet with a stylus has permitted the exploration of novel dynamic handwriting features such as velocity, acceleration, pen pressure, altitude, and azimuth. It might greatly aid in comprehending the dysgraphic handwriting particular, speed up the dysgraphia screening duration, and improve the test accuracy. Digital handwriting has gained popularity as graphic tablets have become more widely available. It enabled the assessment of the handwriting dynamics not just the static handwriting evaluation scores (speed and quality). By using digital handwriting features and artificial neural network (ANN) techniques, this work aims to provide a quick and accurate digital alternative dysgraphia screening tool to the current standard tests.

Table 1. Features of several dysgraphia detection tests

	Age group (y)	Test period (min)	Language	Number of scoring items
DASH	9–16	20	English	5
E scale	6–12	2	French	37
BHK	2–18	5	Multi-language	13
HEE	6–18	5	Hebrew	10

2. RELATED WORKS

According to Rosenblum *et al.* [12] used a digitizing tablet to study disturbances in handwriting legibility and speed in 30 primary school-aged children (15 competent and 15 non-proficient hand writers). Handwriting features such as the standard deviations of letter width (t=2.96, c=0.008), letter height (non-proficient=3.24, c=0.005), and pen elevation (non-proficient=2.91, c=0.008) revealed significant variations between competent and non-proficient hand writers. According to Mekyska *et al.* [13] employed a random forest model to identify children with dysgraphia in the hebrew alphabet. The research comprised 54 third-grade Israeli youngsters who had to repeatedly write one hebrew letter to extract their handwriting features (static, kinematic, pressure, and tilt) as an input of the classifying model.

Drotár and Dobeš [14] used a machine learning technique to identify dysgraphia-affected handwriting by assembling a new handwritten dataset comprised of different handwriting tasks and extracting a diverse set of features to capture various aspects of handwriting. They tested numerous machine learning methods and determined that the adaptive boosting (AdaBoost) approach produced the best results, with nearly 80% accuracy. According to Asselborn *et al.* [15], in an administered experiment, 298 youngsters (242 normal, 56 dysgraphic) were requested to pass the BHK dysgraphia test by writing a long text for 5 minutes on paper posed on a digital tablet. A dataset of BHK test results (quality and writing speed) and 53 digital handwriting features was defined and used to train a random forest classifier to diagnose dysgraphia. This method achieved 96.6% sensibility and 99.2% specificity. Research by Gargot *et al.* [16], examined the influence of age and found the 13 most significant features on dysgraphia screening based on the dataset of their previous work.

3. THE CONCISE ASSESSMENT SCALE FOR CHILDREN'S HANDWRITING

The "BHK test" (the concise assessment scale for children's handwriting) [10] is the gold standard dysgraphia screening tool extensively used in french-speaking nations. It consists of writing a text onto a blank piece of paper for five minutes, starting with simple monosyllabic words and progressing to increasingly complicated words. Thirteen handwriting performance characteristics Table 2 are evaluated to get a final handwriting quality score. The handwriting quality score is a decrement score. Higher scores indicate more faults and lower quality. Each feature is assigned a score between 0 and 5. The first two items (writing size and widening of left-hand margin) are graded on the basis of the subject's age and the full written work. The next 11 items are graded based on whether or not a certain handwriting quality item is present in the first five phrases of the 1st paragraph. When present, a score of 1 is assigned. As a consequence, a quality item score of 0 to 5 is assigned. The total score on all 13 items is used to determine whether the child's handwriting is not dysgraphic (score less than 21), ambiguous (score between 22 and 28), or has a handwriting issue (dysgraphia) (score equal or higher than 29). A speed score is also offered by counting the number of characters written in five minutes and then converted into decile scores based on the child's grade [17].

Table 2. The BHK test thirteen handwriting characteristic items used to evaluate the quality score

BHK test characteristic			
1.	Writing is too large	8.	Inconsistent letter size (of x-height letters)
2.	Widening of left-hand margin	9.	Letter distortion
3.	Bad letter or word alignment	10.	Ambiguous letter forms
4.	Insufficient word spacing	11.	Correction of letter forms
5.	Chaotic writing	12.	Unsteady writing trace
6.	Absence of joins	13.	Incorrect relative height of the various kinds of letters
7.	Collision of letters		

4. MATERIALS AND METHOD

4.1. Participants

The current work was based on a previous work dataset [16] in which 280 students, 231 subjects from various Grenoble schools and 49 from the center for language and learning disorders at Grenoble University Hospital, were asked to take a dysgraphia test, with 62 of them having dysgraphia and 218 having usual developing handwriting. 128 of them were females and 152 males. Their age ranged from 6 to 11 years old, with an SD of 1.47. 247 subjects were right-handed and 33 were left-handed. Exclusion criteria included having a known particular impairment or defined disease, such as any neuro-developmental disorder, or being a non-French native.

4.2. Digital handwriting characteristics dataset

The BHK test was scored by two junior and senior psychomotor therapists. The demographics and clinical characteristics of the youngsters were not disclosed. To extract the handwritten digital characteristics, 280 subjects took the BHK test on blank paper mounted on a Wacom Intuos tablet. Twelve features that fell into four categories Table 3, in addition to the subject's age, gender, laterality, BHK quality score, and BHK speed score, were retrieved in a dataset.

Table 3. Digital handwriting characteristics exactrated using digitizer tablet

Category	Digital feature	Signification
Static features	Space between words	the average distance between words over the whole text
	Standard deviation of	density is the number of points captured in each 300-pixel
	Median of power spectral of tremor Frequencies	High value indicates struggling with the handwriting
Kinematics features	Distance to mean of speed frequencies	Wider distance reveals handwriting difficulties
	Distance to mean of speed frequencies	Wider distance reveals handwriting difficulties
	In-air-time ratio	how long he did not touch the surface while writing
Pressure features	Mean pressure	the mean of all pressure points recorded while writing
	Mean speed of pressure change	calculated using a block of ten recorded pressure's points
	Standard deviation of speed of pressure change	Same speed of pressure calcul method as feature below
Tilt features	Distance to mean of tilt-x frequencies	High distance indicates eclectic handwriting
	Bandwidth tiltx	Handwriting struggling causes wider bandwidth
	Median of power spectral of tilt-y Frequencies	lower value indicates handwriting problems

4.3. Artificial neural network model

Motivated by the intricate functionality of human brains [18], which include hundreds of billions of linked neurons that process information simultaneously. An ANN (or simply neural network) [19] is made up of three sections: an input layer of neurons (or nodes, units), one or two (or more) hidden layers of neurons, and a final layer of output neurons. Determining the number of hidden layers and the number of neurons in each layer is one of the most important considerations in the deployment of ANN. Hyperparameter tuning or hyper tuning refers to the process of finding the best set of layers and neurons for the deep learning model. For hyperparameter optimization, we utilized a Keras-tuner [20], which is a package that supports determining the optimal collection of hyperparameters for TensorFlow applications by iterating over all the possible combinations of the number of hidden layers and the number of neurons in each hidden layer. We changed two parameters: the number of layers (between 2 and 5) and the number of neurons in each layer (min=10 and max=40) [21]. As indicated in Figure 1, we used the python deep learning APi Keras [22] to build the ANN model, which is a simple, fast, and powerful tool for building ANN models.

```
def build_model(hp):
    model = keras.Sequential()
    for i in range(hp.Int('num_layers', 2, 5)):
        model.add(layers.Dense(units=hp.Int('units_' + str(i),
                                            min_value=10,
                                            max_value=40,
                                            step=1),
                               activation='relu'))
    model.add(layers.Dense(1, activation='sigmoid'))
    model.compile(
        optimizer='adam',
        loss='binary_crossentropy',
        metrics=['accuracy'])
    return model
```

Figure 1. Predicting the best number of hidden layers and neurons using keras-tuner method

Activation functions are an important part of the role that ANN may play in dealing with non-linear situations [23]. The rectified linear unit (ReLU) [24] is one of the most often used activation functions in deep neural networks:

$$ReLU\ (z) = \max(0, z) \tag{1}$$

As for the output layer, we used the sigmoid function, which has a range between 0 and 1. It is defined as (2):

$$\sigma(z) = \frac{1}{1 + \exp(-z)} \tag{2}$$

The loss functions [25] are a measure of how accurate the ANN model is at predicting the expected outcome, and since we're dealing with a binary classification model (normal handwriting or dysgraphic), as a loss function, the binary cross-entropy is employed, it can be represented as (3):

$$-(y \log(p) + (1 - y) \log(1 - p)) \tag{3}$$

The ANN model needs to update certain of its parameters, such as weight and biases, to deliver the most precise results. By adjusting the values of the parameters of the ANN model, an optimizer function is utilized to minimize the output of the loss function (learn). The adam optimizer [26] is the one that is most frequently used. Furthermore, to avoid the ANN model overfitting, we applied the dropout strategy [27], which involves briefly removing neurons from the hidden layers during model training. Figure 2 shows the structure of the ANN model with the best hyper-parameters values for the dysgraphia dataset determined by the keras-tuner. Our model is made up of an input layer with 15 neurons representing dataset dimensions and three hidden layers, each with 22 neurons, 26 neurons, and 14 neurons. We used two dropout layers (with p=0.5) after each of the two first fully connected hidden layers, as proposed in the original paper [27], and an output layer with one neuron for binary dysgraphia classifying.

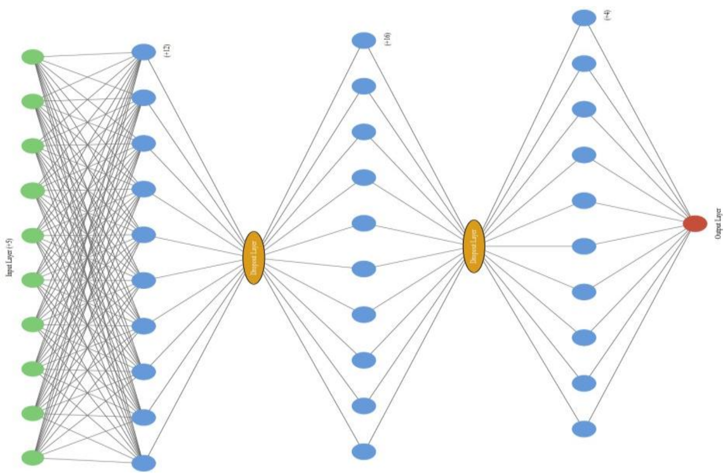
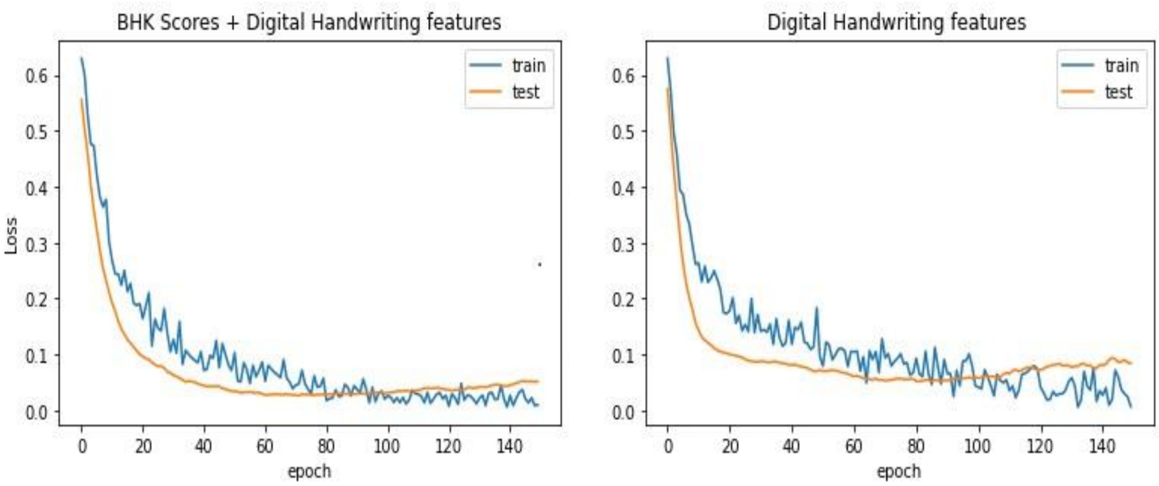


Figure 2. The ANN model employed in dysgraphia prediction with two dropout layers

5. RESULTS AND DISCUSSION

The purpose of this work is to present a dysgraphia detection tool that is quick, reliable, and accurate and only requires digital handwriting as input. The dataset used contains the subject's age, BHK test quality, and speed scores, and digital handwriting characteristics such as static, kinematics, pressure, and tilt features extracted while writing BHK text on a paper attached to a graphic tablet Wacom Intuos. The rest of this section will compare the performance of the ANN model when BHK test scores and digital handwriting characteristics are combined as input and when only digital handwriting features are used. We employed the 'train test split' method to split the dataset in a 75:25 ratio to fit and evaluate the performance of our dysgraphia ANN model, which means that 75% of the data will be used to train the model and 25% will be used as a testing data to analyze the behavior of our model and evaluating its prediction accuracy.

The learning curves may be used to diagnose and analyze a machine learning model's behavior. Underfit learning curves are the results of a model that is unable to learn the training dataset. It happens when the model is unable to get a low enough error value on the training set. Overfitting, on the other hand, refers to a model that has learned the training dataset too well. Overfitting has the disadvantage that the more specialized the model gets in training data, the less successful it can be in generalizing to new data. The use of dropout techniques may aid in obtaining a good fit that exists between the two (overfit and underfit). Figure 3 illustrates the training and validation loss of our model across 150 epochs and a batch size of 10. The blue "training loss" curve measures the model's error on the training set, whereas the orange "validation/test loss" curve assesses the performance of the ANN model on the validation data. Both curves (train and valid/test) were generated using the same set of data by using the random state option in the train test split method. The only difference is that we used all dataset features on Figure 3(a), including the subject's age, as well as his BHK scores for quality and speed, and the 12 digital handwriting features. On Figure 3(b), we only used the age feature and the twelve digital handwriting characteristics. We can see that the model fits well for both inputs, with a slightly higher loss when only digital handwriting features are used as model input.



(a) (b)

Figure 3. Learning curve (a) using all dataset features and (b) learning curve using only digital handwriting characteristics

Regardless of the dataset imbalance, where 78% of the subjects are normal writers (218 subjects) and only 22% are dysgraphic (62 individuals). The classification report in Table 4 illustrates the effectiveness of predictions from our ANN model, similar to the training phase using both inputs. The proposed screening tool performed admirably in dysgraphia predicting attending 96% accuracy using digital handwriting characteristics. Metrics are defined as (4)-(7):

$$P\ recision = \frac{TP}{TP+FP} \tag{4}$$

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

$$F1score = \frac{2\ Precision\ Recall}{Precision + Recall} \quad (6)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

The true positive (TP) represents a dysgraphic subject that was correctly predicted dysgraphic, the false positive (FP) is a dysgraphic subject that was labeled non-dysgraphic, the true negative (TN) is when a non-dysgraphic subject is correctly predicted, and the false negative (FN) is non-dysgraphic children who were mistakenly predicted dysgraphic.

Table 4. Classification report using both Inputs for the ANN model

	BHK scores & digital handwriting features input	Digital handwriting features input
Precision 4	0.929	1.000
Recall 5	0.929	0.786
F1 score 6	0.929	0.880
Accuracy 7	0.971	0.96

Dysgraphia classification results using digital handwriting features appear to be promising. With the ANN model prediction being highly accurate even with a small dataset. This could open the door to new screening methods that do not require a well-trained person in the administration and scoring of standard dysgraphia screening tests such as the BHK test that are automate, quick, and with good accuracy test results.

6. CONCLUSION

Dysgraphia is a learning condition that affects the speed and legibility of a person's handwriting. Because handwriting is important for behavioral and academic development, early detection and rehabilitation can dramatically improve student academic and professional accomplishment. In this work, we provided an automated and quick method for dysgraphia screening employing a digitizer tablet to extract digital handwriting characteristics as input to an ANN dysgraphia classification model. Even with a small dataset, the results of using only digital handwriting characteristics were encouraging, with a precision of 96 percent. Subsequently, new digital handwriting features, as well as the use of various deep learning algorithms, could be considered to improve the classifier's accuracy and automate the dysgraphia identification process. Potentially leading to a novel self-administered dysgraphia screening tool.

REFERENCES

[1] P. J. Chung, D. R. Patel, and I. Nizami, "Disorder of written expression and dysgraphia: definition, diagnosis, and management," *Translational Pediatrics*, vol. 9, no. 1, pp. 46–54, 2020, doi: 10.21037/tp.2019.11.01.

[2] P. Chung and D. R. Patel, "Dysgraphia," *International Journal of Child and Adolescent Health*, vol. 8, no. 1, pp. 27–36, 2015, doi: 10.4324/9780429030277-5.

[3] S. S. Gubbay and N. H. de Klerk, "A study and review of developmental dysgraphia in relation to acquired dysgraphia," *Brain and Development*, vol. 17, no. 1, pp. 1–8, 1995, doi: 10.1016/0387-7604(94)00110-J.

[4] A. Kushki, H. Schwellnus, F. Ilyas, and T. Chau, "Changes in kinetics and kinematics of handwriting during a prolonged writing task in children with and without dysgraphia," *Research in Developmental Disabilities*, vol. 32, no. 3, pp. 1058–1064, 2011, doi: 10.1016/j.ridd.2011.01.026.

[5] V. W. Berninger and M. O'M. May, "Evidence-based diagnosis and treatment for specific learning disabilities involving impairments in written and/or oral language," *Journal of Learning Disabilities*, vol. 44, no. 2, pp. 167–183, 2011, doi: 10.1177/0022219410391189.

[6] V. W. Berninger *et al.*, "Treatment of handwriting problems in beginning writers: transfer from handwriting to composition," *Journal of Educational Psychology*, vol. 89, no. 4, pp. 652–666, 1997, doi: 10.1037/0022-0663.89.4.652.

[7] M. McCloskey and B. Rapp, "Developmental dysgraphia: an overview and framework for research," *Cognitive Neuropsychology*, vol. 34, no. 3–4, pp. 65–82, 2017, doi: 10.4324/9781351020107-1.

[8] A. L. Barnett, S. E. Henderson, B. Scheib, and J. Schulz, "Development and standardization of a new handwriting speed test: the detailed assessment of speed of handwriting," *British Journal of Educational Psychology*, vol. 2, no. 6, pp. 137–157, 2009, doi: 10.1348/000709909X421937.

[9] L. H. -Bletz and A. W. Blöte, "A longitudinal study on dysgraphic handwriting in primary school," *Journal of Learning Disabilities*, vol. 26, no. 10, pp. 689–699, 1993, doi: 10.1177/002221949302601007.

[10] R. Karlsdottir and T. Stefansson, "Problems in developing functional handwriting," *Perceptual and Motor Skills*, vol. 94, no. 2, pp. 623–662, 2002, doi: 10.2466/pms.2002.94.2.623.

[11] N. Erez and S. Parush, "The hebrew handwriting evaluation," School of Occupational Therapy, 1999.

[12] S. Rosenblum, D. Chevion, and P. L. (Tamar) Weiss, "Using data visualization and signal processing to characterize the handwriting process," *Pediatric Rehabilitation*, vol. 9, no. 4, pp. 404–417, 2006, doi: 10.1080/13638490600667964.

[13] J. Mekyska, M. F. -Zanuy, Z. Mzourek, Z. Galaz, Z. Smekal, and S. Rosenblum, "Identification and rating of developmental

dysgraphia by handwriting analysis," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 2, pp. 235–248, 2017, doi: 10.1109/THMS.2016.2586605.

[14] P. Drotár and M. Dobeš, "Dysgraphia detection through machine learning," *Scientific Reports*, vol. 10, no. 1, pp. 1–11, 2020, doi: 10.1038/s41598-020-78611-9.

[15] T. Asselborn *et al.*, "Automated human-level diagnosis of dysgraphia using a consumer tablet," *Npj Digital Medicine*, vol. 1, no. 1, pp. 1–9, 2018, doi: 10.1038/s41746-018-0049-x.

[16] T. Gargot *et al.*, "Acquisition of handwriting in children with and without dysgraphia: a computational approach," *Plos One*, vol. 15, no. 9, pp. 1–22, 2020, doi: 10.1371/journal.pone.0237575.

[17] I. M. B. -Stork, J. Bosga, and R. G. J. Meulenbroek, "Dysgraphic handwriting development and inclusive education: the role of interdisciplinary counseling," *Open Journal of Social Sciences*, vol. 03, no. 08, pp. 35–47, 2015, doi: 10.4236/jss.2015.38003.

[18] P. Sajda, "Neural networks," in *Encyclopedia of the Human Brain*, Netherlands: Elsevier, 2002, pp. 373–383, doi: 10.1016/B0-12-227210-2/00236-3.

[19] A. Amani and D. Mohammadyani, "Artificial neural networks: applications in nanotechnology," in *Artificial Neural Networks - Application*, C. L. P. Hui, Ed. InTech, 2011, pp. 465–478, doi: 10.5772/15551.

[20] Star, "Keras tuner," *Keras Tuner*, 2019. [Online]. Available: https://keras.io/keras_tuner/ (accessed Jul. 20, 2022).

[21] M. I. C. Rachmatullah, J. Santoso, and K. Surendro, "Determining the number of hidden layer and hidden neuron of neural network for wind speed prediction," *PeerJ Computer Science*, vol. 7, 2021, doi: 10.7717/peerj-cs.724.

[22] N. K. Manaswi, "Understanding and working with keras," in *Deep Learning with Applications Using Python*, Berkeley, CA: Apress, 2018, pp. 31–43, doi: 10.1007/978-1-4842-3516-4_2.

[23] N. Kulathunga, N. R. Ranasinghe, D. Vrinceanu, Z. Kinsman, L. Huang, and Y. Wang, "Effects of nonlinearity and network architecture on the performance of supervised neural networks," *Algorithms*, vol. 14, no. 2, pp. 1–17, 2021, doi: 10.3390/a14020051.

[24] A. F. Agarap, "Deep learning using rectified linear units (ReLU)," *Arxiv-Computer Science*, no. 1, pp. 2–8, 2018.

[25] K. Janocha and W. M. Czarnecki, "On loss functions for deep neural networks in classification," *Schedae Informaticae*, vol. 1, 2017, doi: 10.4467/20838476SI.16.004.6185.

[26] D. P. Kingma and J. L. Ba, "Adam: a method for stochastic optimization," *Arxiv-Computer Science*, vol. 1, pp. 1–15, 2015.

[27] K. Sanjar, A. Rehman, A. Paul, and K. JeongHong, "Weight Dropout for Preventing Neural Networks from Overfitting," in *2020 8th International Conference on Orange Technology (ICOT)*, 2020, pp. 1–4, doi: 10.1109/ICOT51877.2020.9468799.

BIOGRAPHIES OF AUTHORS

Mohamed Ikermane is a Ph.D student member of Engineering Science Laboratory at Faculty of Sciences, Ibn Zohr University, Morocco. With a master in distributed systems computing at Faculty of Science, Ibn Zohr University, Agadir (2014). He obtained Bachelor Degree in computer science at Polydisciplinary Faculty of Errachidia. He works as IT teacher at regional center for education and training professions in Guelmim, Morocco. His research interests on autism detection using deep learning techniques. He can be contacted at email: ikermane.mohamed@gmail.com.

Abdelkrim El Mouatasim is a full professor (PES) Faculty Polydisciplinary Ouarzazate, Ibn Zohr University, Morocco coordinator of mathematical and computer sciences (SMI) President of the Moroccan Association of artificial intelligence representative of UNESCO's member states (Morocco) for recommendation on the ethics of AI. His research interest is numerical analyses and optimization, operation research, mathematical modeling and statistical modeling, stochastic processes, algorithms, programming and computer simulation, image processing, data mining, data science and big data, machine learning, and deep learning. He can be contacted at email: a.elmouatasim@uiz.ac.ma.