Igbo handwriting recognition system

**CHAPTER ONE**

**1.1 INTRODUCTION**

The handwriting recognition is an appealing and challenging research area in the field of image processing and pattern recognition. This arises from the need for humans to automate the recognition of handwritten text thereby enabling the computer to receive and interpret those (Oladele *et al.,* 2017). Moreover it helps to convert written text into digital format (Sharma *et al.,* 2015). Handwritten text has some problems such as ambiguities in variation in calligraphy over a period of time, similarity in text and variation in styles of writing. Handwriting data is time consuming and ineffective because the text must be retyped into digital text to be processed by the computer (Sepahvand *et al,* 2017). Handwriting recognition can be divided into the online and offline recognition system. Online recognition converts text written on a digitizer or PDA automatically where the sensor picks up the pen - tip movements and the pen-up/pen-down switching. The signal obtained from the pen - tip movements is converted into digital format. Offline recognition converts a text image to digital format. The image of the written text is scanned and sensed offline by optical scanning or intelligent recognition. In the offline recognition, the fixed static shape of the character is being recognized (Tawde, G.Y. and Kundargi, 2013), while the dynamic motion is recognized in online recognition during handwriting. The character is the basic building block of any language which is used to develop different language structures. Characters are alphabets and the structures developed are the words, strings, sentences, paragraphs and so on. Character recognition also known as optical character recognition is the recognition of optically processed characters. The purpose of character recognition is to interpret input as a sequence of characters from an already existing set of characters.

Handwritten character recognition is the process of converting handwritten text into a form that can be read by the computer, the major problem in handwritten character recognition system is the variation of the handwriting styles of individuals, which can be completely different for different writers.

Igbo language is one of the three main and official languages spoken in Nigeria (Nwachukwu, 2015) and mainly spoken in the south eastern part of the country. The Information Technology (IT) has evolved to the extent of using this language for its operation and creation of data. One can operate (Windows 7 and above) operating system, Microsoft office package and create documents using this language. It is envisaged that in future, as more people are getting used to the new trend, there will likely be large textual data generated with this language together with the growth of Igbo online information, which needs to be managed efficiently. These foreseen challenges have motivated the researcher to introduce an efficient computational approach to represent the language text document for any text-based system like text mining system, information retrieval system, and natural language processing system, owing to the fact that the text representation is one of the key essentials in text-based system. Text representation is a major basic and vital task in any text-based intelligent system. It involves looking for suitable terms to transfer text documents into numerical vectors (Zhang; Yoshida and Tang, 2015).

Igbo language is an agglutinative language, the language in which words are built up stringing different words. The individual meaning of words in a phrase or compound word does not entail the context it is been used for. The text document representation is one of the issues resulting from natural language peculiarities that need to be resolved for the success of any research in the text related fields. According to Soumya and Shibily; (2014), Bag-Of-Words (BOW) model is a traditional and conventional approach for representing text documents in any text-based task. The BOW model treats words independently and does not consider the word ordering, compound words, and collocations for semantic enhancement. When this model is used in an Igbo text document, it will not be dully represented because of compound nature of the language. In Igbo language, compounding is a common type of word formation and many compound words exist. Compound words play high roles in the language. They can be referred as Igbo phrases that make sense only if considered as a whole. Examples are “ụlọ akwụkwọ - school”; “onye nkuzi – teacher”; “kọmputa nkunaka – laptop”. Majority of Igbo terms, key words or features are in phrasal structure. The semantic of a whole is not equal to the semantic of a party (Stavrianou *et al.;* 2017). Some online translators in Igbo language is entangled with this problem of improper text representation. For example if you want to translate “school” in Igbo using most of the systems, it will either display “ụlọ” or “akwụkwọ” which is wrong. This is because BOW model is used in text representation and did not capture the compound word “ụlọ akwụkwọ” as a whole it should be in Igbo for “school”. This paper presents an analysis of Igbo compound word and proposes an efficient computational approach for representing the text for any text-based work using word-based N-gram model. This work is a part of ongoing research to create a text classification system for a text documents in an Igbo language. The model chosen by the researchers will help to discover unidentified facts and concealed knowledge that may exist in the lexical, semantic or relations (Stavrianou *et al.;* 2017) in Igbo text corpus. In Igbo, the semantic of individual words in a compound and combinatory semantics of the larger units are not the same at all (Ndimele, 2010).

**1.2 Background of the study**

Application of offline handwriting character recognition:

* National ID number recognition system,
* Postal office automation with code number recognition on envelop,
* Automatic license plate recognition,
* Bank automation.

**1.3 Statement of the Problem**

Igbo language is the expression and identity of the people of the South East Nigeria and there is no reason why Ndiigbo should not preserve and promote their mother tongue. The mother tongue is the most effective vehicle for cultural transmission. A lot of things, ideas, philosophy which the Igbo people cherish would certainly perish if their language is allowed to disappear and if they stop expressing and codifying those ideas and knowledge. Today there is a paradigm shift in learning method and for that within the education system, there are evidences of change in teaching and learning method. Classrooms and learning materials are no longer situated in one physical environment, but on virtual, online cyberspace. “Online availability of educational materials or digital resources has opened up new avenues for teaching and learning”.

The declining number of scholars yearning to study Igbo language and conduct research in Igbo demands that the approach for the teaching and learning of the language in tertiary institutions in Nigeria needed to be re-examined. Evidences of the positive impacts of information and communication technology in the promotion of teaching and learning English, French, German, Swahili, Yoruba, Hausa and other languages in tertiary institutions abound, and Igbo language, the supreme identity of the people of the South East Nigeria must not be by-passed. The contending issue or question which this paper wish to address is whether there are ICTs tools and programmes that could be applied in teaching and learning Igbo in tertiary institutions in the South East Nigeria; how these equipment can be used; the benefits and drawbacks.

Character recognition has been one of most difficult problem due to the variation in the way different people write, different sizes and also orientation of an individual handwritten varies. In Character recognition, printed documents are transformed into ASCII files for the purpose of editing, compact storage, fast retrieval through the computer. Character recognition can be divided into two: Online character recognition in which text is automatically converted as it is written on digitizer such as PC tablet, where a sensor picks up the pen velocity as characters are scripted. The signals obtained are transformed into a letter code which is usable to a computer and text processing application. The second type is offline character in which handwritten characters are scanned in form of paper document, process and converted to binary or gray scale to make available to a recognition algorithm. Many research works have been done on Chinese (Guarav K Pradeep, 2013) with recognition rate that is more than 98%, English character with recognition rate of 92.5%, English Vowel Character (Onyenwe *et al.;* 2014), Arabic character with recognition rate of 97% and Farsi Language with recognition rate of 85% but little work has been done on African character. Character recognition has many applications. In today’s' technology, Google translate apply character recognition. It recognized characters based on the previous learning by the system. In this work, Igbo vowel characters have been considered because of its importance in the southern eastern of Nigeria. Igbo language is one of the three major languages in Nigeria; it is widely spoken by approximately 18 million people in the south-eastern part of Nigeria and also recognized by minority in Equatorial Guinea (Chih-Fong, 2012). It is widely speak to carrying out daily today activities, it is speaks in customary court, churches, markets, festival activities. Therefore, it is important to develop an automatic vowel recognition system for Igbo language. In this work, we ex-piloted neural network for Igbo vowel handwritten character recognition. Neural network has widely been used in character recognition. In this paper, we present multi-layer feed forward network, trained with back propagation algorithm and develop an intelligent recognition system using neural network.

**1.4 Justification of the study**

The Igbo handwriting recognition system is one of the accuracy most important software in universities across the globe for both students and instructors because of the development in technologies.

**1.5 Significance of the study**

Importance of offline handwriting character recognition:

**1.5 Objective of the Study**

The general objective of this research is to develop a system that can recognize Igbo handwritten recognition system.

The specific objectives are to:

* Data Acquisition.
* Digitization of the collected dataset.
* Pre-processing.
* Creation of database.
* Design of recognition model.
* Develop the design model.
* Evaluate the developed model.

**1.6 Scope of the study**

This research work will covers the performance of sub word recognition for oﬀline Igbo handwritten images. We will also compare the recognition performance with other binarization based features which have been proven to be effective in capturing shape characteristics of handwritten Arabic sub words, such as GSC (a set of gradient, structure, and concavity features) and skeleton based Graph features.

**1.7 Igbo Language**

A language is a method of communication between individuals who share common code, in form of symbols (Ndimele, 2010). The Igbo language is one of the three major languages (Hausa, Yoruba and Igbo) in Nigeria. It is largely spoken by the people in the eastern part of Nigeria. Igbo language has many dialects. The standard Igbo is used formally and is adopted for this research. The current Igbo orthography (Onwu Orthography, 1961) is based on the Standard Igbo. Orthography is a way of writing sentence or constructing grammar in a language. Standard Igbo has thirty-six (36) alphabets (a, b, ch, d, e, f, g, gb, gh, gw, h, i, ị, j, k, kw, kp, l, m, n, nw, ny, ṅ, o, ọ, p, r, s, sh, t, u, ụ,v, w, y, z), consisting of eight (8) vowels and twenty-eight (28) consonants. The 28 consonant characters are “b, ch, d, f, g, gb, gh, gw, h, j, k, kw, kp, l, m, n, nw, ny, ṅ, p, r, s, sh, t, v, w, y, z” and 8 vowels characters are “a, e, i, ị, o, ọ, u, ụ”. There are nine consonants characters that are digraphs: “ch, gb, gh, gw, kp, kw, nw, ny, sh” (Onyenwe et al., *2014).* It uses a Roman Script and it is a tonal language with two distinct tones, high and low. Igbo is an agglutinative language, in which words are built by stringing different morphemes or words together (Ndimele, 2010). Igbo language has a large number of compound words. A compound word is a word that has more than one root, and can made from combination of either nouns, pronouns or adjectives.

2.1: The current orthography of Igbo Alphabets without diacritical marks.

**CHAPTER TWO**

**2.1 Related concept**

Handwriting recognition (HWR), also known as handwritten text recognition (HTR), is the ability of a computer to receive and interpret intelligible [handwritten](https://en.wikipedia.org/wiki/Handwriting" \o "Handwriting) input from sources such as [paper](https://en.wikipedia.org/wiki/Paper" \o "Paper) documents, [photographs](https://en.wikipedia.org/wiki/Photograph" \o "Photograph), [touch-screens](https://en.wikipedia.org/wiki/Touch-screen" \o "Touch-screen) and other devices. The image of the written text may be sensed "off line" from a piece of paper by optical scanning ([optical character recognition](https://en.wikipedia.org/wiki/Optical_character_recognition" \o "Optical character recognition)) or [intelligent word recognition](https://en.wikipedia.org/wiki/Intelligent_word_recognition" \o "Intelligent word recognition). Alternatively, the movements of the pen tip may be sensed "on line", for example by a pen-based computer screen surface, a generally easier task as there are more clues available. A handwriting recognition system handles formatting, performs correct [segmentation](https://en.wikipedia.org/wiki/Segment_(handwriting)" \o "Segment (handwriting)) into characters, and finds the most plausible words.

**2.1.1 Pattern Recognition**

Pattern recognition is the automated recognition of [patterns](https://en.wikipedia.org/wiki/Pattern" \o "Pattern) and regularities in [data](https://en.wikipedia.org/wiki/Data" \o "Data). It has applications in statistical [data analysis](https://en.wikipedia.org/wiki/Data_analysis" \o "Data analysis), [signal processing](https://en.wikipedia.org/wiki/Signal_processing" \o "Signal processing), [image analysis](https://en.wikipedia.org/wiki/Image_analysis" \o "Image analysis), [information retrieval](https://en.wikipedia.org/wiki/Information_retrieval" \o "Information retrieval), [bioinformatics](https://en.wikipedia.org/wiki/Bioinformatics" \o "Bioinformatics), [data compression](https://en.wikipedia.org/wiki/Data_compression" \o "Data compression), [computer graphics](https://en.wikipedia.org/wiki/Computer_graphics" \o "Computer graphics) and [machine learning](https://en.wikipedia.org/wiki/Machine_learning" \o "Machine learning). Pattern recognition has its origins in statistics and engineering; some modern approaches to pattern recognition include the use of [machine learning](https://en.wikipedia.org/wiki/Machine_learning" \o "Machine learning), due to the increased availability of [big data](https://en.wikipedia.org/wiki/Big_data" \o "Big data) and a new abundance of [processing power](https://en.wikipedia.org/wiki/Processing_power" \o "Processing power). These activities can be viewed as two facets of the same field of application, and they have undergone substantial development over the past few decades. Pattern recognition systems are commonly trained from labeled "training" data. When no [labeled data](https://en.wikipedia.org/wiki/Labeled_data" \o "Labeled data) are available, other algorithms can be used to discover previously unknown patterns. [KDD](https://en.wikipedia.org/wiki/Data_mining" \o "Data mining) and data mining have a larger focus on unsupervised methods and stronger connection to business use. Pattern recognition focuses more on the signal and also takes acquisition and [Signal Processing](https://en.wikipedia.org/wiki/Signal_Processing" \o "Signal Processing) into consideration. It originated in [engineering](https://en.wikipedia.org/wiki/Engineering" \o "Engineering), and the term is popular in the context of [computer vision](https://en.wikipedia.org/wiki/Computer_vision" \o "Computer vision): a leading computer vision conference is named [Conference on Computer Vision and Pattern Recognition](https://en.wikipedia.org/wiki/Conference_on_Computer_Vision_and_Pattern_Recognition" \o "Conference on Computer Vision and Pattern Recognition). In [machine learning](https://en.wikipedia.org/wiki/Machine_learning" \o "Machine learning), pattern recognition is the assignment of a label to a given input value. In statistics, [discriminant analysis](https://en.wikipedia.org/wiki/Linear_discriminant_analysis" \o "Linear discriminant analysis) was introduced for this same purpose in 1936. An example of pattern recognition is [classification](https://en.wikipedia.org/wiki/Classification_(machine_learning)" \o "Classification (machine learning)), which attempts to assign each input value to one of a given set of *classes* (for example, determine whether a given email is "spam" or "non-spam"). Pattern recognition is a more general problem that encompasses other types of output as well. Other examples are [regression](https://en.wikipedia.org/wiki/Regression_analysis" \o "Regression analysis), which assigns a [real-valued](https://en.wikipedia.org/wiki/Real_number" \o "Real number) output to each input; (Howard, 2020). [sequence labeling](https://en.wikipedia.org/wiki/Sequence_labeling" \o "Sequence labeling), which assigns a class to each member of a sequence of values (Howard, 2020) (for example, [part of speech tagging](https://en.wikipedia.org/wiki/Part_of_speech_tagging" \o "Part of speech tagging), which assigns a [part of speech](https://en.wikipedia.org/wiki/Part_of_speech" \o "Part of speech) to each word in an input sentence); and [parsing](https://en.wikipedia.org/wiki/Parsing" \o "Parsing), which assigns a [parse tree](https://en.wikipedia.org/wiki/Parse_tree" \o "Parse tree) to an input sentence, describing the [syntactic structure](https://en.wikipedia.org/wiki/Syntactic_structure" \o "Syntactic structure) of the sentence. In sociology, pattern recognition has been determined as a factor in racial and sexist inequality ([https://www.inc.com/kimberly-weisul/vivek-wadhwa-pattern-recognition-another-name racism-sexism.html](https://www.inc.com/kimberly-weisul/vivek-wadhwa-pattern-recognition-another-name-racism-sexism.html)). Pattern recognition algorithms generally aim to provide a reasonable answer for all possible inputs and to perform "most likely" matching of the inputs, taking into account their statistical variation. This is opposed to *[pattern matching](https://en.wikipedia.org/wiki/Pattern_matching" \o "Pattern matching)* algorithms, which look for exact matches in the input with pre-existing patterns. A common example of a pattern-matching algorithm is [regular expression](https://en.wikipedia.org/wiki/Regular_expression" \o "Regular expression) matching, which looks for patterns of a given sort in textual data and is included in the search capabilities of many [text editors](https://en.wikipedia.org/wiki/Text_editor" \o "Text editor) and [word processors](https://en.wikipedia.org/wiki/Word_processor" \o "Word processor).

**2.1.2 Stages in Pattern Recognition**

In [psychology](https://en.wikipedia.org/wiki/Psychology" \o "Psychology) and [cognitive neuroscience](https://en.wikipedia.org/wiki/Cognitive_neuroscience" \o "Cognitive neuroscience), **pattern recognition** describes a [cognitive](https://en.wikipedia.org/wiki/Cognitive" \o "Cognitive) process that matches information from a [stimulus](https://en.wikipedia.org/wiki/Stimulus_(psychology)" \o "Stimulus (psychology)) with information [retrieved](https://en.wikipedia.org/wiki/Recall_(memory)" \o "Recall (memory)) from [memory](https://en.wikipedia.org/wiki/Memory" \o "Memory). Pattern recognition occurs when information from the environment is received and entered into [short-term memory](https://en.wikipedia.org/wiki/Short-term_memory" \o "Short-term memory), causing automatic activation of a specific content of [long-term memory](https://en.wikipedia.org/wiki/Long-term_memory" \o "Long-term memory). An early example of this is learning the alphabet in order. When a carer repeats ‘A, B, C’ multiple times to a child, utilizing the pattern recognition, the child says ‘C’ after they hear ‘A, B’ in order. Recognizing patterns allows us to predict and expect what is coming. The process of pattern recognition involves matching the information received with the information already stored in the brain. Making the connection between memories and information perceived is a step of pattern recognition called identification. Pattern recognition requires repetition of experience. [Semantic memory](https://en.wikipedia.org/wiki/Semantic_memory" \o "Semantic memory), which is used implicitly and subconsciously is the main type of memory involved with recognition. Pattern recognition is not only crucial to humans, but to other animals as well. Even [koalas](https://en.wikipedia.org/wiki/Koala" \o "Koala), who possess less-developed thinking abilities, use pattern recognition to find and consume eucalyptus leaves. The human brain has developed more, but holds similarities to the brains of birds and lower mammals. The development of [neural networks](https://en.wikipedia.org/wiki/Neural_networks" \o "Neural networks) in the outer layer of the brain in humans has allowed for better processing of visual and auditory patterns. Spatial positioning in the environment, remembering findings, and detecting hazards and resources to increase chances of survival are examples of the application of pattern recognition for humans and animals. There are six main theories of pattern recognition: template matching, [prototype-matching](https://en.wikipedia.org/wiki/Prototype-matching" \o "Prototype-matching), feature analysis, [recognition-by-components theory](https://en.wikipedia.org/wiki/Recognition-by-components_theory" \o "Recognition-by-components theory), bottom-up and top-down processing, and [Fourier analysis](https://en.wikipedia.org/wiki/Fourier_analysis" \o "Fourier analysis). The application of these theories in everyday life is not mutually exclusive. Pattern recognition allows us to read words, understand [language](https://en.wikipedia.org/wiki/Language" \o "Language), recognize friends, and even appreciate [music](https://en.wikipedia.org/wiki/Music" \o "Music). Each of the theories applies to various activities and domains where pattern recognition is observed. Facial, music and language recognition, and seriation are a few of such domains. Facial recognition and seriation occur through encoding visual patterns, while music and language recognition use the encoding of auditory patterns (Mattson, 2014).

### Recognition by components theory

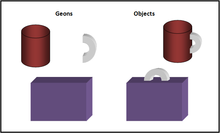
[](https://en.wikipedia.org/wiki/File:Breakdown_of_objects_into_Geons.png)

Image showing the breakdown of common geometric shapes (geons)

Similar to feature detection theory, [recognition by components](https://en.wikipedia.org/wiki/Recognition-by-components_theory" \o "Recognition-by-components theory) (RBC) focuses on the bottom-up features of the stimuli being processed. First proposed by Irving Biederman (1987), this theory states that humans recognize objects by breaking them down into their basic 3D geometric shapes called geons (i.e. cylinders, cubes, cones, etc.). An example is how we break down a common item like a coffee cup: we recognize the hollow cylinder that holds the liquid and a curved handle off the side that allows us to hold it. Even though not every coffee cup is exactly the same, these basic components helps us to recognize the consistency across examples (or pattern). RBC suggests that there are fewer than 36 unique geons that when combined can form a virtually unlimited number of objects. To parse and dissect an object, RBC proposes we attend to two specific features: edges and concavities. Edges enable the observer to maintain a consistent representation of the object regardless of the viewing angle and lighting conditions. Concavities are where two edges meet and enable the observer to perceive where one geon ends and another begins. The RBC principles of [visual object recognition](https://en.wikipedia.org/wiki/Visual_object_recognition" \o "Visual object recognition) can be applied to auditory language recognition as well. In place of geons, language researchers propose that spoken language can be broken down into basic components called [phonemes](https://en.wikipedia.org/wiki/Phoneme" \o "Phoneme). For example, there are [44 phonemes](https://en.wikipedia.org/wiki/English_phonology" \o "English phonology) in the [English language](https://en.wikipedia.org/wiki/English_language" \o "English language).

**2.1.3 Classification of Handwriting Character Recognition System**

Classification and recognition is the important section in the hand written character recognition system. For his artificial neural networks are used. Artificial neural networks are self- learning and self-organized network systems that are similar to that of biological neural network system. These highly parallel distributed network systems are mainly used for the pattern classification and recognition. In the pattern recognition the neural network system is trained to identify the most similar pattern from the data base that similar to the input image. Neural network system consists of several layers of neurons that are inter connected to each other. To activate a neural system input image is given to the input layer and a weight is assigned to each input node. All the node are connected each other so input node will activate the other nodes. It will continue until the output node will activate. After train neural network, the system is ready to recognize the input hand written images. In the proposed system several input images are trained and successfully recognize several hand written characters in different styles. The output from the proposed system is given in the Fig.6 and Fig.7. Fig.6. Input and Output of Character Recognition Fig.7. Input and Output of Character Recognition (Nisha *et al., 2012)*

Optical character recognition or optical character reader (OCR) is the [electronic](https://en.wikipedia.org/wiki/Electronics" \o "Electronics) or [mechanical](https://en.wikipedia.org/wiki/Machine" \o "Machine) conversion of [images](https://en.wikipedia.org/wiki/Image" \o "Image) of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (for example: from a television broadcast) (OnDemand, 2016). Widely used as a form of [data entry](https://en.wikipedia.org/wiki/Data_entry" \o "Data entry) from printed paper data records – whether passport documents, invoices, [bank statements](https://en.wikipedia.org/wiki/Bank_statement" \o "Bank statement), computerized receipts, business cards, mail, printouts of static-data, or any suitable documentation – it is a common method of digitizing printed texts so that they can be electronically edited, searched, stored more compactly, displayed on-line, and used in machine processes such as [cognitive computing](https://en.wikipedia.org/wiki/Cognitive_computing" \o "Cognitive computing), [machine translation](https://en.wikipedia.org/wiki/Machine_translation" \o "Machine translation), (extracted) [text-to-speech](https://en.wikipedia.org/wiki/Text-to-speech" \o "Text-to-speech), key data and [text mining](https://en.wikipedia.org/wiki/Text_mining" \o "Text mining). OCR is a field of research in [pattern recognition](https://en.wikipedia.org/wiki/Pattern_recognition" \o "Pattern recognition), [artificial intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence" \o "Artificial intelligence) and [computer vision](https://en.wikipedia.org/wiki/Computer_vision" \o "Computer vision).

Early versions needed to be trained with images of each character, and worked on one font at a time. Advanced systems capable of producing a high degree of recognition accuracy for most fonts are now common, and with support for a variety of digital image file format inputs.[[2]](https://en.wikipedia.org/wiki/Optical_character_recognition" \l "cite_note-2) Some systems are capable of reproducing formatted output that closely approximates the original page including images, columns, and other non-textual components.

**2.1.4 Machine Language**

In [computer programming](https://en.wikipedia.org/wiki/Computer_programming" \o "Computer programming), machine code is any [low-level programming language](https://en.wikipedia.org/wiki/Low-level_programming_language" \o "Low-level programming language), consisting of machine language [instructions](https://en.wikipedia.org/wiki/Instruction_set_architecture), which are used to control a computer's [central processing unit](https://en.wikipedia.org/wiki/Central_processing_unit" \o "Central processing unit) (CPU). Each instruction causes the CPU to perform a very specific task, such as a load, a store, a [jump](https://en.wikipedia.org/wiki/Jump_instruction" \o "Jump instruction), or an [arithmetic logic unit](https://en.wikipedia.org/wiki/Arithmetic_logic_unit" \o "Arithmetic logic unit) (ALU) operation on one or more units of data in the CPU's [registers](https://en.wikipedia.org/wiki/Processor_register" \o "Processor register) or [memory](https://en.wikipedia.org/wiki/CPU_cache" \o "CPU cache).

Machine code is a strictly numerical language which is designed to run as fast as possible, and may be considered as the lowest-level representation of a [compiled](https://en.wikipedia.org/wiki/Compiler" \o "Compiler) or [assembled](https://en.wikipedia.org/wiki/Assembly_language" \o "Assembly language) computer program or as a primitive and [hardware](https://en.wikipedia.org/wiki/Computer_hardware" \o "Computer hardware)-dependent programming language. While it is possible to write programs directly in machine code, managing individual bits and calculating numerical [addresses](https://en.wikipedia.org/wiki/Memory_address" \o "Memory address) and constants manually is tedious and error-prone. For this reason, programs are very rarely written directly in machine code in modern contexts, but may be done for low level [debugging](https://en.wikipedia.org/wiki/Debugging" \o "Debugging), program [patching](https://en.wikipedia.org/wiki/Patch_(computing)" \o "Patch (computing)) (especially when assembler source is not available) and assembly language [disassembly](https://en.wikipedia.org/wiki/Disassembly" \o "Disassembly).

The majority of practical programs today are written in [higher-level languages](https://en.wikipedia.org/wiki/High-level_programming_language" \o "High-level programming language) or assembly language. The source code is then translated to executable machine code by utilities such as [compilers](https://en.wikipedia.org/wiki/Compiler" \o "Compiler), [assemblers](https://en.wikipedia.org/wiki/Assembler_(computing)" \o "Assembler (computing)), and [linkers](https://en.wikipedia.org/wiki/Linker_(computing)" \o "Linker (computing)), with the important exception of [interpreted](https://en.wikipedia.org/wiki/Interpreted_language" \o "Interpreted language) programs,[[nb 1]](https://en.wikipedia.org/wiki/Machine_code" \l "cite_note-NB_List-1) which are not translated into machine code. However, the *[interpreter](https://en.wikipedia.org/wiki/Interpreter_(computing)" \o "Interpreter (computing))* itself, which may be seen as an executor or processor performing the instructions of the source code, typically consists of directly executable machine code (generated from assembly or high-level language source code).

Machine code is by definition the lowest level of programming detail visible to the programmer, but internally many processors use [microcode](https://en.wikipedia.org/wiki/Microcode" \o "Microcode) or optimise and transform machine code instructions into sequences of [micro-ops](https://en.wikipedia.org/wiki/Micro-operation" \o "Micro-operation). This is not generally considered to be a machine code.

The [MIPS architecture](https://en.wikipedia.org/wiki/MIPS_architecture" \o "MIPS architecture) provides a specific example for a machine code whose instructions are always 32 bits long. The general type of instruction is given by the *op* (operation) field, the highest 6 bits. J-type (jump) and I-type (immediate) instructions are fully specified by *op*. R-type (register) instructions include an additional field *funct* to determine the exact operation. The fields used in these types are:

6 5 5 5 5 6 bits

[ op | rs | rt | rd |shamt| funct] R-type

[ op | rs | rt | address/immediate] I-type

[ op | target address ] J-type

*rs*, *rt*, and *rd* indicate register operands; *shamt* gives a shift amount; and the *address* or *immediate* fields contain an operand directly.

For example, adding the registers 1 and 2 and placing the result in register 6 is encoded:

[ op | rs | rt | rd |shamt| funct]

0 1 2 6 0 32 decimal

000000 00001 00010 00110 00000 100000 binary

Load a value into register 8, taken from the memory cell 68 cells after the location listed in register 3:

[ op | rs | rt | address/immediate]

35 3 8 68 decimal

100011 00011 01000 00000 00001 000100 binary

Jumping to the address 1024:

[ op | target address ]

2 1024 decimal

000010 00000 00000 00000 10000 000000 binary

**2.1.5 Deep Learning**

Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on [artificial neural networks](https://en.wikipedia.org/wiki/Artificial_neural_network" \o "Artificial neural network) with [representation learning](https://en.wikipedia.org/wiki/Representation_learning" \o "Representation learning). Learning can be [supervised](https://en.wikipedia.org/wiki/Supervised_learning" \o "Supervised learning), [semi-supervised](https://en.wikipedia.org/wiki/Semi-supervised_learning" \o "Semi-supervised learning) or [unsupervised](https://en.wikipedia.org/wiki/Unsupervised_learning" \o "Unsupervised learning). Deep-learning architectures such as [deep neural networks](https://en.wikipedia.org/wiki/Deep_learning" \l "Deep_neural_networks), [deep belief networks](https://en.wikipedia.org/wiki/Deep_belief_network" \o "Deep belief network), [deep reinforcement learning](https://en.wikipedia.org/wiki/Deep_reinforcement_learning" \o "Deep reinforcement learning), [recurrent neural networks](https://en.wikipedia.org/wiki/Recurrent_neural_networks" \o "Recurrent neural networks), [convolutional neural networks](https://en.wikipedia.org/wiki/Convolutional_neural_networks" \o "Convolutional neural networks) and [Transformers](https://en.wikipedia.org/wiki/Transformer_(machine_learning_model)" \o "Transformer (machine learning model)) have been applied to fields including [computer vision](https://en.wikipedia.org/wiki/Computer_vision" \o "Computer vision), [speech recognition](https://en.wikipedia.org/wiki/Speech_recognition" \o "Speech recognition), [natural language processing](https://en.wikipedia.org/wiki/Natural_language_processing" \o "Natural language processing), [machine translation](https://en.wikipedia.org/wiki/Machine_translation" \o "Machine translation), [bioinformatics](https://en.wikipedia.org/wiki/Bioinformatics" \o "Bioinformatics), [drug design](https://en.wikipedia.org/wiki/Drug_design" \o "Drug design), [medical image analysis](https://en.wikipedia.org/wiki/Medical_image_analysis" \o "Medical image analysis), [climate science](https://en.wikipedia.org/wiki/Climatology" \o "Climatology), material inspection and [board game](https://en.wikipedia.org/wiki/Board_game" \o "Board game) programs, where they have produced results comparable to and in some cases surpassing human expert performance (Bengio et al., 2015).

[Artificial neural networks](https://en.wikipedia.org/wiki/Artificial_neural_network) (ANNs) were inspired by information processing and distributed communication nodes in [biological systems](https://en.wikipedia.org/wiki/Biological_system" \o "Biological system). ANNs have various differences from biological [brains](https://en.wikipedia.org/wiki/Brain" \o "Brain). Specifically, artificial neural networks tend to be static and symbolic, while the biological brain of most living organisms is dynamic (plastic) and analogue. The adjective "deep" in deep learning refers to the use of multiple layers in the network. Early work showed that a linear [perceptron](https://en.wikipedia.org/wiki/Perceptron" \o "Perceptron) cannot be a universal classifier, but that a network with a nonpolynomial activation function with one hidden layer of unbounded width can. Deep learning is a modern variation which is concerned with an unbounded number of layers of bounded size, which permits practical application and optimized implementation, while retaining theoretical universality under mild conditions. In deep learning the layers are also permitted to be heterogeneous and to deviate widely from biologically informed [connectionist](https://en.wikipedia.org/wiki/Connectionism" \o "Connectionism) models, for the sake of efficiency, trainability and understandability, whence the "structured" part (Marblestone et al., 2016).

**2.1.6 Igbo Orthography**

Orthography is the representation of the sounds in a language by written symbols. Orthography refers to the alphabet of a language. It is not the same as dialect. There is currently one orthography in use in Igbo language. This article is on the history of Igbo orthography and how we arrived at the current orthography. There are five main stages.

**1. Nsibidi**

The earliest known form of writing in Igbo land is the Nsidibi script. It was used by Igbo people and their neighbours in the present-day Cross River and Akwa Ibom states. Nsibidi is an ideogram, and it is independent of any particular language, therefore it was understood by Igbo, Efik, Ibibio and speakers of other languages who were trained to understand it. The use of nsibidi was restricted to the local police and courts. The local police and courts and the attendant nsibidi script were replaced by colonial government and christianity and the Latin script.  
**2. English Alphabet**

The Europeans (mainly British) arrived with the 26-letter latin alphabet:  
  
a b c d e f g h i j k l m n o p q r s t u v w x y z  
  
The alphabet did not capture all the Igbo sounds. For instance, Igbo was spelt as Ibo, because there is no equivalent 'gb' sound in English.

**3. Standard Alphabet**

Karl Lepsius was called upon to adapt the latin alphabet to meet the needs of African languages, including Igbo. In 1854, Lepsius published the ***Standard Alphabet for African Languages***. The Church Missionary Society (CMS) adapted Lepsius’ Standard Alphabet to produce the first Igbo Orthography comprising the following 34 letters:

a b d e f g h i k l m n o p r s t u v w y z gb gh gw kp kw n nw ny ọ š dš tš  
The Standard Alphabet had only six vowels, which was not adequate for the vowel sounds in Igbo language.

**4. Orthography of African Languages**

In 1927, the International Institute of African Languages and Cultures (IIALC) published the ***Practical Orthography of African Languages***. The two main proponents of this orthography were Mr. Adams (Inspector of Education) and Dr. Ward (a language researcher from Britain). The orthography was later known as the Adams-Ward Orthography. The Orthography of African Languages introduced two new letters to represent two vowel sounds: e and ɵ. The new orthography did away with diacritical marks. The orthography was made up of the following 36 letters (8 vowels and 28 consonants):

a b c d e ε f g gb gh h i j k kp l m n ŋ ny o ɔ ɵ p r s sh t u v w y z gw kw nw  
The Catholic mission which had not done lots of publications adopted the Orthography of African Languages, but the Church Missionary Society (Protestant mission) which had done most of their publications using Lepsius’ Standard Alphabet decided not to adopt the Orthography of African Languages. As a result, the Lepsius’ Standard Alphabet and the Orthography of African Languages existed side by side.

**5. Official Igbo Orthography**

As a result of the lack of agreement between the two main stakeholders in education (the Catholic mission and the Protestant mission). The government set up a Committee headed by Dr. S. E. Ọnwụ to recommend a script which would be acceptable to the Catholic mission and the Protestant mission. On 28 November 1953, the committee came up with the following orthography:   
a b gb ch d e f g h gh i ih j k l m n gn o or p kp r s sh t u uh v w y z gw kw nw ny  
The orthography introduced the letters *ih, gn, or, uh* to replace the diacritics. The orthography was neither accepted by the Catholic mission nor the Protestant mission. Consequently, a lot of Igbo scholars came up with their orthographies, and the number of orthographies increased instead of decreasing. The government was keen on establishing and maintaining order. So, the Ọnwụ Committee was reconvened. Whereas one of the goals of the first Ọnwụ committee was to remove the diacritics, one of the goals of the second Ọnwụ committee was to reintroduce the diacritics. In 1961, the Ọnwụ committee published the ***Official Igbo Orthography***. It became known as the ***Ọnwụ Orthography*** and comprises the following 36 letters:  
a b gb ch d e f g gw gh h i ị j k kp kw l m n nw ny ṅ o ọ p r s sh t u ụ v w y z  
In 1973, the Standardization Committee of the Society for Promoting Igbo Language and Culture (SPILC) made additional recommendation for the alphabet to be re-arranged in this order:  
a b ch d e f g gb gh gw h i ị j k kp kw l m n ṅ nw ny o ọ p r s sh t u ụ v w y z

**Vowel Comparison - Orthographies**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **#** | **Ọnwụ** | **IIALC** | **Lepsius** | **English** |
| 1 | a | a | a | a |
| 2 | e | ε | e | e |
| 3 | i | i | i | i |
| 4 | ị | e |  |  |
| 5 | o | o | o | o |
| 6 | ọ | ɔ | ọ |  |
| 7 | u | u | u | u |
| 8 | ụ | ɵ |  |  |

**Ọnwụ Orthography and Dialects of Igbo Language**

The [Ọnwụ orthography](https://www.igbovillagesquare.com/2019/11/igbo-alphabet.html) is the orthography that is currently used for writing Igbo language, and it is used for all Igbo dialects. There are arguments that more letters are needed to cater for the minor dialectical variations in pronunciations.

Although that argument seems reasonable, as the current Igbo alphabet is not perfect, but even the English alphabet we prefer to teach our children is not perfect. Think of how many ways the 't' in 'water' is pronounced from Abuja to Glasgow to London to Texas to Ottawa and so on. There is no value in having more letters. We can have one spelling and a variety of pronunciations.

**2.2 Rreview of related literature**

According to Nwadike (2012) observed that Igbo studies started and passed through thorny roads and setbacks before it migrated into the present era. For example, one of the major problems was that the wordlist was from the beginning beset with the problem of inconsistent orthography and therefore made publishing a comprehensive and reliable Igbo dictionary very difficult. When Adams and Ida C. Ward initiated in an effort to resolve the problem of orthography and in 1929 introduced a new Igbo orthography which generated a lot of controversy.

Obi (2010) cited Kay Williamson to have noted that the delay in producing Igbo dictionary include dialectical variation in Igbo, lack of a powerful Igbo personality to support and implement a standard Igbo language learning, and the disputed orthography. He said that these accounted for the background of poor advancement of Igbo literature. Obi regretted that both Igbo language and their literature could have benefited much through translation, but was not followed up to the present period in Igbo studies. This is because very little attention was paid to the translation in the modern period in spite of the fact that translation is a means of enriching Igbo literature and inspiring more Igbo readers Another setbacks include Igbo dialectical variations, the negative impressions held by the colonial imperialist scholars that Igbo language was difficult to study because it “is a tone language that depends more on the distinction or differences of sound for the meaning of its words, and for the construction of its sentences” (Afigbo, 1972). As many Igbo scholars acquired foreign languages and their culture, there was lack of competent and trained people to handle the job of teaching the language in schools; there was lack of schools and serious disregard and respect for the language by the people. For example, many Igbo parents, unlike their counterparts in Hausa in the Northern or Yoruba in the Western Nigeria, preferred their children to have mastery of English language at the earliest possible opportunity, instead of Igbo language. This problem had its genesis in 1882 when the obnoxious British Education Ordinance made the use of English language as a medium of instruction in schools.

According to Albert (2011) commenting in support of this ugly development said “many parents (in Nigeria) now introduce English language to their children earlier than they introduce their native languages because it is considered to be a source of pride to your (their) children to be able to speak in English language”. He regretted that we got it wrong when we accepted western values without actually knowing why we did it and when we encountered western civilization, we were not able to defend our own culture thinking that we were experiencing something superior, and today we are overwhelmed by that superiority. Oni (2011) in his criticism on this western erosion of African culture said, “The painful reality is that some Nigerian languages are going into extinction…shockingly, some Nigerian parents still believe that communicating with children in a foreign language is a mark of elitism” The repeated failure of early works in Igbo language was attributed to lack of trained people in the language to handle the job of teaching the language, lack of dynamic leader in Igbo who would pioneer genuine interest in the progress of the language, and the uncooperative attitudes of the Igbo people themselves. Others include poor regards and remunerations for graduates of Igbo certificate holders which consequently relegate them into inferiority complexes, lowers the population of students of Igbo studies in tertiary institutions, creates shortage of Igbo lecturers, researchers, writers, publishers, textbooks, and therefore lack of interest in reading Igbo literature. The contemporary problem is the challenges imported by the new information technologies into the teaching and learning field.

However, there were some writers whose work had made great contribution to Igbo language development. One of them was the book titled *Introduction to Igbo Study* written by Miss Ida Ward in 1936. It was one of the best works on Igbo grammar and today it has served as one of the principal source of references for most modern studies on Igbo. The book which was set up to analyze Igbo phonology and Igbo morphology achieved wide distribution and readership. Another great contribution to the development and study of Igbo language was the establishment of tertiary institutions, the Universities of Ibadan and Nsukka, and the creation and commencement of Igbo language studies in the Department of Linguistics and African/Nigeria Languages. The collection and preservation of Igbo books and literature in the Kenneth Dike Library University of Ibadan, and the Nnamdi Azikiwe Library in the University of Nigeria Nsukka, were another great service to early development of Igbo studies in Nigeria.

Researchers in handwriting recognition have been considering recognition in different languages as a dissimilar problem. Each language has its own character sets and the language features which make it difficult to develop a general algorithm that works for all the languages. Many researchers have used Freeman chain code for extracting features of characters or numerals due to its simplicity and the capability for small memory requirement. K-Nearest Neighbors (KNN) was also used because of its simplicity, reduced training time and good performance. Few works that have been done on Igbo character recognition uses Bayes and SVM. There is a need to test the simplicity of Freeman chain code and the less computational complexity training time of KNN on Igbo character and to determine its efficiency over other classifiers. The deep learning was introduced to avert the problem of extracting informative feature from the data using separate algorithm. From literatures, it was observed that deep learning as surpassed the conventional feature extraction algorithm in terms of accuracy. This work presents an approach to recognize Igbo handwritten characters based on Freeman chain code and KNN classifier.

Binarization

STAR

Load Character

Is Character

Grayscale?

Covert Character image to grayscale

Remove background and noise from

Character image (Filtering)

Detect edges from character image

Image dilation

Extract features Using zoning and gradient feature extraction

Is Character

Grayscale?

Yes

No

Input Handwritten Word

Train using KNN

Save training parameters

Segmentation

Stop

Test Image using KNN

Recognize Word

Yes

No

Figure 1: Flowchart of the developed system

**CHAPTER THREE**

**METHODOLOGY**

**3.1 Methodology**

The methods adopted for the proposed model are: data acquisition, pre-processing and classification

**3.2 Data Acquisition**

Data acquisition consisting of alphabets will be gathered . The acquired text will be written in a guided format by literate indigenous writers of Kwara State University, Malete students. Each student will be made to rewrite exactly from a sample copy of all the Igbo alphabets. The offline words were scanned and digitized. Samples of the collected images are shown in Fig 3.2.1. To enhance the number of sample datasets required for training, data augmentation techniques such as flipping and rotating the available dataset, utilizing multiple angles and flipping it in different directions, were used.

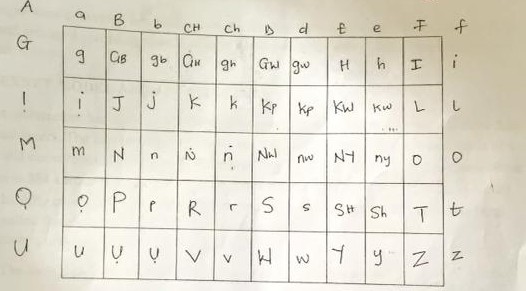


Fig 3.2.1. Original Images

**3.2.0 Pre-processing**

The purpose of preprocessing is to discard irrelevant information in the input data, that can negatively affect the recognition (Huang, B., et al. 2010). This concerns speed and accuracy. Preprocessing usually consists of binarization, normalization, sampling, smoothing and denoising. The second step is feature extraction. Out of the two- or higher-dimensional vector field received from the preprocessing algorithms, higher-dimensional data is extracted. The purpose of this step is to highlight important information for the recognition model. This data may include information like pen pressure, velocity or the changes of writing direction. The last big step is classification. In this step, various models are used to map the extracted features to different classes and thus identifying the characters or words the features represent.

The handwritten vowel characters were obtained by asking friends to write the vowels on the drawing pad software. These were later saved as a picture using Jpeg format. The total number of saved handwritten vowel is 36.This folder was then read into Python for processing. Some of common processing operations performed before recognition are: binarization, morphological operation, edge detection, dilation, skeletonization, rotation of images and feature extractions.

**3.2.1 Binarization**

The images were read into the phython inform of true colour which need to be converted to binary. The conversion depends on the thresholding value of the gray. Thresholding create binary images from gray level by turning all pixels below some threshold value to zero(background) and all pixels above the threshold value to one(foreground)..There are three types of thresholding

(1) Global thresholding

(2) Local thresh- olding

(3) Dynamic or adaptive thresholding. Thresholding may be viewed as an operation that involves test against a function T:

T = T[x,y p(x,y),f(x,y)]…………………….(1) Where f(x,y) –gray level of the point(x,y)  
P(x,y) – some local property of this point  
A thresholded image g(x,y) is defined as  
g(x,y) =1 if f(x,y)>1 or g(x,y) =0 if f(x,y)≤ 1…(2)  
The pixel label 1 correspond to object(foreground) and the  
pixel label 0 correspond to the background.  
Whenever T depend only on f(x,y) i.e. only the gray level val-  
ues,the threshold is known as Global threshold. But if, T depends on both f(x,y) and p(x,y), the threshold is called local. If

T depend on the spatial coordinate of x and y then the thresh- old is called dynamic or adaptive threshold.

**3.2.2 Morphological Operation**

Morphological operations depend on two argument which are binary images and structuring elements. The binary images are the images with two possible values for each pixel. The two colour used in binary images are black and white which is denoted as (1 and 0). The two colour are the foreground colour and the background colour. The structuring element is a shape used to interact with a given image with purpose of drawing conclusion on how this shape fit or misses the shape in the image. It is typically used in morphological operation such as dilation, erosion, opening and closing or as well as hit or miss transformation. The morphological operation compare the structuring element to neighbourhood of each pixel to deter- mine the output of its operation.

**3.2.3 Edge Detection**

Edge detection is a fundamental tool of image processing. It is usually employ for detection and extraction of feature in an image. Edge detection aims identifying the point at which there is a sharp changes in the brightness of a digital image and find discontintinutiesucing the amount of d. The main purpose of using edge detection is that it helps in terms of significantly reducing the amount of data in an image and preserve the structural properties for further image processing.

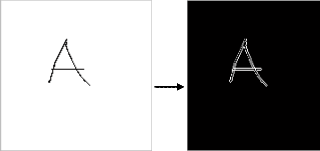
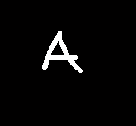
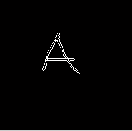


Fig 2 shows conversion of global threshold binary image to edge detection.

### 3.2.4 Dilation

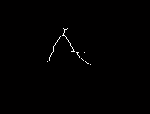
Dilation is a mathematical morphology operator that is basi- cally used in binary image though there is a version that can also be applied to grayscale images. Dilation works by gradually enlarging the boundaries of the region of the foreground pixels. The foreground areas grow in size while the holes within those region becomes smaller. Dilation being one of the morphological operator, it has two arguments which are the binary image to be dilated and the structuring element. The structuring element is usually small and it determines the precise effect of the dilation on the image. The structuring element that was applied in this research work was 3\*3 square with the origin at the centre. To determine dilation of a binary input image by the structuring element. The background pixels in the input image are considered one after the other. For each of the background pixels we place the structuring element on top of the input image so that the origin of the structuring element and the input pixel position coincide. The pixel is set to be foreground, if at least one of the pixel in the structuring element correspond with a foreground pixel in the image underneath. But if all corresponding pixel in the image are background, in this case the input pixel is left at the background value. The diagram below shows the process of dilation. 4a 4b Fig 4a shows the 45 degree rotation of the images, Fig 4bshow 330 degree rotation of the images





**3.2.5 Feature Extraction**

Feature is extracted from the character by cropping the important features. The sub-images have to be cropped sharp to the border of the character in other to standardize the sub- images. The image standardization was done by finding the maximum row and maximum column with 1s and with peak point, increase and decrease the counter until meeting the white space or line with all 0s. The technique of carrying out feature extraction is shown in the figure below

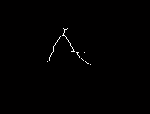


https://www.ijser.org/paper/NEURAL-NETWORK-RECOGNITION-SYSTEM-FOR-IGBO-VOWELS/Image_008.gif

Fig 3 Conversion from edge detection to dilation  
5 b

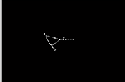
3.3.2 Skeletonization **5 a**

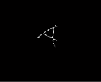




Skeletonization is a morphological operation that provides skeleton of all points belong to the medial axis, but due to pixalization and finite size of the structuring element, it is not necessarily connected nor maximally thin. It reduces the unnecessary pixel from an image and makes the image to left with one pixel thick, it preserve the topology of an object. It makes processing faster and also enhances recognition of images by recognition algorithms. The diagram below is a typical illustration of skeletonization as experiment in this research work.  
Fig 4 Shows conversion from dilation to skeletonization

**3.2.6 Rotation of the images**





Despite the variation in the writing style, shape of an individual handwritten, each handwritten vowels are rotated into two direction, 45 degree and 330 degree was considered in this work. This was done to build into the network some rotational in-variance. This makes the network to be more intelligent in recognizing characters with moderate displacements in orientation. The resampling of the image is shown in the figure be- low Fig. 5a Skeletonization of image, Fig. 5b Cropping of image. After the image has been pre-processed, The next step is to resize the image to meet up with the network input requirement. There is a need to be careful in resizing an image, as the blurriness of an image should be under consideration so as to make all the extracted feature be clear enough for the network to comprehend. In this work, the image was resize to 40 by 40 matrices which comprises background and foreground. The background is made up of 0s(black) and the foreground is made up of 1s(white) pixels. The image is then reshaped to be fed into the neural network. It is reshaped to 1600 \* 1 which was concatenated to 1600 \* 240 with binary matrices(1 and 0).

**3.3**  **Digitization of the collected dataset.**

The feature vector is calculated by converting the pre-processed image into bit mapped version of size 128 × 128. Figure 4 show few examples of the bit map version of different characters utilized in the proposed system. The bit map version preserve the major features of input image in shorter space/ data length. Such that reduces the time elapsed in NN Training without affecting the accuracy of correct character recognition. After that, The bit map images are converted into a single vector of size 128 ×1, which serves as an input vector to the LSTM.

**3.4 Creation of database**

I created an 8 x 9 table with equal squares. Then each alphabet was recorded in each box. A total of 50 students' handwriting was recorded. After that, I cropped each square using CorelDRAW graphics design as a tool to speed up the process. The snipping tool on your computer can also be used in case CorelDRAW doesn't install. Labelling each image was crucial to my next step. I wrote an algorithm that will sort each image based on the label and group them together. After the grouping, I wrote another algorithm that will binarize all the images in each folder.

**3.5.0 Design of recognition model.**

In the equations below, the lowercase variables represent vectors. Matrices {\displaystyle W\_{q}} and {\displaystyle U\_{q}} contain, respectively, the weights of the input and recurrent connections, where the subscript {\displaystyle \_{q}} can either be the input gate  {\displaystyle i}, output gate {\displaystyle o}, the forget gate {\displaystyle f} or the memory cell {\displaystyle c}, depending on the activation being calculated. In this section, we are thus using a "vector notation". So, for example,  {\displaystyle c\_{t}\in \mathbb {R} ^{h}} is not just one unit of one LSTM cell, but contains {\displaystyle h} LSTM cell's units.

### 3.5.1 LSTM with a forget gate

The compact forms of the equations for the forward pass of an LSTM cell with a forget gate are:



where the initial values are {\displaystyle c\_{0}=0} and {\displaystyle h\_{0}=0} and the operator {\displaystyle \odot } denotes the [Hadamard product](https://en.wikipedia.org/wiki/Hadamard_product_(matrices)" \o "Hadamard product (matrices)) (element-wise product). The subscript {\displaystyle t}  indexes the time step.

#### 3.5.2 Variables

*  : input vector to the LSTM unit
* : forget gate's activation vector
*  : input/update gate's activation vector
*  : output gate's activation vector
*  : hidden state vector also known as output vector of the LSTM unit
*  : cell input activation vector
*  : cell state vector
* ,  and  : weight matrices and bias vector parameters which need to be learned during training

where the superscripts {\displaystyle d}  and {\displaystyle h}  refer to the number of input features and number of hidden units, respectively.

#### [Activation functions](https://en.wikipedia.org/wiki/Activation_function)

*  : [sigmoid function](https://en.wikipedia.org/wiki/Sigmoid_function" \o "Sigmoid function).
*  : [hyperbolic tangent](https://en.wikipedia.org/wiki/Hyperbolic_tangent" \o "Hyperbolic tangent) function.
*  :  hyperbolic tangent function or, as the peephole LSTM paper suggests, 

### 3.5.3 {\displaystyle f\_{t}\in {(0,1)}^{h}}Peephole LSTM

The figure 3.3 below is a graphical representation of an LSTM unit with peephole connections (i.e. a peephole LSTM). Peephole connections allow the gates to access the constant error carousel (CEC), whose activation is the cell state.   is not used,   is used instead in most places.



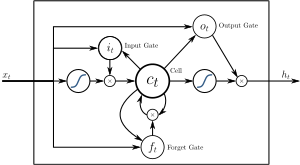


Fig 3.3 A [peephole LSTM](https://en.wikipedia.org/wiki/Long_short-term_memory" \l "Peephole_LSTM) unit with input (i.e. {\displaystyle i}), output (i.e. {\displaystyle o}), and forget (i.e. {\displaystyle f}) gates.

**3.5.4 Peephole convolutional LSTM**

Peephole [convolutional](https://en.wikipedia.org/wiki/Convolutional_neural_network" \o "Convolutional neural network) LSTM.[[20]](https://en.wikipedia.org/wiki/Long_short-term_memory" \l "cite_note-shi2015-20) The {\displaystyle \*}  denotes the [convolution](https://en.wikipedia.org/wiki/Convolution" \o "Convolution) operator.



**3.6.0 Develop the design model**.

An RNN using LSTM units can be trained in a supervised fashion on a set of training sequences, using an optimization algorithm like [gradient descent](https://en.wikipedia.org/wiki/Gradient_descent" \o "Gradient descent) combined with [backpropagation through time](https://en.wikipedia.org/wiki/Backpropagation_through_time" \o "Backpropagation through time) to compute the gradients needed during the optimization process, in order to change each weight of the LSTM network in proportion to the derivative of the error (at the output layer of the LSTM network) with respect to corresponding weight.

A problem with using [gradient descent](https://en.wikipedia.org/wiki/Gradient_descent" \o "Gradient descent) for standard RNNs is that error gradients [vanish](https://en.wikipedia.org/wiki/Vanishing_gradient_problem" \o "Vanishing gradient problem) exponentially quickly with the size of the time lag between important events. This is due to  {\displaystyle \lim \_{n\to \infty }W^{n}=0} if the [spectral radius](https://en.wikipedia.org/wiki/Spectral_radius" \o "Spectral radius) of {\displaystyle W} is smaller than 1.

However, with LSTM units, when error values are back-propagated from the output layer, the error remains in the LSTM unit's cell. This "error carousel" continuously feeds error back to each of the LSTM unit's gates, until they learn to cut off the value.

### 3.6.1. CTC score function

Many applications use stacks of LSTM RNNsand train them by [connectionist temporal classification (CTC)](https://en.wikipedia.org/wiki/Connectionist_temporal_classification_(CTC)" \o "Connectionist temporal classification (CTC)) to find an RNN weight matrix that maximizes the probability of the label sequences in a training set, given the corresponding input sequences. CTC achieves both alignment and recognition.

**3.7 Evaluate the developed model.**

Long short-term memory (LSTM)is an [artificial neural network](https://en.wikipedia.org/wiki/Artificial_neural_network" \o "Artificial neural network) used in the fields of [artificial intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence" \o "Artificial intelligence) and [deep learning](https://en.wikipedia.org/wiki/Deep_learning" \o "Deep learning). Unlike standard [feedforward neural networks](https://en.wikipedia.org/wiki/Feedforward_neural_network" \o "Feedforward neural network), LSTM has feedback connections. . A set of sample is given to the network with its corresponding classes or target. In this research batch training was employed in which all the input training was introduce to the network once with their corresponding target. During training, errors occur due to the difference between the actual output and target output. The dataset that we used in this research work was collected from 50 different people who we employ to help in using word pad to write the Igbo vowel. This is a dataset that contains all Igbo alphabets from start to finish and can be used for character recognition. It was recorded physically and has been binarized. The handwritten of 50 students was captured for both uppercase and lowercase, summing up to 3600 images in total. The data is duplicated to sum up to 7200 images. I created an 8 x 9 table with equal squares. Then each alphabet was recorded in each box. A total of 50 student’s handwriting was recorded. After that, I cropped each square using snipping tool to speed up the process. CorelDraw graphics design on computer can also be used. Labeling each image was crucial to my next step. I wrote an algorithm that will sort each image based on the label and group them together. After the grouping I wrote another algorithm that will binarize all the images in each folder. Also, it prevents the network to stuck at the local mini- ma.

**Chapter four**

**Table 1. LSTM model architect summary.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Layer No (type) | Output Shape | Param | Layer No (type) | Output Shape | Param |
| input (InputLayer) | 128, 128, 3 | 0 | 2 (Conv2D) | (None, 64, 64, 64) | 18496 |
| 1 (Conv2D) | 128, 128, 32 | 896 | batch\_normalization\_1 | (None, 64, 64, 64) | 256 |
| batch\_normalization | 128, 128, 32 | 128 | 1 (Activation) | (None, 64, 64, 64) | 0 |
| activation (Activation) | 128, 128, 32 | 0 | 2 (MaxPooling2D) | (None, 32, 32, 64) | 0 |
| max1 (MaxPooling2D) | 64, 64, 32 | 0 | (Dropout) | (None, 32, 32, 64 | 0 |
| 3 (Conv2D) | 32, 32, 128 | 73856 | 1 (Dense) | (None, 64, 73) | 74825 |
| batch\_normalization\_2 | 32, 32, 128 | 512 | lstm1 | (None, 64, 512) | 675840 |
| 2 (Activation) | 32, 32, 128 | 0 | Lstm2 | (None, 64, 512) | 1574912 |
| max3 (MaxPooling2D) | 32, 16, 128 | 0 | flatten (Flatten) | (None, 32768) | 0 |
| 1 (Dropout) | 32, 16, 128 | 0 | dense2 (Dense) | (None, 73) | 2392137 |
| (Reshape) | 64, 1024 |  | softmax (Activation) | (None, 73) | 0 |

Total params: 4,811,858

Trainable params: 4,811,410

Non-trainable params: 448

4.1. Optimizer and Learning rate

In Deep Learning and computer vision work, the optimization algorithm can change the result and make it prettysufficient. The Adam paper says, "...many objective functions are composed of a sum of subfunctions evaluated at different subsamples of data; in this case, optimization can be made more efficient by taking gradient steps w.r.t. individual sub-functions ..." (Diederik P.et al 2014). The Adam optimization algorithm is straightforward to implement, computationally efficient and little memory requirements that’s why recently adopting most of the computer vision and natural language processing application. Proposed method used ADAM Optimizer.



When using a neural network to perform classification and prediction task we need to calculate the error rate. A recent study shows that cross entropy function performs better than classification error and mean square error (Katarzyna., et al 2017). Proposed method used categorical cross entropy (4) as loss function.

Li = − ∑ t i, j j log(pi, j)

To make the optimizer converge faster and closer to the global minimum of the loss function, using an automatic Learning Rate reduction method (Tom Schaul., et al 2013). Learning rate is the step by which walks through the minimum loss. If the learning rate is too low, it will take more time to reach the global minima, and if the learning rate is too high then the training may not be converging or even diverge. To keep the advantage of the fast computation time we set a high learning rate which is automatically decreased by monitoring the validation accuracy.

**4.2. Data augmentation**

To avoid overfitting, artificially expand the handwritten dataset. This data transformation will create some variance that can occur when someone else writing the digits. For Data augmentation, several methods are chosen:

* Randomly shifting height and width 10% of the images.
* Randomly rotate our training image 10 degrees.
* Randomly 10 % zoom the training image.

**4.3. Training the model**

The proposed model was trained with different training and validation set with the batch size of 32. During training, automatic learning rate reduction formula mentored the validation accuracy and reduce the learning rate if required.

**4.4. Evaluate the model**

The proposed model was applied to the three dataset and gain good result on train, test, validation sets.

**4.4.1. Train, Test and Validation sets**

For Mixed Character(Upper and lower case) 5760 images used as train image and 1400 for the test image. For Upper and lower case characters 2880 images used as train image and 720 for the test image. We take 20% of train images for validation purpose. After training model was tested with 1400 images of Mixed test dataset also with 720 images of upper case dataset and 720 images of the Lower case dataset.

**4.4.2. Model performance**

After 9,9, 8 epochs proposed model gets 98%, 96.81%, 95.71%, and 96.40% validation accuracy respectively for Mixed, Upper and mixed dataset. Also, all of this dataset cross-validate with each other and perform accurately. Fig 4 (a) (b) (c) (d) is showing the accuracy and loss of training and validation set of Mixed, Upper case and Lower case dataset respectively. Table 2 showing the details accuracy on different datasets.

**Table 2. Result comparison in different Dataset.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dataset Name | Tr. Loss | Val. Loss | Tr. Acc. | Val Acc. | Test Dataset | Test Acc. |
| Mixed |  |  |  |  |  |  |
| Upper Case |  |  |  |  |  |  |
| Lower case |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

Table 3 showing a comparison between some the previous work. From this table, we found that our proposed BornoNet got so far, the best accuracy rate for all of the three datasets.

|  |  |  |  |
| --- | --- | --- | --- |
| **Work** | Accuracy | **Work** | Accuracy |
| Recognition of Handwritten Bangla Characters Using Gabor Filter and Artificial Neural Network (Halima Begum et al, 2015) | 79.4% | **Mixed(Upper and lower case ) Igbo Handwriting recognition using LSTM** |  |
| Recognition of Igbo handwritten basic characters and digits using convex hull-based feature set | 87.7% | **Upper case Igbo Handwriting recognition using LSTM** |  |
| Yoruba Handwritten Character Recognition using Freeman Chain Code and K-Nearest Neighbor Classifier(Jumoke Ajao ., 2018) | 85.36% | **lower case ) Igbo Handwriting recognition using LSTM** |  |

**Chapter Five**

**5.1 Error observation**

Analyzing the error from validation set we found that most of the incorrect classification is caused by error labeling on the dataset. It is clear that model performing great for classifying characters. Some of this mistake can also make by humans.

**5.2 Conclusion and Future work**

This research work presented a new LSTM model which performs better classification accuracy in the different dataset for both train and validation set for lesser epochs and less computation time compared to the other LSTM model. LSTM in general costly to train but extremely effective deep learning models. Hence, fast convergence should be viewed as an important front of research. Also, the cross-validation from different distribution’s data proposed model achieve a great result that makes it a robust model that improve any other previous model. Sometimes proposed model confused to understand overwritten character and dataset contained some incorrect labeling images. Also, the model performed poorly if the train on noise-free data. In future work fixing dataset and overcoming the limitation of overwriting Character should fix. Making a benchmark model for Igbo Handwritten all characters that include the Alphabetical letter.

**REFERENCES**

Afigbo, A.E. (1972). The study of Igbo language and the making of a linguistic dilemma, In Igbo Language Development Series.

Albert, Isaac (2011), Thursday March, 3). We need a very strong cultural policy, In Daily Sun 6(2011), 40.

Bengio, Yoshua; LeCun, Yann; Hinton, Geoffrey (2015). "Deep Learning". Nature. **521** (7553): 436–444.

Chih-Fong T.; (2012), “Bag-of-Words Representation in Image Annotation: A Review,” International Scholarly Research Network, ISRN Artificial Intelligence, Volume 2012.

George S. K.; and Joseph S.; (2014) “Text Classification by Augmenting Bag of Words (BOW) Representation with co-occurrence Feature,” IOSR Journal of Computer Engineering (IOSR – JCE) e-ISSN: 2278 – 0661, pp 34 – 38, 2014.

Guarav K; & Pradeep B; ( 2013),”Neural network based Approach for Recognition”, In International Journal of Computer Application,Vol. 62, 2013,pp. 8- 13. [2] LiangbinZheng, Ruqi Chen, XiaojinCheng,”Research on Offline Handwritten Chinese Character Recognition Based on BP Neural Network” In International Conference of Computer Science and Information Technology, Vol. 51, 2012, pp.323- 329

Howard, W.R. (2020). "Pattern Recognition and Machine Learning". Kybernetes. **36** (2): 275. [doi](https://en.wikipedia.org/wiki/Doi_(identifier)" \o "Doi (identifier)):[10.1108/03684920710743466](https://doi.org/10.1108%2F03684920710743466). [ISSN](https://en.wikipedia.org/wiki/ISSN_(identifier)" \o "ISSN (identifier)) [0368-492X](https://www.worldcat.org/issn/0368-492X)

Mattson, M. P. (2014). Superior pattern processing is the essence of the evolved human brain. Frontiers in neuroscience, 8.

Marblestone, Adam H.; Wayne, Greg; Kording, Konrad P. (2016). ["Toward an Integration of Deep Learning and Neuroscience](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5021692)

Nisha V; Hem J; and Singh V; (2012) Offline Character Recognition System Using Artificial

Neural NetworkInternational Journal of Machine Learning and Computing, Vol. 2, No. 4, August 2012

Nwachukwu A.P.; (2015) “Readings on the Igbo verbs,” Isbn 975 175 023 5, 2015.

Nwadike, I.U. (2012). Igbo Studies: from the plantations of West Indies to the forestland of West Africa, 1766-2008. (The) 33 Inaugural Lecture of the UNN, June, 12.

Ndimele O.; (2010), “A first course on morphology & syntax,” ISBN 978-33527-3-3, 2010.

Obi, S.U. (2010). Igbo language study: a historical survey, In Igbo Language Development Series.

Onyenwe I. E.; Uchechukwu C.; and Hepple M. (2014), “Part –of-Speech tagset and corpus development for Igbo, an African language,” The 8th linguistic annotation workshop, Dublin, Ireland, pp 93-98, August 2014.

OnDemand, HPE Haven (2016). ["OCR Document"](https://web.archive.org/web/20160415060125/https:/dev.havenondemand.com/apis/ocrdocument). Archived from [the original](https://dev.havenondemand.com/apis/ocrdocument#overview) on April 15, 2016.

Oni, Duro (2011, Thursday, March 3). Preservation and revival of Nigerian cultural heritage, In Daily Sun 6(2003), 38

Stavrianou A.; Andritsos P.; and Nicoloyannis N.; (2017) “Overview and semantic issues of text mining,” SIGMOD Record, Vol. 36, No. 3, September 2017.

Zhang W.; Yoshida T; and Tang X.; (2015) “Text Classification based on multi-word with support vector machine,” Journal Knowledge-based systems, vol. 21 Issue 8, pp 879- 886, December 2015.

Halima Begum et al, Recognition of Handwritten Bangla Characters using Gabor Filter and Artificial Neural Network, International Journal of Computer Technology & Applications, Vol 8(5),618-621, ISSN:2229-6093