Grenze International Journal of Engineering and Technology, June Issue

A Novel Adaptive Authentication System using Deep Learning and Touch-Dynamics

Ritu Agrawal[1], Pharindra[2] and Ekagra Agrawal[3]
[1]Research Scholar, Nirwan University Jaipur, India
Email: ritu.agrawal@nirwanuniversity.ac.in
[2]Associate Professor, SOEIT, Sanskriti University, Mathura, India
Email: dr.pharindra@gmail.com
[3]Student B.Tech CSE, Manipal University Jaipur, India
Email: ekagraagrawal@gmail.com

***Abstract*— User authentication security is becoming more and more interesting and is an important area of research to address privacy concerns and security threats. Biometric has revolutionized the authentication technology and can be used to identify and authenticate users given their physiological and behavioral traits. Touchscreens have become the leading input medium, and hence touch-dynamics is becoming passive and unobtrusive form of biometric authentication. Though a numerous studies have been done in behavioral biometrics, touch dynamics behavioral biometric still needs to be explored. In our work, we proposed deep learning algorithms based on artificial neural network for user authentication and used touch data set. A customized Convolutional Neural Network (CNN) and Bi-Long Short Term Memory (LSTM) neural network is built for user authentication. By utilizing the Bioident database, our methodology achieved an overall accuracy of 98.7%. After being validated with more varied touch data, the findings of our proposed model demonstrate that it performs superior to current methods and can be employed in real-time authentication mechanisms.**

***Index Terms*— Touch Dynamics; behavioral biometric; deep learning; classification; CNN; LSTM**

I. INTRODUCTION

Fingerprint and face recognition are the most popular physiological biometrics but two primarily issues in adaption of physiological biometric method are (i) slow authentication speed (ii) uncomfortable to use socially. Touch dynamics solve these issues and has certain characteristics that make it easy to implement than other authentication factors. These characteristics are:

(i) By employing existing sensors in the computing devices, it is economical than other biometric authentication methods. Biometric authentication methods based on fingerprint and iris requires additional or specialized devices. Now-a-days mobiles have high resolution sensors which allow the extraction of distinctive features, making touch-dynamics more robust.

(ii) Sometimes there are barrier as lighting conditions and/or background noise levels which may cause problems to some biometrics authentication methods such as face recognition and voice recognition, but it has been reduced in touch-dynamics making it more reliable and usable.

Grenze ID: 01.GIJET.10.2.78

(iii) It is a non-intrusive method and can operate along with user's normal usage activities[1].

As touchscreen devices are becoming more and more popular, the way people interact with them, gives emergence to new types of behavioral biometrics. The potential of utilizing touch behavior for continuous authentication, is suggesting significant promise in this type of biometric.

Touch dynamics biometric involves the measurement and evaluation of the rhythmic patterns of human touch on touchscreen devices, such as smartphones. A digital signature variant is generated upon touching these devices. These signatures are discriminator and unique for each individual, so may be used for individual identification. Touchscreen input features can be classified into various categories, which are as follows[2]:

- Single-Touch (ST). This involves touch press down, and then lifting without any intervening movement.
- Touch Movement (TM). This involves moving along a touch surface, first touch press down followed by dragging and then finally touch press up.
- Multi-Touch (MT). This involves interacting with touchscreen devices with multiple simultaneous, touch points or movements.

Deep learning (DL) is an emerging field and has created new possibilities for this type of authentication method. Machine learning (ML) which is a branch of Artificial Intelligence, uses training data and is used widely in classification, regression and decision modeling techniques [3], [4], [5]. DL which is a subset of ML, uses artificial neural network to improve their accuracy. Deep neural networks exhibit better performance in applications where large volume of data is present(unstructured) where as ML perform better in automation of business activities. Both DL and ML can effectively resolve complex issues but DL neural models can provide enhanced predictive accuracy and efficient performance offer greater prediction accuracy and efficient performance[6] in short span of time. In order to develop soft biometric identification framework for continuous verification, this paper presents DL frameworks that rely on CNNs and LSTMs. The proposed technique is tested using open databases from the Bioident dataset. The following are the highlights of our study.

- We report the work on deep CNNs and LSTM using touch gestures to authenticate the user with high accuracy and precision.
- We did data scaling as pre-processing task to improve the performance of our DL models.
- We compared our models result with the top-performing state-of-the-art techniques.

The remaining content is organised as follows: The concepts presented by different researchers are discussed in Section II. The methodology we employed is shown in Section III. Section IV discussed the configuration and parameters used in setting up the experimental environment for deep learning neural networks used in touch dynamics field. Section V presents the enhancements in the work that can be done and finally conclusion.

II. LITERATURE REVIEW

This section reviews related work of different researchers on touch dynamics.

Feng et. al [7] introduced FAST, a user authentication mechanism based on touch interactions. They constructed 53 features for each touch gesture. The research involved 40 participants, and during the login phase, a False Accept Rate (FAR) of 4.66% and a False Reject Rate of 0.13% were achieved.

Meng et al. [8] presented the paper on touch dynamics and particularly extracted 21 different features using SVM classifier. They proposed a hybrid classifier called PSORBFN and study was performed on 20 users. Using this classifier, they attained an average error rate of 2.92% (with a False Accept Rate of 2.5% and a False Reject Rate of 3.34%).

Frank et al. [9] designed a pilot classification framework and extracted a set of 30 behavioral features using a nearest neighbor classifier and a Gaussian RBF kernel support vector machine (SVM). With these classifiers, they achieved equal error rates (EERs) between 0 and 4 %, depending on the application scenario.

Li et al. [10] used finger movements to develop a re-authentication system. A study was conducted involving 75 users and used 13 features such as initial touch position, initial touch pressure, initial touch area, and initial moving direction. With SVM classifier, a robust accuracy of 95.78% was achieved for swiping up movements.

Meng et al. [8,11] came with an adaptive mechanism to overcome the problem that the choice of classifiers may influence the performance, so they periodically selected a superior classifier to uphold accuracy in authentication process. The authors examined with 50 participants, and their adaptive authentication scheme could attain an average error rate of 2.46%.

Feng et al.[12] named their authentication scheme TIPS. It was a novel, continuous, context-based user authentication system for uncontrolled environment. They implemented TIPS on the Android operating system and evaluated with 123 participants (23 owners and 100 guests) and over 23 different phones as Galaxy S3, Galaxy S4, Nexus 4. Classifiers which were used to gather data at varying times were One Nearest Neighbor (1NN) and

Dynamic Time Warping (DTW). Their system reached 90% accuracy in real time. There was an issue of privacy as there is a need of installing different forms of Android operating system.

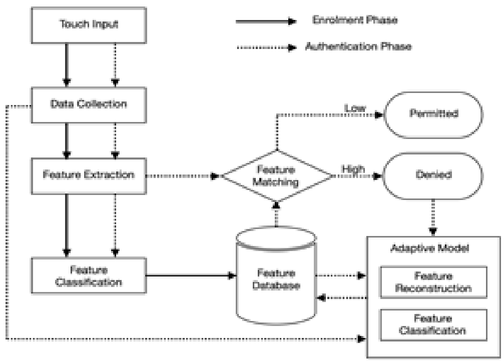


Figure 1: Touch-Dynamics Model

Meng et al. [13] further took the authentication system to a good level by developing a touch authentication system based on movement gesture, called TMGuard, to achieve higher accuracy by evaluating 75 participants. They took the example of 9-dot patterns and found a decline in success rates both for males and females.

Stefania Budulan et al.[14] gathered data from 5 Android phones screens and extracted features based on 41 user's data sets. Touch-down and touch-up is considered as an input sequence action. AdaBoostClassifier (also known as an Adaptive Boost classifier) is used as an ensemble method by adjusting the performances of incorrectly classified instances. GridSearchCV , a tool of Scikit-learn was used for fitting the specific estimator(model) onto the data set. They achieved an accuracy of 83% approximately.

Alfonso et al. [15] presented TGSB, a gender and age-based recognition system, and applied Convolutional Neural Networks provided with pictorial depictions of touch gestures. They tried to understand the most useful gesture and combining multiple gesture strategies. They combined simultaneous touch gestures utilizing a collaborative latent subspace learning procedure. They were able to show that, TGSB demonstrates significantly improved performance in recognizing both gender and age groups.

Tolosana et al. [16] used Bidirectional Recurrent Neural Network (BRNN) to capture information from both past and future contexts. They utilized LSTM and GRU within Siamese network to authenticate an on-line signature. A total of 23 features were derived from the user's signature.They applied their model on BiosecurID dataset. Their model exhibited 2.92% EER.

III. PROPOSED METHODOLOGY

This section offers a thorough scrutiny of deep learning algorithms to model and recognize unique touch dynamics patterns for user authentication. Our proposed design model for touch-based authentication system is shown in figure 1. The suggested model will basically work in two phases: Enrolment phase and Authentication phase. This section offers a comprehensive evaluation of proposed technique through algorithm analysis.

A. Convolution Neural Networks

Convolution neural networks stand out as the foremost advanced DL-based algorithms to autonomously learn robustness and extract discriminative features on their own. The architecture of the suggested CNN is illustrated in Figure 2. It has alternate convolutional and max pooling layer, a fully connected layer, and a soft-max layer. The proposed system employs a one-dimensional convolutional layer since touch data is sequential and represents time series data. This layer focuses on local regions in touch data. The local features extracted may cause an increase in computation time and memory usage, and so to accelerate the efficiency our data, we have used Rectified Linear Unit (ReLU) activation. The ReLU activation function is given by Eq. (1)

$$ReLU(x) = max(0, x) \qquad (1)$$

It outputs the input value same, if the input value is positive otherwise it will output as zero. One common issue with this is, neurons can become inactive and always output zero for any input during training. This is known as dying-ReLU problem. One solution to this problem is Leaky-ReLU, which aim to address the dying ReLU problem by allowing a small, slope for negative inputs(Eq.(2)).

$$LeakyReLU(x) = \begin{cases} x \, if \, x > 0 \\ \alpha x \, otherwise \end{cases} \quad (2)$$

where α is a small positive constant (typically a small fraction like 0.01).

The CNN configuration is outlined as follows: Inputs are processed through successive layers, including Convolution-1D operations with varying feature counts and map sizes, each followed by a ReLU activation. Max Pooling layers with specified strides(stride 3, 2) are interleaved between convolutional layers. The final layers consist of a Flatten operation followed by a Dense layer with 100 units and a SoftMax activation. Notably, superior results were attained through the utilization of the Adam optimization technique in conjunction with categorical cross entropy.

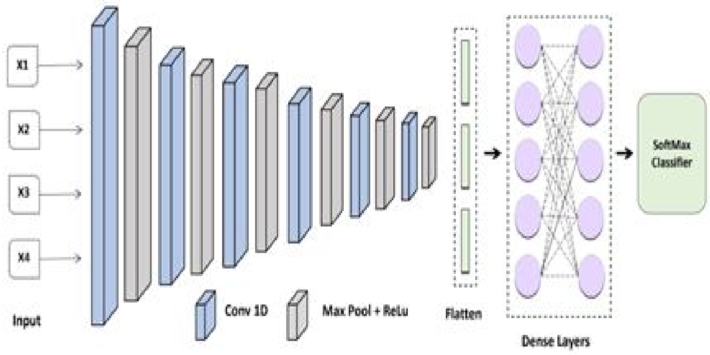


Figure 2. Proposed Convolution Neural Network

B. Long Short-Term Memory Networks

The long-short-term memory (LSTM) is a type of Recurrent Neural Network(RNN) model used for serial and time analysis, and is used in our proposed system as a basic level[17,18].

The design of proposed LSTM consists of the following layers:

Input layer: The input to this layer is raw touch data which is put forward into the next layer as output, where individual row is treated as a sample.

Long Short-Term Memory LSTM Layer: LSTM has intelligent neurons endowed with memory capabilities [19]. Each cell (neurons) in LSTM incorporates three gates Input Gate, Forget Gate, and Output Gate to regulate the flow of information. These gates are used to avoid data redundancy.

Input Gate: this gate is used to control whether new data will retain in the cell state. The activation function used in LSTM is the sigmoid function (Eq. (3)) [19].

$$f(x) = \frac{1}{1+e^{-x}} \quad (3)$$

Forget Gate: this gate is used to discard the irrelevant information from the cell state. This gate uses Sigmoid function (Eq. (3)) and its output value will lie between 0 and 1, 0 signifies the information is not required and is to be discarded whereas 1 signifies the information is necessary[19,20].

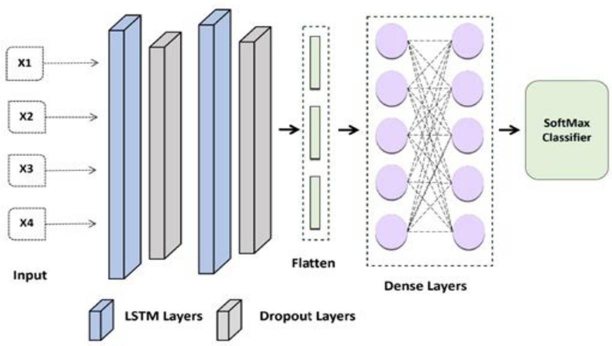


Figure 3. Proposed LSTM Model

Output Gate: this gate is used to show the desired information from the cell state as output. LSTMs produce an output known as the hidden current state, which is based on the updated cell state.
The LSTM's construction is shown in Figure 3.
In our study, we have examined the output at each time-stamp. It has been regularized via a dropout layer. The output is then passed to a fully linked layer. Softmax function(Eq.(4)) is used to determine the probability and the one whose likelihood(probability) is highest is chosen that the touch data sequence belongs to him/her.

$$y_k = \frac{y^{ak}}{\sum_k^k , e^{ak}}, for K = 1 \dots K \tag{4}$$

where k is the number of classes in touch authentication system.
The network architecture was configured as follows: an initial LSTM layer with 64 filters, followed by a Dropout layer with a rate of 0.2, then another LSTM layer with 32 filters, succeeded by a Dropout layer with a rate of 0.2. Subsequently, a Flatten layer was applied, leading to Dense layers culminating in a SoftMax activation. Notably, superior results were achieved by employing the Adam optimization strategy in conjunction with categorical cross-entropy.

C. Hyperparameters and Loss Function

In this section we delved into two critical aspects of training learning models, that is hyperparameters and loss function. In multi-class classification, categorical cross-entropy is used as a loss function to measure the dissimilarity between the predicted and true class distributions. We chose Adams as an optimizer after experimenting with different configurations. This adaptability often results in faster convergence compared to traditional gradient descent methods. The step size during optimization is kept small and are shown as learning rates (LR). The batch size, in our case is of 32 to reduce memory requirements and have faster convergence.

TABLE I: HYPER-PARAMETERS SETTINGS FOR TRAINING.

No.	**Hyperparameters**	**Settings**
1	Loss Function	Binary Cross entropy
2	Optimizer	Adam Optimization
3	Learning rate	0.001
4	Epochs	20
5	Batch size	32
6	Callbacks	Model Check Point

IV. EXPERIMENTAL WORK

The Touch Dynamics authentication system can be used in one of two views: identification (recognition) or verification (authentication).
Verification views: In the verification, it is work to inquire a pre-enrolled identity, which means it is asked to answer the question "is this the legitimate person". Authenticating a mobile user using password or PIN is an example of this mode.
Identification views: In the identification, it is applied to categorize and recognize unknown identity. Here, it is asked to answer the questions like "who are you" or "are you present in the database". Example of this view is utilized in cybercrime or in large security systems.
As touch dynamic authentication system works in two phases, enrolment phase is used to group data and store it as a reference template while in the authentication phase, touch dynamics test samples are matched to reference templates that have been preserved in order to calculate closeness scores that are then compared to a predetermined threshold by machine learning approach.

A. Data Collection

The BioIdent dataset has been used in our proposed system. Android mobiles were used to record the dataset1. While recording the dataset 71 users(56 male and 15 female) were allowed to do reading a text and then answer questions based on it. The users would then select an image from the gallery. All this was done through a client-server mobile application. Besides mobile, other devices were tablets and a total of eight different devices were used, with resolutions in the range 320x480.1080x1205. Each recording includes "Up" and "Down" operations (vertical stroke), "Left" and "Right" operations(horizontal stroke)while reading a text and browsing an image respectively. This data was collected during 4 weeks. The different attributes were extracted from the above mentioned operations associated with touch-based system [21].

B. Feature Engineering

This is an iterative process conducted during both the enrollment and authentication stages. This is a common pipeline of extracting and selecting features, then building and training classification model using identified features. Initial investigations on conventional machine learning algorithms have identified several limitations and challenges. They demand a sophisticated foundation while dealing with authentication approaches. During authentication mechanism numerous features will be selected from signals. An excessive number of features, often referred to as the curse of dimensionality will need more specialized models and techniques. The requirement of extra training time and storage required is indeed a significant challenge when it comes to implement real-time systems. Machine learning algorithms may also encounter challenges when dealing with extensive test datasets. This also can have the problem of overfitting or underfitting, which potentially results in loss of performance. Conventional machine learning methods frequently depend on manually crafted features, a process that is time-consuming and laborious. Deep learning approaches, especially neural networks can self-learn from hierarchical and abstract raw touch signals and therefore offers a more simplified framework for an authentication system. Data normalization which is a step of preprocessing in our proposed system, involves transforming data to achieve a zero mean as in Eq. (5).

$$F_{new} = \frac{F-\mu}{\sigma} \tag{5}$$

where Fnew is the data after normalization and F is the original data, μ is the mean of feature and σ is standard deviation of feature.

When dealing with multi-class classification, converting the class labels Y into one-hot encoding $Y\in\{0,1\}^{k}$ where K is the number of classes, is a common practice. Furthermore, each user in the touch dataset possesses a distinct number of observations compared to others which may lead to class-imbalance problem. Therefore maintaining an equal proportion of examples in each class as observed in the dataset is important. To tackle this issue, we used stratified cross-validation to ensure that each fold maintains a similar class distribution to the overall dataset.

C. Evaluation Metrices

A commonly used tool for assessing the performance of a classification model, particularly in the context of validation or testing data is the confusion matrix (CM). The CM provides comparable columns and rows that display model's predictions and actual class labels. To quantify the performance of biometrics verification system, the metrics that are commonly used are the False Acceptance Rate (FAR), False Rejection Rate (FRR) and Equal Error Rate (EER). False Acceptance Rate (FAR) is the proportion of the number of invalid attempts(imposter attempts) that are accepted by the biometric system. False Rejection Rate (FRR) is the proportion of the number of valid attempts (legitimate users attempts) that are rejected. Equal Error Rate (EER) is used to measure and compare different biometrics authentication methods. It is also known as Crossover Error Rate (CER) which is a compromise between FAR and FRR. This is a trade-off between FAR(security) and FRR(usability). Both FAR and FRR should be low. EER is to equalize FAR and FRR. Accuracy is the overall correctness of the system.(Eq. (6),(7),(8))

$$FAR = \frac{FP}{FP-TN} \tag{6}$$

$$FRR = \frac{FP}{TP-FN} \tag{7}$$

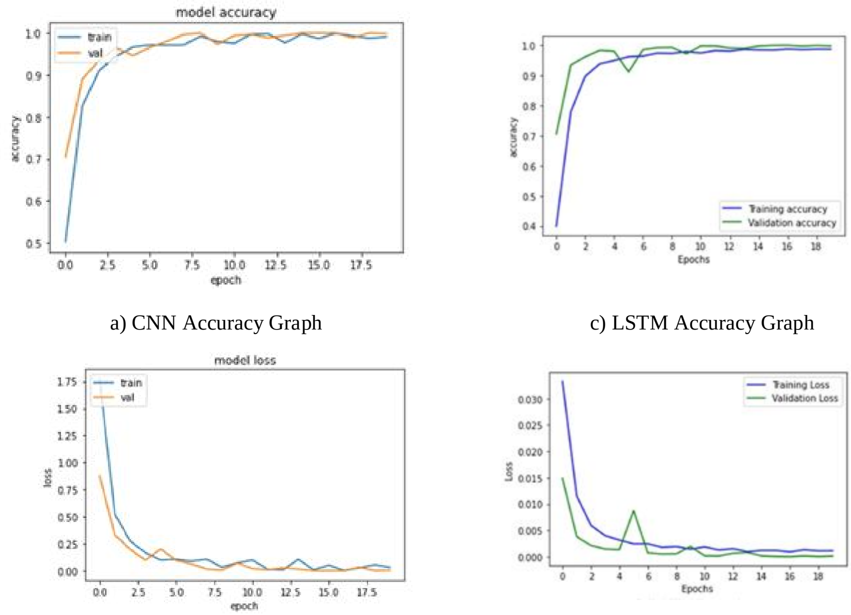
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

V. Results And Analysis

In our proposed system, we studied the DL models against the raw touch data set to obtain the best result that can used to design the touch dynamics authentication system. The outcomes of the training and validation phases are considered here. We trained our suggested model using a range of hyperparameters for improved training. With just a learning phase rate of 0.001, a mini-batch of 32, and the Binary Cross Entropy algorithm, we employed the Adam optimizer. As our chosen classifier, a softmax classifier was used. The Keras API with a TensorFlow backend was used to train Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models. We have split our dataset into training and testing sets, with 80% of the data used for training and 20% for testing. During visualizing the accuracy and loss curves over epochs, it was observed that there was a steady increase in model accuracy as the number of training epochs increased. The accuracy eventually stabilized at a certain number of epochs. The CNN and LSTM performed admirably on the Bioident dataset, achieving high validation accuracy. The CNN model achieved 96.8 percent validation accuracy, while the LSTM model achieved 95.6 percent. The obtained loss and accuracy curves for the suggested CSTM and LSTM model are shown in Figure 4.

VI. Conclusion and Future Work

This paper has a specific focus on exploring the effectiveness of using touch data for user authentication. Not much of the work is done in this field and there is still a scope of further research. Deep Learning algorithms are used in this paper which analyses time-stamped data and can be used as an adaptive authentication. In contrast to prior researches, the models in this research, are able to capture and learn relevant patterns from touch data, leveraging both the spatial features (CNN) and sequential dependencies (LSTM) inherent in the input.



a) CNN Accuracy Graph

c) LSTM Accuracy Graph

b) CNN Loss Curve

d) LSTM Loss Curve

Figure 4: (a) and (b) shows accuracy curves and loss curves for CNN model (c) and (d) shows accuracy curves and loss curves for LSTM model

The end-to-end learning used here eliminates the need for manual feature engineering or fiducial extraction, simplifying the overall model architecture. For Bioident data set best results have been obtained so far. Touch-Dynamics biometric authentication still needs to be improved in terms of availability or security of data sets to validate the system. Feature extraction can be done with different types of auto-encoders or transformer, which will help in further reducing feature engineering time and will improve overall accuracy.

References

[1] Shen, C., Zhang, Y., Guan, X., Maxion, R.A., 2016. Performance Analysis of Touch-Interaction Behavior for Active Smartphone Authentication. IEEE Trans. Inf. Forensics Secur. 11, 498–513.

[2] Y. Meng, D.S. Wong, R. Schlegel, L.-F. Kwok, Touch gestures based biometric authentication scheme for touchscreen mobile phones, in Proceedings of the 8th China International Conference on Information Security and Cryptology (INSCRYPT). Lecture Notes in Computer Science, vol. 7763 (Springer, 2012), pp. 331–350, Beijing.

[3] Hinton, G.E.; Osindero, S.; Teh, Y.-W. A fast learning algorithm for deep belief nets. Neural Comput. 2006, 18, 1527–1554. [Google Scholar] [CrossRef]

[4] Deng, L. Deep Learning: Methods and Applications. Found. Trends Signal Process. 2013, 7, 197–387. [Google Scholar] [CrossRef]

[5] Karhunen, J.; Raiko, T.; Cho, K. Unsupervised deep learning: A short review. In Advances in Independent Component Analysis and Learning Machines; Academic Press: Cambridge, MA, USA, 2015; pp. 125–142. [Google Scholar] [CrossRef]

[6] Ahmed, S.F., Alam, M.S.B., Hassan, M. et al. Deep learning modeling techniques: current progress, applications, advantages, and challenges. Artif Intell Rev 56, 13521–13617 (2023). https://doi.org/10.1007/s10462-023-10466-8

[7] T. Feng, Z. Liu, K.-A. Kwon, W. Shi, B. Carbunar, Y. Jiang, and N. Nguyen, “Continuous mobile authentication using touchscreen gestures,” in IEEE Conference on Technologies for Homeland Security, Nov 2012, pp. 451–456.
[8] Y. Meng, D.S. Wong, L.F. Kwok, Design of touch dynamics based user authentication with an adaptive mechanism on mobile phones, in Proceedings of the 29th Annual ACM Symposium on Applied Computing (ACM SAC), Gyeongju (2014), pp. 1680– 1687.
[9] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, “Touch analytic: On the applicability of touchscreen input as a behavioral biometric for continuous authentication,” IEEE Transactions on Information Forensics and Security, vol. 8, no. 1, pp. 136–148, Jan 2013
[10] L. Li, X. Zhao, and G. Xue, “Unobservable reauthentication for smart phones,” in Proceedings of the 20th Network and Distributed System Security Symposium, 2014.
[11] W. Meng, D.S. Wong, L.F. Kwok, The effect of adaptive mechanism on behavioural biometric based mobile phone authentication. Inf. Manage. Comput. Secur. 22(2), 155–166 (2014).
[12] Feng, Tao; Yang, J., Yan, Z., Tapia, E.M., She, W.: Tips: Context-aware implicit user identification using touch screen in uncontrolled environments (2014)
[13] W. Meng, W. Li, D.S. Wong, J. Zhou, TMGuard: a touch movement-based security mechanism for screen unlock patterns on smartphones, in Proceedings of the 14th International Conference on Applied Cryptography and Network Security (ACNS 2016), Guildford (2016), pp. 629–647.
[14] Budulan, Stefania, Burceanu, Elena, Rebedea, Traian, Chiru, Costin Continuous User Authentication Using Machine Learning on Touch Dynamics DOI - 10.1007/978-3-319-26532-2_65.
[15] Alfonso Guarino, Delfina Malandrino, Rocco Zaccagnino, Carmine Capo, Nicola Lettieri, Touchscreen gestures as images. A transfer learning approach for soft biometric traits recognition, Expert Systems with Applications,Volume 219, 2023, 119614, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2023.119614.
[16] R. Tolosana, R. Vera-Rodriguez, J. Fierrez and J. Ortega-Garcia, “Exploring recurrent neural networks for on-line handwritten signature biometrics,” IEEE Access, vol. 6, pp. 5128–5138, 2018.
[17] J. Brownlee, “How to develop LSTM models for time series forecasting,” 14 11 2018. [Online]. Available: https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/. [Accessed 1 12 2019].
[18] J. Brownlee, “Time series prediction with LSTM recurrent neural networks in python with keras,” July 2016. [Online]. Available: https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/. [Accessed Nov. 2019].
[19] E. Torres, J. Kalita, J. Ventura, R. Lewis and A. Alzubaidi, “Authentication of legitimate users of smartphone based on app usage sequences,” M.S., Univercity of Colorado Colorado Springs, 2018.
[20] F. Gers, J. Schmidhuber and F. Cummins, “Learning to forget: continual prediction with LSTM,” in Proc. of ICANN'99 Int. Conf. on Artificial Neurals Networks, Edinburgh, Scotland, 2000.
[21] M. Antal, Z. Bokor and L. Szabo, “Bioident - touchstroke based biometrics on android platform,” 2014. [Online]. Available: https://ms.sapientia.ro/∼manyi/bioident.html. [Accessed 1 Jan 2020].