

# Personality Prediction from Handwriting using Fine-tuned Transfer Learning Models

Jaishri Tiwari, Ritesh Sadiwala\*

Department of Electronics and Communication Engineering, RKDF University, Gandhinagar, Bhopal, Madhya Pradesh, India.

## ABSTRACT

Experts in computational linguistics have done a number of research to identify and categorize personality characteristics at various aspects, including terms, sentences, paragraphs, and recommendations. In this research work five stages model was proposed and the experimental result was evaluated on handwriting images. This paper presents a comparative analysis of fine-tuning transfer learning convolutional neural network models such as VGG16, ResNet50 and GoogleNet for personality detection. The results of fine-tuned models are assessed using the accuracy, precision, recall and f1\_score measure. From the results it was observed that ResNet50 have achieved best accuracy as compared to GoogleNet and VGG16.

**Keywords:** Personality prediction, Handwriting, Image processing, Deep learning, Transfer learning.

*SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology* (2023); DOI: 10.18090/samriddhi.v15i01.06

## INTRODUCTION

Learning, pondering, perceiving, memorizing, and feeling are some of the several cognitive activities and states of mind that the various disciplines of cognitive science seek to investigate. Among the abovementioned forms of cognition, personality is crucial in determining how people behave in social situations. For a very long period of time, investigation on computer-based personality identification and categorization has been ongoing. Writing, photos, video, and audio are just a few of the several types of media that can be used to identify personalities.<sup>[1]</sup> The field of cognitive-based emotion classification from social media language was inspired by a number of difficulties in cognitive computation research. There has been a significant amount of research in this field, and much more could have been undertaken. Written text cognitive-based sentiment analysis (SA) isn't just a theoretical discipline; there are many areas where it's been put to use, including finance,<sup>[5]</sup> learning,<sup>[4]</sup> wellness,<sup>[2]</sup> and other areas.<sup>[6]</sup> It can bridge the gaps between the abstractions of behavioral science and the more recent field of personality recognition from a people's written response displayed on social networking sites++++<sup>[7]</sup> since it is a fusion of cognitive science and human neurology. Past few years have seen an unanticipated global expansion of social networking services like Twitter, Facebook, and Instagram. For instance, Twitter had approximately 330 million monthly active consumers as of the third quarter of 2019.<sup>[8]</sup> Researchers now have the chance of utilizing big information sources for retrieving and analyzing the written text personality characteristics demonstrated by consumers while utilizing social networking sites due to advancements in natural language processing

---

**Corresponding Author:** Ritesh Sadiwala, Department of Electronics and Communication Engineering, RKDF University, Gandhinagar, Bhopal, Madhya Pradesh, India, e-mail: ritesh14ci@gmail.com

**How to cite this article:** Tiwari, J., Sadiwala, R. (2023). Personality Prediction from Handwriting using Fine-tuned Transfer Learning Models. *SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology*, 15(1), 38-44.

**Source of support:** Nil

**Conflict of interest:** None

---

and textual analytics, provided that the data scientific experts operating on social media subject matter are capable of addressing the difficult problems unique to certain subject matter.

People have been using handwritten text as a form of interaction and interpretation for hundreds of years, but it has only lately been discovered that handwriting has connections to both the psychological and cerebral functions of humans. Graphology is the psychological study of handwriting with the goal of identifying the personality characteristics, psychosocial factors, temperament, or behavioral patterns of the person who writes. It is still an inconclusive field because there is no accepted benchmark, with most handwriting understandings being made arbitrarily by qualified graphologists. Nevertheless, numerous studies have demonstrated the connection between handwriting and the neurophysiological characteristics of people. One such investigation is that of Plamondon,<sup>[9]</sup> which demonstrated that the brain creates personalities relying on having written practices and that every neurological brain sequence creates

a different neuromuscular motion identical for people with much a similar sort of individuality. From this standpoint, writing precisely reflects a person's intellect. The weights of movements [10], the progression of handwriting,<sup>[11]</sup> the manner the letters "t" and "y" are chosen to write,<sup>[12]</sup> as well as many other characteristics linked to how letters or texts are written or how the text is arranged on the page, are just a few of the many handwriting characteristics that graphologists presently examine in order to evaluate the psychosocial factors of the person who writes.

Experts in computational linguistics have done a number of research to identify and categorize personality characteristics at various aspects, including terms, sentences, paragraphs, and recommendations.<sup>[13,14]</sup> Nevertheless, according to several research, personality identification and categorization is a difficult domain in cognitive computations. The earlier research<sup>[13]</sup> on consumer personality recognition only employed a single deep learning convolutional neural network model. The research by<sup>[9]</sup> took advantage of the CNN framework, which only captures the local aspects of sentences and discards their prior context. Hence, additional research is needed to overcome various problems in order to effectively identify and classify a person's personality characteristics.

Researchers suggested a mixed deep-learning approach for identifying individual personalities. The existing research<sup>[13]</sup> on the identification and categorization of consumer personalities served as a model for the presented model. For the classification of personality traits, one deep learning framework has been utilized in earlier investigations. The suggested method uses a combination of deep learning frameworks to effectively categorize the person's personality attributes.

## LITERATURE REVIEW

Chaubey *et al.*<sup>[19]</sup> involved the predicting personality characteristics using handwriting examination. Graphology is the name given to a well-known technique of conducting this examination. It relies on how effective the graphologist is at doing so whether personality characteristics may be predicted from handwriting. The creation of a computer-based method for handwriting evaluation has significant potential for forecasting personality characteristics. In this study, the convolutional neural network (CNN) is employed to forecast the five major personality qualities of the author at the moment of writing.

Sajeevan *et al.*<sup>[20]</sup> suggested a method to evaluate the effectiveness of computer-assisted handwriting assessment and determine the best machine learning categorization method for graphological features. A program is being created that will collect handwriting specimens and produce the writer's personality characteristics. Pen pressure, page margin, and word size are the three key aspects of a handwriting specimen that the app examines. A database of the results from the testing of 300 specimens of handwriting

was created. Support vector machine and K-nearest neighbor models are trained and tested using the datasets. The categorization reports show that the SVM and KNN models have high accuracy rates of 96 and 85%, correspondingly. A smaller group of people was used to test the app, and the results were positive. More than 65% of them expressed satisfaction with the app's features and results. The presented approach is subject to further improvements, such as splitting handwriting samples into words and characters and evaluating their attributes.

Thomas *et al.*<sup>[21]</sup> described the handwriting characteristics associated with neuroticism. Additionally, it suggests a technique to automatically identify handwriting characteristics linked to neurotic characteristics, decreasing the bias and time needed to perform a conventional handwriting examination. The study shows that handwriting traits linked to neuroticism might be identified with machine learning algorithms with a testing precision of more than 95%. With greater than 99% precision, the handwriting characteristics, such as asymmetry in the regions and forceful slashing strokes, were recognized.

Li *et al.*<sup>[22]</sup> suggested an easily interpretable and easy-to-understand model called the vanilla compositional network (VCN) by fusing a convolutional neural network with a sequence modeling layout (*i.e.*, a recurrent neural network or Converter), which takes advantage of the prior contextual knowledge of the handwriting personality. VCN is a two-stage design, although performing significantly better than the prior state-of-the-art SOLHCCR algorithms. Due to its reliance on contextual information, it struggles from high vulnerability when dealing with badly written symbols, including such sloppy handwriting, lacking, or interrupted lines. Furthermore, researchers provided a unique deep spatial and contextual information fusion network to increase the OLHCCR model's robustness (DSCIFN). It heavily incorporates the spatial characteristics of handwritten letters and their prior contextual information in a multi-layer fusion module and uses an autoregressive framework that has been pre-trained on a large-scale phrase corpus.

Joseph *et al.*<sup>[23]</sup> introduced a hybrid technique for character identification of MODI script in this study called WT-SVD (Wavelet Transform-Singular Value Decomposition). The extraction of features process is accomplished using the WT-SVD approach, which combines the singular value decomposition and wavelet transform techniques. For the categorization, Euclidean distance is utilized. Utilizing Symlets and Biorthogonal wavelets, the experimentation is carried out, and the outcomes are contrasted. The most accurate technique was the one utilizing biorthogonal wavelet feature retrieval.

Chaudhari *et al.*<sup>[24]</sup> analyzed various strategies for characteristic extraction to forecast a writer's personality and reveal linkages between handwriting and personality psychology. Personality characteristics can be understood through handwriting characteristics that have psychological

backing. The article links these characteristics and promotes the usage of computer-based graphology for predicting personality. The use of graphology in numerous areas is also covered.

Chitlangia *et al.*<sup>[25]</sup> analyzed and comprehended people's handwriting to categorize the characteristics of the writer. However, the traditional procedure of handwriting examination is time-consuming, expensive, and heavily relies on the graphologists' abilities. In order to automate the procedure, researchers took many characteristics from handwriting specimens and categorized the writer into five personality types: enthusiastic, extroverted, introverted, disorganized, and positive. In order to provide an input for the Support Vector Machine system, which produces an output representing a person's personality attribute, the histogram of oriented gradient (HOG) method retrieves characteristics from a specimen of the writer's handwriting. In order to complete this work, 50 distinct individuals' digital handwriting samples were gathered. The Polynomial kernel used in the presented method forecasts a person's personality attribute with an accuracy of 80%. In this study, researchers presented an automated approach for predicting personality traits from user handwriting. To gauge and contrast the effectiveness of the suggested system, two distinct approaches are used on the identical handwriting specimen information.

Bhade *et al.*<sup>[26]</sup> presented an approach that consists of three primary stages: picture pretreatment, handwriting feature detection, and personality characteristic mapping. The procedure of converting a handwriting specimen into a form that can be analyzed quickly and effectively in subsequent phases is known as picture pre-processing. These procedures include grayscale conversion, thresholding, noise reduction, and picture morphology. Handwriting characteristics are extracted during characteristic recognition. The left margin, right margin, and word gap are retrieved as three characteristics of handwriting. Last but not least, the rule-based approach is used to map retrieved characteristics to personality. Regarding three handwriting characteristics, a writer's personality is shown. The suggested technique has a 90% accuracy rate in predicting a person's personality.

Fatimah *et al.*<sup>[27]</sup> used both of the six characteristic approaches. Four particular letters—"a," "g," "s," and "t"—were examined utilizing a Convolutional Neural Networks (CNN) categorization method as part of the multi-structure study, which included features of margin, a gap between lines, a gap between phrases, slope, and dominant region. The precision of the structured method, according to the outcomes, was up to 82.5–100%, whereas the accuracy of the symbol approach utilizing CNN had a precision of up to 98.03 percent of fresh data with 7–10 minutes of practice.

Valdez-Rodríguez *et al.*<sup>[28]</sup> aimed to infer user personality characteristics from their handwritten messages. In this article, researchers proposed architecture built on a Convolutional Neural Network (CNN) as a classifier and utilized the scanned

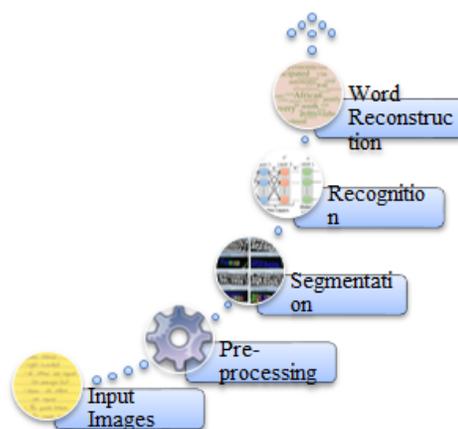


Figure 1: Working flow of Proposed Model

picture of the subject's handwritten article separated into portions to categorize them. The original database consisted of 418 color photos. They then extracted 216 portions of every picture in grayscale, binarized those, and got about 90,000 pictures. Five convolutional layers make up the CNN, which also has three entirely interconnected layers for categorization. The five convolutional layers retrieve characteristics from the regions.

## METHODOLOGY

Pre-processing, segmentation, character recognition, word reconstruction, and conflict resolution are the five primary components that make up the research work that is being suggested. The sequential representation of this work is shown in the following figure, which is labeled in "Figure 1." It is compatible with photographs in any format that may be input.

### Pre-processing

It has been determined that the picture of the document will serve as the input image. The first thing that is done is the pre-processing step, which involves identifying the borders, cropping the borders, applying the transformation to straighten the page, applying the sharpening kernel, identifying and cropping the borders, and eliminating noise, correcting skew and size. In order to get the initial input image ready for further processing, it must first be scanned and examined. Cropping, aligning the page, and checking the borders are all tasks that may be accomplished. Using a Gaussian filter with an Otsu threshold, the noise is eliminated after determining where the page boundary is located. After that, a greyscale image is created from the original picture.

### Character Segmentation

After the image has been pre-processed, the next significant step is the segmentation of the image into words. During the segmentation process, an open-source library known as Tesseract-ocr is utilized. This library generates a stream of characters that are then sent on to the subsequent stage.





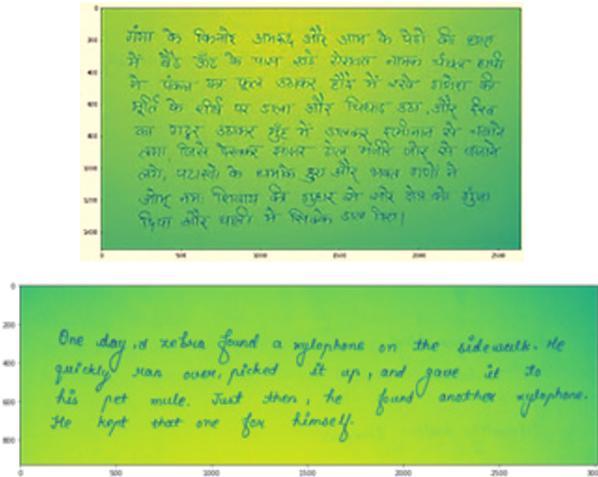


Figure 4: Handwriting samples.

For fine-tuned model, parametric ReLU is adopted because it fine-tunes the learning parameters on its learning rate without any vanishing gradient problem.

### Word Reconstruction

The dictionary’s definition will serve as the basis for the forecast of the word. A confidence rating will be given to the predicted character behavior after it has been predicted. The prediction with the second and third-highest degree of confidence will be chosen as the output rather than the forecast with the highest level of confidence.

## RESULTS AND DISCUSSIONS

The experimental result was evaluated on image dataset collected from Kaggle resource.<sup>[29]</sup> Below in Figure 4, sample of input image for personality prediction is presented. The dataset contains images of the following personalities:

- Openness: Individuals who prefer to try new things and learn new things typically have high openness scores. It possesses traits, including being innovative, insightful, curious, and diverse interests.
- Conscientiousness: People who have self-control, attentiveness, dependability, and are always on time have high Conscientiousness: It has traits like being organized and methodic.
- Extraversion: People gain energy from conversing with others and exchanging ideas. Talkative, energetic,

- sociable, positive feelings and assertiveness are the traits of being extraversion.
- Agreeableness: These individuals are polite, helpful, and compassionate. It possesses traits like kindness, affection, and sympathy.
- Neuroticism: Neuroticism relates with impulse control, intensity of unpleasant emotions, and emotional stability. The traits of being neuroticism are moody, anxiousness, jealousy, and tension.

Further, the results of fine-tuned models are assessed using the accuracy, precision, recall and f1\_score measure as shown in Table 1. To determine whether the model can converge with so few iterations and which suffered from the degrading issue, every experimentation was run for a total of 100 epochs. All fine-tuned networks are hyper-parameterized. The batch size of 64 for the network’s training was used. This study presents a comparative state of art deep learning models for fine-tuned deep learning models. Accuracy, Precision, recall and f1\_score are used for the performance evaluation. Mathematically, they are represented as:

$$Accuracy = \frac{TruePositive + TrueNegative}{TruePositive + TrueNegative + FalsePositive + FalseNegative} \quad (i)$$

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (ii)$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad (iii)$$

$$F1\_score = \frac{2}{1/precision + 1/Recall} \quad (iv)$$

Where true positive, true negative, false positive, false negative are true positive, true negative, false positive and false negative instances, respectively.

Figure 5 shows the training and validation performance of VGG16 considering accuracy, loss and ROC as performance parameters. The epoch varied from 0-100 as the graph converged before that. For accuracy vs epoch, the accuracy converges at around 10 epoch. The average accuracy for VGG16 is 40.2%. For the second performance parameter, the loss first decreases from 0-10 epoch and then becomes constant after 10. For third performance parameter i.e. ROC for classes, were considered from class 0- class 4. The ROC is shown in Figure 5. Figure 6 shows the training and validation performance of ResNet50 considering accuracy, loss and ROC as performance parameters. The epoch varied from 0-100 as the graph converged before that. For accuracy vs epoch, the accuracy converges at around 30 epoch. The average accuracy for VGG16 is 88%. For the second performance parameter,

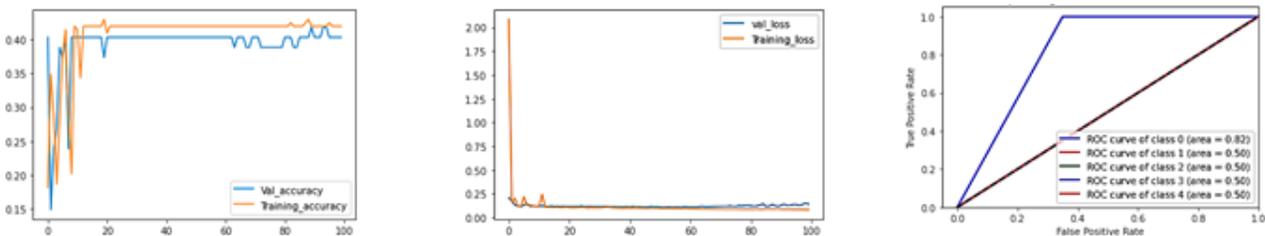


Figure 5: Training and validation performance for VGG16



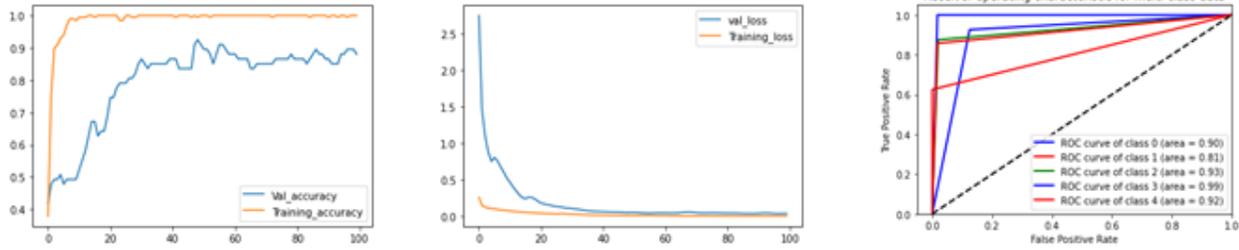


Figure 6: Training and validation performance for ResNet50

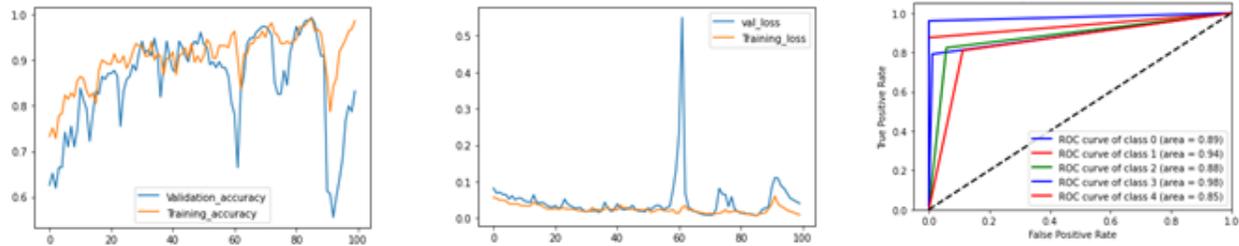


Figure 7: Training and validation performance for GoogleNet

Table 1: Comparative performance evaluation on deep learning models

	Accuracy	Precision	Recall	F1_score
VGG16	0.402	0.27	0.40	0.32
ResNet50	0.88	0.89	0.88	0.88
GoogleNet	0.832	0.88	0.83	0.85

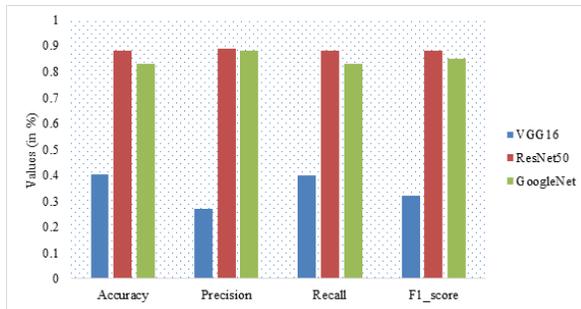


Figure 8: Comparative performance evaluation on deep learning models.

the loss first decreases from 0-30 epoch and then become constant after 30 epoch. For third performance parameter *i.e.* ROC for classes, were considered from class 0- class 4. The ROC is shown in Figure 6. Figure 7 shows the training and validation performance of GoogleNet considering accuracy, loss and ROC as performance parameters. The epoch varied from 0-100 as the graph converged before that. The average accuracy for VGG16 is 83.2%. For second performance parameter, the loss shows peaks at 60 epochs. For third performance parameter *i.e.*, ROC for classes were considered from class 0- class 4. The ROC is shown in Figure 7.

Table 1 compares VGG16, ResNet50, and GoogleNet considering Accuracy, Precision, Recall and F1\_Score as performance parameters. For VGG16 accuracy is 40.2%,

precision is 27%, Recall is 40% and F1\_Score is 32%. For ResNet50 accuracy is 88%, precision is 89%, Recall is 88% and F1\_Score is 88%. For GoogleNet accuracy is 83.2%, precision is 88%, Recall is 83% and F1\_Score is 85%. The accuracy, precision, recall and f1 Score is maximum for ResNet50. Figure 8 shows the comparative performance evaluation on deep learning models VGG16, ResNet50 and GoogleNet. The accuracy, precision, recall and f1 Score is maximum for ResNet50. The accuracy is maximum improved by 47.8%, precision is improved by 61%, recall is improved by 48% and f-1 score is improved by 56% when compared with VGG16 and GoogleNet.

## CONCLUSION

The paper have presented a five-stage model for personality prediction from handwritten images with the help of deep learning and image processing tools. This paper has compared and analyzed the fine-tuning of a cutting-edge deep convolutional neural network for personality detection using image processing tools. Furthermore, the accuracy, precision, recall, and f1 score measures are used to evaluate the outcomes of fine-tuned models. Some existing pre-trained transfer learning models such as VGG16, ResNet50, GoogleNetis fine-tuned and compared their performance. From result analysis, ResNet50 have achieved 88% of accuracy, GoogleNet have achieved 83% accuracy and VGG16 have achieved only 40% accuracy. Therefore,

ResNet50 is most adaptable learning model for personality prediction. In future, the work will be extended to improve the performance of designed models.

## REFERENCES

- [1] H. Ahmad, M. Z. Asghar, A. S. Khan, and A. Habib (2020 July), "A systematic literature review of personality trait classification from textual content", *Open Comput. Sci.*, 10(1), 175–193.
- [2] H. Ahmad, M. Z. Asghar, F. M. Alotaibi, and I. A. Hameed, (2020 Oct), "Applying deep learning technique for depression classification in social media text," *J. Med. Imag. Health Informat.*, 10(10), 2446–2451. doi: 10.1166/jmihi.2020.3169.
- [3] S. Acharya, A.K. Pant, P.K. Gyawali, 2015 "Deep learning based large scale handwritten Devanagari character recognition", *International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, 1–6.
- [4] Rahul and R. Katarya (2019 Dec), "A review: Predicting the performance of students using machine learning classification techniques," in *Proc. 3rd Int. Conf. I-SMAC*, pp. 36–41.
- [5] G. Kou, Y. Xu, Y. Peng, F. Shen, Y. Chen, K. Chang, and S. Kou (2021 January), "Bankruptcy prediction for SMEs using transactional data and two-stage multi objective feature selection," *Decis. Support Syst.*, 140(Art. no. 113429).
- [6] Kazameini, S. Fatehi, Y. Mehta, S. Eetemadi, and E. Cambria (2020), "Personality trait detection using bagged SVM over BERT word embedding ensembles", arXiv:2010.01309.
- [7] L. Liu, D. Preotiuc-Pietro, Z. R. Samani, and M. E. Ungar (2016), "Analyzing personality through social media profile picture choice," in *Proc. Int. AAAI Conf. Social Media (ICWSM)*, 211–220.
- [8] All things Statista. Number of Monthly Active Twitter Users Worldwide From 1st Quarter 2010 to 1st Quarter 2019 Retrieved From. Accessed: Jan. 18, 2020. [Online]. Available: <https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/>
- [9] R Plamondon, (2010 November) Neuromuscular studies of handwriting generation and representation. *International Conference on Frontiers in Handwriting Recognition (ICFHR)*, 261 (2010) Kolkata.
- [10] Y Tang, X Wu, W Bu, Offline text-independent writer identification using stroke fragment and contour-based features (2013 June), *IEEE International Conference on Biometrics (ICB)*, 1–6
- [11] M Naghibol hosseini, F Bahrami, (2008 December) A behavioral model of writing. *International Conference on Electrical and Computer Engineering (ICECE)*, 970–973
- [12] HN Champa, KR Ananda kumar (2010 August), "Automated human behavior prediction through handwriting analysis", *First International Conference on Integrated Intelligent Computing (ICIIC)*, 160–165
- [13] N.Majumder, S.Poria, A.Gelbukh, and E. Cambria (2017 March), "Deep learning based document modeling for personality detection from text," *IEEE Intell. Syst.*, 32(2), 74–79.
- [14] D. Xue, L. Wu, Z. Hong, S. Guo, L. Gao, Z. Wu, X. Zhong, and J. Sun (2018 November), "Deep learning-based personality recognition from text posts of online social networks," *Int. J. Speech Technol*, 48(11), 4232–4246.
- [15] Champa H N and Dr. K R Ananda Kumar (2010 May), "Artificial neural network for human behavior prediction through handwriting analysis", *International Journal of Computer Applications*, 2(2), 0975– 8887.
- [16] H. Devi (2006 Dec), "Thresholding: A Pixel-Level Image Processing Methodology Preprocessing Technique for an OCR System for the Brahmi Script", *Ancient Asia Journal Vol. 1*.
- [17] Mr. Danish Nadeem & Miss. Saleha Rizvi, "Character Recognition Using Template Matching", [https://www.cs.uic.edu/~srizvi/BIT\\_Thesis.pdf](https://www.cs.uic.edu/~srizvi/BIT_Thesis.pdf)
- [18] "Artificial Neural Network", Available at: <http://cogsci.stackexchange.com/questions/8509/dynamic-al-systems-theory-as-a-metaphor-in-psychology-is-it-useful-or-not>
- [19] Specter, A. Z. 1989. Achieving application requirements. in *Distributed Systems*, S. Mullender
- [19] Chaubey, Gyanendra, and Siddhartha Kumar Arjaria. (2022), "Personality Prediction Through Handwriting Analysis Using Convolutional Neural Networks." *Proceedings of International Conference on Computational Intelligence*. Springer, Singapore, 2022.
- [20] Sajeevan, S., and Wiraj Udara Wickramaarachchi. (2022), "Detection of Personality Traits Through Handwriting Analysis Using Machine Learning Approach." *Advances on Smart and Soft Computing*. Springer, Singapore, 79–89.
- [21] Thomas, Sheetal, et al. (2020), "Application of Machine Learning to Detect Neuroticism in Individuals Using Handwriting Analysis." *ICDSMLA 2020*. Springer, Singapore, 521–531.
- [22] Li, Yunxin, et al. (2022), "Fast and Robust Online Handwritten Chinese Character Recognition with Deep Spatial & Contextual Information Fusion Network." *IEEE Transactions on Multimedia*.
- [23] Joseph, Solley, and Jossy George. (2021), "Efficient handwritten character recognition of MODI script using wavelet transform and SVD." *Data Science and Security*. Springer, Singapore, 227–233.
- [24] Chaudhari, Kinjal, and Ankit Thakkar. (2019), "Survey on handwriting based personality trait identification." *Expert Systems with Applications* 124, 282–308.
- [25] Chitlangia, Aditya and G. Malathi. (2019), "Handwriting Analysis based on Histogram of Oriented Gradient for Predicting Personality traits using SVM." *Procedia Computer Science* 165, 384– 390.
- [26] Bhade, Vasundhara, and Trupti Baraskar. (2018), "A model for determining personality by analyzing off-line handwriting." *Advances in Machine Learning and Data Science*. Springer, Singapore, 345–354.
- [27] Fatimah, Sri Hastuti, et al. (2019), "Personality features identification from handwriting using convolutional neural networks." *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*. IEEE.
- [28] Valdez-Rodríguez, José E., Hiram Calvo, and Edgardo M. Felipe Riverón. (2018) "Handwritten Texts for Personality Identification Using Convolutional Neural Networks." *International Conference on Pattern Recognition*. Springer, Cham.
- [29] <https://www.kaggle.com/code/brsdincer/personalityprediction-with-handwriting-end-to-end/data>

