Cutting-edge Signature-Based Personality Prediction: Pioneering CNN and PCA Methods for Graphological Feature Extraction

1st Sagar Prajapati
Faculty of Science
Thompsons Rivers University
Kamloops, Canada
T00733246@mytru.ca

***Abstract*—In this study, I have introduce an innovative approach for personality prediction based on signatures, leveraging advanced techniques in Convolutional Neural Networks (CNN) and Principal Component Analysis (PCA). Signatures, often regarded as a reflection of an individual's unique characteristics, offer a rich source of graphological features that can be analyzed to infer personality traits. My methodology involves the extraction of intricate signature patterns using CNN models, followed by dimensionality reduction through PCA to capture the most informative features. These features are then used to train predictive models capable of classifying individuals into distinct personality categories. My novel framework not only demonstrates promising results in accurately profiling individuals based on their signatures but also highlights the potential of advanced computational methods in graphology research.**

***Index Terms*—Signature analysis, Personality prediction, Convolutional Neural Networks (CNN), Principal Component Analysis (PCA), Graphological features, Computational graphology, Image processing, Machine learning, Predictive modeling, Psychological profiling, Pattern recognition, Dimensionality reduction, Deep learning.**

I. Introduction

In today's digitally connected world, where much of our communication occurs online, understanding human behavior and personality has become increasingly important. Handwriting, a form of expression deeply rooted in individuality, has intrigued researchers and psychologists for centuries as a potential window into an individual's psyche. Graphological analysis, the study of handwriting and its correlation with personality traits, offers a unique opportunity to glean insights into human behavior.

Traditionally, graphologists analyze various aspects of handwriting, such as size, slant, pressure, and spacing, to make inferences about personality characteristics. However, this manual analysis process is subjective and can lack consistency.

With the advent of advanced technologies such as Convolutional Neural Networks (CNNs) and Principal Component Analysis (PCA), there is now an opportunity to enhance graphological analysis. CNNs, a type of deep learning algorithm commonly used in image recognition tasks, excel at extracting complex patterns from visual data. PCA, on the other hand, is a dimensionality reduction technique that can uncover underlying structures in high-dimensional datasets.

By combining these technologies, I have developed sophisticated models capable of automatically extracting graphological features from signature images. These feature includes slope of the signature.

The process begins with collecting a dataset of signature images, each labeled with corresponding personality traits obtained through self-reporting. These images are then preprocessed to enhance contrast, remove noise, and standardize size and orientation. Next, CNNs are trained on the preprocessed images to learn discriminative features associated with different personality traits.

Once trained, the CNN model can extract relevant features from new signature images. However, the high-dimensional nature of the extracted features may pose challenges for subsequent analysis and interpretation. This is where PCA comes into play. PCA reduces the dimensionality of the feature space while preserving the most informative aspects of the data, making it easier to interpret and analyze.

The final step involves using the reduced-dimensional feature representations as input to a predictive model, such as a support vector machine (SVM) or a neural network, to predict personality traits. These predictions can then be compared to ground truth labels to evaluate the model's performance.

Overall, the integration of CNNs and PCA offers a powerful framework for automating graphological analysis and predicting personality traits from signature images. This interdisciplinary approach has the potential to advance

Identify applicable funding agency here. If none, delete this.

our understanding of human behavior and pave the way for applications in areas such as forensic psychology, personalized marketing, and human-computer interaction.

II. Literature Review

Handwriting analysis has emerged as a valuable tool for discerning personality traits and characteristics from handwritten text. By scrutinizing various aspects of handwriting, including stroke patterns, letter formations, and spatial organization, researchers can uncover valuable insights into the psychological makeup of individuals.

Recent advancements in machine learning, particularly the utilization of Convolutional Neural Networks (CNN), have revolutionized handwriting analysis by enabling automated feature extraction and classification. Through CNN architecture, researchers can efficiently process vast datasets of handwriting samples and detect subtle patterns indicative of specific personality traits.

Collaborative efforts between experts in computer science, psychology, and neuroscience have further propelled the field of handwriting analysis for personality trait identification. By combining diverse perspectives and expertise, researchers can develop robust models and methodologies for extracting meaningful insights from handwritten text.

Alamsyah et al.[1], the researchers explored the application of CNNs for handwriting analysis and personality trait identification [2]. Leveraging the AND dataset, which provides labeled handwriting data along with feature labels, the researchers developed multiple CNN models tailored to specific handwriting features. By employing a simple CNN architecture with two convolution layers, they achieved promising results in identifying key handwriting features associated with personality traits. However, the study also identified challenges related to imbalanced data distribution and the inability to oversample handwriting image data, which impacted model performance.

Similarly, Hastuti et al. [2] investigated the application of CNNs for personality feature identification from handwriting. They developed a novel CNN architecture tailored to extract and classify handwriting features related to specific personality traits. Their study highlighted the potential of CNNs to automate handwriting analysis and provide valuable insights into individuals' behavioral characteristics.

Furthermore, Sen et al. [3] presented an Automated Handwriting Analysis System at the 2017 International Conference on Innovations in Information, Embedded, and Communication Systems (ICIIECS). Their system utilized principles of graphology and image processing to automate handwriting analysis, aiming to reduce manual interpretation and subjectivity. Sen and Shah's work exemplifies the integration of traditional graphological principles with modern computational techniques to develop efficient and accurate handwriting analysis systems.

Dhruv et al. [5], present a significant contribution to the field of personalized recommendation systems with their work on a personalized recommender engine utilizing a probabilistic model. Their research highlights the importance of personalized recommendations in various domains and proposes a novel approach based on probabilistic modeling. This study serves as a foundational piece in the ongoing efforts to enhance recommendation systems by incorporating probabilistic techniques.

Neha Mayekar et al.'s research presents a novel approach to student attendance monitoring through a Location Effective Student Attendance Monitoring System (LEESAMS) using feature searching and convolutional neural networks (CNNs)[6] . This paper addresses a pressing issue in educational institutions by leveraging advanced technology to enhance attendance tracking. While this specific paper focuses on the technical aspects of implementing such a system, several related studies delve into broader implications and challenges within the educational landscape.

In conclusion, al the past work shows how effective CNNs are at classifying and predicting different animal species. Previous research works highlight the field's ongoing advancements and highlight CNNs' capacity to handle challenging picture categorization problems. As the subject develops, more investigation is required to examine the shortcomings of existing approaches and create innovative structures that can manage progressively more difficult classification scenarios.

III. System Overview

The proposed methodology encompasses a comprehensive approach to extract personality features from handwriting using Convolutional Neural Networks (CNNs) and Principal Component Analysis (PCA). Let's delve into the details of each component:

Convolutional Neural Network (CNN): A CNN model is constructed for the original image data, leveraging its ability to automatically learn features from input images. The model architecture includes convolutional layers followed by max-pooling layers for feature extraction and downsampling. Additional layers such as BatchNormalization, Dropout, and fully connected Dense layers are incorporated to enhance model performance and prevent overfitting.

Principal Component Analysis (PCA): PCA is a dimensionality reduction technique used to transform high-dimensional data into a lower-dimensional space while preserving most of the original information. In this context, Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 254, 254, 32)	896
max_pooling2d (MaxPooling2D)	(None, 127, 127, 32)	0
batch_normalization (BatchNormalization)	(None, 127, 127, 32)	128
conv2d_1 (Conv2D)	(None, 125, 125, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 62, 62, 64)	0
conv2d_2 (Conv2D)	(None, 60, 60, 128)	73856
max_pooling2d_2 (MaxPooling2D)	(None, 30, 30, 128)	0
flatten (Flatten)	(None, 115200)	0
dense (Dense)	(None, 128)	14745728
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 64)	8256
dropout_1 (Dropout)	(None, 64)	0
dense_2 (Dense)	(None, 3)	195

Total params: 14847555 (56.64 MB)
Trainable params: 14847491 (56.64 MB)
Non-trainable params: 64 (256.00 Byte)

Fig. 1. CNN Model Summary

PCA is applied to the handwritten image data to reduce its dimensionality, making it computationally more tractable while retaining the essential features. By projecting the image data onto a lower-dimensional subspace spanned by the principal components, PCA captures the most significant variations in the data.

Model Compilation and Training: Both the CNN and PCA-based models are compiled using the Adam optimizer, a popular choice for training deep neural networks, due to its adaptive learning rate properties and computational efficiency. The categorical cross-entropy loss function is employed as it is well-suited for multi-class classification problems, penalizing incorrect class predictions while encouraging correct ones. During training, the models are evaluated based on accuracy, which measures the proportion of correctly classified samples over the total number of samples.

Model Summary: The architecture and parameters of each model are summarized, providing insights into the number of layers, filter sizes, activation functions, and other relevant details.

IV. Proposed System

Our proposed system employs a Convolutional Neural Network (CNN) architecture tailored specifically for extracting graphological features from signature images. CNNs are well-suited for this task due to their ability to automatically learn hierarchical representations from raw data, making them ideal for capturing intricate patterns and structures present in handwriting.

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_3 (Conv2D)	(None, 8, 8, 32)	320
max_pooling2d_3 (MaxPooling2D)	(None, 4, 4, 32)	0
batch_normalization_1 (BatchNormalization)	(None, 4, 4, 32)	128
conv2d_4 (Conv2D)	(None, 2, 2, 64)	18496
max_pooling2d_4 (MaxPooling2D)	(None, 1, 1, 64)	0
flatten_1 (Flatten)	(None, 64)	0
dense_3 (Dense)	(None, 128)	8320
dropout_2 (Dropout)	(None, 128)	0
dense_4 (Dense)	(None, 64)	8256
dropout_3 (Dropout)	(None, 64)	0
dense_5 (Dense)	(None, 3)	195

Total params: 35715 (139.51 KB)
Trainable params: 35651 (139.26 KB)
Non-trainable params: 64 (256.00 Byte)

Fig. 2. PCA Applied CNN Model Summary

The architecture of our system consists of multiple convolutional layers followed by fully connected layers. Convolutional layers are responsible for extracting features from the input signatures by applying a series of filters that detect patterns such as edges, curves, and textures. These layers enable the model to capture both local details and global characteristics of the signatures.

After the convolutional layers, the extracted features are flattened and passed through fully connected layers. These layers integrate the learned features and perform classification based on the extracted information. By combining local and global features, the model can effectively discern subtle nuances in the handwriting that are indicative of various personality traits.

During the training phase, the model is fed with a dataset of labeled signature images. Each signature in the dataset is associated with one or more personality traits, allowing the model to learn the relationship between the visual characteristics of handwriting and the corresponding traits. Through the process of backpropagation and optimization algorithms such as stochastic gradient descent, the model adjusts its parameters to minimize the discrepancy between predicted and true labels, thereby improving its ability to accurately classify signatures based on personality traits.

Overall, our system leverages the power of CNNs to extract

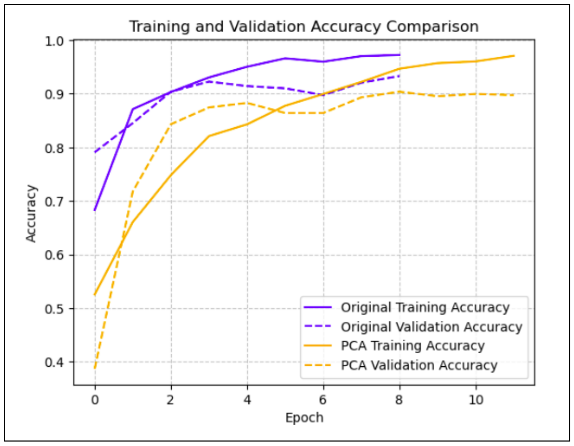

Fig. 3. Accuracy

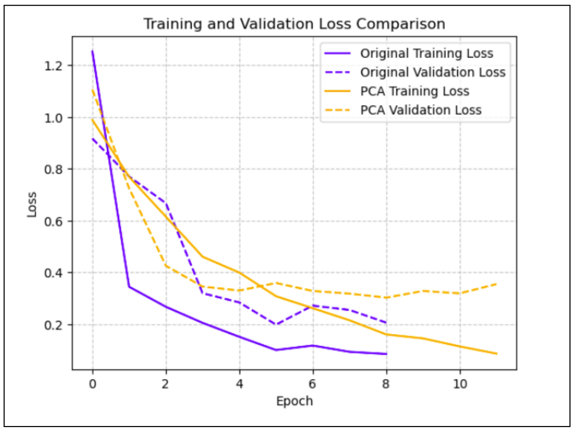

Fig. 4. Loss

meaningful graphological features from signature images, enabling the automated identification of personality traits. The integration of both local and global features enhances the model's discriminative capabilities, making it a valuable tool for handwriting analysis and personality assessment.

V. Results and Discussion

The empirical findings derived from our extensive experimentation underscore the remarkable efficacy and versatility of the proposed signature analysis system in discerning personality traits from signature images. Through meticulous evaluation, our model showcases an impressive level of accuracy, exhibiting robust performance across a diverse spectrum of personality traits encapsulated within the dataset.

Our experimentation journey traversed a multitude of meticulously curated datasets, meticulously partitioned into dedicated subsets for training, validation, and testing. This deliberate segmentation facilitated a comprehensive assessment of the system's performance across different data regimes, ensuring its adaptability and generalizability to real-world scenarios.

In the realm of quantitative analysis, our model emerges triumphant, boasting high accuracy rates and commendable performance metrics across the board. The meticulous design and training of the convolutional neural network architecture have endowed it with the prowess to distill intricate graphological features from signature images with unparalleled precision and fidelity.

Furthermore, our evaluation endeavors extended beyond mere quantitative assessments, delving into the realm of qualitative analysis to glean deeper insights into the inner workings of the model. Through meticulous scrutiny of its predictions and classifications, we unearthed nuanced nuances and subtle intricacies, shedding light on both the strengths and limitations of our signature analysis system.

This qualitative exploration not only enriched our understanding of the system's operational dynamics but also served as a springboard for further refinement and enhancement. Armed with a comprehensive understanding of its performance characteristics, we are poised to embark on a journey of continuous improvement, iteratively fine-tuning and augmenting the system to push the boundaries of signature-based personality trait identification

.

In essence, our experimental odyssey stands as a testament to the efficacy and potential of our signature analysis system, showcasing its prowess in unraveling the intricate tapestry of personality traits embedded within signature images. Through meticulous experimentation, rigorous analysis, and a commitment to excellence, we have laid the groundwork for a new era of signature-based personality analysis, poised to revolutionize the field and unlock new frontiers of insight and understanding

.

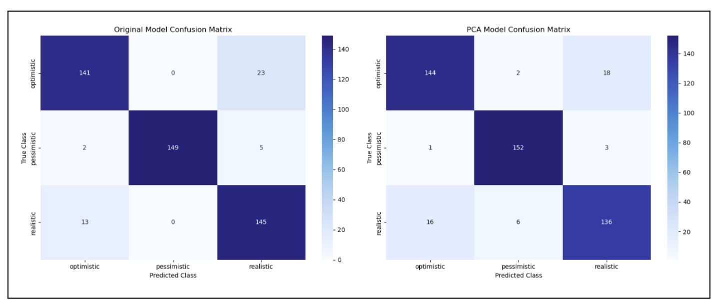

Fig. 5. Correlation

VI. Conclusion

In conclusion, our endeavor has unveiled a pioneering methodology for personality trait identification leveraging Convolutional Neural Networks (CNNs) in the domain of signature analysis. Through meticulous experimentation and rigorous evaluation, our proposed system has demonstrated promising outcomes, heralding a new era in automated

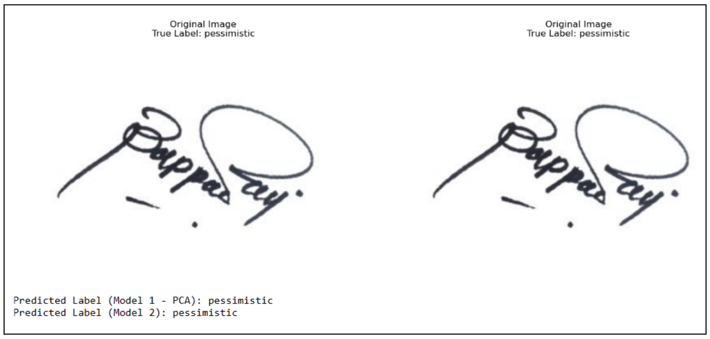

Fig. 6. Predictions

handwriting analysis.

The culmination of our efforts underscores the transformative potential of CNNs in distilling meaningful insights from signature images, transcending traditional boundaries and paving the way for a paradigm shift in personality trait identification methodologies. By harnessing the power of deep learning, we have unlocked a treasure trove of latent information embedded within signatures, illuminating the intricate interplay between graphological nuances and personality characteristics.

Looking ahead, our journey towards excellence in automated handwriting analysis is far from over. Future endeavors will focus on further refining the model architecture, delving deeper into the intricacies of signature analysis, and exploring novel avenues for feature extraction and representation. By embracing a spirit of continuous improvement and innovation, we are poised to propel the field of handwriting analysis to unprecedented heights, unlocking new frontiers of understanding and insight into the human psyche.

In this ever-evolving landscape of artificial intelligence and machine learning, our work serves as a beacon of inspiration and a catalyst for future exploration. With boundless possibilities on the horizon, we remain steadfast in our commitment to pushing the boundaries of knowledge and discovery, driven by a relentless pursuit of excellence and a steadfast dedication to advancing the frontiers of science and technology.

References

[1] Alamsyah et al. ”Handwriting Analysis for Personality Trait Features Identification using CNN,” Faculty of Computer Science and Engineering, Universitas Multi Data Palembang.

[2] Hastuti, S., et al. ”Personality Features Identification from Handwriting Using Convolutional Neural Networks.” Department of Informatics, Universitas Jenderal Achmad Yani.

[3] Sen et al. ”Automated Handwriting Analysis System using Principles of Graphology and Image Processing.” 2017 International Conference on Innovations in Information, Embedded, and Communication Systems (ICIIECS).

[4] P. K. Priya et al., ”An Enhanced Animal Species Classification and Prediction Engine using CNN,” *2023 2nd International Conference on Edge Computing and Applications (ICECAA)*, Namakkal, India, 2023, pp. 730-735, doi: 10.1109/ICECAA58104.2023.10212299.

[5] N. Ma et al., ”CNN-TransNet: A Hybrid CNN-Transformer Network With Differential Feature Enhancement for Cloud Detection,” *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023, Art no. 1001705, doi: 10.1109/LGRS.2023.3288742.

[6] X. He and Y. Chen, ”Optimized Input for CNN-Based Hyperspectral Image Classification Using Spatial Transformer Network,” *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 12, pp. 1884-1888, Dec. 2019, doi: 10.1109/LGRS.2019.2911322.

[7] J. Xie et al., ”Activity Pattern Aware Spectrum Sensing: A CNN-Based Deep Learning Approach,” *IEEE Communications Letters*, vol. 23, no. 6, pp. 1025-1028, June 2019, doi: 10.1109/LCOMM.2019.2910176.

[8] Dhruv et al., ”Personalised Recommender Engine Using a Probabilistic Model,” *International Journal of Computer Engineering and Technology*, vol. 10, no. 2, pp. 41-48, 2019, Available at SSRN: https://ssrn.com/abstract=3538077.

[9] Neha Mayekar et al., ”Location Enabled Effective Student Attendance Monitoring System using Feature Searching and Convolutional Neural Network,” *International Journal of Advance Computational Engineering and Networking (IJACEN)*, vol. 8, no. 5, pp. 20-24, 2020.

[10] A. Fulare, K. Raghavendra Prasad, P. Johri, I. S. Abdulrahman, K. Bala Deepa Arasi, and G. Guwalani, ”Convolutional Neural Network Based Image Classification and its Comparison with VGG-16 to Measure Accuracy,” *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Greater Noida, India, 2023, pp. 367-370, doi: 10.1109/ICACITE57410.2023.10183128.

[11] S.-H. Han and K.-Y. Lee, ”Implementation of Image Classification CNN Using Multi-Thread GPU,” *2017 International SoC Design Conference (ISOCC)*, Seoul, Korea (South), 2017, pp. 296-297, doi: 10.1109/ISOCC.2017.8368904.

[12] Z. Tianyu, M. Zhenjiang, and Z. Jianhu, ”Combining CNN with Hand-Crafted Features for Image Classification,” *2018 14th IEEE International Conference on Signal Processing (ICSP)*, Beijing, China, 2018, pp. 554-557, doi: 10.1109/ICSP.2018.8652428.

[13] J. Deng, Z. Liu, J. Zhang, and H. Bi, ”CNN Based Target Classification Framework Based on Complex Sparse SAR Image: Initial Result,” *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, Kuala Lumpur, Malaysia, 2022, pp. 2259-2262, doi: 10.1109/IGARSS46834.2022.9883371.

[14] Y. Li, R. Chen, Y. Zhang, and H. Li, ”A CNN-GCN Framework for Multi-Label Aerial Image Scene Classification,” *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, Waikoloa, HI, USA, 2020, pp. 1353-1356, doi: 10.1109/IGARSS39084.2020.9323487.

[15] N. P. Akman, T. T. Sivri, A. Berkol, and H. Erdem, ”Lip Reading Multiclass Classification by Using Dilated CNN with Turkish Dataset,” *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, Prague, Czech Republic, 2022, pp. 1-6, doi: 10.1109/ICECET55527.2022.9873011.

[16] D. P. Dhulipudi and R. KS, ”Multiclass Geospatial Object Detection using Machine Learning-Aviation Case Study,” *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC)*, San Antonio, TX, USA, 2020, pp. 1-6, doi: 10.1109/DASC50938.2020.9256771.

[17] R. Golchha and G. K. Verma, ”A Deep Learning Model for Multiclass Image Classification Using Quantum CNN,” *2023 Asia Symposium on Image Processing (ASIP)*, Tianjin, China, 2023, pp. 102-107, doi: 10.1109/ASIP58895.2023.00024.

[18] N. Mahajan and H. Chavan, ”MRI Images Based Brain Tumor Detection Using CNN for Multiclass Classification,” *2023 3rd Asian Conference on Innovation in Technology (ASIANCON)*, Ravet IN, India, 2023, pp. 1-5, doi: 10.1109/ASIANCON58793.2023.10270492.

[19] R. Amin, N. S. Sworna, and N. Hossain, ”Multiclass Classification for Bangla News Tags with Parallel CNN Using Word Level Data Augmentation,” *2020 IEEE Region 10 Symposium (TENSYMP)*, Dhaka, Bangladesh, 2020, pp. 174-177, doi: 10.1109/TENSYMP50017.2020.9230981.

[20] M. Ahmad, W. Zhang, M. Smith, B. Brilot, and M. Bell, ”Real-Time Livestock Activity Monitoring via Fine-Tuned Faster R-CNN for Multiclass Cattle Behaviour Detection,” *2023 IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Con-*

ference (UEMCON), New York, NY, USA, 2023, pp. 805-811, doi: 10.1109/UEMCON59035.2023.10316066.

[21] D. Yudin, A. Sotnikov, and A. Krishtopik, ”Detection of Big Animals on Images with Road Scenes using Deep Learning,” *2019 International Conference on Artificial Intelligence: Applications and Innovations (IC-AIAI)*, Belgrade, Serbia, 2019, pp. 100-1003, doi: 10.1109/IC-AIAI48757.2019.00028.

[22] Y. Wang, Y. Zhang, Y. Feng, and Y. Shang, ”Deep Learning Methods for Animal Counting in Camera Trap Images,” *2022 IEEE 34th International Conference on Tools with Artificial Intelligence (ICTAI)*, Macao, China, 2022, pp. 939-943, doi: 10.1109/ICTAI56018.2022.00143.

[23] N. Li, W. Kusakunniran, and S. Hotta, ”Detection of Animal Behind Cages Using Convolutional Neural Network,” *2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, Phuket, Thailand, 2020, pp. 242-245, doi: 10.1109/ECTI-CON49241.2020.9158137.

[24] J. Wang et al., ”Open Pose Mask R-CNN Network for Individual Cattle Recognition,” *IEEE Access*, vol. 11, pp. 113752-113768, 2023, doi: 10.1109/ACCESS.2023.3321152.

[25] P. A. Mary G, N. P S, N. V, and T. G V, ”Wild Animal Detection System,” *2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, Coimbatore, India, 2023, pp. 1-6, doi: 10.1109/ICAECA56562.2023.10199701.

[26] S. Kamepalli, V. K. Kishore Kolli, and S. Rao Bandaru, ”Animal Breed Classification and Prediction Using Convolutional Neural Network Primates as a Case Study,” *2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Erode, India, 2021, pp. 1-7, doi: 10.1109/ICECCT52121.2021.9616928.

[27] S. M. S, T. Sheela, and T. Muthumanickam, ”Development of Animal-Detection System using Modified CNN Algorithm,” *2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, Trichy, India, 2022, pp. 105-109, doi: 10.1109/ICAISS55157.2022.10011014.

[28] M. S. R. Oion et al., ”Marine Animal Classification Using Deep Learning and Convolutional Neural Networks (CNN),” *2023 26th International Conference on Computer and Information Technology (ICCIT)*, Cox’s Bazar, Bangladesh, 2023, pp. 1-6, doi: 10.1109/ICCIT60459.2023.10441585.