

Neural Network and Information Uses

Subjects: Computer Science, Information Systems

Contributor: Rushit Dave

Recurrent Neural Networks are powerful machine learning frameworks that allow for data to be saved and referenced in a temporal sequence. This opens many new possibilities in fields such as handwriting analysis and speech recognition. This paper seeks to explore current research being conducted on RNNs in four very important areas, being biometric authentication, expression recognition, anomaly detection, and applications to aircraft.

Keywords: recurrent neural network ; biometric authentication ; expression recognition ; anomaly detection ; smartphone authentication ; mouse-based authentication ; aircraft trajectory prediction

1. Introduction

People have always been fascinated by the idea of creating an artificial human brain and these efforts became known as artificial neural networks (ANN). There are numerous variations of specialized ANNs; take convolutional neural networks (CNN), for example, which are adapted to work specifically with image or video data. RNNs are unique because they are comprised of many neural networks chained together, which allows them to process a series of data where a network learns from its previous experiences. RNNs have a wide array of applications, ranging from written language to speech recognition.

RNNs have the potential to improve upon current methods, but also allow advancements in new authentication techniques. However, biometric authentication is so much more than that. What if it were feasible to use biometric authentication to protect cloud data in transit from a mobile device ^[1]? A few examples of biometric authentication are mouse movement authentication, keystroke authentication ^[2], handwritten password authentication ^[3], and even palm print authentication ^{[4][5]}.

Another key implementation of Recurrent Neural Networks is in the field of facial recognition. Facial recognition ranges from identifying one's identity to deciphering their emotions. Expression recognition often relies on a CNN for extraction of important features from image data before that image data can be used by the RNN ^[6]. The ability for software to be able to distinguish different human emotions will be of increasing importance in the future.

Anomaly detection can range from detecting spam emails, to malicious network traffic and maritime vessel traffic. These specialized neural networks can help detect anomalous flight conditions, predict excessive engine vibrations, determine the remaining life of a turbine engine, and aid in landing ^[7]. It looks at which patterns are normal and denotes an event outside of the margin of normal operation as anomalous. The goal of the paper ^[8] is to detect patterns in IoT devices which can then be applied to track unusual patterns in a network of IoT devices.

These are the four main topics that this paper will be reviewing. The goal of this paper is to analyze novel approaches in each of the four applications of RNNs. The remainder of this paper is organized as follows: background discussion of current research, review of biometric authentication, review of facial recognition, review of anomaly detection and aircraft, discussion and analysis of each topic covered in the literature review, discussion and analysis, limitations, conclusion, and future work.

2. Literature Review

Sensors such as iris scanners or fingerprint readers are amongst the most popular forms of smartphone biometric authentication. "Fingerprint and face recognition is based on a physical characteristic, but biometrics can also recognize how a user performs a specific activity" ^[9]. Many different RNN models were tested with varying vector size, number of filters, and fully connected layers. Other novel approaches to smartphone authentication are through ECG signals ^{[10][11]} and holding position combined with touch type authentication ^[12].

An increasingly popular form of biometric authentication is through the recognition of mouse movements or keyboard-based behavioral patterns. [13] is a novel attempt to detect patterns in mouse movements using RNNs and the architecture of this model is represented in **Figure 1**. The paper also describes that data like mouse movement information is easy to collect and contains little privacy-sensitive information. The proposed method involves a fusion of a CNN-RNN, since complex identification tasks benefit from utilizing the fusion of two types of neural networks.

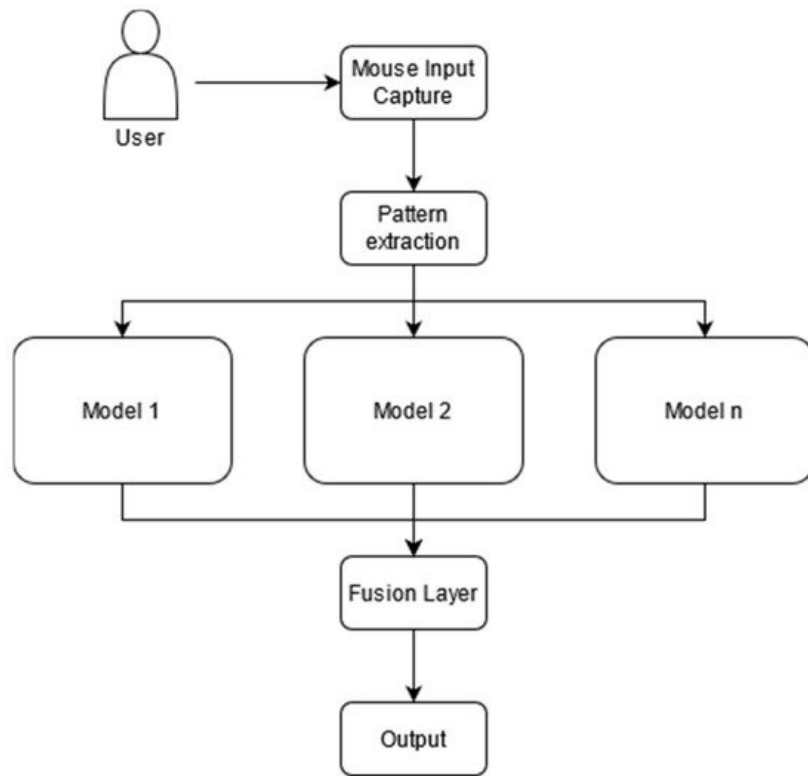


Figure 1. Proposed model for mouse behavior authentication.

This database is comprised of 16 signatures and 12 professional forgeries per user with a total of 400 total users. When this data is fed into the LSTM network the final EER was 6.44% for 1:1 and 5.58% for 4:1 (ratio of number of original signatures to skilled forgeries). Another attempt to authenticate users from their fingerprint data uses handwritten passwords instead of a signature. To collect this data, each user would use their fingers to write out the digits 0–9 a total of four times over two sessions.

These researchers choose to use four different datasets to train and test their model. These datasets are the extended Cohn-Kanade database, which contains 593 image sequences from 123 different subjects, the MMI dataset, which consists of 2885 videos of facial expression from 88 subjects, the Static Facial Expressions in the Wild dataset, which is made up of 663 expression samples, and finally their own dataset, compiled from 80 subjects who each performed the 6 basic emotions. The six basic emotions present in each of these datasets are fear, disgust, anger, happiness, sadness, surprise, and neutral. With their proposed method [14], they were able to attain 99% on CK + dataset, 81.60% on MMI, 56.68% on SFEW (which is highly accurate for that dataset), and 95.21% on their own dataset.

The multimodal approach to expression recognition implements multiple modalities into the RNN framework to improve recognition accuracy. These types of modalities include, but are not limited to, facial expressions, speech, head movements, and body movements. This dataset contains modalities like audio, video, electrocardiogram, and electrodermal activity for each subject, with the emotions of arousal and valence being portrayed. The best results from this proposed model [15] were divided, with the best arousal results coming from the early fusion of all the modalities into the LSTM network that is displayed in **Figure 2**, and the best valence results coming from the late fusion methodology.

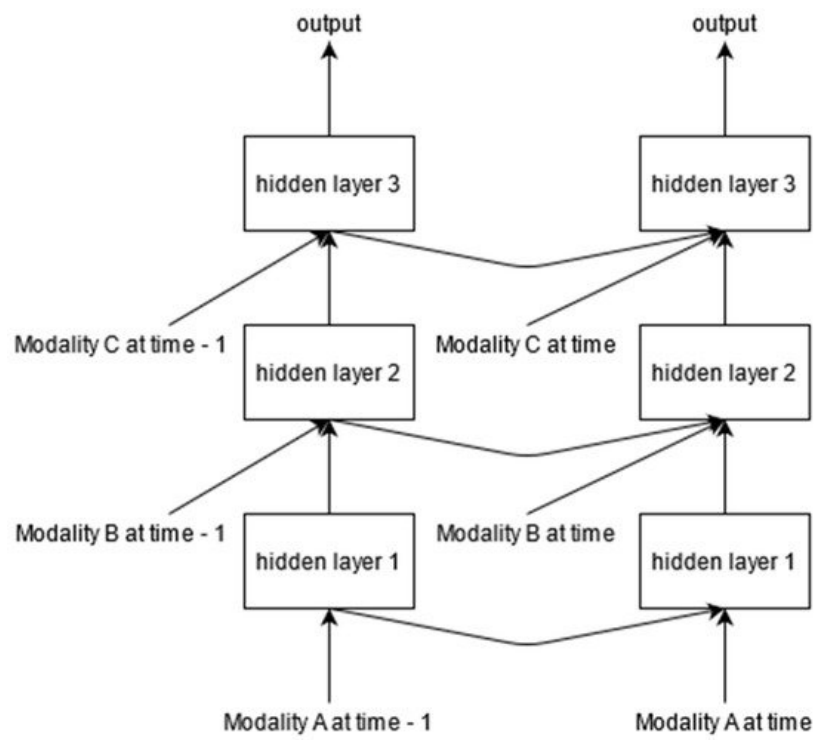


Figure 2. Proposed model for multimodal expression recognition.

Extracting these temporal features was also the goal of [16]. This proposed model extracts the temporal geometry and spatial features, then fuses them to be passed into the LSTM RNN. Both models [17][16] surpass methods that rely solely on a CNN to detect expression. This is where an LSTM becomes helpful in extracting temporal features.

The goal of the research done in paper [18] is to improve transportation and shipping through anomaly detection to increase awareness of all vessels and reduce potential accidents. The researchers use an LSTM RNN architecture to track anomalous vessel movements by feeding it trajectory data shown in **Figure 3**. The RNN was able to detect anomalous course, speed, and route. The anomaly can also be applied to occupancy detection, anomalous exchange rate prices, network anomaly detection, and anomalous stock price detection.

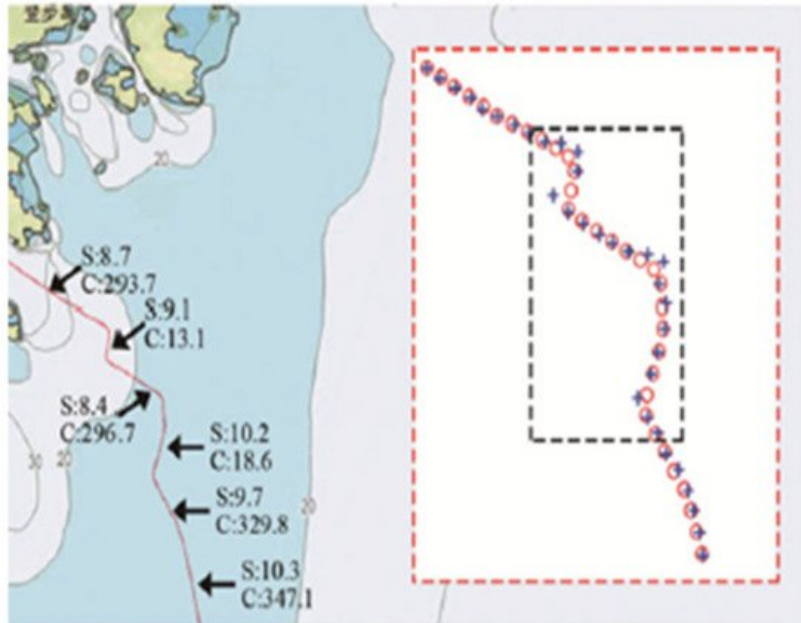


Figure 3. Anomalies in vessels' course [18].

Regulating and monitoring water quality is important for the health and safety of all who rely on that water supply. With a RNN and a dataset collected from real world data [19], it is possible to monitor the quality of water flowing through a water treatment facility. This data consists of temperature, chlorine dioxide levels, acidity (pH) levels, etc. LSTM RNNs can also be used for anomaly detection in network traffic.

Anomaly detection can also apply to tracking and identifying abnormal occurrences surrounding events such as running, loitering, or driving. Each dataset is comprised of multiple videos displaying normal and abnormal events. The sRNN can go frame by frame through these videos and track the anomaly as it progresses through the scene. RNN based strategies can also be useful for detecting anomalies in network traffic.

Flight trajectory prediction is an important tool for planning and executing a safe flight from one destination to another. This type of cost increases even further when multiple aircraft trajectories need to be simulated in real time. The architecture of an LSTM-RNN for predicting flight trajectory can be seen in **Figure 4**. A different LSTM based approach to flight trajectory prediction ^[20] uses data collected from Automatic Dependent Surveillance-Broadcast (ADS-B stations).

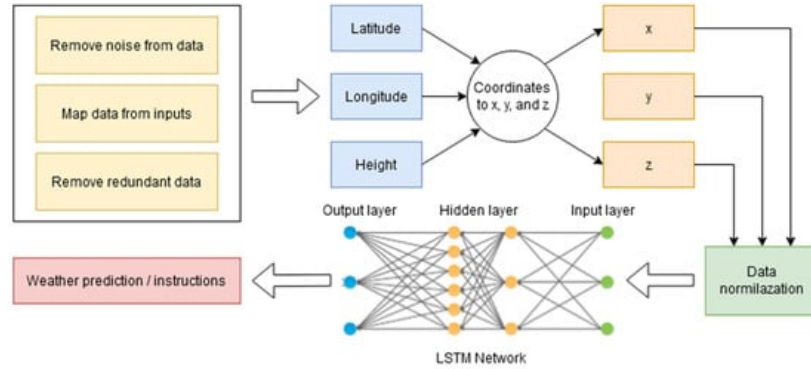


Figure 4. Proposed model for aircraft dynamics simulation.

The data from the flight is collected and the network is trained with normal flight data. This proposed method ^[21] was able to reach an accuracy of 99.7% for forward velocity anomalies and 100% for pneumatic lifting anomalies. A similar methodology can be applied to detecting anomalies in manned aircraft, specifically commercial airline flights. Another group of researchers ^[22] also tried to detect anomalous flight parameters using data generated by X-Plane simulations.

The methodology of ^[23] is to use a LSTM-HMM fusion architecture, which can be seen in **Figure 5**, to predict remaining engine life. LSTM RNNs can also be used to detect excess engine vibration. If a turbine engine has excess vibrations, it can advise engineers that an engine needs maintenance or replacement. The purpose of this model was to predict engine vibrations.

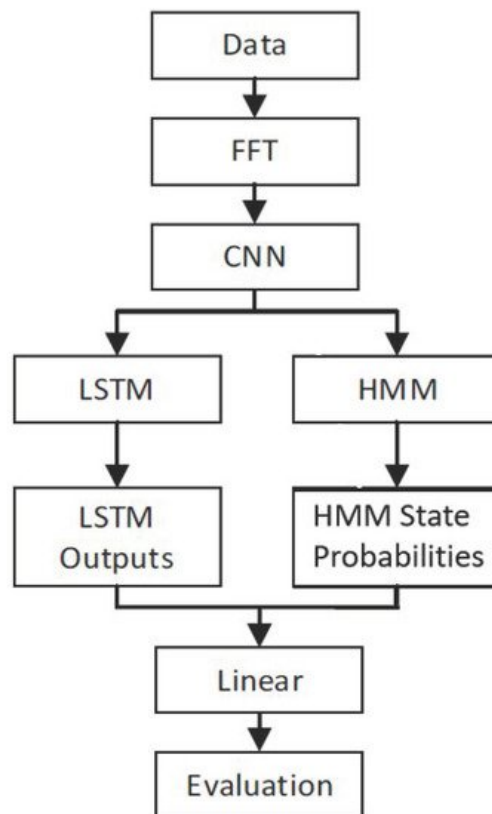


Figure 5. Proposed model for remaining life engine prediction ^[23].

3. Discussion and Analysis

In **Table 1** above each of the three main papers from all four topics are summarized by methodology which includes the structure and data collection strategies, the results of each of the papers along with the dataset used and inference time if available, and finally the pros and cons of each paper. Each method of biometric authentication discussed above has a unique application and one might want to choose a method to better fit their needs, for example the mouse movement authentication technique can be a very simple, portable, and secure method. However, a drawback is that it may take longer for users to configure their information when compared to a fingerprint reader or take less thought like inertial gait authentication. For any authentication technique, there is always a balance speed and security.

Table 1. Comparison of existing RNN applications.

Title	Methodology	Results	Pros and Cons
Novel Smartphone Authentication Techniques [9]	Using an RNN to authenticate users through inertial gait recognition or identify users based on their physical movement patterns. Gait recognition also requires gyroscope and accelerometer sensor data to track movement,	The best performing results obtained an equal error rate of 11.48% using 20% for training, and 7.55% using 70% for training. These results were obtained from the Osaka University Database (OUIDB).	Users can authenticate based on walking patterns. Makes authentication easier, allowing it a wider range of applications. However, sensors are required to collect inertial gait data.
Mouse and Keyboard Based Authentication Methods [24]	Authenticate uses a CNN+RNN fusion to detect behavioral patterns in mouse movement. All this requires is a mouse and a program that can capture the mouse input data.	The proposed model was able to accurately authenticate users 99.39% of the time. The dataset for this paper was provided by Xi'an Jiaotong University of China.	Sensors are not required for biometric authentication; all you need is a mouse. However, authentication could take longer as you may need to perform a longer process to authenticate.
Handwritten Authentication Methods [25]	Employing an LSTM RNN to analyze users' handwriting and confirm or deny them access to a system. To collect user data, there needs to be some sort of device like a tablet for users to write their signature.	The LSTM RNN was able to achieve a final EER of 6.44% for 1:1 and 5.58% 4:1. 1:1 and 4:1 are the ratios of real signatures to skilled forgeries. These researchers generated their own development and evaluation datasets. Each training iteration lasted approximately 30 min with 200 training iterations and 100 testing iterations.	This has more potential than entering a password as it adds an extra layer of security to passwords. However, having to handwrite passwords requires some sort of device or touch screen.
Model for Facial Expression Recognition Using LSTM RNN [14]	Utilizing an LSTM RNN for facial expression recognition against multiple datasets including one developed by the researchers. A camera is needed to collect the necessary video data used to build an expression recognition dataset.	The LSTM RNN was able to reach an accuracy of 99% on CK + dataset, 81.60% on MMI dataset, 56.68% on SFEW dataset, and 95.21% on their own dataset.	The LSTM RNN has shown great promise in expression recognition. However, feature extraction methodologies can hinder recognition accuracies.
Multimodal Expression Recognition Implementing an RNN Approach [15]	Multimodal expression recognition uses multiple modalities like speech, body movement, head movement, etc. All these elements are combined to recognize emotions. Since multiple modalities need to be collected, one might need a camera for video, microphone for audio, and possibly body tracking sensor data.	Results can be seen in Figure 9. The dataset used for this challenge (AVEC2015) was a subset of the larger RECOLA dataset.	Multiple modalities can improve upon expression recognition. However, extracting and processing all these different features can create lots of noise and cause inaccuracies.
Motion History Image Expression Recognition [17]	Using a Cross Temporal LSTM RNN with Motion History Images for facial expression recognition. Motion History images are essentially multiple images stacked on top of each other to form one image that shows a sequence of events so no additional sensors are required.	The proposed method was able to achieve an accuracy of 93.9%, 78.4%, and 51.2% on CK+, MMI, and AFEW datasets, respectively.	Motion History Images allow for all motion in a video to be captured on a still image, leading to easier feature extraction. However, these images can often be cluttered, creating a lot of noise.

Title	Methodology	Results	Pros and Cons
Anomaly Detection of Maritime Vessels ^[18]	Leveraging an RNN to detect anomalies in course, speed, and trajectory of vessels with density-based clustering. Vessel data was gathered from the Automatic Identification System (AIS).	The results can be seen in Figure 10. The dataset for this model was built from the DBSCAN algorithm which was applied to AIS data to generate trajectory points used to train the network.	There is a lot of traffic to sort through within busy ports which, if handled correctly, can significantly improve accuracy. However, there exists room for false positives when dealing with large volumes of data.
Anomaly Detection in Water Quality ^[19]	Using an RNN to monitor the quality of, and detect anomalous traits in, water flowing through a control facility in Germany. Additional sensors needed to measure water quality would contain instruments to measure temperature, acidity, and chlorine dioxide levels plus any other water quality traits.	The proposed model was able to achieve an F1 score of 0.9023. The researchers built their own dataset from a real sensor data taken from Thüringer Fernwasserversorgung public water company.	A high F1 score means not many false alarms were triggered. However, monitoring multiple different qualities in water can make triggering a false positive or false negative more common.
Stacked RNN Strategy for Anomaly Detection in Pedestrian Areas ^[26]	Applying a stacked RNN framework to detect anomalous events and activities in pedestrian areas. For this paper researchers used data collected from cameras that were in public areas.	Using the sRNN, the model was able to achieve accuracies of 81.71% on CUHK Avenue, 92.21% on Pedestrian 2, and 68.00% on their custom dataset. The sRNN model takes about one hour to train on the Avenue dataset and takes 0.02 s to make a prediction from any frame.	Stacked RNN frameworks provide many different cells both vertically and horizontally. However, with this type of anomaly detection it is hard to define what events are anomalies.
Physics Based Aircraft Flight Trajectory Prediction ^[27]	Utilizing a Deep Residual RNN to predict the flight trajectory of aircraft and reduce computational cost of aircraft simulations. This DR-RNN was compared to a more typical LSTM RNN. A tool would be needed to create flight simulations and be able to gather that simulation data.	In case 2, or longitudinal responses, the prediction error was 3.20×10^{-7} . In case 3, or lateral responses, the prediction error was 1.17×10^{-5} . The dataset used to train the DR-RNN was gathered from simulated data of a Boeing 747-100 with introduced anomalies.	Deep Residual RNNs allow for the integration of aircraft dynamics into the simulations used to calculate aircraft trajectories. Both models outperformed previous numerical based simulation methods.
Real Time Anomaly Detection Onboard Unmanned Aerial Vehicles ^[21]	Leveraging and LSTM RNN to detect real time flight data anomalies in UAV drone data. Sensors onboard the drone were able record and log data during the drones flights.	This proposed model was able to get an accuracy of 99.7% for forward velocity anomalies, and 100% for pneumatic lifting anomalies. The dataset used in this model came from the actual flight data logged by drone flights.	Detecting anomalies in real-time can be difficult when using an LSTM RNN architecture. The data coming off the UAV will need to be forwarded through the network quickly to constantly ensure the drone is operating properly.
Prediction of Remaining Life of Jet Turbine Engines ^[23]	The researchers devised a fusion network built from an LSTM-HMM to predict remaining life of a jet turbine engine. Data was gathered from 21 sensors outside and inside of the jet turbine engine to measure vibrations.	The LSTM-HMM network scored an F1 accuracy of 0.781. The dataset used to train and evaluate this model came from the C-MAPSS dataset.	There is often a lot of noise within data coming from engine sensor data, making sure excess vibration anomalies are being correctly identified can be difficult.

Choosing the best method for facial expression recognition might be slightly more straight forward since you would like a method that is both fast and accurate. All the papers reviewed above had great scores, but image processing still takes the most time depending on the pixel density of each frame in the video and given that a 3 s video at 60 fps is 180 frames that need to be propagated through the network.

Now there are different types of RNNs like an LSTM-RNN or a stacked RNN framework. A growing category of anomaly detection is in aviation. Aviation is a newer and growing section of anomaly detection that focuses on all parts of the aircraft from engine vibration to its trajectory. An RNN based approach has also been proven to be the most useful strategy in aviation as well and any new models would greatly benefit from an LSTM-RNN approach if there is any trouble on deciding what model to use.

Recurrent Neural Networks have many benefits over other styles of machine learning methods. This allows for RNN's to process sequential data in time steps which other machine learning models cannot do. However, when you want to do speech recognition, auto generation of captions, or even having a computer generate music, it needs to hold on to that sequential data to help predict the next state. Not only have they seen a benefit over previous machine learning models, but RNN's also open more possibilities for new ways in which machine learning can accomplish a certain task.

4. Conclusions and Future Work

The goal of this paper was to provide insights into current research being done in four similar yet very distinct fields. These areas are biometric authentication, expression recognition, anomaly detection, and aviation. This paper specifically looked at how Recurrent Neural Networks were changing the game and allowing for new innovations. With continued research into these areas, there can be even more improvement in each of these areas: making sure that user data and critical systems are secured with top-level biometric authentication, paving a road for improvement in interactions between man and machine, detecting malicious actors and making sure people stay safe through novel anomaly detection techniques, and making air travel even safer while getting the most use out of aircraft parts.

References

1. Gunn, D.J.; Liu, Z.; Dave, R.; Yuan, X.; Roy, K. Touch-based active cloud authentication using traditional machine learning and LSTM on a distributed tensorflow framework. *Int. J. Comput. Intell. Appl.* 2019, 18, 1950022:1–1950022:16.
2. Tse, K.; Hung, K. User behavioral biometrics identification on mobile platform using multimodal fusion of keystroke and swipe dynamics and recurrent neural network. In *Proceedings of the 2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, Penang, Malaysia, 18–19 April 2020; pp. 262–267.
3. Tolosana, R.; Vera-Rodriguez, R.; Fierrez, J.; Ortega-Garcia, J. BioTouchPass2: Touchscreen password biometrics using time-aligned recurrent neural networks. *IEEE Trans. Inf. Forensics Secur.* 2020, 15, 2616–2628.
4. Shelton, J.; Rice, C.; Singh, J.; Jenkins, J.; Dave, R.; Roy, K.; Chakraborty, S. Palm print authentication on a cloud platform. In *Proceedings of the 2018 International Conference on Advances in Big Data, Computing and Data mining*, Durban, South Africa, 6–7 August 2018.
5. Wang, J.; Li, Y.; Ao, X.; Wang, C.; Zhou, J. Multi-modal biometric authentication fusing iris and palmprint based on GMM. In *Proceedings of the 2009 IEEE/SP 15th Workshop on Statistical Signal Processing*, Cardiff, UK, 31 August–3 September 2009; pp. 349–352.
6. Hung, T.B.; Tien, M.L. Facial expression recognition with CNN-LSTM. In *Research in Intelligent and Computing in Engineering. Advances in Intelligent Systems and Computing*; Kumar, R., Quang, N.H., Kumar Solanki, V., Cardona, M., Pattnaik, P.K., Eds.; Springer: Singapore, 2021; Volume 1254.
7. Mori, R. Anomaly detection and cause analysis during landing approach using recurrent neural network. *J. Aerosp. Inf. Syst.* 2021.
8. Menter, Z.; Tee, W.; Dave, R. A study of machine learning based pattern recognition in IoT devices. In *Proceedings of the 3rd International Conference of Communications and Computing Technologies*, Virtual Format, Paris, France, 27–28 February 2021.
9. Fernandez-Lopez, P.; Liu-Jimenez, J.; Kiyokawa, K.; Wu, Y.; Sanchez-Reillo, R. Recurrent neural network for inertial gait user recognition in smartphones. *Sensors* 2019, 19, 4054.
10. Kim, B.-H.; Pyun, J.-Y. ECG identification for personal authentication using LSTM-based deep recurrent neural networks. *Sensors* 2020, 20, 3069.
11. Salloum, R.; Jay Kuo, C.-C. ECG-based biometrics using recurrent neural networks. In *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, USA, 5–9 March 2017.
12. Chauhan, J.; Seneviratne, S.; Hu, Y.; Misra, A.; Seneviratne, A.; Lee, Y. Breathing-based authentication on resource-constrained IoT devices using recurrent neural networks. *Computer* 2018, 51, 60–67.
13. Sun, Y.; Gao, Q.; Du, X.; Gu, Z. Smartphone user authentication based on holding position and touch-typing biometrics. *Comput. Mater. Contin.* 2019, 61, 1365–1375.
14. Rajan, S.; Chenniappan, P.; Devaraj, S.; Madian, N. Novel deep learning model for facial expression recognition based on maximum boosted CNN and LSTM. *IET Image Process.* 2020, 14, 1373–1381.

15. Chen, S.; Jin, Q. Multi-modal dimensional emotion recognition using recurrent neural networks. In Proceedings of the 5th International Workshop on Audio/Visual Emotion Challenge, Brisbane, Australia, 26 October 2015.
16. Huang, K.; Li, J.; Cheng, S.; Yu, J.; Tian, W.; Zhao, L.; Hu, J.; Chang, C.C. An efficient algorithm of facial expression recognition by TSG-RNN network. In *MultiMedia Modeling Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2019.
17. Wang, H.; Zhou, G.; Hu, M.; Wang, X. Video emotion recognition using local enhanced motion history image and CNN-RNN networks. *Biom. Recognit.* 2018.
18. Zhao, L.; Shi, G. Maritime anomaly detection using density-based clustering and recurrent neural network. *J. Navig.* 2019, 72, 894–916.
19. Muharemi, F.; Logofătu, D.; Leon, F. Machine learning approaches for anomaly detection of water quality on a real-world data set. *J. Inf. Telecommun.* 2019, 3, 294–307.
20. Shi, Z.; Xu, M.; Pan, Q.; Yan, B.; Zhang, H. LSTM-based flight trajectory prediction. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018.
21. Wang, B.; Wang, Z.; Liu, L.; Liu, D.; Peng, X. Data-driven anomaly detection for UAV sensor data based on deep learning prediction model. In Proceedings of the 2019 Prognostics and System Health Management Conference (PHM-Paris), Paris, France, 2–5 May 2019.
22. Nanduri, A.; Sherry, L. Anomaly detection in aircraft data using recurrent neural networks (RNN). In Proceedings of the 2016 Integrated Communications Navigation and Surveillance (ICNS), Herndon, VA, USA, 19–21 April 2016.
23. Bi, J.X.; Fan, W.Z.; Wang, S.B. Remaining life prediction for aircraft turbine engines based on LSTM-RNN-HMM-APPROACH. *Mater. Sci. Eng.* 2021, 1043, 022033.
24. Fu, S.; Qin, D.; Qiao, D.; Amariuca, G.T. RUMBA-mouse: Rapid user mouse-behavior authentication using a CNN-RNN approach. In Proceedings of the 2020 IEEE Conference on Communications and Network Security (CNS), Avignon, France, 29 June–1 July 2020.
25. Tolosana, R.; Vera-Rodriguez, R.; Fierrez, J.; Ortega-Garcia, J. Biometric signature verification using recurrent neural networks. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017.
26. Luo, W.; Liu, W.; Gao, S. A revisit of sparse coding based anomaly detection in stacked RNN framework. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.
27. Yu, Y.; Yao, H.; Liu, Y. Aircraft dynamics simulation using a novel physics-based learning method. *Aerosp. Sci. Technol.* 2019, 87, 254–264.