Gazi University
Journal of Science
PART A: ENGINEERING AND INNOVATION
http://dergipark.org.tr/gujsa

An Investigation of Benford's Law Divergence and Machine Learning Techniques for Intra-Class Separability of Fingerprint Images

Aamo IORLIAM[1*], Emmanuel ORGEM[1], Yahaya I. SHEHU[2]

[1]*Department of Mathematics & Computer Science, BSU, Makurdi, Nigeria*
[2]*Shehu Shagari College of Education, Sokoto, Nigeria*

Keywords

Benford's Law
Divergence Values
Machine Learning Techniques
Intra-Class Separability
Fingerprint Images

Abstract

Protecting a biometric fingerprint database against attackers is very vital in order to protect against false acceptance rate or false rejection rate. A key property in distinguishing biometric fingerprint images is by exploiting the characteristics of these different types of fingerprint images. The aim of this paper is to perform an intra-class classification of fingerprint images using Benford's law divergence values and machine learning techniques. The usage of these Benford's law divergence values as features fed into the machine learning techniques has proved to be very effective and efficient in the intra-class classification of biometric fingerprint images. The effectiveness of our proposed methodology was demonstrated on five datasets resulting in a total of 367 samples. All the machine learning techniques used in this experiment were trained using the k-fold cross validation and the dataset was split into ten times (10-folds). The models achieved high intra-class classification mean accuracies of 99.72% for the Convolutional Neural Networks (CNN), and 95.90% for the Naïve Bayes. Again, the Decision Tree and Logistic Regression, achieved accuracies of 95.62%, and 94.47%, respectively. These results showed that Benford's law features and machine learning techniques, especially the CNN and Naïve Bayes can be effectively applied for the intra-class classification of fingerprint images. The implication of these results is that the different types of fingerprint images can be effectively discriminated using Benford's law divergence values and machine learning technique for forensics and biometrics applications.

Cite

Iorliam, A., Orgem, E., & Shehu, Y. I. (2022). An Investigation of Benford's Law Divergence and Machine Learning Techniques for Intra-Class Separability of Fingerprint Images. *GU J Sci, Part A*, *9*(3), 211-224.

Author ID (ORCID Number)	Article Process	
A. Iorliam, 0000-0001-8238-9686	**Submission Date**	22.02.2022
E. Orgem, 0000-0003-3211-876X	**Revision Date**	26.04.2022
Y. I. Shehu, 0000-0001-8924-9344	**Accepted Date**	08.09.2022
	Published Date	26.09.2022

1. INTRODUCTION

Biometric experts have been dependent on fingerprints over the years for verification and identification purposes. There exist different types of fingerprint images which include contact-less acquired fingerprints, optically acquired fingerprints, and synthetically generated fingerprints (Hildebrandt et al., 2013; Maltoni et al., 2009). Since these fingerprints are used for different purposes, they should not be intentionally or unintentionally used for another purpose as this may cause a serious security threat (Iorliam et al., 2016). Therefore, this paper performs an investigation of machine learning techniques for intra-class classification of fingerprint images to classify fingerprint images of different types that have the same modality.

It has been reported in the literature that since 1938, Benford's law has proved beyond reasonable doubt that it possesses the capability to detect/classify original/untampered data from fake/tampered data (Benford, 1938; Hill, 1998; Iorliam et al., 2016). This interesting law (Benford's law) is therefore adopted for the intra-class separability of fingerprint images. Firstly, the gray-scale fingerprint images are used to calculate the first digit distribution of the JPEG coefficients using Equation 2. Furthermore, Equation 3 is used to calculate Benford's law divergence values. These divergence values are fed as inputs into the machine learning techniques such as the Naïve Bayes, Decision Tree, Logistic Regression, and Convolutional Neural Networks.

*Corresponding Author, e-mail: aamoiorliam@gmail.com

The applicability of this research is that it can serve as a preliminary forensic tool in classifying different types of fingerprint images for forensics and biometric applications. The major contributions of this work are summarized as follows:

i. We propose a novel use of Benford's law divergence values to improve the intra-class classification of fingerprint images.

ii. We provide a detailed analysis of the novel intra-class classification of fingerprint images using an empirical study based on theoretical and empirical perspectives.

iii. We used only six Benford's law features (reduced features) and performed intra-class classification with a high-performance evaluation.

The rest of the paper is organized as follows. Related works are described in Section 2. Section 3 describes our experimental setup, need for intra-class separability of fingerprint images, datasets used, divergence metric determination, data pre-processing for intra-class separability of fingerprint images, and evaluation metrics used in our paper. Results and discussions are presented in Section 4. Conclusion and future work are presented in Section 5.

2. RELATED WORKS

Benford's law has proved to be very effective in detecting forged/tampered images (Fu et al., 2007; Iorliam 2016; Iorliam et al., 2017). The Benford's law was first discovered in 1881 by Simon Newcomb, where he noticed that the first pages of the logarithm table containing the first digits were worn more than the last pages of the logarithm table, which meant that people were looking up for numbers starting with 1 more often than numbers starting with 2, and so on (Hill, 1998). Unfortunately, Newcomb could not prove why the theory and formula worked. Then in 1938, Frank Benford proposed the Benford's law, also referred to as the first digit law, which states that multi-digit numbers beginning with 1, 2, or 3 appear more frequently than multi-digit numbers beginning with 4, 5, 6, 7, 8 and 9 (Benford, 1938; Iorliam et al., 2016). Therefore, original/untampered data is expected to follow Benford's law, which is illustrated in Figure 1.

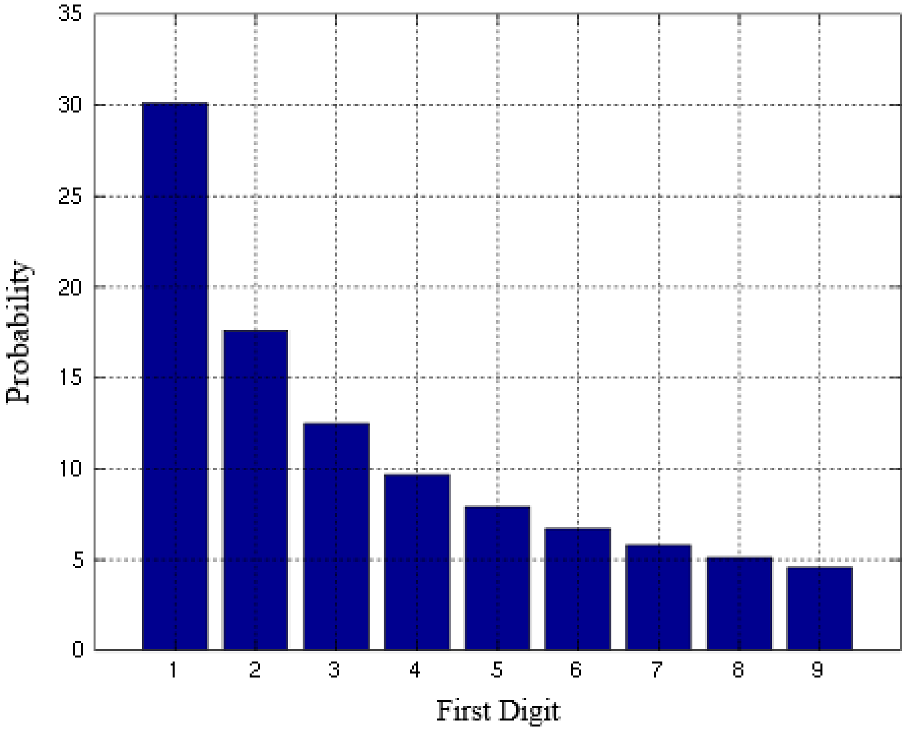

***Figure 1**: The first digit probability distribution of Benford's law (Benford, 1938; Iorliam et al., 2016).*

Taking into consideration the Most Significant Digit, where 0 is not included, and a dataset under investigation satisfies Benford's law, the standard Benford's law is expressed in Equation 1.

$$p(x) = \log_{10}(1 + \frac{1}{x}), x = 1, 2, 3, \ldots, 9 \quad (1)$$

where *x* is the first digit of the number and *p(x)* refers to the probability distribution of *x*.

Since the inception of Benford's law, it is expected that naturally generated datasets should obey this law, whereas tampered or randomly generated datasets should deviate from this law. This inherent characteristic of the Benford's law can lead to important applications in forensics such as detecting anomalies or fraud in a given dataset (Iorliam & Shangbum, 2017; Satapathy et al., 2020) or classifying different types of biometric images (Iorliam et al., 2016; 2017).

Intra-class separability of biometric images means the classification of biometric images that appear to be identical (closely related) (Iorliam et al., 2017). For instance, fingerprint images such as contact-less acquired latent fingerprints, optically acquired fingerprints and synthetically generated fingerprints are closely related due to the fact that they cannot be easily classified based on their physical appearance, hence classifying them could be referred to as intra-class classification of fingerprint images (Iorliam et al., 2016). Even though intra-class classification of biometric images seems to be a novel area, Table 1 summarizes the related work in this area.

***Table 1**: Summary of Related Works*

S/No	Author (s)	Implementation Strategy	Advantages
1	Hildebrandt and Dittmann (2015)	The model employed the use of Benford's law and WEKA's Bagging classifier in 10-fold stratified cross-validation.	Differentiated between real latent fingerprints and printed fingerprints using Benford's law in the spatial domain.
2	Iorliam et al. (2016)	Applied Benford's law and neural networks for the classification of biometric images.	A novel approach for the source identification of captured biometric images.
3	Iorliam and Shangbum (2017)	Used Benford's law with SVM in the biometric fingerprint tampering detection and separability of fingerprint images.	A novel approach to fight against insider attackers and hackers for securing biometric fingerprint images.
4	Hildebrandt (2020)	A thesis that contributed to digital forensics, latent fingerprint processing, and latent fingerprint forgery detection.	A novel contribution to digitized forensic and latent fingerprints.
5	Bonettini, et al. (2021)	Used Benford's law features with a simple Random Forest classifier.	Discriminated GAN-generated images from natural photographs.
6	Our Proposed Method	Used Benford's law features with Naive Bayes, Decision Tree, Logistic Regression, and CNN.	Effectively reduced features and achieved high intra-class separability of fingerprint images for forensics and biometrics applications.

To the best of the researcher's knowledge and review presented, this paper presents for the first time the novel use of Benford's law features with machine learning techniques to accurately classify the fingerprint images.

3. EXPERIMENTAL SETUP

The goal of this experiment is to utilize the acquired Benford's law divergence values from fingerprint images as proposed by Fu et al. (2007) and Iorliam et al. (2017) for the separability of fingerprint images.

In essence, the intra-class classification of biometric fingerprint images is performed on Benford's law divergence values of DB1, DB2, DB3, DB4, and the artificially acquired contact-less latent fingerprints images. The 10-fold cross-validation is applied on the extracted Benford's law divergence values. These fingerprint features are then fed into the Naive Bayes, Decision Tree, Logistic Regression, and Convolutional Neural Networks (CNN) algorithms as input data for separability purposes. The need for the separability of fingerprint images is discussed in Section 3.1.

3.1. Need for Intra-Class Separability of Fingerprint Images

The two key uses of biometrics are verification and identification. Verification is usually a 1-to-1 matching, whereas identification is a 1-to-many matching. For more than a century, fingerprints have been used for identification purposes (Jain et al., 1997; Iorliam et al., 2017). Fingerprints are used for different purposes as explained by Iorliam et al. (2017). Therefore, it is very important to avoid using a particular type of fingerprint for another purpose either intentionally or unintentionally. This can be achieved by studying the characteristic of each type of the different fingerprints and as such identifying the source of the captured fingerprint images. This could be possible if the source hardware that captured the fingerprint image is identified (Bartlow et al., 2009). One way to do this is by utilizing Benford's law divergence values with machine learning techniques to achieve the fingerprint images separability.

3.2. Data Sets Used

The FVC2000 (2000) fingerprint datasets which consists of four different fingerprint databases (DB1, DB2, DB3, and DB4) are used in this paper. Furthermore, artificially printed contact-less acquired latent fingerprint images are used for this research. Therefore, a total of five different datasets are used for testing our proposed model. The first four (4) sets of datasets each contain 80 grayscale fingerprint images (FVC2000, 2000). While the artificially printed contact-less acquired latent fingerprint images have 48 fingerprint biometric images (Hildebrandt et al., 2013). Other details about the datasets used are provided in Table 2.

***Table 2**: Summary Description of Datasets*

Source	Dataset	Sensor Type	Sample No
FVC2000 (2000)	DB1	Low-cost Optical Sensor captured by "Secure Desktop Scanner".	80
	DB2	Low-cost Optical Capacitive Sensor captured by "TouchChip"	80
	DB3	Optical Sensor "DF-90"	80
	DB4	Synthetically generated images from Synthetic Generator	80
Hildebrandt et al., 2013	DB5	Artificially printed contact-less acquired latent fingerprint images	48

3.3. Divergence Metric Determination

Benford's law divergence values are obtained based on the biometric fingerprint dataset used in this paper described in Section 3.2. The divergence metric is used to show how close or far a particular dataset is using the standard or generalized Benford's law. In any case, smaller divergence yields a better fitting. In this paper, the first digit distributions of the JPEG coefficients are extracted from the gray-scale images as demonstrated by Iorliam et al. (2016).

Fu et al. (2007), extended the standard Benford's law to the Generalized Benford's law which closely follows the logarithmic law as expressed in Equation 2.

$$p(x) = N \log_{10}(1 + \frac{1}{s + x^q}) \tag{2}$$

where N is the normalization factor which makes *p(x)* a probability distribution. The model parameters s and q describe the distributions for different fingerprint images and different compressions of the Quality Factor (QF). Through experiments, Fu et al. (2007), provided values for N, s, and q using the Matlab toolbox, which returns the Sum of Squares due to Error (SSE). The N, s, and q values are as shown in Table 3 for the Generalized Benford's law experiments.

***Table 3**: Model Parameters Used for the Generalized Benford's law (Fu et al., 2007)*

Q-factor	**Model Parameters**			**Goodness-of fit (SSE)**
	N	q	s	
100	1.456	1.47	0.0372	$7.104e - 06$
90	1.255	1.563	−0.3784	$5.255e - 07$
80	1.324	1.653	−0.3739	$3.06838e - 06$
70	1.412	1.732	−0.337	$5.36171e - 06$
60	1.501	1.813	−0.3025	$6.11167e - 06$
50	1.579	1.882	−0.2725	$6.05446e - 06$

To test for conformity of a particular dataset (fingerprint images) to Benford's law, one of the most common criteria used is the chi-square goodness-of-fit statistics test (Acebo & Sbert, 2005; Li et al., 2012; Iorliam et al., 2016). The chi-square divergence is expressed in Equation 3.

$$x^2 = \sum_{x=1}^{9} \frac{(P'x - Px)^2}{Px} \tag{3}$$

where $P'x$ is the actual first digit probability of the JPEG coefficients of the fingerprint biometric images and *Px* is the logarithmic law (Generalized Benford's law) as given in Eq. (2). In this study, the fingerprint datasets are singly compressed at a QF of 50 to 100 in a step of 10 (Iorliam et al., 2016). The divergence is calculated as an average on all the datasets earlier described in Table 2.

3.4. Data Pre-Processing for Intra-Class Separability of Fingerprint Images

In this paper, the fingerprint datasets are transformed into a format that can be easily interpreted by the machine learning techniques under consideration. The available biometric fingerprint images are transformed and this results in 368 instances with 6 features and a Class Label as shown in Table 4. The first 6 attributes are compression Quality Factors (QFs) ranging from 50 to 100 in a step of 10. The 7th attribute is the class label, which is represented as 0, 1, 2, 3, and 4 for DB1, DB2, DB3, DB4, and the artificially printed contact-less acquired latent fingerprint images (contact-less), respectively. The pre-processed values are achieved with the help of Benford's law divergence which is explained in Sections 3.3.

***Table 4**: The Pre-processed Dataset*

QF-50	QF-60	QF-70	QF-80	QF-90	QF-100	Class Label
10.49719	10.80152	9.790624	8.71157	7.745091	7.761463	0
11.30665	10.93996	9.681446	8.675576	7.871395	8.270308	0
10.12913	10.14469	9.176165	8.658021	7.656819	7.184597	0
8.979308	9.382513	8.908421	8.30894	7.466352	6.91005	0
11.77114	11.00361	9.720688	8.637295	8.06052	8.575394	0
11.79356	11.31105	9.884672	8.865339	8.049554	8.608477	0
12.07603	11.50856	9.993754	8.806326	8.007224	8.698664	0
11.61334	11.30085	9.944549	8.701875	8.010162	8.350574	0
10.33387	10.20274	9.487949	8.407865	7.774554	7.828174	0
10.26303	9.98995	9.359262	8.556169	7.858129	8.139676	0
8.815993	9.117034	8.976531	8.549314	7.675995	7.306121	0
10.13061	9.755888	8.945211	8.413265	7.927866	8.046223	0

The divergence values (pre-processed dataset) therefore serve as inputs into the machine learning techniques (Naive Bayes, Decision Tree, Logistic Regression and CNN) algorithms considered in this paper. The Python programming language virtual platform (Google Colaboratory) is used to implement the proposed algorithms. The goal of our proposed method is to train these machine learning techniques to carry out the intra-class separability of fingerprint images. Therefore, to avoid any case of over-fitting of the pre-processed data used in this experiment, the 10-fold cross validation is applied on both the training and testing of the model. These algorithms are selected for usage because they are well suited for the labeled datasets considered in this paper. Figure 2 summarises the schematic diagram of the proposed model.

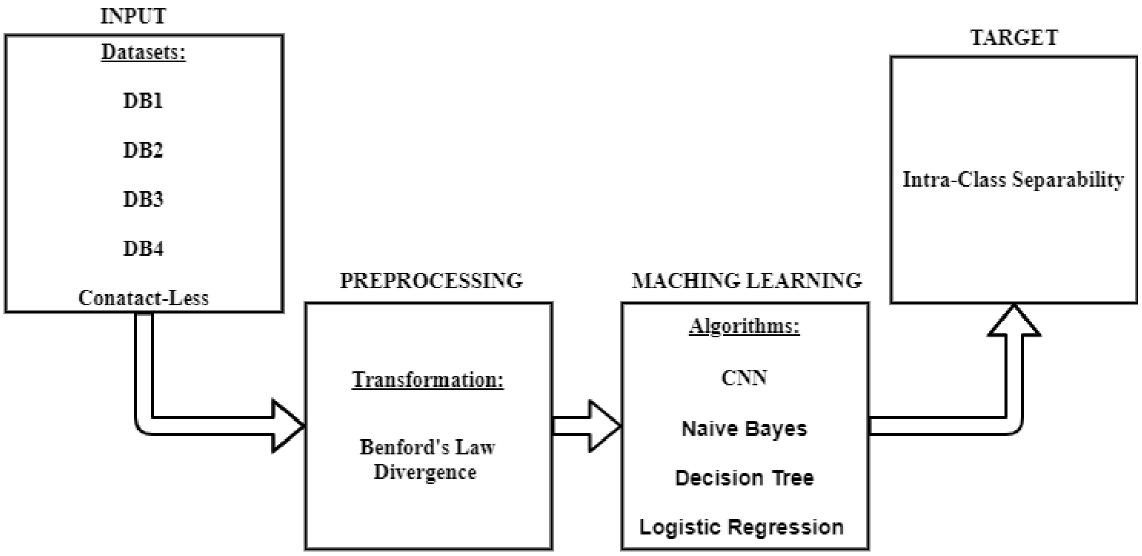

***Figure 2**: Schematic Diagram of the Proposed Model*

3.5. Evaluation Metrics

To evaluate the proposed model, the following evaluation metrics are used:

i. **Accuracy**: This is mathematically expressed by the formula:
Accuracy = (TP + TN) / (TP + TN + FP +FN)

Where:

TP (True Positive): The outcome where the model correctly predicts the positive class.
TN (True Negative): The outcome where the model correctly predicts the negative class.
FP (False Positive): The outcome where the model incorrectly predicts the positive class.
FN (False Negative): The outcome where the model incorrectly predicts the negative class.

ii. **Precision**: This is shown mathematically as:
Precision = (TP) / (TP + FP)

iii. **Recall**: This is shown mathematically as:
Recall = (TP) / (TP + FN)

iv. **F1-Score**: This is mathematically expressed as:
F1-Score = (2 x Precision x Recall) / (Precision + Recall).

4. RESULTS AND DISCUSSIONS

A. Naive Bayes Performance Results

The pre-processed fingerprint data is split ten times (10-folds) and fed into the Gaussian Naïve Bayes algorithm. As shown in Table 5, the Naïve Bayes algorithm achieved a mean accuracy of 95.90%, with a mean F1 score of 96.15%, the mean Precision value of 96.75%, and a 96.25% mean Recall value.

***Table 5**: Naive Bayes Results*

Mean Accuracy:	95.90%
Mean F1 Score:	96.15%
Mean Precision:	96.75%
Mean Recall:	96.25%

Again, Figure 3 shows the confusion matrix result for the Naïve Bayes algorithm. The resulting confusion matrix shows that DB1, and contact-less classes were excellently classified at an accuracy of 100 percent.

However, for DB2, 11% of the fingerprint images were misclassified as DB3, and 89% of the fingerprint images were accurately classified as DB2. Considering DB3, 12% of the fingerprint images were misclassified as DB2, and 88% of the fingerprint images were accurately classified as DB3. For the DB4, 1% of the fingerprint images was misclassified as DB1 and 99% of the fingerprint images were accurately classified as DB4.

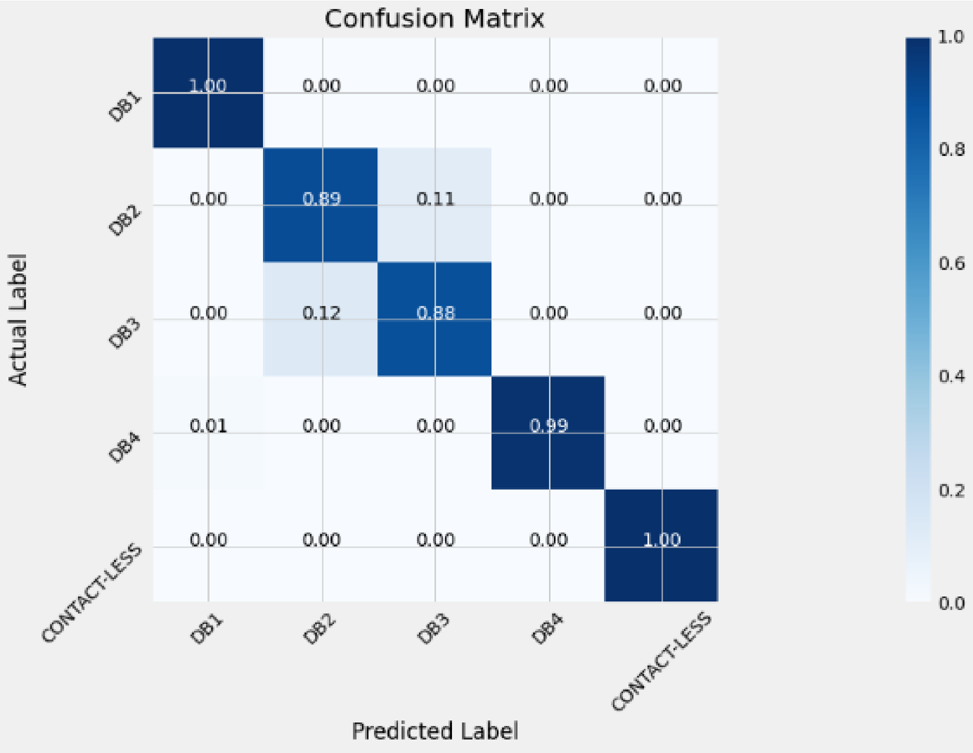

***Figure 3**: The Confusion Matrix for the Naïve Bayes Algorithm*

B. Decision Tree Performance Results

The pre-processed fingerprint data is split ten times (10-folds) and fed into the Decision Tree algorithm. As shown in Table 6, the Decision Tree algorithm achieved a mean accuracy of 95.62 %, with a mean F1 score of 96.42%, the mean Precision value of 96.87%, and a 96.25% mean Recall value.

***Table 6**: Decision Tree Results*

Mean Accuracy:	95.62 %
Mean F1 Score:	96.42%
Mean Precision:	96.87%
Mean Recall:	96.25%

The Decision Tree confusion matrix shows that DB1, DB4, and contact-less classes were excellently classified at 100% accuracy as shown in Figure 4. However, 10% of the fingerprint images in DB2 were misclassified as DB3, and 7% of the fingerprint images in DB3 were misclassified as DB2.

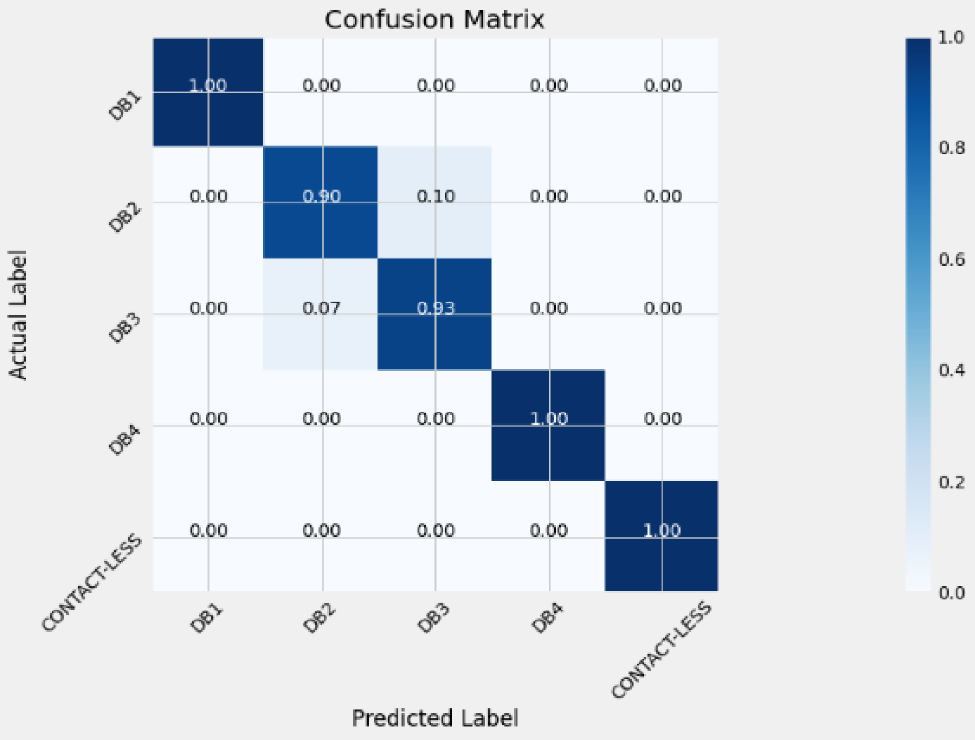

***Figure 4**: The Confusion Matrix for the Decision Tree Algorithm*

C. Logistic Regression Performance Results

The Logistic Regression is also used to classify the fingerprint images into DB1, DB2, DB3, DB4, and the artificially acquired contact-less latent fingerprints images, taking into consideration the independent variables as QF's ranging from 50 to 100 in a step of 10. The pre-processed fingerprint data is split ten times (10-folds) and fed into the Logistic Regression algorithm. In the Logistic Regression, the "max_iter=4000" to enable the algorithm converge properly.

The Logistic Regression algorithm achieved a mean accuracy of 94.47%, with a mean F1 score of 94.47%, the mean Precision value of 94.47%, and a 94.47% mean Recall value as shown in Table 7.

***Table 7**: Logistic Regression Results*

Accuracy:	94.47%
F1 Score:	94.47%
Precision:	94.47%
Recall:	94.47%

The Logistic Regression confusion matrix shows that DB1, and contact-less classes were accurately classified with an accuracy of 100%. For the DB2, 11% of the fingerprint images were misclassified as DB3. Considering DB3, 12% of the fingerprint images were misclassified as DB2, and for the DB4, 1% of the fingerprint images were misclassified as DB1 as shown in Figure 5.

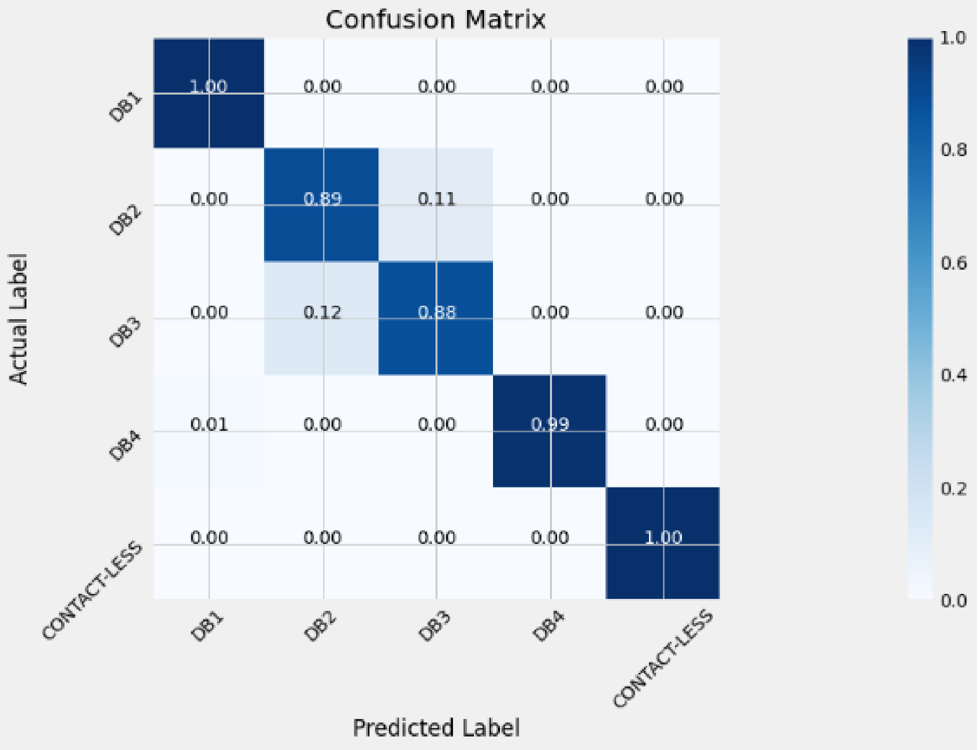

***Figure 5**: The Confusion Matrix for the Logistic Regression Algorithm*

D. Convolutional Neural Network (CNN) Results

The pre-processed fingerprint data is split ten times (10-folds) and fed into the CNN. The sequential model “sequential ( )” is used as the first layer. Furthermore, the first hidden layer had 6 input parameters, and 480 neurons. The rectified linear activation function (ReLu) is first chosen due to its ability to achieve higher performance. Another dense layer is added with 240 neurons. Again, the next dense layer is added with 120 neurons. The model is concluded with 5 dense layers, and a sigmoid activation function.

The binary_crossentropy is used as the loss function, the adam is used as the optimizer, and the accuracy is used as the metrics for the compilation of the CNN model. Two hundred (200) epochs are used in this experiment with a batch size of sixteen (16). Based on these parameters, Figure 6 shows the training loss versus epochs for the intra-class classification of fingerprint images for the CNN technique.

Based on the optimization capacity of the CNN, we always expect a lower loss to produce a better model, especially when considering the training vs. epochs.

We can see in Figure 6 that the loss tends more towards zero around 80 epochs and more. This shows that our proposed model performed well for the training and validation datasets, especially after 80 epochs.

Furthermore, the training accuracy vs. epochs considered in this experiment, for both the training and validation datasets, shows that epochs above 80 show more consistent and higher results that are closer to 1, as shown in Figure 7.

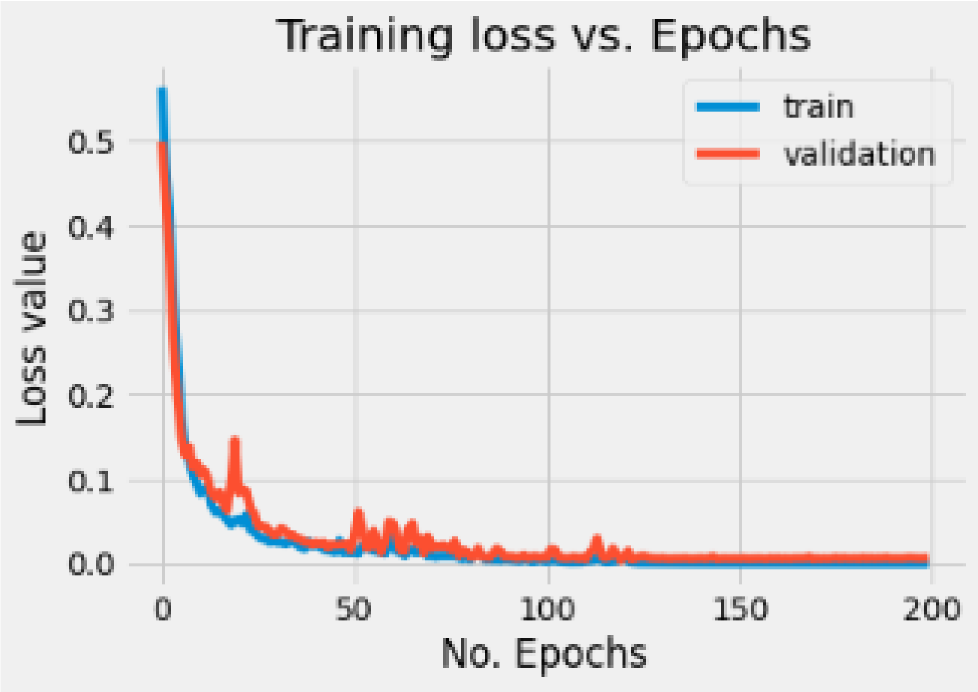

***Figure 6**: The CNN Training Loss Vs Epochs Graph*

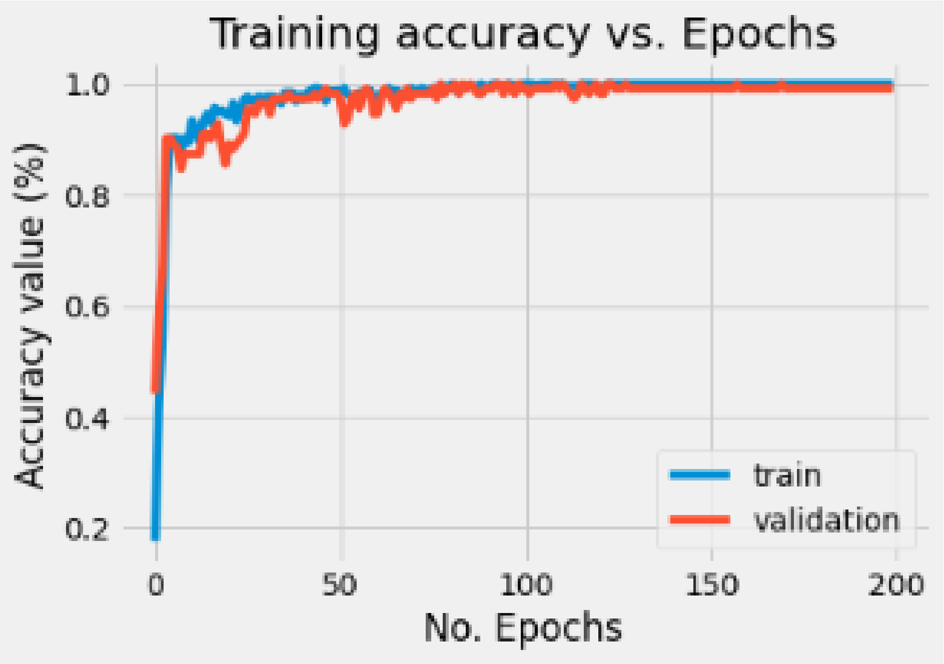

***Figure 7**: The CNN Training Accuracy vs Epochs Graph*

Again, Figure 8 shows the confusion matrix for the intra-class classification using CNN. The confusion matrix shows that all the fingerprint images in DB1, DB4, and contact-less classes were correctly classified at 100% accuracy. While the DB2 fingerprint images were correctly classified at 90% accuracy, 10% of the DB2 fingerprint images were misclassified as DB3. Furthermore, the DB3 fingerprint images were correctly classified at 91% accuracy, and 9% of the DB3 fingerprint images were misclassified as DB1.

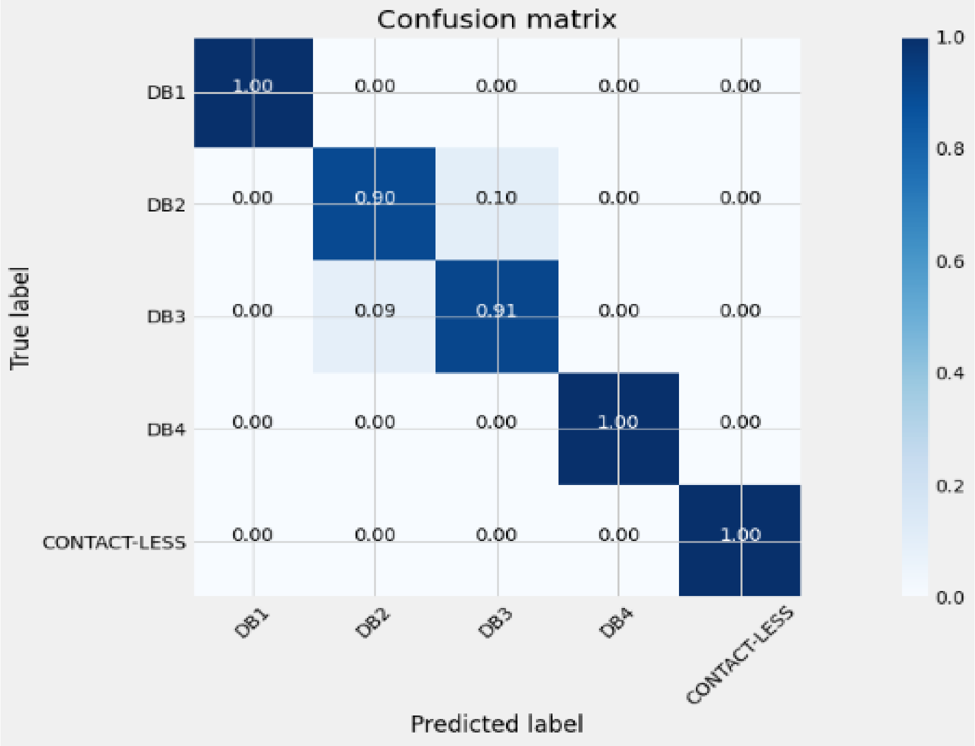

***Figure 8**: The Confusion Matrix for the CNN Algorithm*

The novel proposed CNN method indeed performed the intra-class classification of fingerprint images accurately with an average classification accuracy of 99.72%.

For comparative purposes, our proposed method accuracy based on the Naïve Bayes (95.90%) outperformed that of Mishra and Maheshwary (2017) where they used the Naïve Bayes and SVM classifier for the classification of fingerprint images and achieved accuracies of 87.4% and 76.06%, respectively. Furthermore, we achieved similar results to that of Baştürk et al. (2018). Baştürk et al. (2018) recognised different types of fingerprint images using different machine learning techniques and observed that the deep neural network was more suited for the effective recognition of fingerprint images. Lastly, our proposed method when considering the CNN (99.72%) outperformed that of Qi et al. (2022). Qi et al. (2022) performed a gender-related classification based on fingerprint images using dense dilated convolution ResNet Autoencoder and achieved an average accuracy of 96.5%.

5. CONCLUSION AND FUTURE WORK

This paper proposed the novel use of Benford’s law divergence values and machine learning techniques for the classification of fingerprint images by characteristics and/or sensor sources using Naive Bayes, Decision Tree, Logistic Regression, and CNN algorithms. It was shown that the Naive Bayes, Decision Tree, Logistic Regression and CNN algorithms successfully classified the fingerprint images with mean accuracies of 95.90%, 95.62%, 94.47%, and 99.72%, respectively.

This shows that our proposed method can effectively reduce features and achieve high intra-class separability results especially using the CNN and Naive Bayes algorithms. For future work, we plan to investigate other classification techniques such as Long Short-Term Memory (LSTM) for the intra-class classification and source identification of fingerprint images.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Acebo, E., & Sbert, M. (2005, May 18-20). *Benford's law for natural and synthetic images*. In: Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging (Computational Aesthetics'05) (pp. 169-176).

Bartlow, N., Kalka, N., Cukic, B., & Ross, A. (2009, June 20-25). *Identifying sensors from fingerprint images*. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (pp. 78-84). doi:10.1109/CVPRW.2009.5204312

Baştürk, A., Baştürk, N. S., & Qurbanov, O. (2018). A comparative performance analysis of various classifiers for fingerprint recognition. *Omer Halisdemir University Journal of Engineering Sciences*, *7*(2), 504-513. doi:10.28948/ngumuh.443160

Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society*, *78*(4), 551-572. URL

Bonettini, N., Bestagini, P., Milani, S., & Tubaro, S. (2021, January 10-15). *On the use of Benford's law to detect GAN-generated images*. In: 25th International Conference on Pattern Recognition (ICPR) (pp. 5495-5502).

Fu, D., Shi, Y. Q., & Su, W. (2007, January 28 - February 1). *A generalized Benford's law for JPEG coefficients and its applications in image forensics*. In: E. J. Delp III & P. W. Wong (Eds.), Steganography, and Watermarking of Multimedia Contents IX (SPIE 6505, Forensics III, pp. 65051L). doi:10.1117/12.704723

FVC2000. (2000), Fingerprint Verification Competition Databases. URL

Hildebrandt, M. (2020). *On digitized forensics: novel acquisition and analysis techniques for latent fingerprints based on signal processing and pattern recognition*. PhD Thesis, Otto-von-Guericke-University of Magdeburg.

Hildebrandt, M., & Dittmann, J. (2015, February 8-12). *Benford's Law based detection of latent fingerprint forgeries on the example of artificial sweat printed fingerprints captured by confocal laser scanning microscopes*. In: A. M. Alattar, N. D. Memon & C. D. Heitzenrater (Eds.), Media Watermarking, Security, and Forensics 2015, (SPIE 9409, Biometric, pp. 94090A). doi:10.1117/12.2077531

Hildebrandt, M., Sturm, J., Dittmann, J., & Vielhauer, C. (2013, September 25-26). *Creation of a public corpus of contact-less acquired latent fingerprints without privacy implications*. In: B. Decker, J. Dittmann, C. Kraetzer & C. Vielhauer (Eds.), Communications and Multimedia Security, 14th IFIP TC6/TC11 International Conference (CMS 2013) (pp. 204-206). doi:10.1007/978-3-642-40779-6_19

Hill, T. P. (1998). The first digit phenomenon: A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data. *American Scientist*, *86*(4), 358-363. URL

Iorliam, A. (2016). *Application of power laws to biometrics, forensics and network traffic analysis*. PhD Thesis, University of Surrey (United Kingdom).

Iorliam, A., & Shangbum, C. F. (2017). On the use of benford's law to detect jpeg biometric data tampering. *Journal of Information Security*, *8*(3), 240-256. doi:10.4236/jis.2017.83016

Iorliam, A., Ho, A. T. S., Waller, A., & Zhao, X. (2016, September 17-19). *Using Benford's law divergence and neural networks for classification and source identification of biometric images*. In: Y. Q. Shi, H. J. Kim, F. Perez-Gonzalez & F. Liu (Eds.), Digital Forensics and Watermarking, 15th International Workshop (IWDW 2016) (pp. 88-105). doi:10.1007/978-3-319-53465-7_7

Iorliam, A., Ho, A. T. S., Poh, N., Zhao, X., & Xia, Z. (2017). Benford's law for classification of biometric images. In: C. Vielhauer (Eds.), *User-Centric Privacy and Security in Biometrics* (pp. 237-256). IET Digital Library. doi:10.1049/PBSE004E_ch11

Jain, A. K., Hong, L., Pankanti, S., & Bolle, R. (1997). An identity-authentication system using fingerprints. *Proceedings of the IEEE*, *85*(9), 1365-1388. doi:10.1109/5.628674

Li, X. H., Zhao, Y. Q., Liao, M., Shih, F. Y., & Shi, Y. Q. (2012). Detection of the tampered region for JPEG images by using mode-based first digit features. *EURASIP Journal on Advances in Signal Processing*, *2012*, 190. doi:10.1186/1687-6180-2012-190

Maltoni, D., Maio, D., Jain, A. K., & Prabhakar, S. (2009). *Handbook of fingerprint recognition*. Springer Science & Business Media.

Mishra, A., & Maheshwary, P. (2017). A novel technique for fingerprint classification based on naive bayes classifier and support vector machine. *International Journal of Computer Applications*, *169*(7), 58-62. doi:10.5120/ijca2017914806

Satapathy, G., Bhattacharya, G., Puhan, N. B., & Ho, A. T. S. (2020, October 7-9). *Generalized Benford's Law for Fake Fingerprint Detection*. In: D. Dey, S. Dalai, S. Ray & B. Chatterjee (Eds.), 2020 IEEE Applied Signal Processing Conference (ASPCON) (pp. 242-246). doi:10.1109/ASPCON49795.2020.9276660

Qi, Y., Qiu, M., Lin, H., Chen, J., Li, Y., & Lei, H. (2022). Research on Gender-related Fingerprint Features, Extracting Fingerprint Features Using Autoencoder Networks for Gender Classification. [Preprint] doi:10.21203/rs.3.rs-1399918/v1