Covenant Journal of Engineering Technology (CJET) Vol.3 No.1, June 2019
ISSN: p 2682-5317 e 2682-5325 DOI: 10.20370/2cdk-7y54

An Open Access Journal Available Online

A Review of Voice-Base Person Identification: State-of-the-Art

Folorunso C. O.[*1], Asaolu O. S.[2] & Popoola O. P.[3]

[1, 2, 3]Department of Systems Engineering, University of Lagos, Akoka, Lagos, Nigeria
[*1]comfortfolorunso@gmail.com, [2]oasaolu@unilag.edu.ng,
[3]toyin_net@yahoo.com

Received: 22.10.2018 Accepted: 11.04.2019 Date of Publication: June, 2019

Abstract - Automated person identification and authentication systems are useful for national security, integrity of electoral processes, prevention of cybercrimes and many access control applications. This is a critical component of information and communication technology which is central to national development. The use of biometrics systems in identification is fast replacing traditional methods such as use of names, personal identification numbers codes, password, etc., since nature bestow individuals with distinct personal imprints and signatures. Different measures have been put in place for person identification, ranging from face, to fingerprint and so on. This paper highlights the key approaches and schemes developed in the last five decades for voice-based person identification systems. Voice-base recognition system has gained interest due to its non-intrusive technique of data acquisition and its increasing method of continually studying and adapting to the person's changes. Information on the benefits and challenges of various biometric systems are also presented in this paper. The present and prominent voice-based recognition methods are discussed. It was observed that these systems application areas have covered intelligent monitoring, surveillance, population management, election forensics, immigration and border control.

Keywords: Person-identification, biometrics, information and communication technology, authentication, voice-base person identification.

1. Introduction

Person identification is the process of recognizing an individual based on specific characteristics. Identification by name is the most common method of distinguishing between persons. Conventional person identification technologies are developed based on "something we know" (password and personal identification number (PIN)),

"something we have" (Token, an Identity Card, an Access Control Card or a Radio Frequency Identity (RFID) card), unlike something we are (i.e. bodily features like fingerprint, finger-vein, voice, gait). These identification schemes have their limitations. For instance, a password or a PIN can easily be forgotten or even guessed by an impersonator while a token or an identity card can be lost, stolen or even cloned. As such, the conventional forms of identification may not really distinguish between an authorized user and an impersonator [1].

The need to verify the identity of an individual before such individual can have access to some restricted places is very important as so many means of breaking into one's privacy is being developed every day by hackers. The use of the state-of-the-art person identification technology which uses biometrics technology cannot be over-emphasized, due to so many advantages this type of recognition system proffers. Biometrics is the science and technology of measuring, analysing and comparing some human physiological and sociological related data [2]. Many biometrics-based identification systems have been employed both for domestic and commercial applications. The use of biometrics has shown increasing significance in our daily life and as such, many books, journals and conference proceedings have enumerated advances in theory and application of this technology. In all, the use of biometrics has emerged as the best way of identifying an individual since no two persons share entirely the same biometric traits [3].

With numerous developments in technology, the security sector has encountered some significant advancement. We can safely say that technology has transformed security. The challenge was how to make computers identify or recognise an individual. Biometric being based on measuring physical characteristics uses features which are common and readily available to all class of people. These features are distinct, easily collected and tested, and have high variability to carter for repetition of data class. Some of the biometric features currently used include: Facial thermogram, hand vein, gait, keystroke, odour, ear, hand geometry, fingerprint, face, retina, iris, palm-print, voice, signature and DNA.

In voice-based recognition system, the features of an individual's voice are based on the physical characteristics of their vocal tract, nasal cavities and the articulators (which include mouth, lips, teeth and so on) which are used in generating sound. These features are unchanging for an individual though the behavioural features may change over time due to age, location, medical conditions and/or emotional state [4]. Voice-based recognition methods are classified into two ways: Automatic Speaker Verification (ASV) and Automatic Speaker Identification (ASI) system [4]. The limitations of the conventional person identification system have always been challenging, hence the need to research into the state-of-the-art biometric technologies novel

biometric features that can be proposed.

2. Review of related works

So much research has been done on person identification using biometrics technology. Biometrics traits or characteristics are divided into two major categories. The first category is physiological characteristics, which are based on how one is, using data (features) that are derived from direct measurements of parts of the human body (fingerprints, face, finger geometry, iris scans, retina scans, hand geometry, finger-vein among others) as shown in Figure 1. The second category, which is the behavioural characteristics are based on what one does (a person's action), it uses data that are derived from indirect measurement of a person's action (voice recognition, keystroke scans, signature, gait and so on as shown in Figure 2 [1], [5].

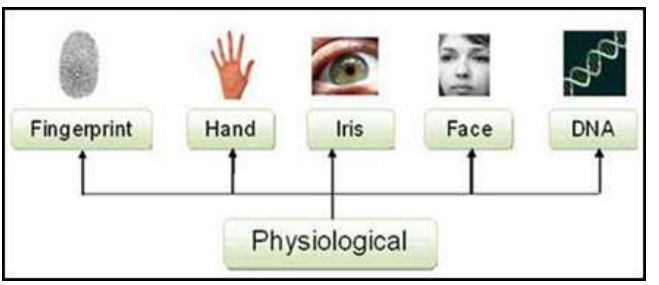

Figure 1: Selected physiological biometrics in use [3]

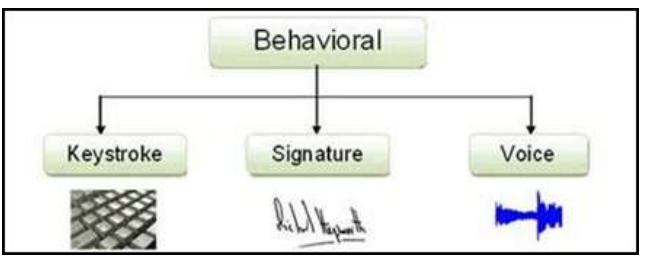

Figure 2: Selected behavioural biometrics in use [3]

A biometric authentication system can be further classified either as identification or a verification system. Identification system is a one-to-many system where biometrics is used to determine a person's identity from a number of some people's data stored in a database. For example, the police may try to identify an individual's fingerprint or face from a Forensic database. Verification on the other hand, is a one-to-one system of identification where biometrics is used to confirm a person's identity for access control [6], [7].

Kaur and Kaur [8], presented a brief survey of different voice biometric for speaker verification in an attendance

system. They proposed the use of voice, having considered different methods that have been employed for automatic attendance for students and as such the use of voice for this purpose is a very important and highly welcome phenomenon. They used Gammatone filter bank instead of Mel filter bank, after which the discrete cosine transform was applied to separate overlaying signals. The use of Gammatone frequency cepstral coefficient (GFCC) with the Gaussian Mixture Model as well as Artificial Neural Networks was incorporated for training and matching task respectively.

Tran et al [9] introduces normalization technique which depends on fuzzy set theory to improve the performance of voice-based verification system. In order to authenticate a claimed personality, a likeliness value was evaluated with a threshold in order to allow or reject the person. The use of noise clustering as well as the c-means clustering membership function was introduced to eradicate the problem of ratio-type scores which affects the false acceptance rate. Result however, showed great reduction in the false acceptance and false rejection rate.

Delac and Grgic [4], and Dugelay et al [10] presented a concise summary of different biometric methods which include single as well as multiple biometric systems. Voice-based biometric system uses some of the features of human-speech that are not changing for a particular individual. Though the behavioural features of the same human speech varies over time due to age medical, emotional as well as environmental conditions. The voice-based biometric system is classified into automatic speaker verification (ASV) and automatic speaker identifications (ASI). The former uses voice as validation characteristic in a one to one verification scenario. While the latter uses voice to recognise who a person truly is. A particular voice feature of an individual is matched against a stored pattern in a database. A typical voice feature can be formants or any other sound characteristics which are unique to each individual's vocal tract.

Belin et al [11] working from a neuro-cognitive point of view looked at neural association of voice perception. Having considered the ability to evaluate a person's gender and age bracket from listening to their voices was a very strong motivation behind this work. They tried to look at the voice as an auditory face in comparison with the face recognition system. They then proposed the use of Bruce and Young's model of face perception as a structure for understanding the perceptual as well as the cognitive processed contained in voice perception. This purpose model predicts functional detachment similar to those perceived for faces. In future, the use of face archetype to test the prediction of the proposed model was suggested. Similarly, more research is needed in order to ascertain some of the following; -

- Are voices more noticeable than any other sound around irrespective of the fact that such sound contains speech or not.

- To what degree does the voice perception system permit us to excerpt identity and emotional information from the expression of other species such as cats, dog and so on.
- What is the perceptual emergent of voices?

Aronowitz [12] considering the rapid development of smart phones and mobile internet banking, there is no doubt the imperative need of a very strong authentication system. While the current use of password or pattern to unlock phones is not sufficient for the smart phones having considered the high threat of the device getting to the custody of an un-authorized user. Current developments in voice-based biometric system present a huge potential for a robust authentication of a mobile smart phone using voice. This aspect is very crucial for the financial and banking industry, where financial organizations are considering a flexible mobile customer services and easy authentication and at the same time maintaining security and most importantly, reducing fraudulent usage.

They presented the use of the current technology text-independent as well as text-dependent speaker authentication skills for user verification speaker specific digit strings, global digit strings, prompted digit strings as well as text independent were evaluated adapting the Joint Factor Analysis (JFA), Gaussian mixture models with nuisance attribute projection (GMM-NAP) and Hidden Markov Model (HMM) with NAP to improve security as well as reliability of the system.

Korshunov and Marcel [13] considered the fact that most biometric technology systems are vulnerable to spoofing which reduces their wide use sometimes hence they presented the need to develop anti-spoofing detection methods also referred to as presentation attack detection (PAD) systems. They presented an integration of PAD and Automatic Speaker Verification (ASV) systems using i-vector and local binary pattern histograms based ASV-PAD hybrid system. The AVspoof database which contains practical presentation attack was used to validate this method and result obtained show a significantly enhanced resistance of this hybrid system to attacks at the detriment of trivial degraded implementation for scenarios without spoofing attacks.

The use of cascading scheme for score synthesis and evaluating it with the joint ASV-PAD system was suggested for future research. In addition, the state of the art spoofing database can be employed to explore multi-model system considering ASV and PAD systems of different modalities for instance speech and image may be combined to enhance the performance in both 'licit' and 'spoof' scenarios.

Krawczyk and Jain [14] considering the large evolution from paper-based medical records towards electronic medical records, guaranteeing the security of such private and highly sensitive data cannot be over-emphasized since the health care giver only need to edit and update patient's record on the tablet, personal

computers (PC) or even smart phones, hence the need to protect the patient's privacy as required by all governmental regulations. A safe authentication system must be put in place anytime such records are to be accessed, hence the need for biometric-based access cannot be disputed. Research showed that on-line signature integrated with voice modalities is the most appropriate means for the users in such verification system since tablet PC is built with the associated devices. Evaluation of hybrid system of online signature and voice-based biometrics was carried out in this work. Dynamic programming method of string matching was incorporated for the signature verification while off-the-shelf commercial software development kit was used for voice verification. Fusion of these two methods were carried out at the matching score level after normalization of the data. The prototype of these hybrid systems was tested using a small truly multi-modal database of fifty users and an EER of 0.86% was reported.

Mazaira-Fernandez et al [15] came from the view of looking at the environment which has been so much improved by technology of social media whereby some other people uses these social media as a form of terrorism to transmit their message. In such a situation a typical biometrics recognition method such as face or fingerprints have been substituted by another biometric traits such as voice as this may be readily available in such scenario. They proposed a gender-dependent extended biometry factors (GDEB). The GDEB factors classify features extracted from voice source and tract factors and other pertinent features such as format data, having in mind that male and female voices show both acoustic-phonetic variations as well as physiological differences. The main idea was to improve classification rate in speaker recognition using few parameters.

Scheffer et al [16] worked on two of the challenges facing voice biometrics technology. These challenges include non-ideal recording conditions (these are often operational situation problems such as noise, echoes, voice channels and so on) and audio compression. Adapting the SRI International's innovations which arise from the Intelligence Advances Research Project Activity (IARPA) Biometric Exploration Science and Technology (BEST) and the Defence Advances Research Projects Agency (DARPA) Robust Automatic Transcription of speech (RATS) projects. The SRI's approach uses various features excerpted from speech which are demonstrated using latest machine learning algorithm. They recommended a general audio classification system that excerpts metadata facts using i-vector. The logistic regression and the cross entropy was used to combine the various features that are automatically extracted from the audio signal at the i-vector level.

Bhogal et al [17] research showed that voice is a more natural means of communication and verbal communication is quicker and more efficient than textual communication. Considering a virtual environment,

they evaluated the use of virtual universe (VU) residents also known as Avatars in online service employing audio biometrics. The method for example can include, spurring consumer applications with a demand for an utterance, managing the response to the demand and developing a voiceprint or voice summary of the speaker or participant. This voiceprint can be connected to an avatar and when an utterance is obtained, the avatar can be recognized by evaluating the utterance with the voiceprint. Such evaluation can be used to recognise avatars as they go from one online service to another and as such, voice biometrics can be used to verify an avatar for a particular activity. The main idea of their research was that they used voiceprint to approve an operation limited to an authorized user. In other word, they demonstrated the possibility of using biometric in internet based activities.

Zhang et al [18] came from the view of the inability of a system to recognize an individual from a non-frontal view. They presented the people in photo Albums (PIPA) corpus with very high differences in pose, clothing, image resolution, illumination and camera viewpoint. They proposed the pose-invariant Person Recognition (PIPER) algorithm to tackle the problem identified earlier. PIPER combines the signals of pose-let level individual identifiers trained by Deep Convolution Neural Network (DCNN) to reduce the differences in pose and then combine with a global recognizer and face recognizer. Result obtained showed that this algorithm out-perform one of the best face recognizer (Deep face)

Khitrov [19] at the speech technology centre reviewed the use of voice-based biometrics for data access and security. He identified some challenges and advantages of the system.

However, voice can be used with any other biometric technology system in order to achieve 100% accurate recognition and verification. Voice-based person identification system can be used in the following areas to safeguard security and expedient user verification. These areas include:-

1. Call centres and interactive voice response systems (IVRs) such as to get account balance, medical result etc. which are private and highly sensitive information in order to save time while delivering these services to the rightful owners once the system can recognize their voices.
2. Mobile voiced based verification system helps in securing our sensitive and private information stored on the smart devices.
3. For cloud computing and Bring Your Own devices (BYOD) applications, employees prefer the use of their personal smart devices to access organization resources. There is a dire need to secure the organizational information of these smart devices. Hence the use of voice signature presents a natural, appropriate and legally binding option to hand-written signatures.
4. Voice-based authentication system can be employed to secure and enhance security of our social

networking platforms such as Facebooks, LinkedIn and so on.

Vatsa et al [20] having researched into various biometric technology, they identified some serious problems affecting this technology and they categorise them into accuracy, computational speed, security, cost, real-time attacks and scalability. They also identified the various possible attacks on biometric technology and these include impersonation, coercive, replay attack, as well as the attack on feature extractor, template database, matcher and matching results to mention a few. In improving the performance of biometric technology they identified two major ways one can protect the biometric information from such attack. These include encryption as well as watermarking. They introduced a three level Redundant Discrete Wavelet Transform (RDWT) biometric watermarking algorithm to implant the voice biometric MFCC in a colour face image of the same person for enhanced robustness, security and accuracy. In biometric watermarking, a particular quantity of information also known as watermark is implanted into the initial cover image using a private key such that the constituents of the cover image are not altered. The result obtained was over 90% verification accuracy.

2.1 Biometric system architecture

Biometrics system for identification encompasses different stages as shown in Figure 3.

In recognition mode, the system executes a one-to-many comparison against a biometric database in an attempt to establish the identity of an unknown individual. The system will be successful in identifying the individual if the comparison of a biometric sample to a template in the database falls within a previously set threshold. This mode can be used for positive recognition (so that the user does not have to provide any information about the template to be used) or for negative recognition of the person "where the system establishes whether the person is who he or she (implicitly or explicitly) denies to be". The latter function can only be achieved through biometrics since other methods of personal recognition such as passwords, PINs or cards are ineffective [6].

When an individual uses a biometric system for the first time, it is referred to as enrolment. During this process, biometric data from the individual is captured and saved. In subsequent uses, biometric data is sensed and matched up to data saved at the time of enrolment. The storing and recovery of such systems must be critically protected if the biometric system is to be robust. In Figure 3, the data acquisition phase is made up of some sensors, which act as link between the real world and the system; it has to obtain the entire required data e.g. image or fingerprint. All necessary pre-processing of the raw audio signal in order to remove sensor noise and unwanted sections are carried out in this phase using some kind of standardization. The second stage is the feature extraction phase, where highly descriptive and discriminative features are extracted and represented. This stage is a very important stage of

a biometric system architecture. At the third stage, a vector of numbers or an image with specific characteristic is used to build a template. A template is a representation of most descriptive features. It is often created and then stored in a database for future use. During the enrolment phase, the template is either stored on a card, within a database or both. At the matching stage, the obtained input features are passed to a matcher that evaluates it with other current templates and estimate the distance between them using a metric (e.g. Hamming distance). Decision is based on the outcome of their evaluation such as granting access or denial from a restricted area [6].

The use of biometrics as a verification system consists of three phases. In the first phase, a sample of a biometric feature is captured, processed and stored in the database as a reference model for future use. In the second phase, some samples are compared with a reference model in order to see if the newly presented biometrics matches or

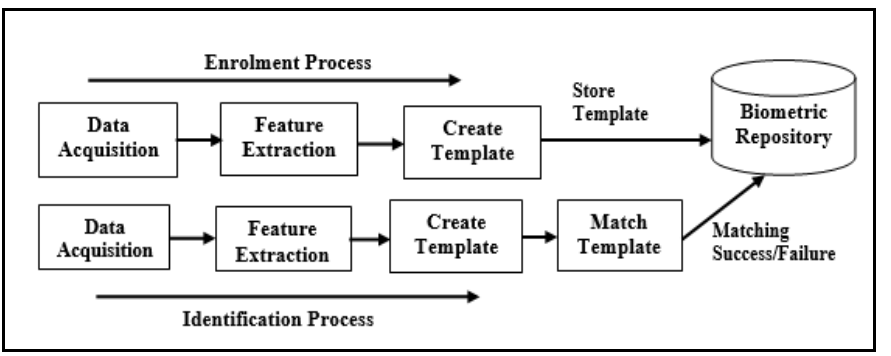

Figure 3: Biometrics system architecture, for identification process

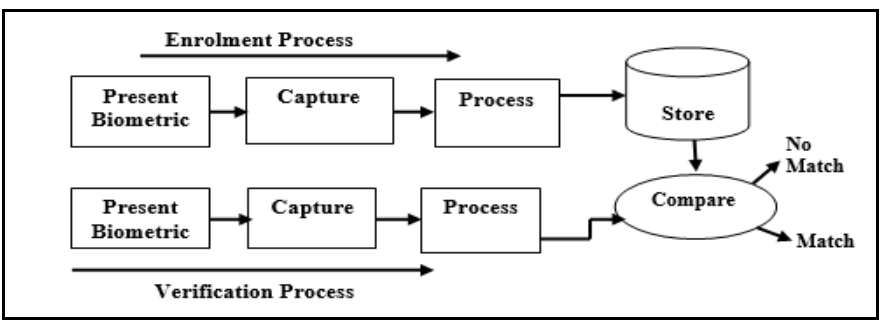

Figure 4: Biometrics System Architecture, for Verification Process

not, after which a threshold is computed. During the third phase, access is given if the feature matches or access is denied if the features do not match. Positive recognition is a common use of the verification mode,

"where the aim is to stop many people from registering into the system by using similar identity but separate biometric data each time they enrol" [21].

Figure 5: Word cloud of popular biometric signatures

Figure 5 shows the popularity of various biometrics used to develop person identification technologies. The popularity is determined by these key factors: universality, permanence, collectability, uniqueness, acceptability, reducibility, circumvention, privacy, performance and inimitability. Brief explanations of these factors are presented as follows:

i. Universality implies that every person must possess this characteristic,
ii. Permanence implies that the attributes should not change with time,
iii. Collectability requires that the properties must be suitable for capture and measured quantitatively without delay,
iv. Uniqueness means no two people should have the same characteristics,
v. Acceptability implies that the system must be accepted by the majority,
vi. Reducibility means the extracted data should be reduce-able to a file,
vii. Circumvention requires that the property should not be masked or manipulated or even fooled,
viii. Privacy, demands that capturing process should not defy the privacy of the individual,
ix. Performance requires that the identification accuracy should be very high, and
x. Inimitability means the attribute should be irreproducible by other means.

2.2 Commonly used algorithm in biometrics technology

i. Artificial Neural Networks (ANN): ANN are computational tools used for analysing many complex real world problems by using the biological neuronal networks based on the structure and functions of neurons. It is made up of layers of computing

neurons that are interconnected by some weighted lines that are capable of performing very large parallel computations for data processing. ANN has been used in various areas ranging from speech recognition to prediction as well as in medicine [22]–[24].

ii. Gaussian Mixture Model (GMM): The GMM is a parameter based probability density function, which is represented as a weighted sum of Gaussian element densities. It is majorly used for measuring continuous distribution features in biometrics as well as vocal tract associated spectral features in speaker recognition system. It can also be used in automatic laughter recognition system as well as hand geometry detection [25], [26].

iii. Support Vector Machine (SVM): SVM is an algorithm for solving a binary classification and regression analysis problem. This algorithm maximize margin by defining an optimal hyper-plane. It is used for non-linearly separable problems. It also maps data to higher component area where it easily classifies linear decision surfaces. Its uses ranges from handwritten digit recognition to face recognition to medicine [25], [27].

iv. Hidden Markov Model (HMM): HMM is a statistical tool used for demonstrating generative sequences described by a set of observable sequences. HMM is made up of two stochastic processes which include the invisible process of hidden states and a visible process of observable symbols. The hidden state makes up the Markov chain and the probability distribution of the observed symbol depend on the fundamental state. It is used for modelling time-varying spectral vector sequences [28]–[30].

v. Deep Neural Network (DNN): Deep learning is a machine learning algorithm that is being introduced to the area of large vocabulary speech recognition system. It has many hidden layers as compared to the artificial neural network (ANN) and it incorporates new method for training its data [30], [31].

The discussion on these algorithms pros and cons are presented in Table 1.

Table 1: Selected person identification algorithms pros and cons

Algorithm	Pros	Cons
i. ANN	It has been used to train very complex models. It is easy to conceptualize. It has lots of libraries for easy implementations. It can be used extensively	It requires more data for training Choosing the right parameters may be very complex. Final model may use up a lot of memory. Multi-layered neural networks are

	for academic research work.	usually harder to train and require tuning of lots of parameter [32].
ii. GMM	It is a lot more flexible with regards to cluster covariance. GMM allows mixed membership of points to cluster.	Learning may be a little difficult Few libraries for implementation [33].
iii. SVM	It reduces overfitting in the course of training. It manages high dimensional problems. It is used for linearly and non-linearly separable data. Its optimality is guaranteed due to its convex optimization. It can be implemented in various programming languages (Python, Matlab), and its libraries are easily accessible.	It can be time consuming to train. It may not be explanatory enough as it does not return a probabilistic confidence value. It does not support structured representations of text which gives high performance in Natural Language Processing (NLP) [34].
iv. HMM	It is highly flexible. It is suitable for better data fitting. It is highly statistical in nature. It can learn from raw data.	It has a huge numbers of unstructured parameters. It can only represent a small fraction of the entire distributions. Dependencies between layers may not be explicitly represented [35], [36].
v. DNN	It continuously grows data to improve its training. It creates its new features.	It requires very large amount of unlabeled data for training. It requires very high performance hardware for its implementation. It requires more time to train [37].

3. Timeline of Biometrics technology

The advancement of biometric recognition is still in progress, hence the need for continuous improvements in performance and usability.

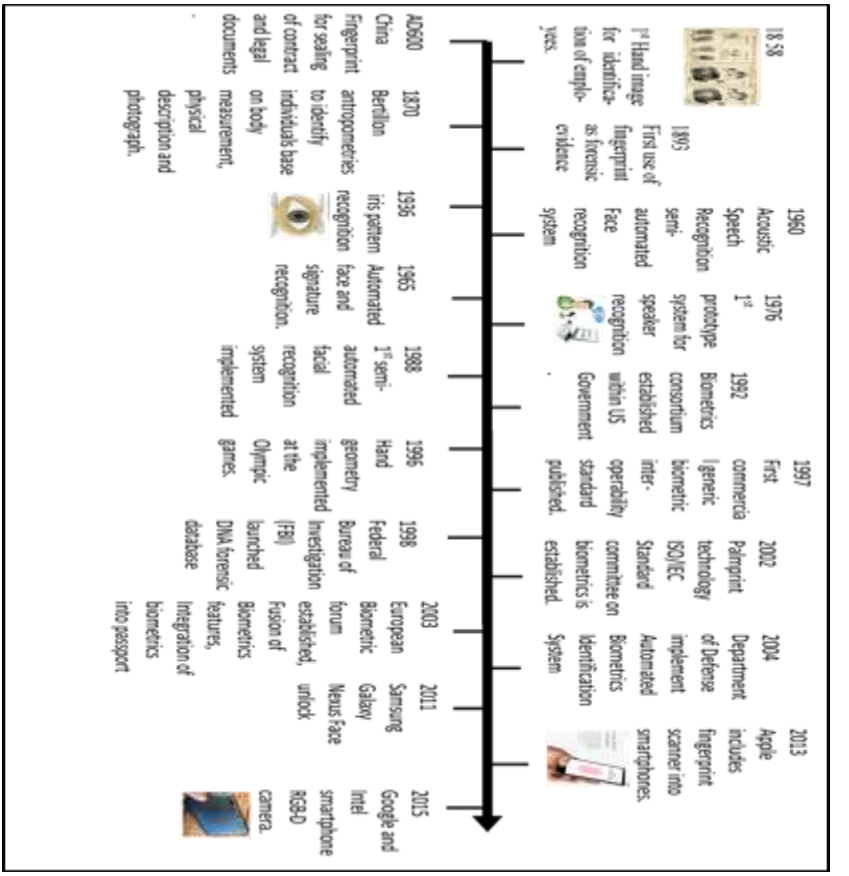

Figure 6: Biometrics timeline over the past one and a half centuries

3.1 Importance and application of biometrics technology timeline

The importance of biometrics technology cannot be over emphasized. There are various areas of applications of this technology around the world. Among the various applications are:

- As far back as 1792-1750BC, the kings of Babylon used the imprint of their right hand for authenticating the written codes of law printed on clay and also for business transactions [38], [39].
- 618-907 Chinese use both fingerprint and handprint on clay as authentication. They also employed palm and footprints to differentiate one child from another [39], [40].
- 1800, Thomas Bewick, an English naturalist, used the printing of his fingerprint to identify his published books [41].
- 1856, Sir William Herschel, a British officer who worked for the Indian civil service, commenced the use of thumbprint on documents as alternative for written signature mostly for illiterate and other people [39].

- 1892, Juan Vucetich, an Argentinean police researcher used fingerprint to identify a mother in the murder of her children, hence, Argentina was the first Country to substitute anthropometry with fingerprint [42], [43] .
- 1920s, fingerprint identification was employed by the law enforcement all over the world including the US military and the Federal Bureau of Investigation (FBI) as a form of identification [39].
- 1974, first commercial hand geometry system for access control, time, attendance and personal identification became available [40].
- 1975, FBI sponsors the development of devices and minutiae extracting expertise [44].
- 1996, Iris scanning was implemented in prison in USA [45]. More so, the hand geometry recognition system was implemented at the Olympic games where over 65,000 people were successfully enrolled [46].
- 1998, Nationwide Building Society in Britain introduced iris recognition within its cash dispensing machine in substitute of the Personal Identity Number (PIN)[45].
- 1999, National Bank United in USA introduces the iris recognition technology to three ATM outlets in Houston, Dallas and Ft Worth [45].
- Late 1990s and early 2000s biometrics was introduced for verification of electorates during election. As research in this area intensifies, different biometric traits were used to verify electorates during election. Somaliland was the first to use the iris biometrics voting system in the world [47]. Also some other countries throughout the world also implement biometrics technology for election purpose, among them are Kenya, Nigeria, Ghana, Angola which implements the fingerprint biometrics system to identify registered voters. So many other biometric technology have been used over the years all over the world[48].
- 2001, Face Recognition was used at the Super Bowl in Tampa Florida [49]. Also, Malaysia was one of the first countries to implement thumbprint on her national Identity card. Australia included chips for biometric identification on the international passports. In addition, Canada used biometric for anti-terrorism measure [39].
- 2002, the enhanced Border security and Visa Reform Act was put into law. This act include biometric data in the passports of VISA waiver program travelers like Belgium and other Union Member States [39].
- 2003, US Government National Science and Technology Council, inaugurated a subcommittee on biometrics to coordinate biometrics research and development, policy, outreach and international collaboration [50]. Also, in December 2003, the United Linkers Company was born in India. This company offers

biometric solutions in car security implementing various biometrics system ranges from voice to fingerprint [51].

- 2004, US implemented biometrics into immigration services and it was tagged- US-Visit. Also, US government called for compulsory government-wide personal identification card for all federal workers and contractors [50]. In addition, Europe adopted the use of biometric system for passports and travel documents using both facial images and fingerprints [39].
- 2005, OMRON Corporation show case the world's first face recognition technology at the security show tagged Japan 2005. This technology can be deployed on Personal Digital Assistants (PDAs), mobile phones and other mobile devices with camera function [52]. Moreso, seven countries in Europe consisting of Austria, Belgium, The Netherlands, France, Germany, Luxembourg and Spain signed an agreement on the improvement of cross border information exchange and data comparison using DNA profiles, fingerprints and vehicle registration data [39].
- 2008, US Government commenced the organization of the use of biometric database [40]. Also [53] presented an experimental result on fusion of iris and palm print for multi-biometric system. The result obtained showed a very great improvement on a single biometric system. In addition, FBI set up a 1-billion dollar database which consist of fingerprint, iris, facial images and DNA samples [39].
- 2010, US national security apparatus employs biometrics for terrorist identification [40]. In addition, SBTel and devices commenced the selling of biometrics devices for attendance and access control solutions in Nigeria [50].
- 2011, Biometrics identification was employed in the identification of Osama Bin Laden. Also. SBTel, a biometric device vendor, launched web based biometrics attendance for a cloud-based attendance system [50].
- 2013, Apple included fingerprint scanner into smartphones [40].
- In January 2017, Australia's Department of Immigration and Border Protection announced strategy to execute a novel system at the country's international airport by 2020. This biometric system makes use of face, iris and/or fingerprint recognition system. The biometric system is intended to replace the existing paper identity card passports [54]. In addition, in January of the same year, President Donald Trump ordered that all non-citizens be subjected to biometric checks while entering or leaving the United States [54]. Moreover, in August 2017, First-Biometrics announces the use of next biometrics flexible fingerprint sensor for smart payment and identity card [54].

3.2 Evolution of voice-based recognition system

The initial objective of speech recognition was to develop a system which imitate a person's speech communication ability.

- 1773 – The Russian scientist Christian Kratzenstein, a professor of physiology in Copenhagen created vowel sound with the use of resonance tubes which was attached to the organ pipes [55].
- 1791 – Wolfgang Von Kempelen in Vienna developed the acoustic mechanical speech machine [55].
- Mid 1800s – Charles Wheatstone constructed Von Kempelen's speaking machine employing resonators from leather such that different speech-like sounds could be produced by using hand to change the configurations [55].
- 1879 – Thomas Edison discovered the first dictation machine [56].
- 1930s – Homer Dudley developed a speech synthesizer called Voice Operating Demonstrator (VODER) which was an electrical prototype of Wheatstone's work [55].
- 1952 – Davis, Biddulph and Balashek of Bell laboratories developed Audrey a system for segregated digit recognition for a particular talker [55].
- 1950s – Oslon and Belar of RCA laboratories constructed a ten (10) syllable recognizer for a particular talker[55].
- 1959 – Fry and Denes of University College in England constructed a phoneme recognizer which identifies four (4) vowels and nine (9) consonants [55].
- 1960s – Atal and Itakura invented the basic theory of Linear Predictive Coding (LPD) which abridged the approximation of the vocal tract reponse from speech waveform [55].
- 1962 – IBM shoebox was developed and it can understand 16 English words [56].
- 1971 – Alexander Waibel, developed the Harpy machine at Carnegie Mellon University. This machine can understand 1,011 words with some phrases [56].
- 1970s – Tom Martins started the leading speech recognition company named Threshold Technology incorporation building the first Automatic Speech Recognition (ASR) product known as VIP-100 system. Also, Lenny Baum of Princeton University discovered a mathematical speech recognizer called the Hidden Markov Model (HMM). In addition, the use of pattern recognition technology to speech identification system using LPC was introduced [55].
- 1980s – The use of Artificial Neural Network was introduced for speech hence; speech recognition tends towards prediction.
- 1986 – IBM Targora was developed employing the HMM to predict the next phonemes in

speech. This system was the world's fastest typist at the time as it was able to recognize up to 20,000 English words and a number of sentences [56].

- 1990s – The emergence of Automatic Speech Recognition (ASR) [55].
- 1997 – The world's leading uninterrupted speech recognizer was announced as a Dragon's naturally speaking software. This software was able to understand 100 words per minute without any break in between the words [55].
- 2006 – The National Security Agency (NSA) commenced the use of speech recognition to identify key terms in recorded speech [56].
- 2008 – Google introduces a voice exploration app on mobile devices [56].
- 2011 – Apple launches Siri, for voice based digital assistant [56].

3.3 Benefits and challenges of biometrics systems

Despite this progress, a number of challenges continue to restrain the full potential of biometrics to automatically recognize humans. Some specific challenges, in terms of biometrics groups, are presented in Table 2 [6].

The need for a reliable system for person identification cannot be over-emphasized, while the state-of-the-art biometric identification systems are taking over access control and many security establishments. These systems have been in existence for over five decades, though it started gaining attention in recent time. Although, there are many challenges that are needed to be addressed in order for these systems to be robust and more accurate (Table 2). A major challenge of these systems is spoofing attack. It is a condition whereby a person effectively impersonates another person. This is often done in order to gain an illegal access to a facility or system.

3.4 Areas of implementation of Voice-based recognition system

1. For attendance system
2. In mobile phone for auto texting
3. Forensic
4. Banking and financial institutions
5. Using one's indigenous language for recognition purpose.
6. In medicine.
7. Assisting the aged in text composition, bank transaction and many more
8. According to Boyd [56], "cloud-based computers have entered millions of homes and can be controlled by voice, even offering conversational responses to a wide range of queries".
9. Voice activated home speakers.
10. Purchases can be made over the internet using voice enabled machines.

According to Akhtar et al [57], the US National Institute of Standards and Technology enumerate the susceptibility of biometrics to spoofing in their national vulnerability database. Most of the existing biometrics systems are susceptible to this attack. For instance, the iris, finger print and face images taken from impersonators were identical with that of the real users. In order to

cub this attack, there is a need to introduce a liveness detection into biometrics systems. This is an area of future research. In addition, the use of different biometric systems combination that is a multi-modal biometric technology has the potential to reduce this attack [58-60]. Thus, more studies are required in this direction. Finally, biometric trait and databases are needed in order to validate new biometric systems such as the use of tongue, ear, sneeze, cough, laughter and so on.

Table 2: Benefits and challenges of Biometrics systems

Biometrics systems	Benefits	Challenges
Hand based: Fingerprint or finger scan Hand geometry Palm print	Easy to use and non-intrusive. High accuracy and long term stability. Ability to enrol various fingers. Mature technology	Fingerprint of those working in chemical production companies, or farming is frequently affected. Affected by skin condition. Affected by a few cuts, blister or wound either short term or permanent.
Face/Eye based Facial recognition Retina scan Iris scan Facial thermograph	Ability to operate secretly. Non-intrusive. Not affected by environmental condition.	. Twins may have similar face features The eyes of people with diabetes may get affected hence resulting in differences. It may change as the person gets older
Behavioural characteristics. Voice recognition Signature recognition Key stroke scan Speaker recognition. Gait recognition	It uses current telephony setup. It is easy to use and non-intrusive. The voice-based biometrics further enjoys the following benefits: - It does not need to be in contact with the individual in order to be captured, hence it can be taken remotely. It can be used while driving, from another room, through mobile device, while a person is busy with some other activity and so on. It is simple and easy to use Recorded voices cannot be used to spoof this system.	Signature is affected by inconsistency in signing which makes it difficult to use. The voice is affected by cold. Environmental noise. It may change as the person gets older. Channel effects such as inference and distortion, frequency response, channel encoding. Presentation effects such as speech sample duration, psycho physiological state of the speaker (such as illness, emotions), effects of vocal strain and so on.

4. Conclusions

Even though there are some challenges facing biometrics technology, it has tremendous usage in this information technology age. Choice of biometric type depends on characteristics measurement and user requirement. Other factors which affect the choice of a biometric based are sensor and device availability, computational time and reliability, cost, sensor size and power consumption. In addition, cultural disposition is also a factor that affects biometric technology selection.

Currently, due to divers' nature of biometric applications, no single biometric trait is likely to be ideal and satisfy all the requirements of all applications, hence the need to use multiple biometric features in order to ensure very high reliability is essential. Investigating new biometric features for identification is an area where more research is needed to be carried out.

Having considered the advancement in technology, moving from the use of paper to electronics and online base medical records to banking and financial industry as well as social media. The huge and sensitive data involved must be adequately protected from being hacked into. Voice-based recognition system has however, been considered as a very appropriate biometrics for this task having considered the speed, efficiency and customer relation characteristics it possesses.

According to [56], "Voice is the future. The world's technology giants are clamouring for vital market share, with ComScore agitating that 50% of all searches will be voice searches come 2020."

Acknowledgements
The authors wish to acknowledge Ighravwe D. E. for taking his time to peer review this work.

References

[1] A. K. Jain, K. Nandakumar, and A. Ross, "50 years of biometric research: Accomplishments , challenges , and opportunities," vol. 79, pp. 80–105, 2016.

[2] T. Reuters, "Global directory," 2017. [Online]. Available: https://blogs.thomsonreuters.com /answerson/biometrics-technology-convenience-data-privacy/. [Accessed: 03-Oct-2018].

[3] D. Carlson, "Biometrics - Your Body as a Key," 2014. [Online]. Available: http://www.dynotech.com/articles/biometrics.shtml. [Accessed: 04-May-2017].

[4] K. Delac and M. Grgic, "A survey of biometric recognition methods," 46th Int. Symp. Electron. Mar., 2004.

[5] A. M. Bojamma, B. Nithya, and C. N. Prasad, "An Overview of Biometric System," Int. J. Comput. Sci. Eng. Inf. Technol. Res. ISSN 2249-6831, vol. 3, no. 2, pp. 153–160, 2013.

[6] R. B. Jadhao and A. Bakshi, "An Overview of Biometric System," Int. J. Sci. Res. Educ., vol. 3, no. 7, 2015.

[7] B. P. Salil and W. L. Damon, "Biometric authentication and identification using keystroke dynamics: A survey," J. Pattern Recognit. Res., vol. 7, no. 1, pp. 116–139, 2012.

[8] J. Kaur and S. Kaur, "A Brief Review: Voice Biometric For Speaker Verification in Attendance Systems," Imp. J. Interdiscip. Res., vol. 2, no. 10, 2016.

[9] D. Tran, M. Wagner, Y. W. Lau, and M. Gen, "Fuzzy methods for voice-based person authentication," IEEJ Trans. Electron. Inf. Syst., vol. 124, no. 10, pp. 1958–1963, 2004.

[10] J.-L. Dugelay, J.-C. Junqua, C. Kotropoulos, R. Kuhn, F. Perronnin, and I. Pitas, "Recent advances in biometric person authentication," in International Conference on Acoustics, Speech, and Signal Processing (ICASSP), IEEE, 2002, vol. 4, p. 4060-4063.

[11] P. Belin, S. Fecteau, and C. Bedard, "Thinking the voice: neural correlates of voice perception," Trends Cogn. Sci., vol. 8, no. 3, pp. 129–135, 2004.

[12] H. Aronowitz, R. Hoory, J. Pelecanos, and D. Nahamoo, "New developments in voice biometrics for user authentication," in Twelfth Annual Conference of the

International Speech Communication Association, 2011.

[13] P. Korshunov and S. Marcel, "Joint operation of voice biometrics and presentation attack detection," 8th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS), ieeexplore.ieee.org, 2016.

[14] S. Krawczyk and A. K. Jain, "Securing electronic medical records using biometric authentication," in International Conference on Audio-and Video-Based Biometric Person Authentication, 2005, pp. 1110–1119.

[15] L. M. Mazaira-Fernandez, A. Álvarez-Marquina, and P. Gómez-Vilda, "Improving Speaker Recognition by Biometric Voice Deconstruction," Front. Bioeng. Biotechnol., vol. 3, Sep. 2015.

[16] N. Scheffer, L. Ferrer, A. Lawson, Y. Lei, and M. McLaren, "Recent developments in voice biometrics: Robustness and high accuracy," in International Conference onTechnologies for Homeland Security (HST), IEEE, 2013, pp. 447–452.

[17] K. S. Bhogal, A. Rick, D. Kanevsky, and Pickover C A, "Using voice biometrics across virtual environments in association with an avatar's movements," US Patent, Google Patents, 2012.

[18] N. Zhang, M. Paluri, Y. Taigman, R. Fergus, and L. Bourdev, "Beyond frontal faces: Improving person recognition using multiple cues," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 4804–4813.

[19] M. Khitrov, "Talking passwords: voice biometrics for data access and security," Biometric Technol. Today, vol. 2, pp. 9–11, 2013.

[20] M. Vatsa, R. Singh, and A. Noore, "Feature based RDWT watermarking for multimodal biometric system," Image Vis. Comput., vol. 27, no. 3, pp. 293–304, 2009.

[21] P. Ghosh and R. Dutta, "A new approach towards biometric authentication system in palm vein domain," Int. J. Adv. Innov. IJAITI, vol. 1, no. 2, pp. 1–10, 2012.

[22] O. N. A. Al-Allaf, "Review of face detection systems based artificial neural networks algorithms," arXiv Prepr. arXiv1404.1292, pp. 1404–1292, 2014.

[23] F. Amato et al., "Artificial neural networks in medical diagnosis," J. Appl. Biomed., vol. 11, pp. 47–58, 2013.

[24] A. S. George, E. Roy, A. Antony, and M. Job, "An Efficient Gait Recognition System for Human Identification using Neural Networks," Int. J. Innov. Adv. Comput. Sci., vol. 6, no. 5, 2017.

[25] J. R. Pinto, J. S. Cardoso, A. Lourenço, and C. Carreiras, "Towards a Continuous Biometric System Based on ECG Signals Acquired on the Steering

Wheel," Sensors, vol. 17, p. 2228, 2017.

[26] J. S. Edwards, "Using Gaussian Mixture Model and Partial Least Squares regression classifiers for robust speaker verification with various enhancement methods," Rowan University, Rowan Digital Works, 2017.

[27] C. D. Smyser et al., "Prediction of brain maturity in infants using machine-learning algorithms," Neuroimage, vol. 136, pp. 1–9, 2016.

[28] U. Jerome, "Acoustic Laughter Processing," TCTS Lab. Publ. Univ. Mons., 2014.

[29] Ç. Hüseyin, U. Jérôme, and D. Thierry, "Synchronization rules for HMM-based audio-visual laughter synthesis," in IEEE International Conference on Acoustics Speech and Signal Processing ICASSP, 2015.

[30] A. Hannun et al., "Deep speech: Scaling up end-to-end speech recognition," arXiv Prepr. arXiv1412.5567, 2014.

[31] F. Lingenfelser, J. Wagner, J. Deng, R. Bruckner, B. Schuller, and E. Andre, "Asynchronous and Event-based Fusion Systems for Affect Recognition on Naturalistic Data in Comparison to Conventional Approaches," IEEE Trans. Affect. Comput., 2016.

[32] L. Matthew, "Deep Learning," 2014. [Online]. Available: https://www.quora.com/What-are-the-pros-and-cons-of-neural-networks-from-a-practical-perspective-Personal-comments-from-heavy-users-welcome. [Accessed: 03-Jul-2018].

[33] O. Wibisono, "Deep Learning," 2016. [Online]. Available: https://www.quora.com/What-are-the-advantages-to-using-a-Gaussian-Mixture-Model-clustering-algorithm. [Accessed: 03-Jul-2018].

[34] A. Ng, "Deep Learning," 2016. [Online]. Available: https://www.quora.com/What-are-some-pros-and-cons-of-Support-Vector-Machines. [Accessed: 03-Jul-2018].

[35] P. J. Rani, "Advantages and disadvantages of hidden markov model.," 2016. [Online]. Available: https://www.slideshare.net/joshiblog/advantages-and-disadvantages-of-hidden-markov-model. [Accessed: 04-Jul-2018].

[36] K. Tuzcuoglu, "Deep Learning," 2018. [Online]. Available: https://www.quora.com/What-are-the-advantages-and-disadvantages-of-Hidden-Markov-Models-in-forecasting-values-of-a-time-series-compared-to-other-methods-e-g-ARIMA-Can-we-say-that-HMM-is-a-Non-Statistical-learning-methods. [Accessed: 03-Jul-2018].

[37] O. Maslovska, "Deep Learning: Definition, Benefits, and Challenges," 2017. [Online]. Available: https://stfalcon.com/en/blog/post/deep-learning-benefits-and-challenges. [Accessed: 03-Jul-2018].

[38] J. Ashbourn, "The social implications of the wide scale implementation of biometric and related technologies," Backgr. Pap. Inst. Prospect. Technol. Stud. DG Jt. Res. Centre, Eur. Comm., 2005.

[39] E. J. Kindt, "Privacy and data protection issues of biometric applications," Springer, 2016.

[40] S. Mayhew, "History of Biometrics," 2016. [Online]. Available: http://www.biometricupdate.com/201501/history-of-biometrics. [Accessed: 09-May-2017].

[41] K. Fauci, "Prezi Presentations," 2014. [Online]. Available: https://prezi.com/5nzrgdiroqzn/1800s-thomas-bewick-an-english-naturalist-used-engraving/.

[42] USA-Govt, "Visible Proofs, U.S. National Library of Medicine," vol. 2017. U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894, 2014.

[43] D. A. Glaeser, "The Fingerprint, in the Service of the SWISS Confederation." Federal Office of Police fedpol, 2013.

[44] S. M. Martinez, "THe FBI-Federal Bureau of Investigation," 2013. [Online]. Available: https://archives.fbi.gov/archives/news/testimony/overview-of-fbi-biometrics-efforts.

[45] K. Chadwick, J. Good, G. Kerr, F. McGee, and F. O'Mahony, "Biometric Authentication For Network Access And Network Applications," 2001. [Online]. Available: http://ntrg.cs.tcd.ie/undergrad/4ba2.02/biometrics/now.html.

[46] N. Duta, "A survey of biometric technology based on hand shape," Pattern Recognit., vol. 42, no. 11, pp. 2797–2806, 2009.

[47] TeleSur, "Somaliland: 1st in World to Use Iris Scanner Technology to Stem Voter Fraud," 2017. [Online]. Available: https://www.telesurtv.net/english/news/Somaliland-1st-in-World-to-Use-Iris-Scanner-Technology-to-Stem-Voter-Fraud-20171114-0006.html.

[48] P. Wolf, A. Alim, B. Kasaro, P. Namugera, M. Saneem, and T. Zorig, "Introducing Biometric Technology in Elections," Voter Information, Commun. Educ. Netw. Glob. Knowl. Netw. Voter Educ., 2017.

[49] L. K. Nadeau, "Tracing the History of Biometrics," 2012. [Online]. Available: http://www.govtech.com/Tracing-the-History-of-Biometrics.html.

[50] Solutions-Broadway-Digest, "History of Biometrcs," 2017. [Online]. Available: https://www.sbtelecoms.com/history-of-biometrics-conclusion/.

[51] United-Linkers-Biometric-and-Robotic-Solutions, "World's First Face Recognition Biometric for Mobile Phones," 2017. [Online]. Available: http://automobile-security.com.

[52] World's-First-Face-Recognition-Biometric-for-Mobile-Phones, "World's First Face Recognition Biometric for Mobile Phones," 2005. [Online]. Available:

https://phys.org/news/2005-03-world-recognition-biometric-mobile.html.

[53] V. C. Subbarayudu and M. V. N. K. Prasad, "Multimodal biometric system," in Emerging Trends in Engineering and Technology, 2008. ICETET'08. First International Conference on, 2008, pp. 635–640.

[54] J. Lee, "Biometric Update.com," 2016. [Online]. Available: http://www.biometricupdate.com/201701/biometrics-to-replace-passports-at-australian-airports.

[55] B.-H. Juang and L. R. Rabiner, "Automatic speech recognition–a brief history of the technology development," Georg. Inst. Technol. Atlanta Rutgers Univ. Univ. California. St. Barbar., vol. 1, p. 67, 2005.

[56] C. Boyd, "The startup," 2018. [Online]. Available: https://medium.com/swlh/the-past-present-and-future-of-speech-recognition-technology-cf13c179aaf. [Accessed: 03-Oct-2018].

[57] Z. Akhtar, C. Micheloni, and G. L. Foresti, "Biometric liveness detection: Challenges and research opportunities," IEEE Secur. Priv., vol. 13, no. 5, pp. 63–72, 2015.

[58] K. O. Okokpujie, S. N. John, E. Noma-Osaghae, C. Ndujiuba, I. P. Okokpujie. AN ENHANCED VOTERS REGISTRATION AND AUTHENTICATION APPLICATION USING IRIS RECOGNITION TECHNOLOGY. International Journal of Civil Engineering and Technology (IJCIET). 2019 Feb 28;10(2):57-68.

[59] K. Okokpujie, E. Noma-Osaghae, O. Okesola, O. Omoruyi, C. Okereke, S. John, I. P. Okokpujie. Fingerprint Biometric Authentication Based Point of Sale Terminal. In International Conference on Information Science and Applications 2018 Jun 25 (pp. 229-237). Springer, Singapore.

[60] K. Okokpujie, E. Noma-Osaghae, O. Okesola, O. Omoruyi, C. Okereke, S. John, I. P. Okokpujie. Integration of Iris Biometrics in Automated Teller Machines for Enhanced User Authentication. In International Conference on Information Science and Applications 2018 Jun 25 (pp. 219-228). Springer, Singapore.