Recurrent Neural Network (RNN) for Igbo Handwritten Character Recognition

Jumoke F. Ajao[1]*; Oluwasogo A. Okunade[2]; and AbdulaIzzez O. Ajao[3]

[1]Department of Computer Science, Kwara State University, Malete, Nigeria.
[2]Department of Computer Science, National Open University of Nigeria, Abuja, Nigeria.
[3]Department of Computer Engineering Technology, Federal Polytechnic, Offa, Nigeria.

E-mail: jumoke.ajao@kwasu.edu.ng*
aokunade@noun.edu.ng
ajaodoc@yahoo.com

ABSTRACT

Language is an important medium of communication between human beings and likewise men and machines. The field of computer science, Artificial Intelligence (AI), and linguistics, mainly focus on the interactions between computers and human or natural languages. Likewise, letters/hand-written as a means of communication, due to the differences in the styles, sizes, and orientations of each person's handwriting, character recognition has proven to be one of the most challenging problems. There must be a way of presenting and processing message(s) received/sent in a non-understandable language to a machine understandable language.

Not much work on Igbo Languages have been explored and this may result into exclusion of those Igbo Language. Data were acquired from literate Igbo indigenous writers. The Igbo Language recognition was carried out using Recurrent Neural Network (RNN) Model, where the input data was convolved with 3x3 filters and the maxpooling to reduce the dimension of the convolve result and this was done on the complicated computation utilizing variables with random initializations (called weights and biases), to generate a predicted outcome through the network. The network then checks for error by comparing that result to the intended value, and it propagates that error back over the same path to the variables to be changed.

The following parameters were obtained by implementing this network; after 10 epochs, the validation loss is 4.2798, the validation accuracy is 0.0127, the loss is 4.2756 and the accuracy is 0.0141. The predicted accuracy of the test dataset is 0.0080. This research helps to reduce this gap in the language technology system and enhance Igbo language to be recognized in communication.

(Keywords: handwritten character set, Natural Language Processing, NLP, neural network, personal handwriting, human machine communications)

INTRODUCTION

Language was initially known as a means of communication between two or more people. Later it was understood that language can communicate between living things. Today language is no longer limited to means of communication between two or more people or among living things alone, but also serves as a means of communication between living and non-living things, such as humans and machines. For effective communication between living (Human) and non-living things (Machine), there must be a way of presenting and processing message(s) received in a non-understandable language to an understandable language such as the conversion of human-understandable language (also known as a Natural Language) to machine understandable language using compiler or interpreter and vice-versa. This process is called Natural Language Processing (NLP). Natural languages are human languages that are spoken by people (Abhimanyu, *et al.,* 2013).

Natural language processing is the process of computer to understand, interpret, and generate human language. Natural language processing (NLP) according to Abhimanyu, *et. al*., (2013) and Chopra, *et al*., (2013) is a field of computer science, Artificial Intelligence (AI), and linguistics, that mainly focuses on the interactions between computers and human or natural languages.

NLP focuses on the area of human-computer interaction. It is a means of communication between humans and machines. Natural language processing attempts at building machines that understand and respond to text or voice data and respond with text or speech of their own in the same way as human does. Natural Language Processing (NLP) also known as the synchronized model of language (Wang, *et al*., 2021), according to Sethunya, *et al*., (2016) is a research and application area focusing on computer understanding and analyzing natural language and speech. Various research conducted in the field has discussed NLP in the following listed context according to (Sethunya, *et al*., 2016):

- Natural Language Understanding
- Natural Language Generation
- Speech or Voice recognition
- Machine Translation and
- Spelling Correction and Grammar Checking

NLP manipulates the supplied text and speeches in increasingly sophisticated manners to structure the text in such a way that can promote processing. Natural language will likely be the main form of communication between people and machines in the future, according to recent advancements in the field of natural language processing. It is fascinating to see how computers mimic some of the most fundamental human communication abilities. For example, a significant amount of contemporary communication is conducted between people and computers (Onyenwe, 2017). Humans no longer need to exchange cash with other people in order to deposit or withdraw funds. The Automatic Teller Machines (ATM) has overtaken these exchanges.

A human being may use the digital assistants built into smartphones to make Internet searches, set alarms and calendar reminders, find and connect with friends, and more. Humans can even converse with others in various languages as a result of "speak and translate" technology functioning as a mediator (for example, Apple iTunes free live voice and text translator that speaks 42 languages and holds written chats in 100 languages) (Onyenwe, 2017).

When a human interacts with a computer, the ability of a device to respond becomes the next level of interaction. The machine’s response to transmissions varies such as indicator light; the battery is low, the filter is clogged, a container needs to be emptied, something is broken, front is not clear, there is a barrier ahead, a door is open, access is unauthorized, the water-level is low, an alarm sound, and so on. The devices started to initiate communication when something went wrong. (Bylieva, 2020).

Due to the differences in the styles, sizes, and orientations of each person's handwriting, character recognition has proven to be one of the most challenging problems. Character recognition converts printed documents into ASCII files for editing, efficient storage, and quick computer retrieval. Character recognition falls into one of two categories: In the first - Online character recognition - text is automatically translated as it is written on a digitizing device like a computer tablet, where a sensor detects the pen velocity as characters are penned. The impulses are converted into a letter code that a computer and text processing software may use.

The second type - Offline character recognition, involves making the handwritten characters available to a recognition system. They are scanned as paper documents, processed, and transformed into binary or grayscale (Olaniyi, *et al.*, (2018). Language barriers are being broken down with language technology systems, but most African languages are under-resourced and have not been featured in this line of research due to a lack of resources. It is likely that if nothing is done, the speakers might be excluded from communication in the world using their languages. This research helps to reduce this gap in the language technology system and enhance this language to recognize Igbo Language. Many studies have been conducted on handwritten characters or word recognition, but there are few studies on Igbo recognition. The few studies that exist also concentrate on character recognition, and the studies conducted on word recognition focused on identifying specific words, such as medical terms, because the system was trained using these selected words.

The Igbo people are also referred to as the Ibo. The Igbo people are one of the largest single ethnic groups in Africa. Most of the most fluent Igbo speakers reside mostly in the south-eastern part of Nigeria. They are made up 17% of Nigeria's population. The Igbo language has Thirty-six (36) alphabets (see Table 1 and 2), nineteen (19) consonants, eight (8) vowels, and nine (9) diagraphs.

Table 1: The Igbo Alphabet in Uppercase.

A	B	CH	D	E	F	G	GB	GH
GW	H	I	Ị	J	K	L	M	N
Ṅ	NW	NY	O	Ọ	P	R	S	SH
T	U	Ụ	V	W	Y	Z		

Table 2: The Igbo Alphabet in Lowercase.

a	b	ch	d	e	f	g	gb	gh
gw	h	i	ị	j	k	l	m	n
ṅ	nw	ny	o	ọ	p	r	s	sh
t	u	ụ	v	w	y	z		

The earliest Igbo words and phrases were discovered in the book Geschichte der Mission der Evangelischen Bruder auf der Carabischen Insel by German missionary G.C.A. Oldendorp. Published in 1777, "History of the Mission of the Evangelical Brothers in the Caribbean". The interesting narrative of Olaudah Equiano's life was published in 1789 in London, England, and contained 79 words. After that. J.F. Schon and Samuel Ajayi Crowther, two missionary linguists from the CMS (Church Missionary Society) staff in Freetown, were part of a Norris expedition on the Niger in 1841, along with twelve interpreters (including Igbo) who were descendants of freed slave families who had moved in Freetown. Schon had a passion for Igbo and Hausa and attempted to speak with Igbo people in their native language but was disappointed when they were unable to understand him (perhaps due to his accent). After that, he had a twenty-year break from studying Igbo.

The first African Anglican Bishop, linguist, and ex-slave Samuel Ajayi Crowther created the first Igbo textbook, *Isoama-Ibo A Primer*, in 1857. The first Igbo grammar, *Oku Ibo: Grammatical Elements of the Ibo Language*, was published by J. F. Schon in 1861 and was written in the Isuama dialect using Lepsius orthography (Onyenwe, 2017).

The research provides an inclusive opportunity for the African language in the language technological system for feature opportunity and technology inclusive. The research focused on the use of neural networks for Igbo language alphabets character recognition, using a Recurrent Neural Network. A Recurrent Neural Network (RNN) is a type of Artificial Neural Network (ANN) that uses sequential data or time series data. It is an ANN in which there is at least one loop in the propagation of signals (Haykin, 1994; Medsker and Jain, 2000 in Nikolic, 2017).

It is a deep learning technique commonly used for temporal problems like language translation, Natural Language Processing (NLP), speech recognition, and image captioning. They are being used by popular applications such as Siri, Voice Search, and Google Translate. Like other neural networks (for example feedforward and Convolutional Neural Networks (CNNs)), RNN utilizes training data to learn given tasks. Many neural network-based methods have also been applied to natural language processing (NLP) (Lin, *et al.*, 2015).

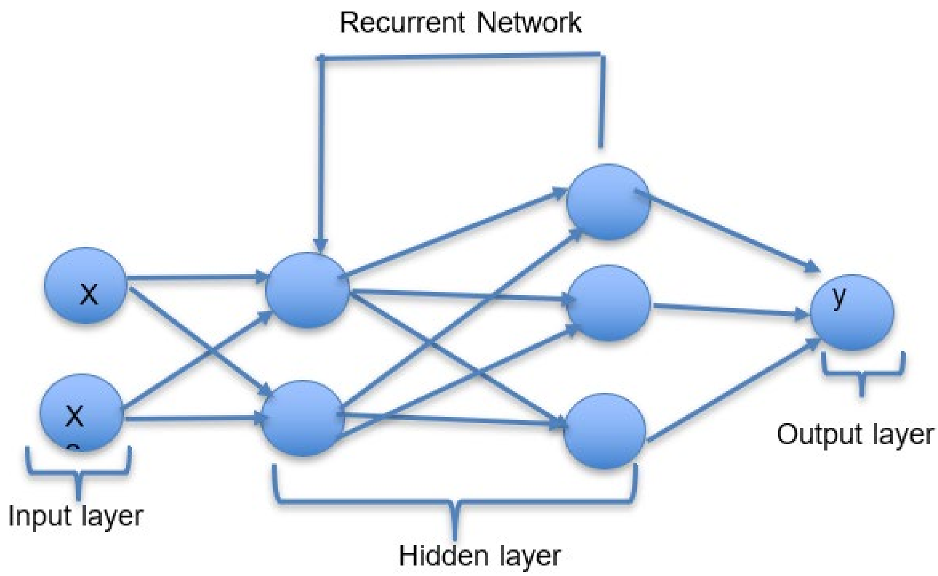


Figure 1: Recurrent Neural Network (source: Author, 2022).

There are factors that determined the architecture and performance of recurrent neural networks such as: number of layers, number of nodes in a layer, number placement and time delay of the recurrent connections. Recurrent Neural Network (RNN) that consists of input, hidden and output layers is illustrated in Figure 1. It is more powerful compared to feedforward neural networks. RNNs have demonstrated great success in sequence labelling and prediction tasks such as handwriting recognition and language modelling (Sak, et. al., 2014.).

Recurrent Neural Network language models demonstrated outstanding performance in a variety of natural language processing tasks. is the general class of a neural network that is the predecessor (Sherstinsky, 2021). RNN has been widely used in modeling online streaming data such as natural language texts and speech and online learning tasks such as time series prediction, language modeling, text generation, machine translation, speech recognition, text-to-speech generation, and so on (Sak, *et al*., 2014 and Zaremba, *et al*., 2014 in Diao, *et al*., 2019).

RNN at each time step consumes one input token, updates its hidden state vector, and predicts the next token by generating a probability distribution over all permissible tokens (Chen, *et al*., 2018). According to Abdulkarim (2016) RNN has been used with success in grammar/language processing, sentences of natural language are more than just a linear arrangement of words. Artificial deep-learning neural networks include convolutional neural networks that are used in image recognition and computer vision. These neural networks will be investigated, trained, and applied to recognize Igbo alphabets based on their beneficial properties and domain of usage. The neural network with the best accuracy will then be adopted.

METHODOLOGY

An Igbo language handwritten character set was put into practice and evaluated from a variety of different handwritten entries by both male and female students. 100 sheets of paper were used to record Igbo alphabets uppercase and lowercase (Figure 2) for each student and scanned to convert the dataset into softcopy using an HP Scanjet Pro 4500 scanner. This resulted in a total of 3,600 images, that is, 1,800 for uppercase letters and 1,800 for lowercase letters/case.

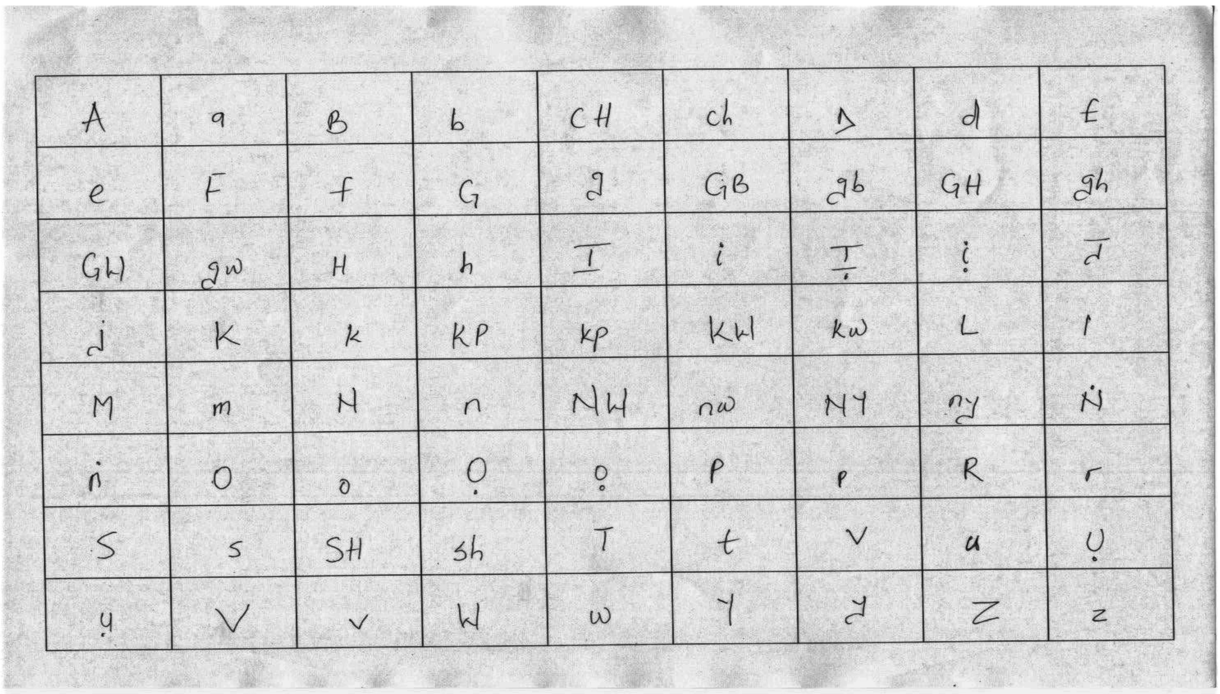


Figure 2: Dataset Acquisition Template.

Pre-Processing

The images were turned into grayscale (Figure 2) using an algorithm, the algorithm was then modified to loop through all folders and turn every image it finds into grayscale resized into 50 x 50.

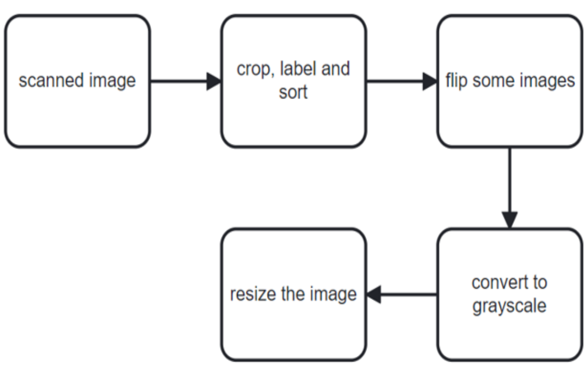
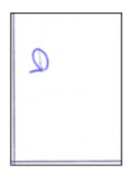
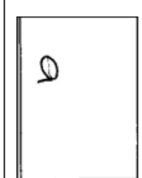


Figure 3: Pre-Processing of the Dataset.

Table 3: Stages of Dataset Processing

Dataset	Grayscale	Resized
		

Recurrent Neural Network (RNN) Model

RNN input data from the dataset as a complicated computation utilizing variables with random initializations (called weights and biases), to generate a predicted outcome through the network. The network then checks for error by comparing that result to the intended value, and it propagates that error back over the same path to the variables to be changed. Up till the variables are clearly defined, this procedure is repeated.

Applying these factors to a fresh, unforeseen input results in a forecast.

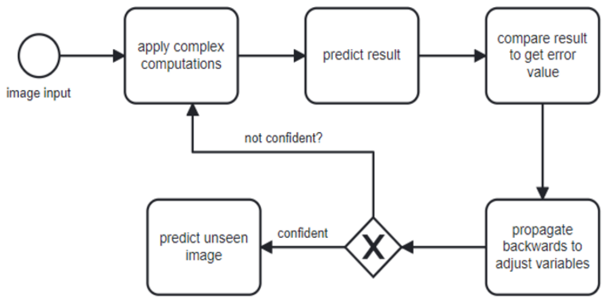


Figure 4: The RNN Process.

The architecture shown in Table 4 is the basic equation that defines the RNN as shown:

$$a^{(t)} = b + Wh^{(t-1)} + Ux^{(t)} \quad \textbf{(1)}$$

$$h^{(t)} = tanh(a^{(t)}) \quad \textbf{(2)}$$

$$o^{(t)} = c + Vh^{(t)} \quad \textbf{(3)}$$

$$y^{(t)} = softmax(o^{(t)}) \quad \textbf{(4)}$$

where a^t is the output, b is the bias associated with hidden layer, w_h is the weights associated with hidden unit of the recurrent layer U_x is the vector that stores the hidden unit at time t.

RESULTS AND DISCUSSION

Accuracy

This section discusses the accuracy that was achieved from both neural networks from training with the same recorded dataset as mentioned in the architecture.

The architecture adopted for this neural network is shown in Table 4.

Table 4: The RNN Architecture

Model: "sequential_1"

Layer (type)	Output Shape	Param #
lstm_2 (LSTM)	(None, 120, 120)	115680
dropout_3 (Dropout)	(None, 120, 120)	0
lstm_3 (LSTM)	(None, 120)	115680
dropout_4 (Dropout)	(None, 120)	0
dense_2 (Dense)	(None, 60)	7260
dropout_5 (Dropout)	(None, 60)	0
dense_3 (Dense)	(None, 72)	4392

Total params: 243,012
Trainable params: 243,012
Non-trainable params: 0

None

Table 5: Training of the Epochs.

```
Epoch 1/10
155/155 [==============================] - 126s 685ms/step - loss: 4.2892 - accuracy: 0.0127 - val_loss: 4.2779 - val_accuracy: 0.0163

Epoch 2/10
155/155 [==============================] - 112s 721ms/step - loss: 4.2781 - accuracy: 0.0135 - val_loss: 4.2763 - val_accuracy: 0.0163

Epoch 3/10
155/155 [==============================] - 86s 556ms/step - loss: 4.2770 - accuracy: 0.0101 - val_loss: 4.2776 - val_accuracy: 0.0163

Epoch 4/10
155/155 [==============================] - 93s 600ms/step - loss: 4.2767 - accuracy: 0.0143 - val_loss: 4.2788 - val_accuracy: 0.0163

Epoch 5/10
155/155 [==============================] - 65s 417ms/step - loss: 4.2764 - accuracy: 0.0145 - val_loss: 4.2780 - val_accuracy: 0.0127

Epoch 6/10
155/155 [==============================] - 53s 345ms/step - loss: 4.2760 - accuracy: 0.0153 - val_loss: 4.2785 - val_accuracy: 0.0127

Epoch 7/10
155/155 [==============================] - 54s 347ms/step - loss: 4.2761 - accuracy: 0.0151 - val_loss: 4.2795 - val_accuracy: 0.0127

Epoch 8/10
155/155 [==============================] - 50s 324ms/step - loss: 4.2757 - accuracy: 0.0135 - val_loss: 4.2796 - val_accuracy: 0.0127

Epoch 9/10
155/155 [==============================] - 47s 301ms/step - loss: 4.2757 - accuracy: 0.0157 - val_loss: 4.2797 - val_accuracy: 0.0127

Epoch 10/10
155/155 [==============================] - 46s 300ms/step - loss: 4.2756 - accuracy: 0.0141 - val_loss: 4.2798 - val_accuracy: 0.0127
```

RNN

For this neural network, 'categorical_crossentropy' is chosen, a validation split of 10% and 10 epochs (Table 5).

Upon completion of the training, the network achieved 0.0141 accuracy and a validation accuracy of 0.0127.

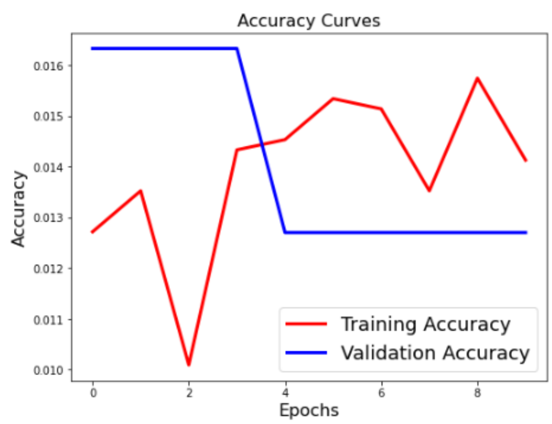


Figure 5: The Accuracy Curve for RNN.

RESULTS

The neural networks were trained on the same PC.

DISCUSSION OF RESULT

The research dataset was collected from 100 literate students on 100 distinct sheets of paper, ensuring a variety of handwriting styles. The dataset was then trimmed and sorted according to each class of the alphabets. OCR was developed in this study utilizing RNN. The dataset underwent pre-processing steps that included grayscale conversion and image scaling to ensure that both neural networks would perform well.

RNNs are made to receive an input, process it, and use variables with random weights or biases to carry out complex computations. It produces a predicted result, compares it to the desired value, and then propagates any errors backward. The following parameters were obtained by implementing this network; after 10 epochs, the validation loss is 4.2798, the validation accuracy is 0.0127, the loss is 4.2756 and the accuracy is 0.0141. The prediction accuracy of the test dataset is 0.0080.

RECOMMENDATION

The majority of pattern recognition tasks, including object identification, segmentation, and classification, are built on convolutional neural networks. Regardless, recurrent neural networks also specialize in speech recognition, Prediction problems, Text summarization, and so on. Therefore, it follows that if this model is used for a Character Recognition System, a usually reliable and efficient performance would be attained.

AUTHOR CONTRIBUTION

Oluwasogo A. Okunade prepared the literature review and oversaw the article writing.

Jumoke F. Ajao wrote the research methodology and performed fieldwork.

Abdulalzzez O. Ajao conducted the statistical analysis and interpreted the results.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Abdulkarim, S.A. 2016. “Time Series Prediction with Simple Recurrent Neural Networks”. *Bayero Journal of Pure and Applied Sciences*. 9(1): 19 – 24.

2. Abhimanyu, C., P. Abhinav, and S. Chandresh. 2013. *International Journal of Technology Enhancements and Emerging Engineering Research (IJTEEE)*. 1(4): 131-134.

3. Bylieva, D. 2020. “The Language of Human-Machine Communication”. *Technology and Language. Technology and Language Технологии в инфосфере*. 1(1): 16-21.

4. Chen, Y., S. Gilroy, A. Maletti, J. May, and K. Knight. 2018. “Recurrent Neural Networks as Weighted Language Recognizers”. *Proceedings of NAACL-HLT*. Association for Computational Linguistics: New Orleans, LA. 2261–2271.

5. Chopra, A., A. Prashar, and C. Sain. 2013. “Natural Language Processing”. *International Journal of Technology Enhancements and Emerging Engineering Research, IJTEEE*. 1(4): 131-134.

6. Diao, E., J. Ding, and V. Tarokh. 2019. “Restricted Recurrent Neural Networks”. IEEE.

7. Haykin, S. 1994. *Neural Networks*. Macmilan College Publishing Company: New York, NY.

8. Lin, R., S. Liu, M. Yang, M. Li, M. Zhou, and S. Li. 2015. “Hierarchical Recurrent Neural Network for Document Modeling”. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. 899–907.

9. Medsker, L.R. and L.S. Jain. 2000. *Recurrent Neural Networks: Design and Applications*. The CRS Press. International Series on Computational Intelligence.

10. Nikolic, K.P. 2017. “Training Neural Network Elements Created from Long Shot Term Memory”. *Oriental Journal of Computer Science & Technology. An International Open Free Access, Peer Reviewed Research Journal*. Oriental Scientific Publishing Co., New Delhi, India. 10(1), 01-10.

11. Olaniyi, E.O., O.K. Oyedotun, and K. Adnan. 2018. “Neural Network Recognition System for Igbo Vowels”. *International Journal of Scientific and Engineering Research.* 5(12): 435–440.

12. Onyenwe, I.E. 2017. “Developing Methods and Resources for Automated Processing of the African Language Igbo”. April.

13. Rodriguez, P., J. Wiles, and J.L. Elman. 1999. “A Recurrent Neural Network that Learns to Count”. *Connection Science*. 11(1): 5-40.

14. Sak, H., A. Senior, and F. Beaufays. 2014. “Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modelling”. In: Fifteenth Annual Conference of the International Speech Communication Association.

15. Sethunya, R.J., H. Hlomani, L. Keletso, K. Freeson, and S. Kutlwano. 2016. “Natural Language Processing: A Review”. *International Journal of Research in Engineering and Applied Sciences. IJREAS*. 6(3): 207-210.

16. Sherstinsky, A. 2021. “Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network”. *Physica D: Nonlinear Phenomena*. Special Issue on Machine Learning and Dynamical Systems. Vol. 404.

17. Wang, Y., E. Yue, and Y. Zhong. 2021. “Understanding Differences Between Human Language Processing and Natural Language Processing by the Synchronized Model”. *International Conference on Education, Language and Art (ICELA 2021). Advances in Social*

Science, Education and Humanities Research. 637.

18. Zaremba, W., I. Sutskever, and O. Vinyals. 2014. "Recurrent Neural Network Regularization". arXiv Preprint. arXiv:1409.2329.

SUGGESTED CITATION

Ajao, J.F., O.A. Okunade, and A.O. Ajao. 2023. "Recurrent Neural Network (RNN) for Igbo Handwritten Character Recognition". *Pacific Journal of Science and Technology*. 24(2):38-46.

Pacific Journal of Science and Technology