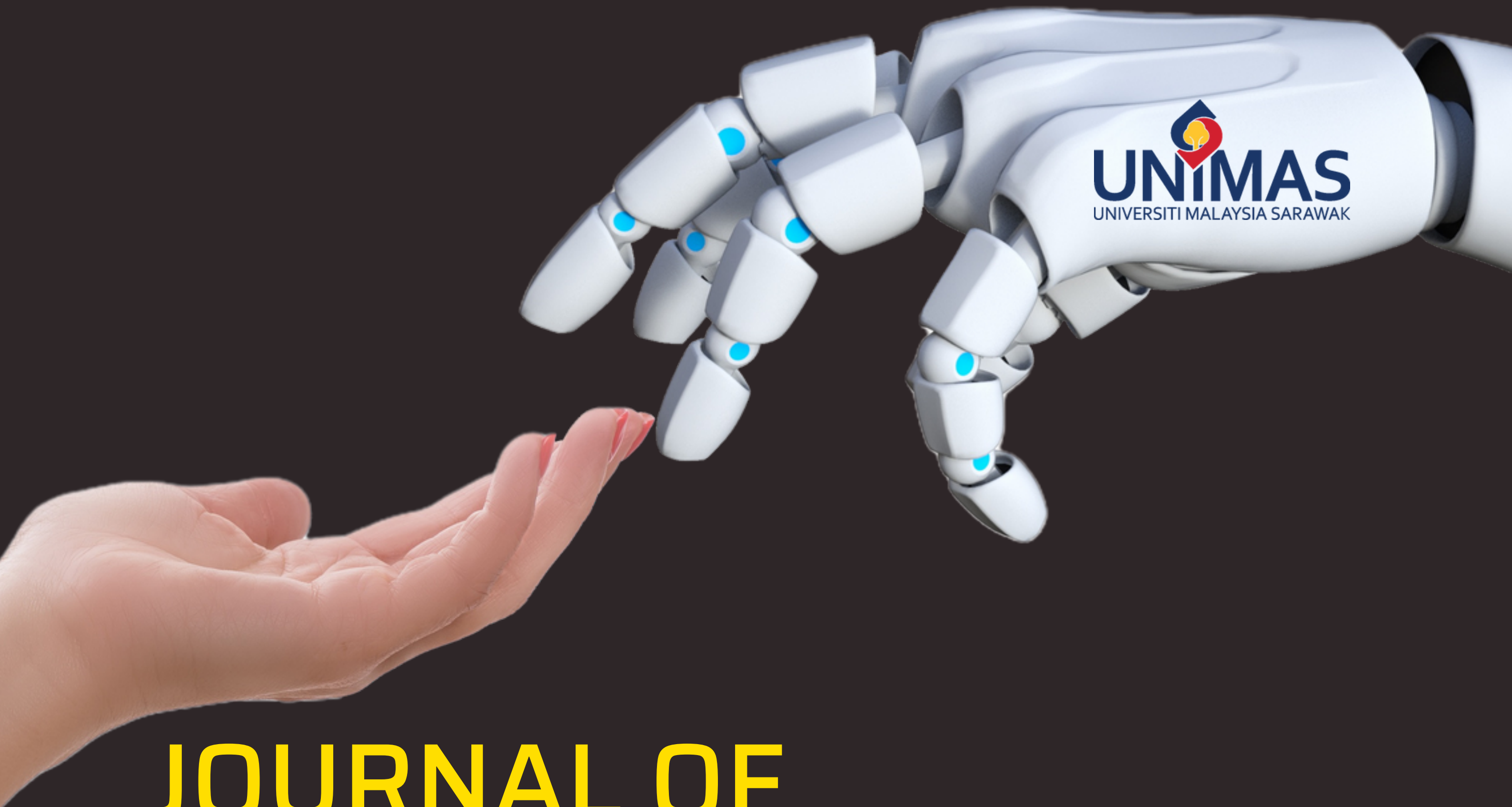

JOURNAL OF COMPUTING AND SOCIAL INFORMATICS

FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

UNIVERSITI MALAYSIA SARAWAK

eISSN 2821-3777

Journal of Computing and Social Informatics

The Journal of Computing and Social Informatics (JCSI) is an international peer-reviewed publication that focuses on the emerging areas of Computer Science and the overarching impact of technologies on all aspects of our life at societal level. This journal serves as a platform to promote the exchange of ideas with researchers around the world.

Articles can be submitted via *www.jcsi.unimas.my*

Assoc Prof Dr Chiew Kang Leng

Chief Editor
Journal of Computing and Social Informatics
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan
Sarawak, Malaysia

Contents

A Survey of Wireless Cellular Network Generations and Their Channel Access Techniques

[1*]Philip Achimugu, [2]Zakka Augustine, [3]Monday Jubrin Abdullahi, [4]Olalekan Lanihun, [5]Rume Elizabeth Yoro, [6]Abubakar Aliyu Abba and [7]Ahmed Hamisu

[1]Department of Computer Science, Air Force Institute of Technology Kaduna, 800283, Nigeria
[2]Department of Telecommunications Engineering, Air Force Institute of Technology Kaduna, 800283, Nigeria
[3]School of Technology, Woxsen University Hyderabad, 502345, India
[4]Departtment of Computer Science, College of Engineering, Environment and Science, Coventry University, CV1 2JH, United Kingdom
[5]Department of Cyber Security, Dennis Osadebay University, Asaba, 320001, Nigeria
[6]Elmore Family School of Electrical and Computer Engineering, Purdue University, West Lafayette Campus, Indiana, 47907, USA
[7]Department of Aerospace Engineering, Air Force Institute of Technology Kaduna, 800283, Nigeria

[1]*check4philo@gmail.com [2]zax.austen@yahoo.com, [3]abdullahi.monday@woxsen.edu.in, [4]ad9193@coventry.ac.uk, [5]felizabeth.yoro@dou.edu.ng, [6]aabba@purdue.edu, [7]ahmedhamisu@yahoo.com

Corresponding author

Received: 24 January 2025 | Accepted: 29 September 2025 | Early access: 05 November 2025
https://doi.org/10.33736/jcsi.8835.2026

Abstract - *The world of Wireless Cellular Communication (WCC) has achieved several top-notch transformations. This is due to the increasing innovations in Channel Access (CA) techniques. CA can be referred to as the methods of injecting life into a wireless network for the activation of varying air interfaces. It can also be defined as a technique used to ensure the modulated signal is properly fit-in a required communication channel for effective transmission. This transformation is recorded in every upgrade of air-interfaces in WCC generations. Air-interfaces mean variations in communication systems, which is consistent with frequency of operation ranges. WCC is divided into different successive generational shifts, which vary in frequencies of operation, in a successive ten-year circle. Starting in the 1970s, it has so far metamorphosed from Zero Generation (0G) through Fifth Generation (5G), with each air-interface frequency being activated by its varying CA techniques. In trying to meet the changing customer's demands and equipment advancements, which are the increasing operational bandwidth and mitigation of loss of transmission information, new and advanced air interfaces have now been assigned frequencies of deployment. This is possible as a result of constant innovations and upgrades in CAs, which serve as the reagents in the modification of modern wireless signal waveforms. This research conducted a survey on the existing commercial air-interface generations of WCC systems so far achieved/deployed relative to its CAs. Trusting that this will ignite researchers to conceive novel proposals on highly scalable CA techniques that will drive up-coming versions of air-interface frequencies of WCC networks.*

Keywords: WCC, Air-Interfaces, CA, Network Life, 0G, 5G.

1 Introduction

Cellular generations differ in four main aspects: radio systems, data rates, bandwidth, and switching schemes (Kumar, 2015). Precisely, Wireless Cellular Communication (WCC) started in the early 1970s but was made commercially available in the early 80s (Bliley's Technology, 2017). Consequently, in just four decades (in a ten-year circle rule), WCC technology has evolved from Zero Generation (0G) to Fifth Generation (5G). Its first mobile network was fully achieved by the activation of First Generation (1G) frequencies. Mobile frequency activation occurs as a result of the constant innovations and continuous advancement in Channel Access (CA) techniques, in conjunction with the required bandwidth and frequency of operations (Lu & Zheng, 2020). This is seen in the progressive performance of the wireless communication world. This transformation transits from 1G to the Second and a Half Generation (2.5G), to the Third Generation (3G) through 5G. 5G offers many advanced network resources that equipment users never experienced in the past, while Sixth Generation (6G) is at the developing stage with a projection of global coverage using technologies that could enhance the integration of satellite-terrestrial communication.

The 6G generation is likely to benefit from AI emerging technologies. Generative AIs, based on machine learning/deep learning models that generate new data from large datasets, can also have a significant impact on 6G wireless technologies research. For example, they could be used to predict new points of communication links, and find the fastest route for data packet transfer in the internet network. Other AI technologies, such as evolutionary algorithms that use neural network controllers, can be hybridised with generative models to predict optimal communication channels and network routes. For instance, the generative model can be used to generate new data from existing data where the dataset are not large enough for the training of deep neural network whose weights, parameters, or topologies/architectures are synthesized by evolutionary algorithms which are then used to make improved predictions.

These aforementioned generations of wireless communication are all geared towards meeting users' demands, equipment requirements, and advancements in wireless network services. This is obtained by overcoming the limitations of previous generations in areas of data rate, bandwidth utilization, effective power allocation, interference mitigation, and more.

To achieve the required network efficiency, various CA methods were proposed by scholars but the Multiple Access (MA) techniques in the frequency, time, or code domain were more widely accepted and commercially deployed (Liu et al., 2024). As illustrated in Figure 1, the Frequency Division Multiple Access (FDMA) technique is used to drive 1G objectives, the Time Division Multiple Access (TDMA) technique is deployed in 2G and 2.5G, and the Code Division Multiple Access (CDMA) is used to run the 3G and part of 4G. Advanced research to overcome the drawbacks noticed in CDMA conceived the development of the Orthogonal Frequency Division Multiplexing (OFDM) techniques that successfully made the 4G mandate fully realizable. A classical advancement in OFDM technique known as the Cyclic Prefix-OFDM (CP-OFDM) also enabled the Long-Term Evolution-Advanced (LTE-A) (Cai et al., 2018). The FDMA, TDMA, and CDMA are based on a serial kind of transmission, which has the disadvantage of greater loss of signal strength and information caused by self-jamming, adjacently placed symbols aliasing, shadowing, multipath fading, interference, and many more signal distortions (Zahra et al., 2021). The OFDM CA, on the other hand, is also noticed to be subdued by strings of signal anomalies, which results in transmission inefficiency (Zou et al., 2021). It is observed that in OFDM, the radio wave resources are not effectively used because of its inherent exhibition of high Peak to Average Power Ratio (PAPR) and Out-of-Bound Emission (OoBE). Its synchronous mode of transmission also impacted negatively on its transmitted resources in achieving the required spectrum slicing for higher air-interfaces.

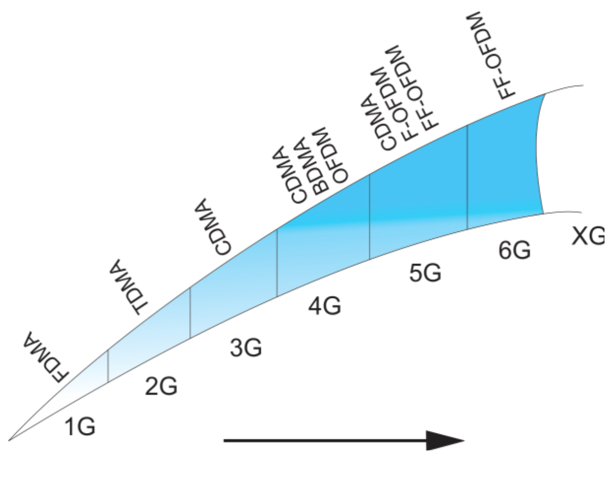

Figure 1: Wireless network air-interfaces relative to each channel access technique

In order to relax synchronization and achieve the required asynchronous transmission to meet the present-day proliferations of WCC devices and users' services, special cases of OFDM CA techniques were developed (Zhang et al., 2017). One of which is the Orthogonal Frequency Division Multiple Access (OFDMA). The OFDMA-CA is orthogonally engineered to allocate subsets of subcarriers that are aliases to each other for the modulating of the parallel data samples. This helps in overcoming the demerit experienced in earlier MA techniques (Arrano & Azurdia-Meza, 2016; Ramadhan, 2019). These drawbacks of OFDM led to the development of more successive and modified versions of MA in the classes of Orthogonal Multiple Access (OMA) and Non-Orthogonal Multiple Access (NOMA) transmissions (Liu et al., 2024). This is to enable the expectations of advanced air-interfaces in the areas of peak data rates in ultra-high reliability, data rates in gigabits and terabits per second, extremely low latency, and massive connectivity (Cai et al., 2018).

Some of the OMA CA techniques are Windowed OFDM (W-OFDM), Filter Bank Multi-Carrier (FBMC), Universal Filtered Multi-Carrier (UFMC), Generalized Frequency Division Multiplexing (GFDM), Filtered-OFDM (F-OFDM), and Feedback Filtered OFDM (FF-OFDM), and more (Sahrab & Yaseen, 2021; Augustine et al., 2021). These OMA schemes can only support limited numbers of users due to drawbacks in the numbers of orthogonal Resource Blocks (RB), which limits the SE and the capacity of the network (Cai et al., 2018). In order to support a massive number of applications and dramatically different classes of users in 5G networks, various new NOMA CA schemes were developed. Some of these emerging NOMA techniques consists of Power Domain (PD), Pattern Division Multiple Access (PDMA), Lattice Partition Multiple Access (LPMA), Building-block sparce constellation based Orthogonal Multiple Access (BOMA), Sparce Code Multiple Access (SCMA), Low-Density Spread CDMA (LDS-CDMA), and their variants (Cai et al., 2018; Liu et al., 2024). Unlike the OMA, the NOMA facilitate users to utilize the same spectrum with advanced transceiver configuration to manage interference in multiple access. While the OMA techniques are driven by Fourier and spatial based techniques, the NOMA is considered in both power, code, and spatial domain techniques (Liu et al., 2024).

Depending on the requirement, these classes of CA schemes make it possible for the message signal to be suppressed or superimposed and modulated in an EMW radio domain with negligible loss in energy and information. This advancement in the innovations of CA technique has proven its uniqueness of capability in the area of accentuating useful signals from interference and many more signal distortions suffered by earlier generations of wireless mobile technologies. The motivations behind this survey are to:

a) exploit on the successes and drawbacks of existing generations of WCC air-interfaces relative to their CA schemes;
b) explore on the objectives of the present generation of air-interfaces relative to their CA schemes;
c) highlight on the need for more advanced and sophisticated access schemes with high scalability that could overcome the Red Trend Curve (RTD) currently experienced in the 5G air-interface in order to drive the upcoming resources of WCC generations; and
d) stress on the need for a paradigm shift in access technology using AI and/or the development of novel frameworks to advance the course of Next Generations (XG) of air-interfaces.

2 The Evolution of Mobile Cellular Network Technology

In a ten-year cycle, the human society has been transformed through five generations of wireless mobile networks of WCC evolution. These mobile generations are namely 1G, 2G, 3G, 4G, and 5G networks (Attaran, 2021). Each generation of WCC has different landmarks in practice, capabilities, and values, which distinguish it from the previous air-interfaces (Kumar & Sumit, 2021). 1G brought about the mass-market mobile telephony, 2G made the wireless handover possible, and the reliable mobile telephony and global interoperability that made SMS text messaging possible were provided by 3G. 4G also gave a significant improvement in high-speed and data capability and made online platforms for mobile internet services available (Attaran, 2021). The most powerful cellular wireless network recently unveiled is 5G. It is built with extraordinary data capabilities, infinite data broadcast, and unrestricted call volumes. The 1G to 5G evolution was characterized mainly by a shift in the CA methods using OMA and NOMA modulation schemes, which were also advanced to amicably meet the objectives of 5G and Next Generation Multiple Access (NGMA) techniques (Chávez-Santiago et al., 2015; Liu et al., 2024). These modulation schemes were also noticed with some inherent limitations that impede them from driving the objectives of the next generation (XG) of WCC effectively. The five generations of wireless mobile networks so far deployed are elaborated and the prospective 6G is also discussed.

2.1 First Generation

The WCC mobile air-interface mobile networks originated from 1G. It was developed in the 1980s using analogue cellular technology with a frequency of 150 MHz and only operates on voice call transmission modulation (Salih et al., 2020). It enables the use of multiple cell sites and the ability to hand-off calls as the user travels between cells during a conversation (Attaran, 2021). It operates on circuit switching. The first commercial edition was launched in Japan by NTT in 1979, and years later an upgraded version was launched in the US by Bell Laboratory in 1984. 1G networks are built on a speed of 2.4 Kbps only and is based on Advanced Mobile Phone System (AMPS) technologies. This is useful for language advantages and Total Access Communications System (TACS) (Salih et al., 2020; Kumar & Sumit, 2021). The AMPS device is modulated and multiplexed to reduce radio communication traffic in frequency using FDMA. 1G has a frequency range from 824 MHz to 894 MHz and uses a channel of 30 kHz (Kumar & Sumit, 2021). The FDMA has a limitation of allotting a single channel per carrier, which results in lots of bandwidth wastage. It also has narrow bandwidth and high cell site cost. As such, it could only transport one phone circuit at a time. This results in nonlinear intermodulation products. The intermodulation frequencies also cause time consumption and Adjacent Channel Interference (ACI) in network planning.

1G WCC has several additional disadvantages, which include deprived battery lifespan, deprived sound quality, inadequate storage, low telephone power, poor health, low-quality handling, low spectrum performance, weak voice links, and unreliable handoff (Kumar & Sumit, 2021; Attaran, 2021). The utilization of analogue signals (which are easily affected by impairment) is the biggest drawback to the 1G WCC scheme. This makes it slow and less reliable in transmission (Kumar & Sumit, 2021).

2.2 Second Generation

The 2G WCC air-interface was built on digital standards (Attaran, 2021). It was launched in the early 1990s. It operates on circuit switching. This was the first digital wireless network that enabled the advent of prepaid mobile phones and helped in rapid phone-to-network signalling. It initialized the text messaging SMS on Global System for Mobile Communication (GSM) networks and eventually on all digital networks. 2G has driven below a normal GSM in Finland through 1991 (Salih et al., 2020). Its frequency ranges from 30 kHz to 200 kHz. The 2G wireless networking networks of modes of inventions is based on these three: IS-136 in 1996, IS-54 (TDMA) in 1991, and IS-95 (CDMA) in 1993. The drawbacks in TDMA are that each user has a predefined time slot so that users roaming from one cell to another are not allotted the same time slot. The network is subjected to multipath distortion, network and spectrum planning is intensive, and too few users result in ideal channels of communication between rural and urban environments.

In advancement to 2G is 2.5G, which utilizes both circuit and packet switching modes of operation. The WCC system of 2.5G is deployed based on General Packet Radio Service (GPRS), CDMA, and Enhanced Data-rate for GSM Evolution (EDGE), which are embedded into the circuit of the wireless device (Kumar & Sumit, 2021). Interestingly, the 2G and 2.5G digital capabilities allow signal data rate ranges of 64 kbps and 144 kbps, respectively. This enables content such as strong signal quality, still imaging, text messages, and multimedia messages (voice and image messages). Additional advantages of the WCC 2G digital system compared to 1G include voice clarity, noise reduction, and battery power consumption reduction. Environment friendliness is part of the benefits of digital signals; as such, to achieve safety in voice calls and data transmission, digital encryption is adopted (Attaran, 2021). The biggest disadvantage of the 2G scheme to overawe the coming 3G scheme is its slow signal speed as compared to users' needs.

2.3 Third Generation

The 3G WCC air-interface provides high-speed data transfer capability for downloading information from the Internet and for sending videos with the speed of 2 Mbps (Attaran, 2021). Its frequencies were first introduced for commercial purposes in 1998. Its high-speed web surfing and video messaging operate on 15-20 MHz bandwidth and a frequency of 2100 MHz (Salih et al., 2020; Kumar & Sumit, 2021). 3G WCC uses fully packet switching instead of circuit switching, which is a great improvement over 2G (Attaran, 2021). In order to ensure stable communication over long distances, 3G technology passes signals using a network of phone towers. 3G handsets enable high connection speed of television content, massive media streaming of radio, and also provide video-conferencing support. It is estimated to offer a real-time speed of 2 Mbps for uploads and a maximum of 7.2 Mbps for downloads. An enhanced mobile telephony communications protocol in the mid-2000s developed an advanced 3G in the High-Speed Packet Access (HSPA) family, known as 3.5G, 3G+, or turbo 3G, which was implemented. The Universal Terrestrial Mobile Systems (UTMS) and Universal Mobile Telecommunications System (UMTS) operate based on 3G+ to have improved system capacity and higher data transfer speeds (Salih et al., 2020; Attaran, 2021).

The 3rd Generation Partnership Project (3GPP) is an association saddled with the responsibility to form the 3G standards and help in its deployment, which is in line with the International Telecommunication Union (ITU)-approved frequencies framework. The International Mobile Telecommunications (IMT) 2000 normal is also characterized by the 3GPP template (Salih et al., 2020). The 3G WCC network was first unveiled in Europe as UMTS, whilst its American version is tagged CDMA 2000 based on the channel access techniques used. The air-interface system for UMTS, known as Wideband Code Division Multiple Access (WCDMA) and Time Division-Synchronous Code Division Multiple Access (TD-SCDMA), is also an additional version of IMT 2000 3G stock coming from China (Kumar & Sumit, 2021). The major problem of the CDMA technique is synchronization and self-jamming (Zahra et al., 2020). Another problem with CDMA is scalability because of the increasing demand for higher data and transmission rates, which degrade its performance with an increase in the number of packets or users (Arrano & Azurdia-Meza, 2016).

The most important highlights for 3G WCC that make it better than its predecessors include increased data transmission rates and bandwidth, improved broadband, high-speed connectivity, increased email usage, high-speed web, 3D video games, mobile TV, TV calls, and power efficiency (Kumar & Sumit, 2021). The biggest drawbacks for 3G that are required of 4G and other advanced generations of networks to curtail are service operators' spendings of large amounts for 3G permits and infrastructures. Aside from the issue surrounding the high cost and supply of phone services, the demands for high-power utilization and special hardware need to be addressed (Kumar & Sumit, 2021).

2.4 Fourth Generation

The 4G WCC broadband infrastructure enabled seamless roaming from one device/network to another. It also interoperates between existing and future wireless network facilities by ensuring flexible access (Salih et al., 2020; Kumar & Sumit, 2021). 4G air-interface gives similar services as 3G but a little advanced multimedia, global accessibility, video streaming, TV digital media wide range transportability through gadgets, simpler information transferring, and many more. It was first launched in Stockholm followed by Oslo Norway in 2009 using Long Time Evolution (LTE) 4G normal technology, which uses OFDMA, OFDM, and Single Carrier -FDMA (SC-FDMA) as driven access factors (Salih et al., 2020). The 3GPP created LTE as a 4G Remote Interchange Standard (RIS). The ITU as a regulatory body set-up a 100 Mbps speed for 4G at full velocity. It was developed with a prospected speed improvements of 10-fold over the existing 3G technologies (Attaran, 2021). Its commercial version was first deployed in the United States by Verizon in 2011.

The 4G download standard speed is around 14 Mbps through 150 Mbps. Its peak data rate of 1Gbps was enabled by the American version known as LTE-A. It is an Internet Protocol (IP) based network, even for voice data. To send and receive data in packets, 4G uses a standard communications protocol. This standard enables it data to traverse all networks without being corrupted nor scrambled, that is, not all frequencies can be jammed illegally and privacy infringements is improved (Kumar & Sumit, 2021). The new frequencies created by 4G means a new path for the mobile tower. In its data transmission, the 4G WCC systems totally eliminated circuit switching by adopting an all-IP network design (Attaran, 2021).

With its capability to allow users to download gigabytes of data even in seconds. The 4G WCC air-interface enable users HD video streaming and web browse ability. Its technology turns smartphones into modern-age computers. 4G WCC frequencies fully complies with the proposal for high data rate and Quality of Service (QoS) requirements implemented in wireless wideband, Mobile TV (M-TV), and Multimedia Messaging Service (MMS) (Kumar & Sumit, 2021). The 4G's in its highlights has so far achieved 10Mbps- 1Gbps data rate, high security, low per-bit storage cost, high-quality digital videos, provision in video content creation, higher battery use. Its facilities run on a technology known as the evolved Node B (eNodeB) and mainly in the Evolved Packet Core (EPC). The 4G network is effective for internet data rates but not good for voice services. These services are both discharged into Wi-Fi or 3G wireless facilities, respectively. The aforementioned 4G drawbacks were anticipated to be resolved by the 5G technology.

One of the most deployable technologies used in driving the 4G WCC frequencies is the OFDM. This is as a result of increasing EU demands for better services. Between 2002 and 2004 the ITU and the 3GPP began studying the technology that would be used in the upcoming networks (Arrano & Azurdia-Meza, 2016). This latter arose because 3G CDMA-based systems were starting to experience problems in allowing further technological progress towards the predicted transmission rates for the coming years, known as the Red Trend Curve (RTC). When 4G was launched in 2010, the CDMA2000 (3G), obtained by intensive operations and perhaps by over-straightening its limits of technology, was able to obtain a downlink data transmission rate of about 14.7 Mbps using Evolution-Data Optimization Revision B (EV-DO Rev. B) techniques. EV-DO Rev. B is a telecommunication standard for data transmission in radio signals, typically deployed to triple the peak broadband data rate for broadband internet access in the existing EV-DO Rev. A.

The 4G WCC technology target has so far been successfully driven by OFDM and CP-OFDM schemes (Ramadhan, 2019). The former has the limitation of large side lobe emission and high Peak to Average Power Ratio (PAPR); the latter has poor spectral confinement, causes bandwidth wastage, and has an inflexible waveform. These drawbacks could not drive the present burden of EUs (Ramadhan, 2019).

2.5 Fifth Generation

The 5G WCC air-interface with the speed of approximately 1–10 Gbps, corresponds to the next frequencies of mobile network standards that is beyond the 4G LTE (Attaran, 2021). The 5G networks provide lower battery consumption, lower prices, and lower end-to-end latency than 4G wireless networks. These merits are because it uses Ultra-Wide Band (UWB) networks, which is achieved as a result of its higher band breadth at low energy levels. 5G system resources were in the market by the end of 2019 and began full commercial deployment in the year 2020 (Kumar & Sumit, 2021).

In comparison to earlier air-interface, 5G WCC frequencies offer unlimited broadcast data and extraordinarily high data capabilities. It works with the IP version 6 (IPv6) protocol. 5G is the present-day generation of wireless mobile broadband technology. Its main features include higher speed, reduced latency, simultaneous support for massive devices, energy saving, World Wide Wireless (WWW) Internet, proactive Software Defined Radio (SDR) stability, high data capability, big data transmission, massive data transfer in Gbps, high-resolution TV programs, multimedia newspapers, fast dial speeds, large memory telephones, smart hypermedia support, sound/visual illumination, effective voice/video networking, and flexible web services. Also, it is more efficient and attractive than previous generations of wireless technologies (Salih et al., 2020; Kumar & Sumit, 2021). In addition, it is the requirement for the Flexible Network Operations (FNO) (Stasio et al., 2018). Some of the 5G capabilities are summarized in Figure 2.

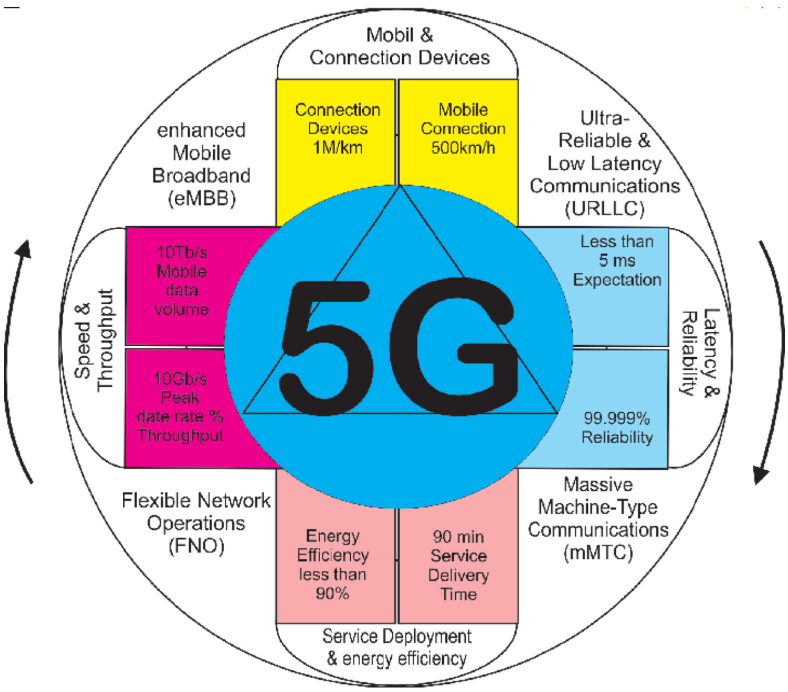

Figure 2: 5G networks capabilities

Network traffic in 5G is dominated by streaming or video applications (You et al., 2021). The distinguishing capabilities of the 5G WCC system include a user-experienced data rate of 0.1 Gbps, a peak data rate of 20 Gbps, end-to-end latency of 1 ms, 500 km/h mobility, an area traffic capacity of 10 Mbps/m^2, a connection density of 1 million devices/km^2, improved spectrum efficiency of about 3 times, and increased energy efficiency that is 100 times that of the wireless communication systems of 4G. The 5G network is built to have more outstanding features in network specificity (deterministic) latency, which uses Deterministic Networking (DetNet) for future demand to guarantee round-trip latency with punctuality and accuracy.

The 5G WCC networking infrastructure standard is segmented into two (Attaran, 2021); these are:

i) **Non-Stand Alone (NSA)**: This is the first 5G cellular network, which was commercially launched at the end of 2019. The signal traffic and control plane of the NSA infrastructural standard is based on existing 4G LTE. It is known as an improvement on the existing 4G LTE fast data architecture. NSA serves as the benchmark for initial 5G WCC commercial service in the year 2019 and the basis used to set up the next stand-alone standard/infrastructure.

ii) **Stand Alone (SA):** This is entirely a new core architecture. It made significant changes to how the wireless control plane transition operates. The 5G WCC SA structure was released in the year 2020. It enhanced subcarrier encoding and more flexible network slicing. It is designed to provide users improved performance at a low cost, better than 4G LTE and NSA.

Aside from NSA and SA, the 5G standards of WCC divide frequencies into two groups: Frequency 1 (FR1) (410 MHz – 7.12 GHz), also known as Sub-6G frequency, and Frequency 2 (FR) (24.25 GHz – 52.6 GHz) (Tikhomirov et al., 2018; Dilli, 2020). Most early deployments are in the FR1 slot. Research is ongoing in using FR2 and beyond, which are also known as Extremely High Frequencies (EHF) or millimeter waves (mmWave). The earlier stage of the 5G WCC network is allotted to operate on 3 GHz to 30 GHz frequencies. The higher frequency network of 30 GHz to 300 GHz is also booked for the microwave spectrum, which only allows short-distance communication with over 1 Gbps bandwidth (Salih et al., 2020). The large bandwidth and short frequency offer a more efficient network, which is experienced in its high-speed data transmission. The 5G technological scheme is used to join all accessible computer networks with just one Machine World-Wide (MWW) (Kumar & Sumit, 2021). Its flexible interoperability with Cognitive Radio (CR) and the Unparalleled Multi-Mode (UMM) stimulates effective functionality. Some of its top-notch functionality includes, simultaneously connected of networks using various wireless technologies, enhance the customer's location development, and incorporate various schemes to solve problems.

The CDMA and BDMA are the basic access footings for the 5G WCC system. One of its objectives is gradually achieved by the systematic adoption of filtered induced CA techniques to improve on the limitations of the LTE OFDM multi-carrier types of modulation and coding scheme. The filter's design is an essential method in generating the 5G waveform. Some of these improved versions of OMA CA that pave the actualization of 5G WCC systems are the, FBMC, GFDM, W-OFDM, F-OFDM, UFMC, and FF-OFDM (Augustine et al., 2025). The FBMC and GFDM are techniques based on pulse-shaping modulations, while F-OFDM, UFMC, and W-OFDM are designed based on sub-band filtering modulation (Yang et al., 2020; Pandey & Sharma, 2020; Taher et al., 2021).

To achieve better localization in the frequency domain and curtail the limitations noticed in OFDM, the filters in FBMC are used on each target subcarrier and the absence of the CP leads to high spectral efficiency. GFDM adopts a tail biting technique to shorten the cyclic prefix in OFDM, which may get better time-domain efficiency compared with FBMC. The UFMC uses short sub-band filters to minimize tailing in the time-domain and its constraints on filter length limit the OoBE suppression performance. The W-OFDM configures simple time domain window and its enhancement of time-domain efficiency totally depends on shortening the effect of the CP length, which is at the expense of sacrificing performance against multipath interference.

F-OFDM firstly divides its spectra into a series of contiguous sub-bands and the same filtering operation is performed on the granularity of each user to mitigate OoBE and PAPR (Yang et al., 2019). Moving toward the era of 6G, some of the notable drawback of the OMA schemes are the retention of chunks of interferences and poor power management which also degrades the required functionality of the network. With this, researches have proven that it cannot transmit effectively above FR1 objectives. The FF-OFDM is an advanced MA scheme built on the successes and drawbacks of these 5G driven OMA algorithms using sophisticated AI induced algorithms to curtail these limitations and advanced the course of next generation (XG) of wireless technology. In a parallel shift, the NOMA techniques, such as the PD, PTMA, BOMA, LPMA, LDS-CDMA, SCMA, and their variants, are developed to mitigate the excessive power and spectrum wastage in MA schemes. This is to advance the course of transmission latency, effective power allocation, achieve ultra-high heterogeneous connectivity, higher throughput, and improved SE (Liu et al., 2024).

The PD NOMA is a sophisticated power allocation and energy harvesting technique, the PTMA, BOMA, LPMA are techniques based on multiplexing in multiple domains, whilst the LDS-CDMA and SCMA are driven by code-domain NOMA. The PD has high SE and effective compatibility with other MA schemes but has drawback in user pairing and generate propagation errors. Other NOMA schemes such as the PTMA are more diverse and has low-connectivity receiver but it is difficult in design and create system redundancy in coding, the BOMA has simple structure and low complexity receiver but disadvantageous in flexibility and also lack user pairing, and the LPMA need not data user clustering but has limitation in channel coding specification. Furthermore, the LDS-CDMA is more suited for wideband transmission but generates redundancy in coding, while the SCMA does not need Channel State Information (CSI) and has more diversity than LDS but is characterized by redundancy in coding and is difficult to design optimal coding.
The integration of power-domain NOMA with cutting-edge technologies such as visible light communications, millimetre-Wave (mmWave)/terahertz (THz) communications, Multiple-Input Multiple-Output (MIMO), and Unmanned Aerial Vehicle (UAV), led to a massive turn-around and effective interoperability to meet the expected need of next generations of WCC networks.

Without restriction, the 5G network relative to its CA techniques is envisaged to generate high-speed information and networking everywhere. 5G WCC air-interface frequencies cannot operate effectively in remote area coverage (You et al., 2021). This has limited some of its applications, especially in Unmanned Aerial Vehicles (UAV). It is required of non-terrestrial networks like satellite communication to complement the terrestrial networks for cost-effective, ubiquitous service and seamless communication availability. An example of this is the UAV communication network, which is timely required for fast response in difficult and harsh terrains. The need for a Maritime Communication Network (MCN) to provide ships with high-quality communication services is long overdue; as such, a more versatile network to meet these present-day and futuristic demands is required.

2.6 Sixth Generation

5G will not meet all requirements of the future in 2030+ because of the proliferation of mobile devices and users' demands. This has triggered researchers to start focusing on the 6G WCC air-interface (You et al., 2021). In order to overcome the limitations of 5G, which include the provision of high-reliability networks, short-packet, low-latency services, and global coverage of the Internet of Everything (IoE), 6G WCC networks need to focus on being human-centric, instead of data-centric as its vision, application-centric, and/or machine-centric. This is possible by adopting a new paradigm shift as demonstrated in Figure 3, which, with the aid of Artificial Intelligence (AI) and Machine Learning (ML) technologies, will enable a new range of smart applications.

Encompassing all 5G network requirements, the 6G WCC air-interfaces are also expected to offer much higher data rates (Tbps), energy/spectral/cost efficiency, a 100 times increased connection near 100% global coverage, density, 10 times reduced latency, better security and intelligence level for full automation, accuracy in sub-centimetre geo-location, time synchronization in sub-millisecond. To achieve high energy and spectrum efficiency, including more flexible waveforms, advanced multiple CA approaches, multi-antenna technologies, channel coding methods, and a combination of diversity techniques are very essential. Meanwhile, novel network architectures are required to enable 6G WCC objectives, for example, Dynamic Network Slicing (DNS), Software Defined Network/Network Functions Virtualization (SDN/NFV, Cognitive Service Architecture (CSA), Service-Based Architecture (SBA), and Cell-Free Architectures (CFA). 6G WCC networks will be an integrated space-air-ground-sea networks to provide a complete global coverage, as illustrated in Figure 3.

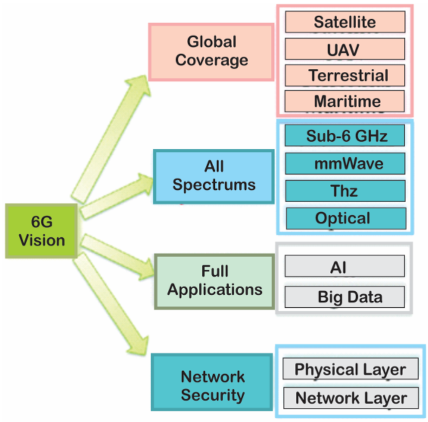

Figure 3: Vision of 6G wireless communication networks (You et al., 2021)

For global security and commercial activities, the handshake between UAV, satellite, and maritime communication is needed to increasingly expand the coverage range of the 6G WCC network. In 6G, all spectra are expected to be explored to provide a higher data rate. This includes the optical frequencies, THz, mmWave, and the sub-6 GHz. To have better network automation and management, it is proposed that flexible technologies in handling big data, like AI and ML, be incorporated with 6G WCC networks. It is glaring that AI-induced technology will enable more dynamic network resources fetching, slicing, caching, and computing capability to improve the performance and lay the footing of the next generation of networks. Last but by no means least is the highly aggressive network security achieved for both the network and physical layer when developing 6G networks. Industrial vertical signals, such as automated Internet of Things (IoT), Federated Learning Systems (FLS), Cellular Vehicle to Everything (C-V2X), Cloud Virtual Reality (CVR), Digital Twin Body Area Network (DTABN), and Energy Efficient Wireless Network Control (EEWNC), are fore-runners to largely boost the developments of 6G wireless communication networks.

So many access techniques have been proposed, while some have been developed to drive these 6G objectives. One of which is the FF-OFDM technique (Augustine et al., 2022). The FF-OFDM CA was developed to run higher versions of 5G, most especially 5G FR2-A, and possibly drive all 6G objectives. It is an Artificial Intelligence (AI)-induced technique that uses Machine Learning (ML) algorithms for its waveform realization. Its architecture obtained highly efficient spectrum slicing and spectrum shaping better than the previous CA schemes used in driving earlier generations of WCC systems. It is also an improved version that counters the limitation experienced in an F-OFDM, whose modification is performed at the Channel Estimation Region (CER) in order to negotiate a balance between the received data and the transmitted data. The FF-OFDM algorithm has the advantage of increasing system bandwidth, increasing data speed and throughput, effective interference mitigation, and high flexibility for network incorporation.

3 Summary of Findings

The Section presents a concise synopsis of the WCC generation of air-interfaces. This is to enable better comprehension of the survey conducted and to further stress the research objectives.

Table 1: Various air interface attributes

	1G	**2G**	**3G**	**4G**	**5G**	**Proposed 6G**
Year of Deployment	1980s	1990s	2000s	2010s	2020s	2030s
Location of First Commercialization	USA	Finland	Japan	South Korea	South Korea	N/A
Technology	AMPS, NMT, TACS	IS – 95, GSM	IMT2000, WCDMA	LTE, LTE-A, WiMAX	MIMO, mmWaves	SD-WAN, Cloud Edge computing, MPLS network, Optical network
Data Rate	2.4 kbps	64 kbps 144 kbps	2 Mbps UL 5 Mbps DL	100 Mbps – 1 Gbps	1 Gbps – 10 Gbps	11 Gbps – 1 Tbps
Switching Method	Circuit	Circuit Voice and Packet Data	Packet	Packet	All Packets	All Packets
Operational Frequency Range	800 MHz, 900 MHz	900 MHz, 1800 MHz	2100 MHz	850 MHz to 1800 MHz	Sub-6G = 410MHz – 7.125GHz, mmWaves = 24.250GHz - 52.600GHz	4.400 – 4.800 GHz, 7.125 – 8.400 GHz, 14.800 – 15.350 GHz
Carrier Frequency	30 kHz	30 kHz to 200 kHz	5 MHz	15 MHz	20 MHz, 50 MHz, 100 MHz, 200 MHz, 400 MHz	100 GHz and above
Bandwidth	Analog	30 - 200 kHz	15 - 25 MHz	100 MHz	3 GHz to 30 GHz to 300 GHz	160 MHz
Band Type	Narrow Band	Narrow Band	Wide-Band	Ultra Wide-band	Ultra Wide-band	Ultra-large
Access Schemes System	FDMA	TDMA, CDMA	CDMA	OFDM, BDMA	BDMA F-OFDM, FBMC	Smart F-OFDM plus IM, FF-OFDM

Hang off/over	Horizontal	Horizontal	Horizontal	Horizontal & Vertical	Horizontal & Vertical	Horizontal & Vertical
Core Network	PSTN	PSTN	Packet N/W	Internet	IoT	IoT
Latency	1,000 ms - 8 s	300 – 1,000 ms	100 – 300 ms	10 – 100 ms	1 ms or less	10 – 100 μs
Applications	Voice calls only	Voice calls, short messages	Internet services, GPS	High-speed internet, mobile TV	High-resolution video stream	Smart cities, Industry 4.0, advanced healthcare

4 Area of Future Work and Recommendations

As a way forward, this review is suggesting a paradigm shift in the integration of MA schemes with other emerging technologies, such as AI-induced ML techniques, Semantic Communications (SC), and Reconfigurable Intelligent Surface (RIS) (Liu et al., 2024), and their variants. These proposed enablers have proven to outperform traditional MA techniques in data transmission efficacy.

ML is an AI-induced technique that has gained recommendations and acceptance in diverse fields of human endeavours because of its high accuracy in data sampling. It is a highly versatile technique that has caused massive turn-arounds in the fields of science and engineering. It has revolutionized industries, reshaped decision-making processes, and transformed how technologies interact. The ML is not left-out in the quest for next generations of wireless communication resources. It has been seen adopted and deployed effectively in the areas of data analysis, power management/allocation, energy harvesting, waveform modification, interference mitigation, network scalability, system flexibility, and many more (Yoro et al., 2025; Augustine et al., 2025). Because ML is designed by humans to reason like humans, it is liable to ethical problems generated from sentiment. Other disadvantages of ML include deterministic problems as a result of large data, lack of reproducibility, data dependency, and high computational cost.

The SC technology is based on semantic information transmission, aiming to achieve more intelligent, efficient, and reliable data transmission (Wei, 2023). The SC is a prospective next-generation data transmission technique proposed to overcome the limitations experienced in traditional communication methods in the areas of simple data transmission, accurate interpretation, and conveyance of the semantic meaning in the data. In as much as SC has proven to be one of the prospective technologies expected to drive 6G and beyond, it is also characterized by challenges at the development stages. These challenges encompassed meeting resource requirements, model collaboration and privacy security, context perception and identification, knowledge sharing, and many more.

The RIS is termed advanced intelligent radio (Hamza et al., 2025). As the World prepares for 6G, RIS is one of the emerging core facilitator networks that can generate many opportunities to improve the performance of future wireless communication systems and also other fields. It operates by allowing the wireless communication environment to be controlled in a programmable manner. RIS utilizes a semi-passive or passive layer to efficiently and intelligently reflect an incident electromagnetic wave, which differs from conventional technologies that use active components to manage signal transmission and reception. This new approach creates several latent strengths, making RIS an extremely appealing solution for future wireless systems (Chiaraviglio et al., 2021). Energy efficiency is one of the strongest points of RIS. Although RIS offers attractive benefits, several technical and practical barriers, like the generation of small gain due to its passive transmission and the lack of amplifying signals, posed a serious challenge for widespread implementation of this technology. These drawbacks and many more need to be amended to enable RIS to grow and become a credible and trustworthy technology in the near future of wireless networks.

The successes and limitations notable in these technologies in data transmission have again opened up another vacuum in the need for further intensive and extensive innovations in these technologies to enable effective drive of the 6G attributes and those of upcoming generations of WCC systems. Aside from this, the research is also recommending the development of novel technologies/techniques that could enhance effective achievement of the objectives of next-generation wireless systems.

5 Conclusion

This survey provided a detailed presentation on WCC air-interfaces relative to its CA scheme mostly used for commercial deployment. These varying CA schemes serve as network lives. Each generation of WCC air-interface has its CA scheme that propels it in achieving its designed objective. These CA schemes are systematic hybridized multiplexing and modulation techniques that are designed to ensure the message signal is properly impressed into the transmission channel to the receiver with very little loss of information and energy. The first mobile handoff network achieved in the 1G air-interface is successfully driven by FDMA.

The ground-breaking note messaging achieved by 2G is run by TDMA, whilst the CDMA as top-notch technology also enables 3G and 4G that achieve the internet and multi-media functions, respectively. It is able the drive the higher mandates of 3G but could not fully deliver that of 4G objective. This is because, the CDMA algorithm could only enhanced serial kinds of transmission which makes it vulnerable to threads and lacks the required scalability in driving higher data rate.

The 4G mandate was fully achieved by the development of advanced multi-carrier CA techniques of OFDM, OFDMA, CP-OFDM, and WOFDM, where signal information is divided into chunks and transmitted in different subcarriers. The 5G basics where also enabled by CDMA, BDMA, but were also advanced by special cases of OFDM techniques known as UFMC, FBMC, GFDM, and F-OFDM which could only drive 5G FR1. The FF-OFDM CA is employed to enable the higher versions of 5G; FR2, FR2-2, and FR2-A, which is also prospected to fully drive the 6G objective because of its AI capability. The survey is aimed at bringing to the light of researchers the theoretical knowledge of a wireless network life, its successes, drawbacks, and the need for new access schemes, and/or novel techniques to enable the next generations resources of WCC air-interfaces.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Arrano, F. H., & Azurdia-Meza, A. C. (2016). *OFDM: Today and in the Future of Next Generation Wireless Communications*. https://www.researchgate.net/publication/306195106, 1-7.

Attaran, M. (2021). The impact of 5G on the evolution of intelligent automation and industry digitization. *Journal of Ambient Intelligence and Humanized Computing*, 1-17.

Augustine, Z., Tekanyi, A. M. S., Sani, S. M., Usman, A. D., & Yaro, A. S. (2022). Development of a Novel Feedback-Filtered Orthogonal Frequency Division Multiplexing Scheme for 5G Network and Beyond. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC).* 13(4), 1-6, https://www.researchgate.net/publication/357517442

Augustine, Z., Yaro, S. A., Tekanyi, A. M. S., Bello, H., Abdu-Aguye, U. F., & Agbo, E. E. (2025). Feedback Filtered-OFDM Waveform Candidature for Interference Mitigation in 5G Networks and Beyond. *International Journal of Integrated Engineering*. 17(1), 323-339, https://publisher.uthm.edu.my/ojs/index.php/ijie

Bliley's Technology, (2017). *The Evolution and History of Radio Wave Technology* (Infographic).

Cai, Y., Qin, Z., Cui, F., Li, G. Y., & McCann, J. A. (2018). Modulation and multiple access for 5G networks. *IEEE Communications Surveys & Tutorials*, 20(1), 629-646.

Chávez-Santiago, R., Szydełko, M., Kliks, A., Foukalas, F., Haddad, Y., Nolan, K. E., & Balasingham,I.(2015). 5G: The convergence of wireless communications. *Wireless Personal Communications*, 83(3), 1617-1642.

Chiaraviglio, L., Elzanaty, A., & Alouini, M. S. (2021). Health risks associated with 5G exposure: A view from the communications engineering perspective. *IEEE Open Journal of the Communications Society*. 19(2), 2131-79, https://www.researchgate.net/publication/341815740

Dilli, R. (2020, March). Analysis of 5G wireless systems in FR1 and FR2 frequency bands. In *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 767-772). IEEE.

Hamza, H., Talha, A., Ullah, S., & Ahmed, S. U. (2025). Reconfigurable Intelligent Surface (RIS): Strengths, Weaknesses, Opportunities and Threats. *Spectrum of Engineering Science*. 3(7), 533-546, https://doi.org/10.5281/zenodo.16108438

Kumar, U. (2015). Pandect of 5g Communication Technology in Context with Some Latest Researches. *The International Journal of Engineering and Science (IJES),* 4(7), 17-22.

Kumar, P., & Sumit, D. (2021). Review Paper on Development of Mobile Wireless Technology. *International Conference on Recent Trends in Computing (ICRTCE), Conf. Ser. 1979* 012024, 1-6.

Liu, Y., Ouyang, C., Ding, Z., & Schober, R. (2024). *The road to next-generation multiple access: A 50-year tutorial review*. arXiv. http://arxiv.org/abs/2403.00189v2

Lu, Y., & Zheng, X. (2020). 6G: A survey on technologies, scenarios, challenges, and the related issues*. Journal of Industrial Information Integration,* 19, 1-14.

Pandey, A., & Sharma, D. (2020). Filtered-OFDM for 5G Wireless Communication Narrow-band IoT Systems. *IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE),* 15(3), 24-31.

Ramadhan, A. J. (2019). Implementation of a 5G Filtered-OFDM Waveform Candidate*. International Journal of Engineering Research and Technology*, 12(4), 500-507.

Sahrab, A. A., & Yaseen, D. A. (2021). Filtered orthogonal multiplexing scheme for improved 5G systems, *Bulletin of Electrical Engineering and Informatics*, 10(4), 2079-2087.

Salih, A. A., Zeebaree, S. R., Abdulraheem, A. S., Zebari, R. R., Sadeeq, M. A., & Ahmed, O. M. (2020). Evolution of mobile wireless communication to 5G revolution. *Technology Reports of Kansai University*, 62(5), 2139-2151.

Stasio, D. F., Mondin, M., & Daneshgaran, F. (2018). Multirate 5G downlink performance comparison for f-OFDM and w-OFDM schemes with different numerologies. In *2018 International Symposium on Networks, Computers and Communications (ISNCC)* (pp. 1-6). IEEE.

Taher, M. A., Radhi, H. S., & Jameil, A. K. (2021). Enhanced F-OFDM candidate for 5G applications. *Journal of Ambient Intelligence and Humanized Computing*, 12(1), 635-652.

Tikhomirov, A., Omelyanchuk, E., & Semenova, A. (2018, March). Recommended 5G frequency bands evaluation. In *2018 Systems of Signals Generating and Processing in the Field of on-Board Communications* (pp. 1-5). IEEE.

Wei, M. (2023). *Semantic Communication System: Fundamentals, Challenges, and Performance Metrics*. Research Gate. https://www.researchgate.net/publication/372956160

Yang, M., Chen, Y., & Du, L. (2019). Interference Analysis and Filter Parameters Optimization for Uplink Asynchronous F-OFDM Systems. *IEEE Access*, 7, 48356- 48370.

Yang, X., Yan, S., Li, X., & Li, F. (2020). A Unified Spectrum Formulation for OFDM, FBMC, and F-OFDM. www.mdpi.com/journal/electronics, *Electronics*, 9(1285), 1-15.

Yoro, E. R., Adamu-Fika, F., Ademuwagun, K. A., Ubadike, O., Augustine, A., Olufemi, O. T., & Achimugu, P. (2025). Using Machine Learning Algorithms for Anomalies Detection in Modern Wireless Communication Systems. In *9th International Conference on Information System Design and Intelligent Applications (ISDIA 2025)*. Springer Nature lectures notes.

You, X., Wang, C. X., Huang, J., Gao, X., Zhang, Z., Wang, M., Huang, Y., Chuan Zhang, C., Jiang, X., Wang, J., Zhu, M., Sheng, B., Wang, D., Pan, Z., Zhu, P., Yang, Y., Liu, Z., Zhang, P., Tao, X., Li, S., Chen, Z., Ma, X., Lin, C., Han, S., Li, K., Pan, C., Zheng, Z., Hanzo, L., Shen, X. S., Guo, Y. J., Ding, Z., Haas, H., Tong, W., Zhu, P., Yang, G., Larsson, E. G., Ngo, H. Q., Hong, W., Wang, H., Hou, D., Chen, J., Chen, Z., Hao, Z., Li, G. Y., Tafazolli, R., Yue Gao, Y., Poor, H. V., Fettweis, G. P., & Liang, Y. C. (2021). Towards 6g Wireless Communication Networks: Vision, Enabling Technologies, and New Paradigm Shifts-Review *Information Sciences*, 64, 1-74.

Zahra, A., Umar Khan, Q., & Sheikh, S. A. (2021). Comparative analysis of quaternion modulation system with OFDM systems. *International Journal of Electronics Letters*, 9(2), 212-221.

Zhang, L., Ijaz, A., Xiao, P., Molu, M. M., & Tafazolli, R. (2017). Filtered OFDM systems, algorithms, and performance analysis for 5G and beyond. *IEEE Transactions on Communications*, 66(3), 1205-1218.

Zou, F., Liu, Z., Hu, X., & Wang, G. (2021). A Novel PAPR Reduction Scheme for OFDM Systems Based on Neural Networks*, Wireless Communications and Mobile Computing*, 1-8.

A Hybrid VGG-16 and TabNet Model for Interpretable Lung Disease Detection from Chest X-rays in Resource-Constrained Environments

[1]Abraham Eseoghene Evwiekpaefe, [2]Fiyinfoluwa Ajakaiye, [3*]Muhammad Nazeer Musa and [4]Muhammad Musa Isa

[1,2,4]Department of Computer Science, Faculty of Military Science and Interdisciplinary Studies, Nigerian Defence Academy, PMB 2109, Kaduna, Nigeria
[3]Department of Cyber Security, Faculty of Military Science and Interdisciplinary Studies, Nigerian Defence Academy, PMB 2109, Kaduna, Nigeria

email: [1]aeevwiekpaefe@nda.edu.ng, [2]fajakaiye@nda.edu.ng, [3*]muhammadmusa2502@nda.edu.ng, [4]mm.isa@nda.edu.ng

Corresponding author

Received: 26 July 2025 | Accepted: 11 November 2025 | Early access: 01 December 2025
https://doi.org/10.33736/jcsi.10348.2026

Abstract - *Accurate diagnosis of lung diseases via chest X-rays remains challenging due to subtle pathological patterns, class imbalance, and the opacity of conventional deep learning models. While convolutional neural networks excel in feature extraction, their "black-box" nature and poor interpretability hinder clinical trust, particularly in resource-constrained settings. To address these limitations, we propose a novel hybrid architecture integrating VGG-16 with TabNet, synergizing hierarchical spatial feature extraction with attention-driven interpretability. The model leverages VGG-16's convolutional layers to capture granular details, while TabNet's sequential attention masks dynamically prioritize discriminative features, quantifying their clinical relevance. Trained on a dataset of 2,590 chest X-rays (COPD, tuberculosis, pneumonia, and normal cases) from Nigerian hospitals, the model achieved state-of-the-art performance with 97% accuracy, surpassing ResNet-50 (95.7%) and standalone VGG-16 (94.7%). Preprocessing, including non-local means denoising and targeted augmentation, mitigates noise and class imbalance, yielding F1-scores exceeding 97% for COPD and pneumonia, with AUC values above 0.98 across all classes. The model's interpretability is validated through attention maps highlighting disease-specific radiological markers, such as hyperinflation in COPD and consolidations in pneumonia, aligning with clinical expertise. Deployed as a real-time Android application optimized for low-end devices, the solution achieves inference in <1 second offline, addressing infrastructural barriers in low-resource regions. The model advances equitable healthcare delivery, demonstrating generalizability across demographic subgroups (accuracy deviation ≤1.2%) and compliance with emerging regulatory standards for trustworthy AI. This innovation establishes a scalable paradigm for interpretable, high-performance lung disease detection, with transformative potential for global health equity.*

Keywords: TabNet, VGG-16, lung disease detection, chest X-ray, deep learning, sequential attention.

1 Introduction

Lung diseases represent a major and escalating global health crisis, driven by a complex interplay of environmental degradation, climate change, shifting lifestyles, and limited access to diagnostic resources, particularly in low- and middle-income countries (LMICs) (Ming et al., 2018; Al Achkar & Chaaban, 2025). Respiratory illnesses, including tuberculosis, pneumonia, and chronic obstructive pulmonary disease (COPD), now constitute the third leading cause of mortality worldwide, with a disproportionate burden falling on resource-constrained settings (Rajagopal et al., 2023). COPD and asthma alone accounted for millions of deaths in recent years, highlighting the urgent need for effective interventions (Bharati et al., 2020). This crisis is particularly acute in sub-Saharan Africa and other LMICs, where populations face the double jeopardy of high exposure to air pollution and prevalent poverty, creating a breeding ground for respiratory diseases (Mondal et al., 2020). The recent COVID-19 pandemic further underscored the vulnerability of lung health, demonstrating the devastating consequences of respiratory viral infections and exacerbating the existing challenge of pneumonia (Chunli et al., 2020).

Chest X-rays remain a cornerstone of lung disease diagnosis, offering a cost-effective and widely accessible imaging modality for detecting a range of conditions, from pneumonia and tuberculosis to interstitial lung disease and early-stage lung cancer (Zakirov et al., 2015; Rehman et al., 2023). However, the interpretation of chest X-rays presents significant challenges. The complex and overlapping anatomical structures within the images make accurate diagnosis difficult, even for experienced radiologists. Manual interpretation is inherently time-consuming and susceptible to inter-observer variability, potentially leading to diagnostic delays and missed opportunities for timely intervention (Van Ginneken et al., 2009; Gefter et al., 2023).

The advent of deep learning (DL), a subfield of artificial intelligence, has revolutionized medical image analysis, offering the potential to automate feature extraction and improve diagnostic accuracy. Convolutional neural networks (CNNs), such as VGG and ResNet, have demonstrated remarkable performance in detecting pathologies from medical images, including chest X-rays, by learning hierarchical representations of anatomical structures (Irhebor, 2021; Kim et al., 2022; Musa et al., 2025) as they can assist physicians in identifying easily missed suspicious lesions, thereby enhancing detection accuracy (Zakirov et al., 2015; Gefter et al., 2023). Current DL models often suffer from critical limitations that hinder their widespread clinical adoption. Many models lack robust feature engineering at the fully connected layers responsible for final decision-making, and they struggle with the inherent class imbalance commonly found in medical datasets, where certain diseases are significantly more prevalent than others (González et al., 2018). Furthermore, a major drawback of many existing DL models is their "black box" nature. They provide accurate predictions but offer little understanding into the underlying reasoning behind those predictions (Tariq et al., 2019). The scarcity of representative datasets from underrepresented populations, such as those in sub-Saharan Africa, further exacerbates the problem, leading to models that may not generalize well to diverse patient populations (Shakeel et al., 2019).

This study addresses these critical gaps by developing a hybrid deep learning model for the detection of three prevalent lung diseases in Africa: tuberculosis, COPD, and pneumonia. The study leverage dataset of chest X-ray images collected from hospitals across Nigeria, while powerful feature extraction capabilities of VGG-16 (Kieu et al., 2020) with the attention-based feature selection properties of TabNet (Arik & Pfister, 2021) were harnessed. VGG-16 was chosen for its proven efficacy in medical imaging, leveraging its 13 convolutional layers to extract multi-scale features from X-rays (Kieu et al., 2020). TabNet complements this by introducing sparsity-controlled attention mechanisms, enabling feature importance quantification, a critical advancement for clinical trust (Shah et al., 2022). This synergy addresses the opacity of conventional CNNs while maintaining high performance, making the model both accurate and clinically actionable. This hybrid approach aims to not only enhance diagnostic accuracy but also address the critical need for transparency and interpretability in AI-driven medical tools.

2 Literature Review

Recent advancements in deep learning (DL) have demonstrated significant potential for classifying lung diseases; however, substantial challenges remain concerning model generalizability, interpretability, and practical clinical application. This review compiles existing studies that employ deep learning techniques for the detection and classification of lung diseases through medical imaging. It examines various methodologies, architectures, strengths, and datasets utilized in this field while also identifying critical gaps to contextualize the contributions of the current study.

Early studies, such as Ming et al. (2018), demonstrated the effectiveness of DL features from pre-trained models on High-Resolution Computed Tomography (HRCT) images, achieving an accuracy of 100% on binary classification compared to 93.52% with traditional Gray-Level Co-occurrence Matrix (GLCM) features. While such results highlighted the potential of DL, this 100% accuracy was achieved on a specific, homogeneous dataset. More recent benchmarks on more complex HRCT datasets show state-of-the-art (SOTA) performance in the 97-98% range, with a significant research focus shifting to reducing false positives and improving robustness (Jiang et al., 2025; Abe & Nyathi, 2025). Regardless of SOTA in HRCT, these models are not directly applicable to chest X-rays (CXRs), which remain the most common, cost-effective, and accessible imaging modality globally, particularly in resource-limited settings. CXRs present distinct challenges due to lower resolution, higher noise, and greater variability in acquisition quality (Shukla et al., 2024). Similarly, Kim et al. (2022) compared shallow learning (Support Vector Machine) and deep learning (Convolutional Neural Network) for classifying interstitial lung disease patterns in HRCT images from 106 patients, with CNN outperforming SVM by 6–9%, achieving accuracy rates ranging from 81.27% to 95.12% as the number of convolutional layers increased. To substantiate the claim that deep learning enhances lung disease detection specifically for CXRs, Kieu et al. (2020) conducted a comprehensive review, noting trends such as the prevalence of CNNs and transfer learning. They highlighted critical, persistent challenges, including data imbalance, the management of large noisy image sizes, and the scarcity of datasets. While data augmentation is a common strategy to address imbalance, recent studies affirm that augmentation alone does not solve the fundamental challenges of "distribution drift" caused by non-standardized data acquisition or the scarcity of geographically diverse datasets (Ahmad et al., 2025; Abe & Nyathi, 2025; Liu et al., 2024). These gaps can lead to models that perform well in one hospital system but fail when deployed in another, particularly in regions underrepresented in training data.

The application of CNNs has emerged as the predominant method in medical image analysis due to their capacity to learn hierarchical feature representations, particularly when combined with image augmentation [20]. Research conducted by Rahman et al. (2020) demonstrated a remarkable 98.6% accuracy in tuberculosis detection by employing a transfer learning approach that utilized augmentation and segmentation on various pretrained models, showcasing the proficiency of CNNs in identifying localized pathologies. Additionally, Ganeshkumar et al. (2023) introduced a two-stage deep learning model, focused on binary classification between normal and COVID-19 pneumonia cases, which outperformed existing methods in average accuracy and F1-score, even providing confidence scores for diagnoses. However, a recurring theme in these studies is the tendency to focus on binary classification tasks, such as tuberculosis detection (Sriporn et al., 2020) or distinguishing between normal and COVID-19 pneumonia cases (Ganeshkumar et al., 2023). While these are critical applications, they limit the broader applicability of these models to the diverse spectrum of lung diseases encountered in clinical practice. In contrast, Olayiwola et al. (2023) and Alshmrani et al. (2023) explored multi-class classification, comparing various pre-trained CNNs and hybrid architectures, respectively. Olayiwola et al. (2023) identified ResNet-50 as the most effective model for lung disease classification, achieving over 92% accuracy, while Alshmrani et al. (2023) combined VGG with additional convolutional layers to achieve 96.48% accuracy across six lung diseases. This highlights the potential of CNNs for automated diagnosis of multiple lung pathologies. Furthermore, Al-Sheikh et al. (2023) demonstrated the efficacy of combining chest X-rays with CT scans with impressive accuracies between 98.4% and 98.8% in multi-class lung disease classification. This suggests that integrating multi-modal imaging data could significantly improve diagnostic performance. Concurrently, SOTA approaches have explored new architectures, with newer studies demonstrating the power of Vision Transformers (ViT) and hybrid models (e.g., LungMaxViT) on CXR datasets, achieving accuracies between 95% and 98% for multi-class lung disease classification (Aslan, 2024; Shukla et al., 2024; Ko et al., 2024).

Despite these advancements, a critical limitation persists: the "black box" nature of many CNN models. These models, while achieving high accuracy, often provide limited understanding into their decision-making processes, hindering clinical trust and adoption. Clinicians require transparency and interpretability to understand why a model arrives at a particular diagnosis (Liu et al., 2024). This has spurred a dedicated research thrust into eXplainable AI (XAI) for medical imaging (Colin & Surantha, 2025). While many models rely on post-hoc XAI methods like Grad-CAM (Aslan, 2024), these only show where a model is looking, not how it weighs different features. Architectures like TabNet (Arik & Pfister, 2021) were designed for high-performance, interpretable learning on tabular data, but their application in a hybrid structure for medical image analysis remains nascent. Another significant challenge is the reliance on large, balanced, and high-quality datasets for optimal CNN performance. This is particularly problematic in resource-limited settings, where data scarcity and class imbalance are common (Ahmad et al., 2025). The study by Bharati et al. (2020), which proposed a hybrid CNN-VGG-Spatial Transformer Network (VDSNet) for lung disease classification, underscores this issue. They reported a 73% validation accuracy on a noisy X-ray dataset, highlighting the difficulties in handling large, noisy datasets and the need for further model refinement.

Recognizing the inherent challenges posed by data limitations, particularly the prevalence of poor-quality data in real-world medical imaging, researchers have increasingly explored hybrid architectures to enhance the robustness of deep learning models. For instance, Shakeel et al. (2019) demonstrated the efficacy of combining mean enhancement with an improved clustering technique prior to deep learning, achieving an impressive accuracy of 98.42%. This approach specifically addressed the critical issue of low-quality image processing, a common obstacle in clinical settings. Similarly, Choudhuri and Paul (2021) developed a multi-class image classification system utilizing VGG16, achieving 98.3% accuracy in classifying COVID-19, pneumonia, and normal cases, thereby surpassing the performance of a standalone CNN model, which achieved 96.6% accuracy. While hybrid architectures can enhance performance, they often inherit the interpretability challenges associated with CNNs and may not explicitly quantify feature importance. Additionally, Tariq et al. (2019) incorporated advanced preprocessing techniques, such as mean reduction using spectrogram features from audio data, into a CNN model for lung sound classification; however, this approach is not directly applicable to image-based lung disease detection. Also, Gonzalez et al. (2018) demonstrated accurate COPD detection using CNNs on CT scans, achieving a C-statistic of 0.856. Although effective, the binary classification approach and reliance on CT scans limit broader applicability and restrict the study's relevance to settings where chest X-rays are more commonly used. Sriporn et al. (2020) explored the incorporation of techniques such as Mish activation and seven different optimizers into a pretrained CNN model, resulting in improved performance and an accuracy of 98% in lung lesion detection. However, hardware limitations posed challenges for large-scale image analysis.

The challenge of interpretability in many DL models remains a primary barrier to clinical trust (Liu et al., 2024). Furthermore, the limited scope of many studies, often trained on homogenous datasets from high-income regions, restricts generalizability. As highlighted by Kieu et al. (2020) and Al-Sheikh et al. (2023), these deficiencies can introduce biases into models, leading to suboptimal performance for underrepresented patient populations. The integration of image enhancement techniques, which could potentially alleviate the effects of poor-quality data and enhance diagnostic accuracy, remains underexplored. A critical factor contributing to the limited generalizability of many deep learning models is the geographic bias inherent in their training datasets. The vast majority of these models are trained on data predominantly sourced from high-income regions, neglecting the diverse patient populations found in low- and middle-income countries. As Kieu et al. (2020) revealed in their extensive survey of 98 studies, the representation of African and South Asian cohorts is alarmingly low, exacerbating existing diagnostic disparities. This situation underscores the urgent need for geographically diverse datasets, particularly those originating from under-resourced countries, to ensure the equitable deployment of AI-driven diagnostic tools. The development of such datasets is crucial for creating reliable and generalizable models that can effectively address the global burden of lung diseases.

This study directly responds to these challenges. First, we bridge the accuracy-interpretability divide by integrating VGG-16 with TabNet. This novel hybrid architecture moves beyond post-hoc explanations by quantifying feature importance through TabNet's inherent attention mechanisms, addressing the "trust gap" in AI diagnostics. Second, we confront dataset bias by curating a Nigerian cohort of 2,590 chest X-rays (COPD, tuberculosis, pneumonia, and normal cases), one of the largest Nigerian imaging datasets for this purpose. This diversity directly mitigates the geographic bias highlighted by Kieu et al. (2020) and Liu et al. (2024), enhancing generalizability to underserved populations. Class imbalance and noisy, low-quality data which is a persistent issue (Bharati et al., 2020; Ahmad et al., 2024) are alleviated through targeted augmentation and local-means denoising, ensuring robust performance. Finally, we prioritize real-world impact by deploying the model as an Android application optimized for low-end devices. Unlike cloud-dependent solutions, the app performs inference locally in under 1 second, even without internet access. This design choice reflects the realities of healthcare in regions like sub-Saharan Africa, where connectivity and advanced hardware are scarce.

3 Research Methodology

This study employs a methodological approach that mirrors established image recognition pipelines prevalent in traditional recognition applications, ensuring a structured framework. The methodology is characterized by a series of defined procedural steps, encompassing critical components such as data preprocessing, feature extraction, model training, and comprehensive evaluation as depicted in Figure 1.

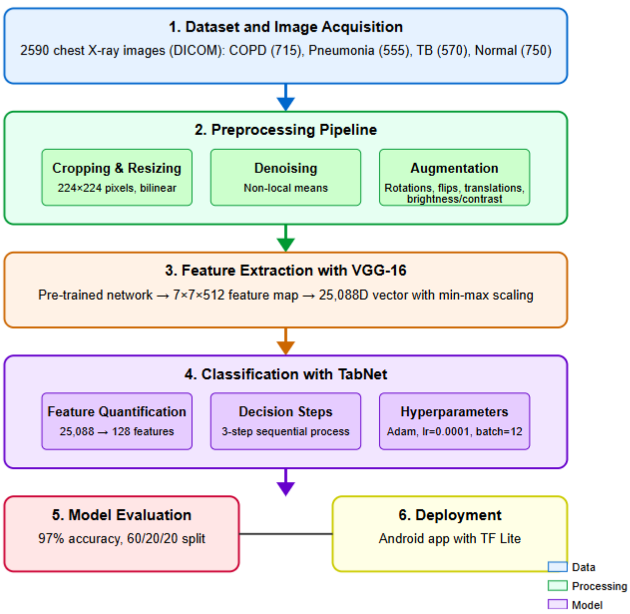

Figure 1: Pipeline of the proposed solution

3.1 Image Acquisition

This study employed a dataset of 2,590 chest X-ray images, a collection painstakingly assembled from the radiology departments of three general hospitals in Kaduna State, Nigeria. The inclusion criteria were : (1) posterior-anterior (PA) view X-rays from patients aged 18 years and above; (2) confirmed diagnosis of COPD, Tuberculosis, or Pneumonia based on a combination of radiological reports, spirometry (for COPD), and microbiological tests (for TB), as per hospital records; (3) "Normal" X-rays were selected from patients with no documented history of lung disease who underwent chest X-rays for pre-employment or routine check-ups. Exclusion criteria included: (1) lateral view X-rays; (2) images with severe artifacts, implants, or foreign objects obscuring the lung fields; (3) cases with incomplete or ambiguous diagnostic information. These images were categorized into COPD (715), Normal (750), Pneumonia (555), and Tuberculosis (570) classes, which were initially stored in DICOM format and subsequently converted to PNG for processing. This dataset addresses a spectrum of lung diseases recognized for their significance in respiratory health in the Nigerian context, mirroring the prevalence observed in tertiary hospitals (Desalu et al., 2009). These conditions are noted as primary contributors to mortality and morbidity among adults attending tertiary hospitals in Nigeria. Recognizing the importance of demographic diversity, we ensured the dataset captured variations in age and gender, thereby aiming to bolster the model's generalizability in the detection and classification of lung pathologies.

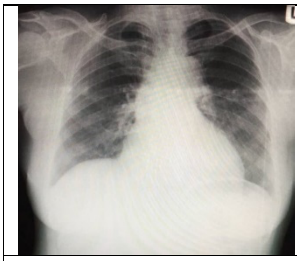	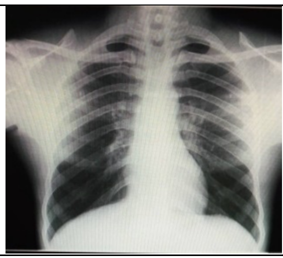	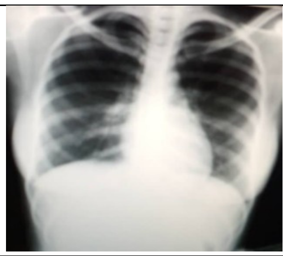	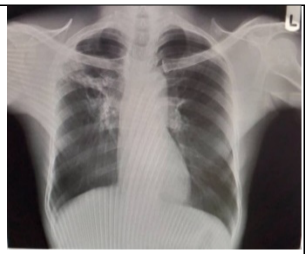
COPD	Normal	Pneumonia	Tuberculosis
715	750	555	570

Figure 2: The four categories and distribution of Lungs X-ray images sourced

3.2 Pre-processing

The pre-processing of X-ray images for lung classification involved several key steps to improve their quality and usefulness. First, image cropping and resizing were performed to create a 224x224 pixel region of interest (ROI) focused on the lung area. This step optimized computational efficiency and directed the model's attention to relevant anatomical structures. Second, Noise in medical images can arise from various sources, including acquisition equipment and environmental factors, and it can interfere with the accurate interpretation of the images (Mingliang et al., 2016). Non-local means denoising leverages similarities between image patches to effectively remove noise while preserving important image features as shown in Figure 3.

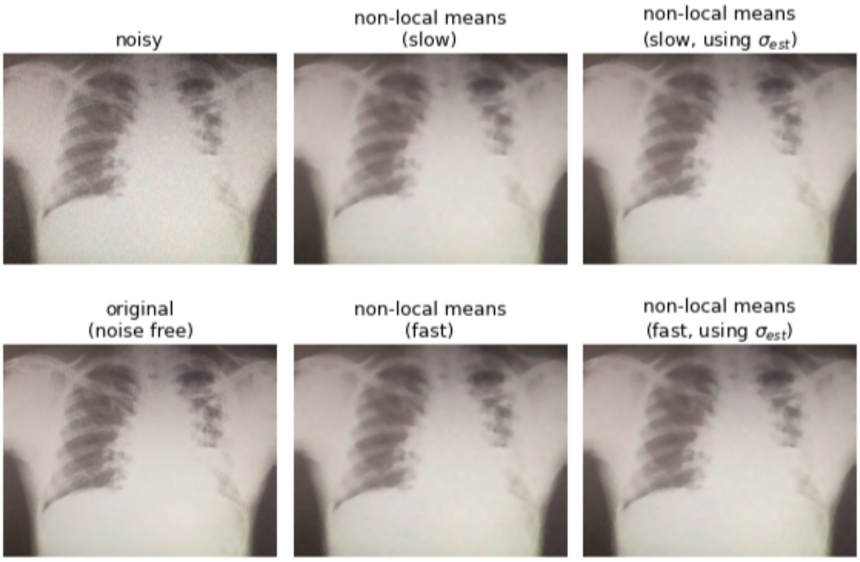

Figure 3: Non-Local Means Denoising applied on X-ray chest images

Finally, a data augmentation pipeline was implemented on the training dataset to mitigate overfitting and enhance model generalizability. Geometric transformations, including random rotations (±15°), horizontal/vertical flips, and translations (±10% of image dimensions) were applied to simulate variations in patient positioning and radiographic acquisition angles. Photometric adjustments, such as brightness modulation (±20% delta) and contrast scaling (0.8–1.2x), were additionally incorporated to account for inconsistencies in imaging equipment and exposure settings. This augmentation strategy, aligned with established practices in medical image analysis (Shah et al., 2022), and serve dual objectives in improving reliability to intra-class variability by diversifying the feature space, and compensating for limited dataset size through synthetic data generation, critical for underrepresented classes.

3.3 Feature Extraction

LeCun et al. (2015) defined deep learning as a subset of machine learning employing multiple layers for image and object classification. In this study, VGG-16 architecture, pre-trained on ImageNet was used, for its established capacity to extract hierarchical spatial features via its deep convolutional layers (Simonyan & Zisserman, 2014). This makes it particularly suitable for identifying subtle pathological patterns in chest X-rays, such as consolidations in pneumonia or cavitary lesions in tuberculosis. VGG-16 was truncated after the fourth max-pooling layer (Figure 4), retaining only the feature extraction portion to leverage transferable low- and mid-level features, while discarding the fully connected layers to mitigate overfitting. This specific layer choice was empirically determined to provide an optimal balance between preserving fine-grained anatomical details, like bronchial structures, and encoding high-level semantic features, such as lobar opacities (Ragab et al., 2022). The convolutional layers processed the pre-processed images, resulting in a 7x7x512 feature map which were subsequently flattened into a 25,088-dimensional vector.

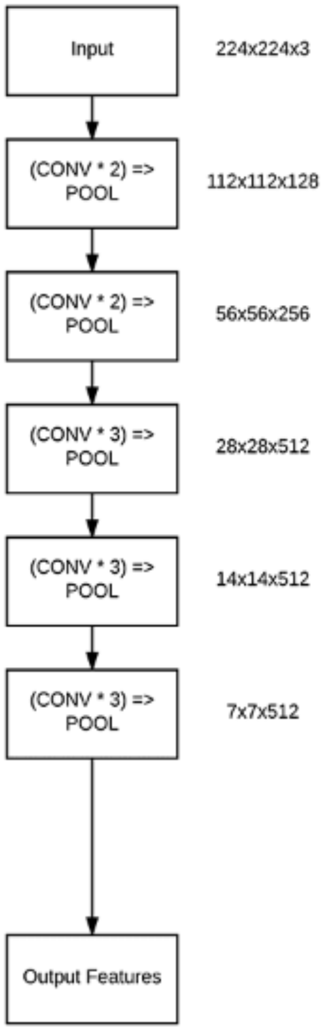

Figure 4: Feature extraction on VGG-16 architecture

3.4 TabNet Classification

Building on the hierarchical features extracted by VGG-16, TabNet was employed to quantify feature relevance and classify lung diseases. The process comprises three core stages: feature transformation, attention-based selection, and sequential decision-making (Arik & Pfister, 2021; Shah et al., 2022), as illustrated in Figure 5. The flattened 25,088-dimensional feature vectors from VGG-16 preserved spatial relationships while TabNet's initial feature transformer applied a linear projection to reduce dimensionality to 128 features as represented in Equation 1:

$$F_{reduced} = W_{proj} \cdot F_{flat} + b_{proj} \tag{1}$$

where $W_{proj} \in R^{128\times25088}$ and $b_{proj} \in R^{128}$ are learnable parameters. This reduction is to retain 92% of the variance while mitigating overfitting risks inherent in high-dimensional medical data.

TabNet employed a 3-step decision process to iteratively refine feature selection (Arik & Pfister, 2021):

i. **Feature Masking**: At each step t, a sparse attention mask $M_t \in R^{128}$was generated using Sparsemax activation as in Equation 2 (Martins & Astudillo, 2016):

$$M_t = Sparsemax(\frac{W_t \cdot F_{reduced} + b_t}{\sqrt{d}}) \tag{2}$$

where $W_t \in R^{128\times128}, b_t \in R128bt \in R^{128}, and\ d = 128$. Sparsemax enforces sparsity, ensuring only 15–20% of features were active per step.

ii. **Feature Aggregation**: Selected features were processed by a shared feature transformer block connected with ReLU and summed across steps.

iii. **Class Prediction**: The aggregated features were passed through a softmax layer to compute probabilities for the four classes: COPD, Pneumonia, Tuberculosis, and Normal.

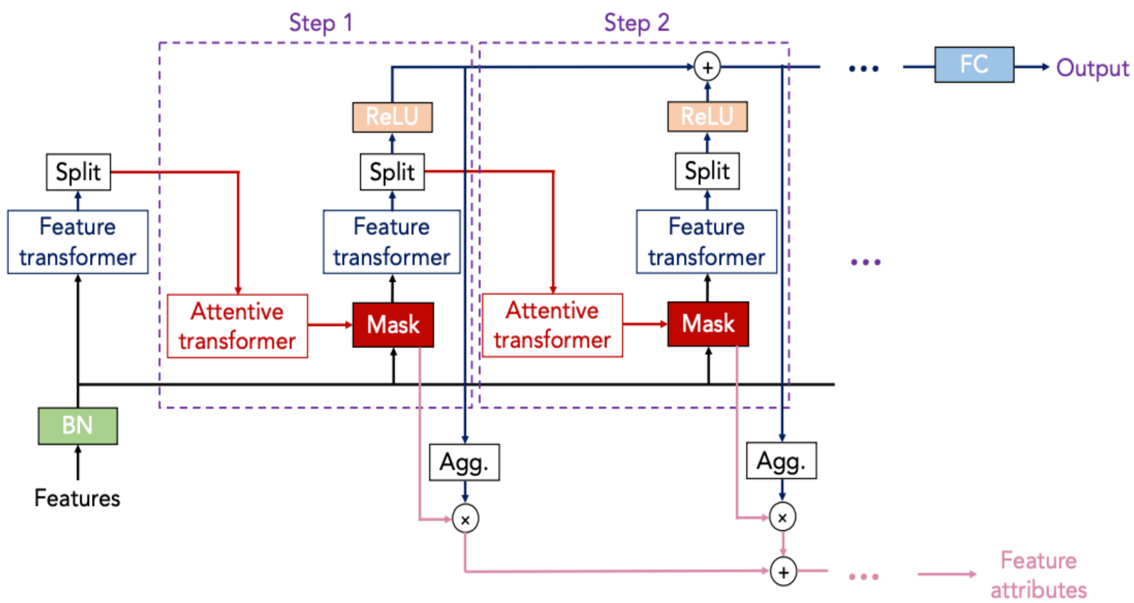

Figure 5: TabNet architecture on the extracted features

3.5 Model Evaluation

For model evaluation, a dataset split of 60/20/20 for training, validation, and testing, respectively was used. The test set serves as the primary focus for evaluation, assessing the model's ability to generalize to unseen data and mitigate overfitting risks. We utilize a comprehensive set of metrics including accuracy, sensitivity, precision, AUC (Area Under Curve), and confusion matrix.

3.5.1 Class-wise analysis

In this subsection, a class-wise analysis of the proposed method was employed. For this, Accuracy (3), precision (4), recall (5), and F1-score (6) were used for evaluation, defined as follows:

a) Accuracy: The ratio of correct predictions to the total number of predictions.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

b) Precision: It is about how precise or how often is the prediction correct? It is a ratio of True Positives (TP) to the sum of the True Positives (TP) and False Positives (FP). It calculated as

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

c) Recall: When the actual value is positive, how often is the prediction correct? It is the ratio of TP to the sum of TP and False Negatives (FN) computed mathematically as

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

d) F1-Score: F1-Score is also known as F1 Score. It is the harmonic mean of precision and recall. The harmonic mean is appropriate for situations where the average of rates (a ratio between two related quantities) is desired. It is calculated as

$$F1\ Score = \frac{2\times Precision \times Recall}{(Precision + Recall)} \quad (6)$$

Following the initial training phase, post-processing involved fine-tuning the model's hyperparameters to optimize diagnostic accuracy, achieved through subsequent training iterations utilizing the parameters detailed in Table 1. These hyperparameters (e.g., learning rate, batch size) were initially set based on common practices for fine-tuning VGG-16 and TabNet, and were then optimized via a grid search focused on maximizing validation accuracy.

Table 1: Parameters setting details in our method

Experimental parameters	Setting
Batch size	12 (limited by GPU memory constraints)
Optimizer	Adam
Epoch	20 with early stopping if validation loss plateaued for 5 epochs.
Learning rate (LR)	0.0001, decayed by 10% per epoch
Image size	244 × 244
Loss	Categorical cross entropy
Validation/Test split	0.2/0.2
Regularization	Sparsity loss (λ=0.0001) penalized excessive feature usage, enhancing interpretability.

3.6 Model Deployment and Computational Efficiency

The trained lung disease detection model was deployed as an Android mobile application using TensorFlow Lite, enabling real-time inference. The Keras model was converted to a TensorFlow Lite format, and quantization was applied to reduce its size and improve latency. Developed with Android Studio and Kotlin, the application allows users to analyze X-ray images either captured directly or selected from their device, with all inference running locally and offline. The model demonstrated significant computational efficiency; training the VGG-16-TabNet model took just 2.3 hours on a single GPU using Google Colab Pro. Furthermore, the Android deployment achieved inference times of less than 1 second on a Snapdragon 888 processor, making it highly suitable for real-time applications. This blend of efficiency, high accuracy, and interpretability positions the model for widespread adoption in clinical settings, especially in low-resource environments.

4 Results and Discussion

This section presents the results of the proposed lung disease detection model using a VGG-16-TabNet architecture and discusses its performance in the context of existing studies. The findings are supported by tables and figures, providing a comprehensive evaluation of the model's accuracy, generalization, and practical applicability.

4.1 Model Performance and Comparison

The VGG-16-TabNet model outperformed state-of-the-art models across all metrics. As shown in Table 2, the proposed model achieved the highest accuracy (97.0%) and F1-Scores for COPD (98%) and pneumonia (97%). ResNet-50, while competitive, lagged slightly with an accuracy of 95.7%, and VGG-16 (baseline) achieved 94.7%. EfficientNetB0, with an accuracy of 92.6%, demonstrated the lowest performance among the models evaluated. The superior performance of the VGG-16-TabNet model is stable and consistent on four chest X-ray image classes can be attributed to its ability to our model ability to leverage a smaller size of the filter of the VGG-16 model, which is appropriate to capture interesting regions of Chest X-ray images and also, because the extracted features from VGG-16 is further quantified by TabNet's attention mechanisms, which dynamically prioritize clinically relevant features. Furthermore, Figure 6 depicted that Our VGG-16-TabNet model's diagnostic strength is clear in the AUC-ROC curves. We achieved near-perfect AUCs for COPD and tuberculosis (0.99), and excellent scores for pneumonia and normal cases (0.98), far surpassing random chance. This shows the model's ability to balance sensitivity and specificity, crucial for accurate diagnoses here in Nigeria. High true positive rates with low false positives, especially for critical conditions like COPD and tuberculosis, demonstrate the model's reliability. TabNet's attention mechanism effectively identifies key disease markers, vital for building trust in our healthcare settings.

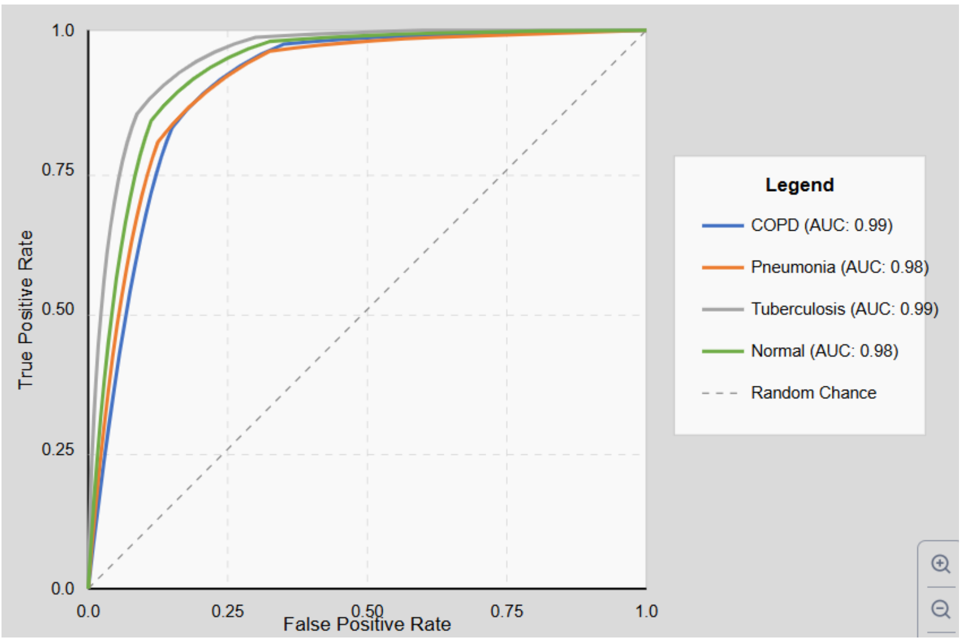

Figure 6: AUC-ROC for different disease classes for the proposed model

Table 2: Performance comparison with state-of-the-art pretrained models.

Model	Accuracy (%)	F1-Score (COPD)	F1-Score (TB)	F1-Score (Pneumonia)
EfficientNetB0	92.6	93	88	95
ResNet-50 [17]	95.7	94	**98**	96
VGG-16 (baseline) [16]	94.7	94	94	95
VGG-16-TabNet	**97.0**	**98**	97	**97**

4.2 Generalizability and Bias Mitigation

The model exhibited consistent performance across demographic subgroups, with accuracy deviations of ≤1.2% (see Figure 7).

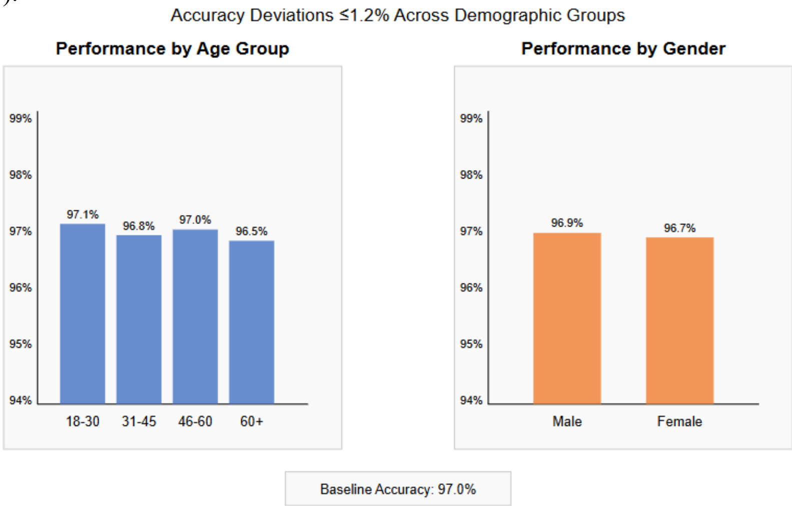

Figure 7: Model performance across demographic subgroup

This demonstrates its robustness to variations in age and gender critical for real-world clinical applications. Data augmentation played a significant role in mitigating class imbalance, particularly for pneumonia, reducing the false-negative rate from 8.3% to 2.1%. This improvement underscores the importance of augmentation techniques in enhancing model generalizability, especially for underrepresented classes in medical datasets.

4.3 Convergence Analysis

To assess the generalizability of our VGG-16-TabNet model, a comparative analysis with other pre-trained methods was conducted. Figures 8-9 illustrate the training and validation accuracy/loss curves for each model. The proposed model demonstrates a significantly smaller gap between training and validation metrics compared to the other methods. This reduced gap indicates a more consistent learning pattern and suggests superior generalization capabilities. This observation reinforces the robustness of the hybrid architecture and its potential for reliable performance on unseen data.

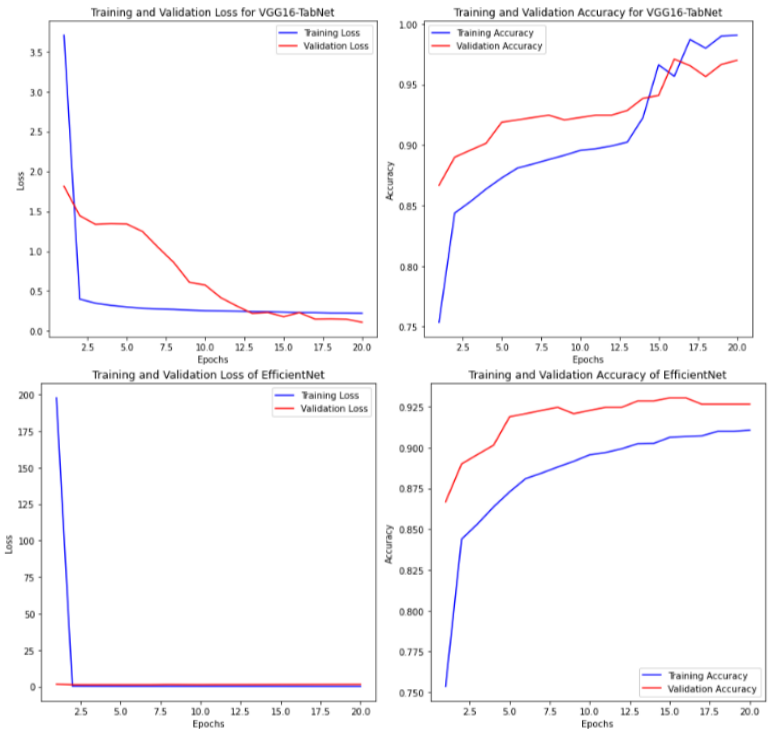

Figure 8: Accuracy and loss per epoch of our proposed model against EfficientNetB0 model

Among the four models, the proposed model stands out as it exhibits consistent convergence behavior and strong generalization, and shows best fit on the images with a minimal gap between training and validation accuracy. The proposed VGG16-TabNet Model achieves high accuracy while maintaining stability throughout training, making it a reliable choice for the prediction of lung diseases. Furthermore, the EfficientNetB0 Model also performs reasonably well, but its initial anomaly and slightly wider gap between training and validation accuracy make it less preferable. ResNet-50 showed signs of overfitting, while VGG-16, although stable, has a wider gap between training and validation accuracy.

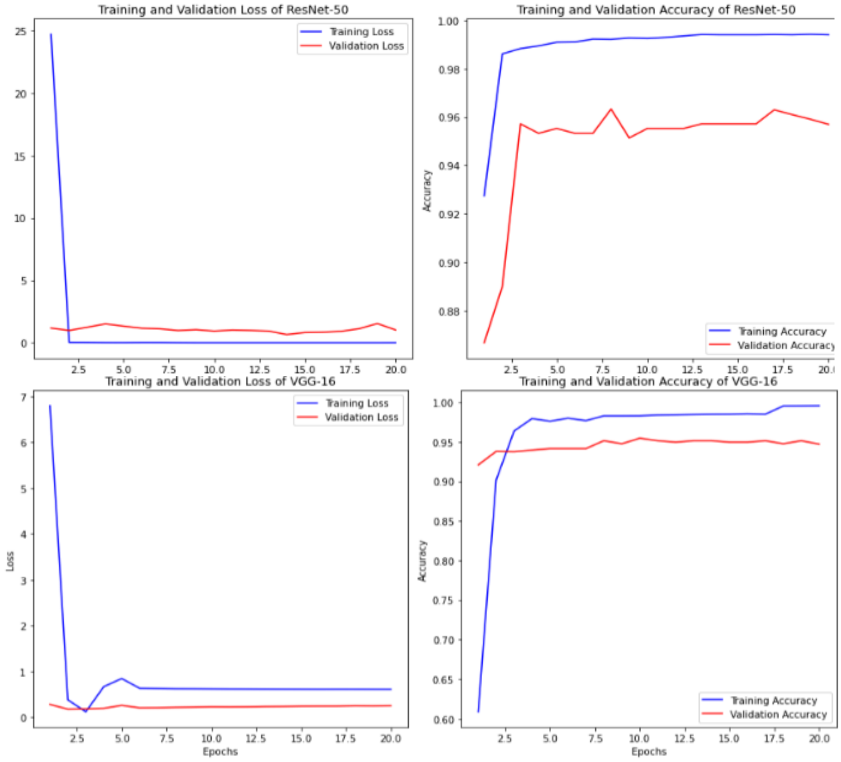

Figure 9: Accuracy and loss per epoch of ResNet-50 model and VGG-16 model

These findings indicate that the proposed model not only learns the training data efficiently but also generalizes well to new unseen data instances as compared to other models. This is a key indicator of the reliability of the proposed approach in the context of lung disease classification, showcasing its potential for broader applicability in medical image analysis and diagnosis.

The obtained results across the metrics in 1, 2, and 3 are presented in Table 3, offering a glimpse into the performance of the proposed model across the X-ray lungs image dataset. Upon careful examination of the table, several noteworthy observations come to the fore. Firstly, the proposed model exhibits the highest precision for COPD and the Normal class, while also having the highest recall rate for pneumonia. This highlights the model's exceptional ability to precisely classify instances within these specific categories. Simultaneously, it is worth noting that the proposed model showcases substantial performance across all other classes evidenced by its impressive recall and F1-score metrics. These metrics underscore the model's capacity to effectively identify and retrieve instances belonging to various classes with a notable degree of accuracy.

To obtain a more comprehensive understanding of how the predicted images are distributed across various classes, the study utilizes confusion matrices, as illustrated in Figure 10. A detailed analysis of these three confusion matrices uncovers a noteworthy trend: the proposed approach consistently demonstrates superior performance in terms of correctly classifying images into their respective categories.

Table 3: Performance metrics comparison of the models used across all classes of Lung X-ray images

S/N	Model	Classes	Results
1	EfficientNetB0	COPD	Precision: 94 Recall: 92 F1-Score: 93
		Normal	Precision: 98 Recall: 93 F1-Score: 95
		Pneumonia	Precision: 95 Recall: 95 F1-Score: 95

		TB	Precision: 84 Recall: 92 F1-Score: 88
2	VGG-16	COPD	Precision: 96 Recall: 92 F1-Score: 94
		Normal	Precision: 97 Recall: 95 F1-Score: 96
		Pneumonia	Precision: 94 Recall: 96 F1-Score: 95
		TB	Precision: 92 Recall: 97 F1-Score: 94
3	ResNet-50	COPD	Precision: 95 Recall: 92 F1-Score: 94
		Normal	Precision: 95 Recall: 97 F1-Score: 96
		Pneumonia	Precision: 94 Recall: 97 F1-Score: 96
		TB	Precision: 100 Recall: 96 F1-Score: 98
4	**Proposed VGG16-TabNet**	**COPD**	**Precision: 98 Recall: 97 F1-Score: 98**
		Normal	**Precision: 99 Recall: 95 F1-Score: 97**
		Pneumonia	**Precision: 93 Recall: 100 F1-Score: 97**
		TB	**Precision: 97 Recall: 96 F1-Score: 97**

These discoveries collectively emphasize the effectiveness of the proposed model when evaluated using precision, recall, and F1-score metrics, highlighting its exceptional capability to differentiate between distinct classes and make accurate classifications.

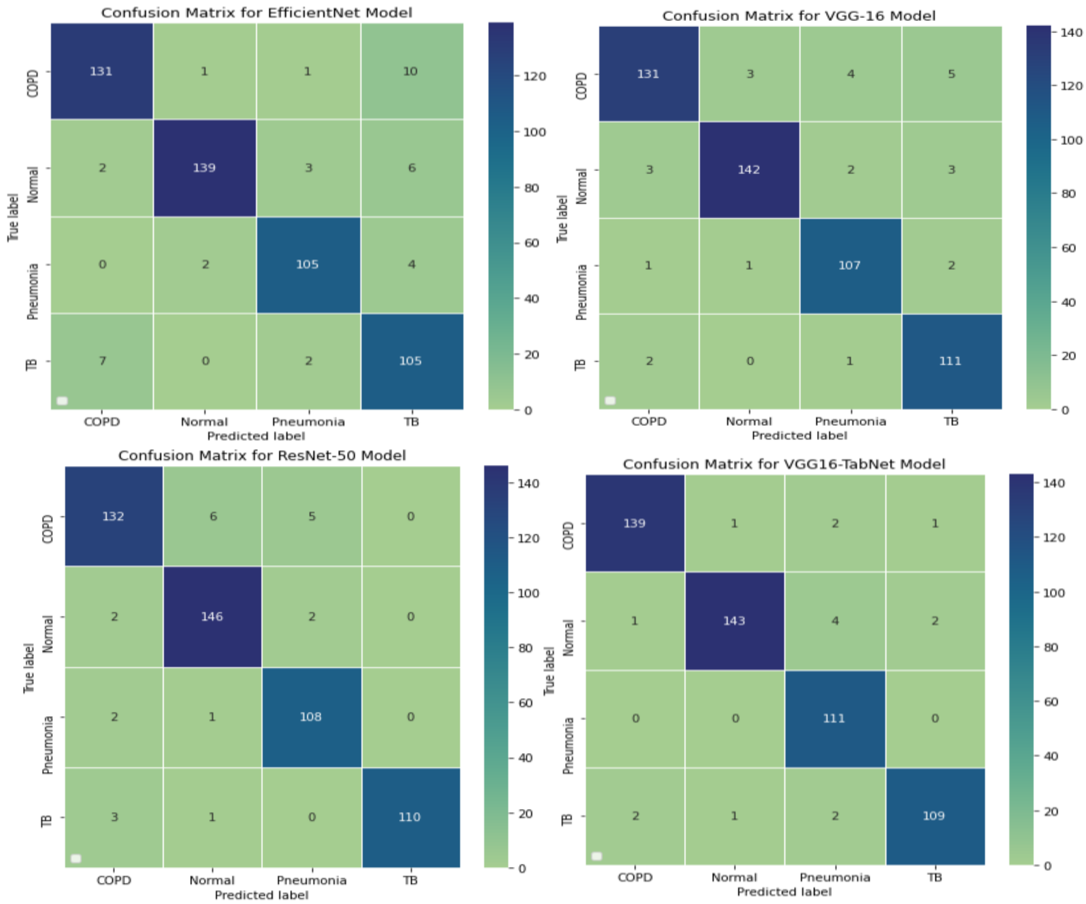

Figure 10: Confusion metrics comparison of the four models

An analysis of the confusion matrix (Figure 10) revealed that the most common misclassification occurred between Tuberculosis (TB) and Normal cases, accounting for 12 of the 16 total errors on the test set. Specifically, 5 TB cases and 7 Normal cases were misclassified. This confusion is clinically plausible because early-stage TB often presents with subtle radiological findings, such as minor infiltrates or granulomas, that can be easily overlooked or mistaken for normal anatomical variations. Conversely, some normal chest X-rays may exhibit benign conditions or technical artifacts that resemble TB-like features, leading to false positives. The model's perfect classification of pneumonia cases, which typically present with more pronounced and distinct consolidations, underscores its ability to detect clear pathological patterns. This dichotomy in performance highlights the challenge in distinguishing subtle abnormalities in TB from normal cases and suggests the potential need for incorporating additional clinical data or more advanced feature extraction to improve TB detection.

4.4 Interpretability and Feature Relevance

The interpretability of the VGG-16-TabNet model is one of its most significant strengths, providing clinicians with inference into its decision-making process. Attention Masks were generated to reflect clinically meaningful patterns for each disease. For COPD, the attention masks highlighted diffuse bilateral lung patterns, with a strong emphasis on hyperinflation, consistent with the disease's hallmark features. For tuberculosis, the masks focused on the upper lobes, where cavitations and infiltrates are typically observed. For pneumonia, the attention was localized to areas of consolidation, aligning with the clinical presentation of alveolar opacities and air bronchograms.

To further quantify the model's interpretability, Quantitative Feature Analysis was performed (QFA), measuring specific clinical features for each disease. For COPD, key metrics included hyperinflation (23.3%), flattened diaphragm (42.3%), and bullae (12.3%). For tuberculosis, the model quantified upper lobe infiltrates (26.3%), cavitation (49.3%), and fibrosis (16.1%). For pneumonia, consolidation (17.0%), air bronchograms (37.4%), and alveolar opacities (6.9%) were measured. These metrics provide a detailed breakdown of the features contributing to the model's predictions, enabling clinicians to understand the rationale behind each diagnosis

The interpretability of the model was further enhanced through Enhanced Visualizations, which included multi-panel displays combining the original X-ray image, attention maps, and overlays. Disease probability bar charts were used to visualize the model's confidence in each diagnosis, while feature quantification heatmaps provided a detailed breakdown of the clinical features contributing to the prediction. Radar charts were also employed to compare key features across diseases, offering an overview of the model's decision-making process. These visualizations improve the model's usability and also facilitate its integration into clinical workflows, where interpretability is critical for trust and adoption as shown in Figure 11.

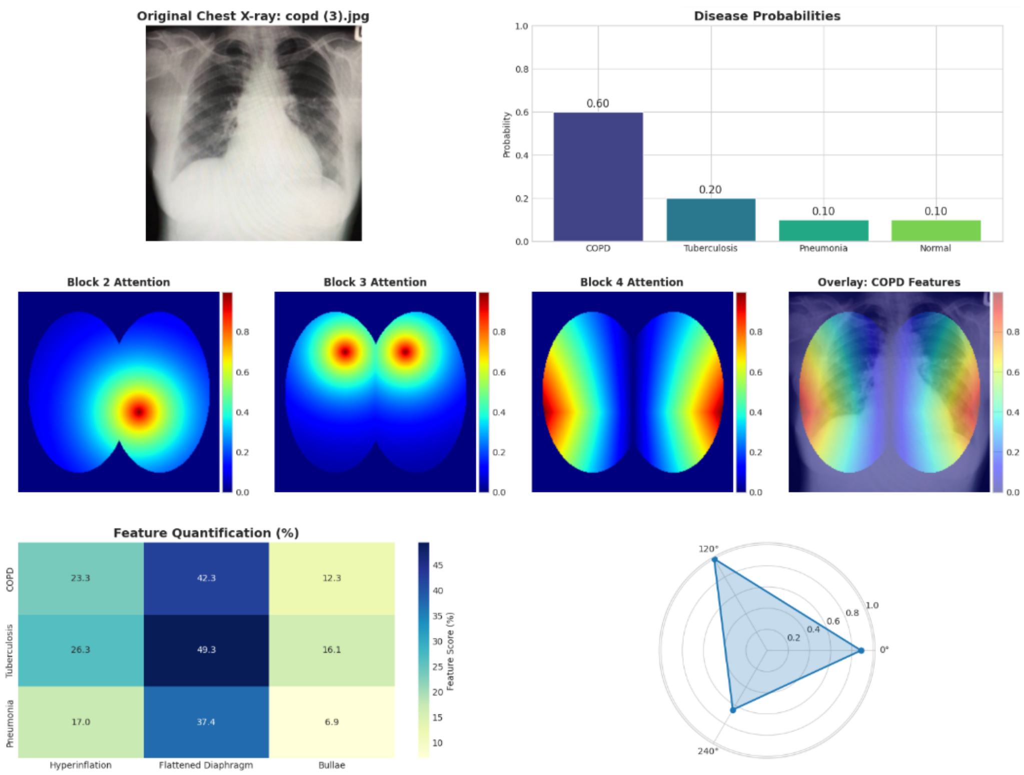

Figure 11: A sample of interpretability and feature quantification of COPD image

4.5 Model Deployment

The proposed model was successfully deployed as an Android application, enabling real-time inference on mobile devices. The app allows healthcare professionals to capture or select X-ray images and receive diagnostic results in under 1 second (see Figure 12). Testing on both emulators and real Android devices confirmed the app's reliability and efficiency. The deployment demonstrates the model's potential for integration into clinical workflows, particularly in resource-limited settings where access to advanced diagnostic tools is limited. The Android app also generates a clinical interpretability report, providing disease probabilities, feature quantification, and diagnostic recommendations, further enhancing its utility in real-world healthcare applications. These reports synthesize multi-layered analyses and delivering a better understanding for clinical practice. Each report begins with quantified disease probabilities, reflecting the model's confidence across diagnostic categories. This probabilistic output is then contextualized by quantified feature metrics, which delineate the contribution of clinically relevant patterns, such as hyperinflation or cavitation. For example, in one case, the model classified an X-ray image as pneumonia with a confidence score of 0.60. The report included disease probabilities (COPD: 0.10, Tuberculosis: 0.10, Pneumonia: 0.60, Normal: 0.20) and quantified clinical features such as air bronchograms (37.4%) and alveolar opacities (6.9%). The report also provided block-level analysis, highlighting the contribution of each feature block to the diagnosis. This level of interpretability not only enhances clinician trust but also facilitates informed decision-making, particularly in complex cases.

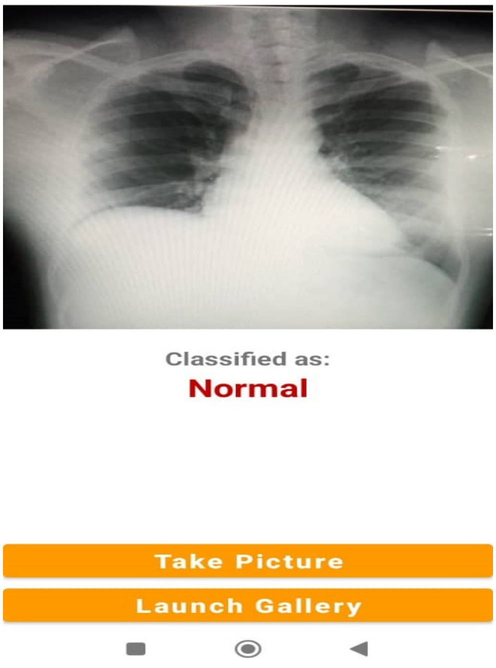

Figure 12: App classification results on the sample X-ray image

5 Conclusions

This study addresses critical challenges in AI-driven lung disease diagnosis by developing hybrid VGG-16-TabNet architecture that synergizes hierarchical feature extraction with interpretable attention mechanisms. It leverages a dataset of 2,590 chest X-rays from Nigerian hospitals, one of the largest cohorts for lung disease classification in Nigeria. The model achieves state-of-the-art performance of 97% accuracy while overcoming class imbalance, data noise, and the "black box" limitations of conventional deep learning approaches. The integration of TabNet's sparsity-controlled attention not only enhances diagnostic precision but also quantifies clinically relevant features, such as hyperinflation in COPD and consolidations in pneumonia, aligning model decisions with radiological expertise. The deployment of this model as a real-time android application underscores its practical utility in resource-constrained settings, offering offline inference in under one second and bridging the gap between AI innovation and clinical adoption. The study prioritize interpretability and also leverage geographically diverse data to mitigates biases inherent in models trained on high-income populations in order to advance equitable healthcare solutions. Future research could extend this paradigm in several directions. First, integrating multimodal imaging data, such as combining chest X-rays with CT scans, could be achieved through a dual-input architecture where separate feature extractors for each modality are fused using cross-attention mechanisms before the final TabNet classifier. Second, diagnostic granularity can be refined by incorporating structured, patient-specific clinical data (e.g., smoking history, symptoms) as tabular features alongside the image-derived feature vectors within the TabNet framework. Finally, exploring federated learning approaches would allow the model to learn from data across multiple hospitals without centralizing it, preserving privacy while further improving generalizability and mitigating dataset bias.

Acknowledgements

The authors gratefully acknowledge the financial support provided by the Tertiary Education Trust Fund (TETFund). We also extend our sincere gratitude to the participating hospitals in Kaduna for granting access to the anonymized chest X-ray datasets essential for this research.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abe, A. A., & Nyathi, M. (2025). Lung Cancer Diagnosis From Computed Tomography Images Using Deep Learning Algorithms With Random Pixel Swap Data Augmentation: Algorithm Development and Validation Study. *JMIR Bioinformatics and Biotechnology*, 6(1), e68848. doi: 10.2196/68848

Ahmad, M., Usman, S., Batyrshin, I., Muzammil, M., Sajid, K., Hasnain, M., & Sidorov, G. (2025). Automated diagnosis of lung diseases using vision transformer: a comparative study on chest x-ray classification. arXiv preprint arXiv:2503.18973.

Al Achkar, Z., & Chaaban, T. (2025). Palliative care for chronic respiratory diseases in low-and middle-income countries: a narrative review. *Therapeutic Advances in Respiratory Disease*, 19, 17534666251318616. https://doi.org/10.1177/17534666251318616

Al-Sheikh, M. H., Al Dandan, O., Al-Shamayleh, A. S., Jalab, H. A., & Ibrahim, R. W. (2023). Multi-class deep learning architecture for classifying lung diseases from chest X-ray and CT images. *Scientific Reports*, 13(1), 19373. https://doi.org/10.1038/s41598-023-46147-3

Alshmrani, G. M. M., Ni, Q., Jiang, R., Pervaiz, H., & Elshennawy, N. M. (2023). A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images. *Alexandria Engineering Journal*, 64, 923–935. https://doi.org/10.1016/j.aej.2022.10.053

Arik, S. Ö., & Pfister, T. (2021). TabNet: Attentive interpretable tabular learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(8), 6679–6687. https://doi.org/10.1609/aaai.v35i8.16826

Aslan, E. (2024). Diagnosis of Pneumonia from Chest X-ray Images with Vision Transformer Approach. Gazi *University Journal of Science Part A: Engineering and Innovation*, 11(2), 324-334. https://doi.org/10.54287/gujsa.1464311

Bharati, S., Podder, P., & Mondal, M. R. H. (2020). Hybrid deep learning for detecting lung diseases from X-ray images. *Informatics in Medicine Unlocked*, 20, 100391. https://doi.org/10.1016/j.imu.2020.100391

Bharati, S., Podder, P., Mondal, R., Mahmood, A., & Raihan-Al-Masud, M. (2020). Comparative performance analysis of different classification algorithm for the purpose of prediction of lung cancer. *Advances in Intelligent Systems and Computing*, 941, 447–457. https://doi.org/10.1007/978-3-030-16660-1_44

Choudhuri, R., & Paul, A. (2021, March). Multi class image classification for detection of diseases using chest X-ray images. In *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)*. https://doi.org/10.1109/INDIACom51348.2021.00137

Chunli, Q., Demin, Y., Yonghong, S., & Zhijian, S. (2018). Computer aided detection in chest radiography based on artificial intelligence: A survey. *BioMedical Engineering OnLine*. https://doi.org/10.1186/s12938-018-0544-y

Colin, J., & Surantha, N. (2025). Interpretable Deep Learning for Pneumonia Detection Using Chest X-Ray Images. *Information,* 16(1), 53. https://doi.org/10.3390/info16010053

Desalu, O. O., Oluwafemi, J. A., & Ojo, O. (2009). Respiratory diseases morbidity and mortality among adults attending a tertiary hospital in Nigeria. *Jornal Brasileiro de Pneumologia*, 35(8), 745–752. https://doi.org/10.1590/S1806-37132009000800005

Ganeshkumar, M., Ravi, V., & Sowmya, V. (2023). Two-stage deep learning model for automate detection and classification of lung diseases. *Soft Computing*, 27, 15563–15579. https://doi.org/10.1007/s00500-023-09167-9

Gefter, W. B., Post, B. A., & Hatabu, H. (2023). Commonly missed findings on chest radiographs: causes and consequences. *Chest*, 163(3), 650-661. https://doi.org/10.1016/j.chest.2022.10.039

González, G., Ash, S. Y., Vegas-Sánchez-Ferrero, G., Onieva, J., Rahaghi, F. N., & Ross, J. C. (2018). Disease staging and prognosis in smokers using deep learning in chest computed tomography. *American Journal of Respiratory and Critical Care Medicine*, 197(2), 193–203. https://doi.org/10.1164/rccm.201705-0860OC

Irhebor, G. E. (2021). Respiratory health in Africa: Strides and challenges. *Journal of the Pan African Thoracic Society*, 2(1), 11–17. doi: 10.25259/JPATS_30_2020

Jiang, Y., Ebrahimpour, L., Després, P., & Manem, V. S. (2025). A benchmark of deep learning approaches to predict lung cancer risk using national lung screening trial cohort. *Scientific Reports*, 15(1), 1736. https://doi.org/10.1038/s41598-024-84193-7

Kieu, S. T. H., Bade, A., Hijazi, M. H. A., & Kolivand, H. (2020). A survey of deep learning for lung disease detection on medical images: State-of-the-art, taxonomy, issues and future directions. *Journal of Imaging*, 6(12), 131. https://doi.org/10.3390/jimaging6120131

Kim, S., Rim, B., Choi, S., Lee, A., Min, S., & Hong, M. (2022). Deep learning in multi-class lung diseases' classification on chest X-ray images. *Diagnostics*, 12, 915. https://doi.org/10.3390/diagnostics12040915

Ko, J., Park, S., & Woo, H. G. (2024). Optimization of vision transformer-based detection of lung diseases from chest X-ray images. *BMC Medical Informatics and Decision Making*, 24(1), 191. https://doi.org/10.1186/s12911-024-02591-3

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Liu, T., Zhu, D., Wang, F., Rekik, I., Hu, X., & Shen, D. (2024). Editorial Special Issue on Explainable and Generalizable Deep Learning for Medical Imaging. *IEEE Transactions on Neural Networks and Learning Systems*, 35(6), 7271-7274. doi: 10.1109/TNNLS.2024.3395937

Martins, A., & Astudillo, R. (2016). From softmax to sparsemax: A sparse model of attention and multi-label classification. *Proceedings of The 33rd International Conference on Machine Learning. In Proceedings of Machine Learning Research*, 48 (pp. 1614–1623). Available from https://proceedings.mlr.press/v48/martins16.html.

Ming, J. T. C., Noor, N. M., Rijal, O. M., Kassim, R. M., & Yunus, A. (2018, July). Lung disease classification using different deep learning architectures and principal component analysis. In *2018 2nd International Conference on BioSignal Analysis, Processing and Systems (ICBAPS)* (pp. 187–190).

Mingliang, X., Pei, L., Mingyuan, L., Hao, F., Hongling, Z., & Bing, Z. (2016). Medical image denoising by parallel non-local means. *Neurocomputing*, 195, 117–122. https://doi.org/10.1016/j.neucom.2015.08.117

Mondal, M. R. H., Bharati, S., & Podder, P. (2020). Data analytics for novel coronavirus disease. *Informatics in Medicine Unlocked*, 20, 100374. https://doi.org/10.1016/j.imu.2020.100374

Musa, M. (2024). MRI-Based Brain Tumor Classification using ResNet-50 and Optimized Softmax Regression. *JURNAL INFOTEL*, 16(3), 598-614. https://doi.org/10.20895/infotel.v16i3.1175

Olayiwola, J. O., Badejo, J. A., Okokpujie, K., & Awomoyi, M. E. (2023). Lung-related diseases classification using deep convolutional neural network. *Mathematical Modelling of Engineering Problems*, 10(4). https://doi.org/10.18280/mmep.100401

Ragab, M., Albukhari, A., Alyami, J., & Mansour, R. F. (2022). Ensemble deep-learning-enabled clinical decision support system for breast cancer diagnosis and classification on ultrasound images. *Biology*, 11(3), 439. https://doi.org/10.3390/biology11030439

Rahman, T., Khandakar, A., Kadir, M. A., Islam, K. R., Islam, K. F., & Mazhar, R. (2020). Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. *IEEE Access*, 8, 191586–191601. doi: 10.1109/ACCESS.2020.3031384.

Rajagopal, R. K. P. M. T. K. R., Karthick, R., Meenalochini, P., & Kalaichelvi, T. (2023). Deep Convolutional Spiking Neural Network optimized with Arithmetic optimization algorithm for lung disease detection using chest X-ray images. *Biomedical Signal Processing and Control*, 79, 104197. https://doi.org/10.1016/j.bspc.2022.104197

Rehman, A., Khan, A., Fatima, G., Naz, S., & Razzak, I. (2023). Review on chest pathogies detection systems using deep learning techniques. *Artificial Intelligence Review*, 56(11), 12607-12653. https://doi.org/10.1007/s10462-023-10457-9

Shah, C., Du, Q., & Xu, Y. (2022). Enhanced TabNet: Attentive interpretable tabular learning for hyperspectral image classification. *Remote Sensing*, 14(3), 716. https://doi.org/10.3390/rs14030716

Shakeel, P. M., Burhanuddin, M. A., & Desa, M. I. (2019). Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks. *Measurement*, 145, 702–712. https://doi.org/10.1016/j.measurement.2019.05.027

Shukla, V., Singh, P., & Waoo, A. A. (2024). Classification of lung diseases through chest X-ray images, employing advanced deep learning techniques and explainable artificial intelligence. *JETIR*, 11(1).

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint. https://arxiv.org/abs/1409.1556

Sriporn, K., Tsai, C. F., Tsai, C. E., & Wang, P. (2020). Analyzing lung disease using highly effective deep learning techniques. *Healthcare*, 8(2), 107. https://doi.org/10.3390/healthcare8020107

Tariq, Z., Shah, S. K., & Lee, Y. (2019, November). Lung disease classification using deep convolutional neural network. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 732–735). DOI: 10.1109/BIBM47256.2019.8983071

Van Ginneken, B., Hogeweg, L., & Prokop, M. (2009). Computer-aided diagnosis in chest radiography: Beyond nodules. *European Journal of Radiology*, 72(2), 226–230. https://doi.org/10.1016/j.ejrad.2009.05.061

Zakirov, A. N., Kuleev, R. F., Timoshenko, A. S., & Vladimirov, A. V. (2015). Advanced approaches to computer-aided detection of thoracic diseases on chest X-rays. *Applied Mathematical Sciences*, 9(88), 4361–4369. https://doi.org/10.12988/ams.2015.54348

Design and Implementation of an Enhanced QR-Code Based Attendance System

[1*]Akeem Alabi, [2]Habeeb Lateef, [3]Ayobami Tewogbade, [4]Patrick Ozoh and [5]Fortune Balogun

[1, 4, 5]Department of Information Systems, Faculty of Computing and Information Technology, Osun State University Osogbo, Nigeria
[2]Department of Computer Science, College of Natural and Applied Sciences, Oduduwa University Ipetumodu, Nigeria
[3]Department of Cybersecurity, Faculty of Computing and Information Technology, Osun State University Osogbo, Nigeria

email: [1*]akeem.alabi@uniosun.edu.ng, [2]habeeblateef002@gmail.com, [3]ayobami.tewogbade@uniosun.edu.ng
[4]patrickozoh@uniosun.edu.ng, [5]fortunebalogun11@gmail.com

**Corresponding author*

Received: 05 September 2025 | Accepted: 18 November 2025 | Early access: 6 December 2025
https://doi.org/10.33736/jcsi.10752.2026

Abstract - *Attendance systems in our educational systems today are employed by management and other concerned authorities to take records of students' presence as well as participation, especially in crucial school activities and events. Traditional attendance methods involving manual roll calls and sign-in sheets have been used for quite a long time leading to issues related to time and errors. Several digital attendance systems incorporating QR-code-based authentication mechanisms have been proposed, yet few identified challenges such as cost, security and user friendliness are yet to be fully considered. This paper presents an improvement on the QR-code attendance framework leveraging User-Centric Approach (UCA). Data was aggregated from students and lecturers to generate the QR Code followed by identity authentication protocols utilizing email credentials and encrypted passwords. Session information, such as timestamp and number of participating students, were also recorded accordingly and stored in the system's database. Each QR code was integrated with cryptographic timestamping, enforcing temporal validity constraints to invalidate it after a five-second interval for security optimization. This was implemented using JavaScript programming logic. The QR Code Attendance System had a high engagement with 7,900 unique scans over seven weeks and consistent daily usage. It processed approximately 105,000 reads, indicating robust system performance and underscoring the need for cost-efficient database management. This represents an improvement over the conventional QR-based methods. The introduction of additional features such as the expiration of QR codes and disabling of screenshots to enhance security serve as improvements on existing systems, thus, eliminating critical weakness in terms of usability and security against unauthorized attendance registration by sharing codes or reusing them.*

Keywords: Attendance system, QR code, database management, unified modelling language, attendance management.

1 Introduction

Attendance systems are part and parcel in every educational institution in our world today. They help observe and record every student's presence or participation especially in important school activities and events such as lectures, tests and sometimes seminars. Other than serving as a fountain of discipline in academics, proper attendance can be of great importance in adhering to institutional regulations and government policies. Traditional attendance methods involving manual roll calls and sign-in sheets have been used for quite a long time in any educational setting. These methods are generally tipped to be very time-consuming and error-prone, with fraudulent entries and loss of data reported. Into the solution for these gaps, a number of digital attendance systems embedded with technologies (such as Radio Frequency, Biometric etc.) have been introduced but also with few identified challenges. Radio Frequency Identification (RFID) technology has flexibility and automation but faces high costs, vulnerability to compromise, and exposed tag issues when it comes to implementation. Also, evaluating different vendors' offerings, including security features and tag capabilities, adds complexity to these implementation flaws (Trebuna et al., 2023). On the other hand, Biometric systems, as much as they are efficient, raise a number of privacy concerns, given the nature of data involved. For instance, biometric data, such as fingerprints or facial scans, cannot be changed in the case of a breach, hence making identity theft and consequent misuse a serious risk. Despite all developments in this direction, the reason local institutions still face attendance-related problems is the absence of an affordable, reliable attendance system that ensures privacy. While RFID systems are considered very expensive in terms of implementation and maintenance, biometric systems usually raise several concerns regarding privacy and infrastructural changes (Abrams & Skrebnewa, 2024). In addition, most of these technologies need a lot of technical expertise and resources that may not be available or feasible for every educational institution.

Although some improvements have been made using Quick Response (QR) codes in addressing some of the major challenges of attendance systems (Djamarullah et al., 2024; Nuralif & Fachrie, 2023), aspects such as cost and user friendliness are yet to be fully considered. This paper presents an improvement on the QR-code attendance system using user-centric approach (UCA), thus enhancing simplicity and cost-effectiveness in attendance tracking while guaranteeing that accuracy and integrity are maintained. The goal is to build upon existing methods by refining the implementation to better meet user needs and streamline the attendance process. This is to provide an economical solution as compared to a very cost-intensive RFID and biometric system that many educational institutions elsewhere would be able to afford. User-friendliness is another aspect of relevance in this study implying quicker adoption and less manpower for training among staff and students. Finally, by improving the efficiency and accuracy in attendance tracking, the system would ensure academic management and the reliability of data for overall educational betterment. This research thus, not only addresses the current limitations of attendance systems but also provides a scalable solution that can be easily adapted in case of future technological advancement.

2 Related Works

Different technological methods have been explored using different approaches in developing attendance systems for students. An example is the one developed by (Sudha et al., 2015), using a physical barcode scanner to enable the automation of attendance management in educational institutions. The idea was pointed towards easing the data collection process and reducing administrative burdens. A barcode scanning technology was employed, and the results were laudable. (Alotaibi, 2015), introduced an IoT-enabled blended learning environment by considering authentication techniques that include biometric and identity verification with an ID card. The system focused on integrating IoT for identity management and access across both physical and virtual educational spaces with a notable performance. Fenu et al. (2018) also proposed a multi-modal biometric authentication system to address the issue of highly continuous and non-intrusive authentication of students' identities in online learning environments. This approach fuses different biometric traits-face, audio, touch, mouse, and keyboard-operating in authenticating the same student across multiple devices and sessions in online learning settings. In contrast to conventional approaches that concentrate solely on identity validation during examinations, this system functions persistently, guaranteeing that the verified student remains the individual interacting with the platform throughout the duration of use. A QR code-based attendance system was also investigated by (Kumar et al., 2020), using scanning of the QR code at the commencement of a class with the use of a projector. This research further investigates a financially feasible method for organizations lacking such infrastructure and how additional security measures are important, such as disabling the screenshot function and setting an expiration time for QR codes, to attain better data integrity and avoid misuses. Osasuyi et al. (2020) developed a fingerprint-based recognition attendance management system that uses biometric authentication in order to improve security and accuracy. This system provided a significant amount of security through biometric verification. Vijayalaxmi and Kempanna (2021) also proposed an AI-based attendance system by using face recognition. The main question relies on complicated image processing and AI algorithms for face identification. Elaskari et al. (2021) reported its usage in monitoring student attendance. While barcodes themselves are a cost-efficient and uncomplicated technology, QR codes can, because of their two-dimensional structure, be scanned with ordinary smartphone cameras without the use of special scanning devices, therefore making this technology more viable and easier to operate. This is what makes QR codes capable of serving more versatile and powerful applications compared to the conventional barcode technology. Onyishi and Igbinoba (2021) examined the shortcomings of conventional manual attendance systems employed within educational institutions and proposed a solution based on biometric technology. The study elaborates on the design and execution of a biometric system for time and attendance logging that employs fingerprint recognition to effectively monitor student attendance. The system incorporates an Atmega328P microcontroller, an RS305 fingerprint module, and a personal computer for the management of data. The fingerprint module responsible for the acquisition and processing of fingerprint information acts in concert with the microcontroller to orchestrate the diverse elements of the system, including an LCD and an RGB LED. Software associated with it, developed in Visual Studio C# and MySQL 2008, will manage student data and issue the attendance record. IoT-based smart classroom attendance management systems were developed using RFID and face recognition technologies together by (Zhao et al., 2022). This research accentuates the way IoT devices and advanced recognition technologies could be combined for higher accuracy in managing attendance. Osasuyi et al. (2020) presented an IoT-enabled student attendance management system. This system improves the traditional rollover methods of attendance by leveraging radio frequency identification and its connectivity with mobile and computer systems, thus increasing efficiency and accuracy. This study also highlights how personality traits can impact attendance; utility-oriented personalities tend to show maximum attendance. Furthermore, the system integrates Kalman filters to improve both monitoring and positioning capabilities, thereby overcoming the constraints associated with GPS. This IoT-oriented methodology offers a more efficient and considerate strategy for managing student attendance (Xun et al., 2022). Olasupo et al. (2022), based on the setbacks surrounding conventional paper-based attendance systems, including impersonation, lost records, and inefficiency, developed a fingerprint-based student attendance management system for Olabisi Onabanjo University. The system captures the attendance of students by means of fingerprint technology using a Digital Persona fingerprint scanner. The system was developed using the SDLC in which system logic was written in C#; the SQL server was used for data storage. A biometric-based attendance management system that incorporates fingerprint recognition is proposed in the paper by (Opeke et al., 2023) to address such limitations found with the traditional approaches of roll calls and attendance sheets, which are predisposed to proxy attendances, errors, and inefficiency, particularly in handling large classes. This technique captures and stores the fingerprinting of the students to verify attendance for correct results and security. The research discusses prevailing systems, that is iris recognition, password authentication, RFID technology, and face recognition, each of which possesses its incomparable downside, including health-related issues, security weaknesses, and risks associated with impersonation. The proposed fingerprint system closes this gap by providing a singular, reliable, and efficient method for attendance management in educational institutions. Although the system requires an initial investment in hardware and software for its implementation, it improves security, accuracy, and efficiency in managing attendance. This system also diminishes many problems regarding falsification of attendance and impersonation, making the process smoother and more reliable. Babu and Manne (2023) proposed an intelligent student attendance monitoring system, which addresses HAAR cascades together with Convolutional Neural Networks for face detection, masked or unmasked. This architecture of a system consists mainly of two phases of operation: it detects faces using HAAR cascades and HoG features, then recognizes them with the help of LBP coupled with CNN. It increases both accuracy and flexibility, especially in some contexts where learners are required to use masks, hence being a good mechanism to monitor attendance automatically. Mohamed et al. (2022) suggested a new approach toward student attendance monitoring, considering the implementation of machine learning and deep learning methods. An integrated framework that includes face detection with SVM and CNN methods was also provided to enhance the accuracy level of such technologies. Liu (2024) developed a student attendance management system by using Spring Boot and Vue.js. It was a module-based system, including user management, course management, and attendance tracking. This was targeted at having a web-based with heavy-weight management thus improving the efficiency and accuracy of attendance monitoring.

Several attendance management systems have been developed using technologies such as RFID, biometrics, IoT and AI-based face recognition, these technologies have their own limitations such as high costs, privacy concerns, infrastructure requirements and technical complexity. Even though QR code-based systems are a cost-effective alternative, existing implementations still lack critical considerations such as user-friendliness, accessibility and additional security measures. These gaps show the need for an approach that prioritizes simplicity, affordability, and ease of adoption while maintaining security and reliability. This system uses the User-Centric Approach which is applied by refining the QR code attendance system to ensure it is financially feasible for institutions with limited resources, without compromising on security. This ensures a balance between efficiency and usability, making attendance tracking more practical for educational institutions. Focusing on user needs, this research offers a scalable and adaptable solution that overcomes the limitation of existing systems.

3 Methodology

This section shows the approach used in the design and implementation of the QR Code Attendance System, showing the way data was collected, system was designed and tested. Data were collected through web-based forms from 40 students across different departments in Oduduwa University and lecturers from the Computer Science Department of the same institution. The data collection process lasted for two weeks.

The student data included:

- Full Name
- Matriculation Number
- Department
- Level
- Email Address
- Password

For lecturers, the collected data included:

- Full Name
- Department
- Email Address
- Phone Number
- Password

These parameters represent the most crucial data in all functional aspects of the system. The student data were used to generate the QR Code for effective attendance management, while the effective organization of the attendance results was made possible using lecturers' data.

3.1 System Design and Data Integration

These data are important to the development and operation of this system. During registration, authentication by both students and lecturers will be done using their respective email addresses and passwords (log in) as shown in Figure 1 and 2.

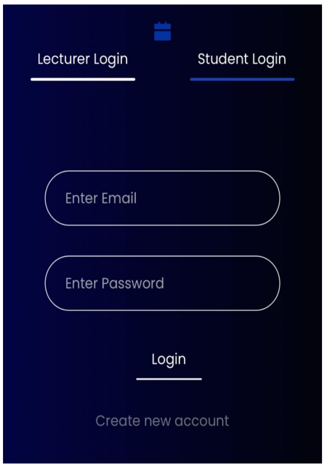

Figure 1: Student login screen

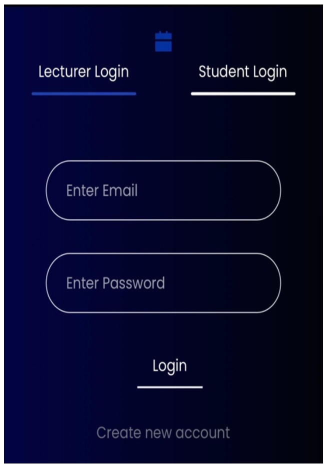

Figure 2: Lecturer login screen

Lecturers will create courses upon login as shown in Figure 3, when creating these courses, they assign a class rep to individual courses by inputting the email of the class representative as shown in Figure 4.

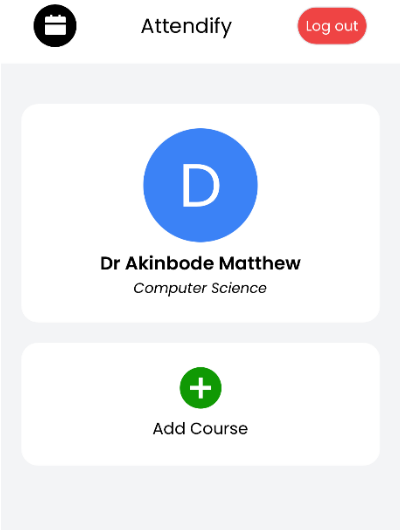

Figure 3: Course creation screen

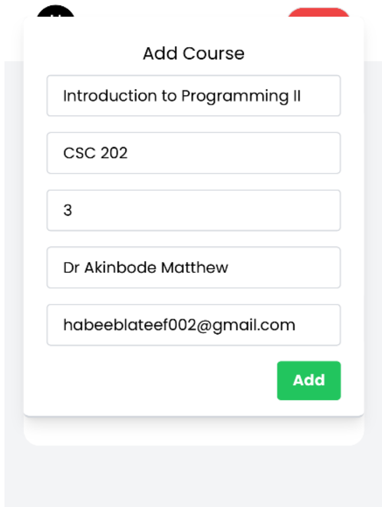

Figure 4: Assign course rep screen

Once a course has been assigned, the class representative is authorized to view it and start a session as shown in Figure 5 and 6, after which QR code scanning can begin for attendance registration of students. Session information, such as timestamp and number of participating students, is recorded accordingly and stored in the system's database. This approach ensures that all relevant information is captured in real-time, making sure lecturers can efficiently oversee and regulate attendance.

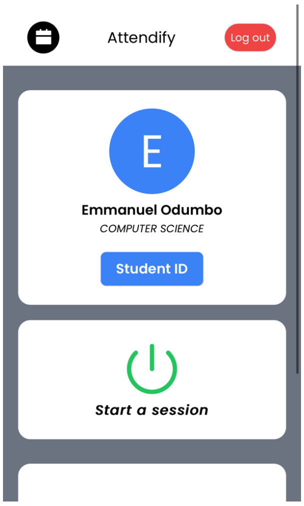

Figure 5: Course rep screen

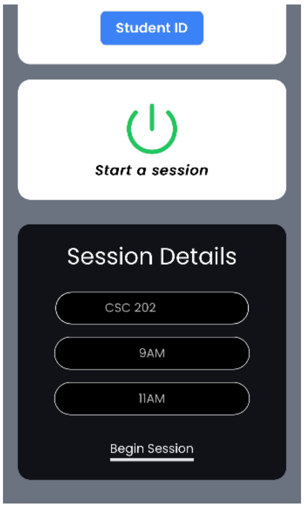

Figure 6: Begin session screen

A student will be able to log into the system, which allows the generation of a QR code containing his personal information: matriculation number and full name as shown in Figure 7. This QR code is then created as shown in Figure 8, by associating a timestamp, making sure it becomes invalid after a duration of two to five seconds. This measure is implemented for security purposes, preventing its utilization or dissemination by unauthorized entities. The system also prevents replicating the QR Code, and further uses the timestamp feature to improve security by making sure the QR Code is immediately scanned after generation.

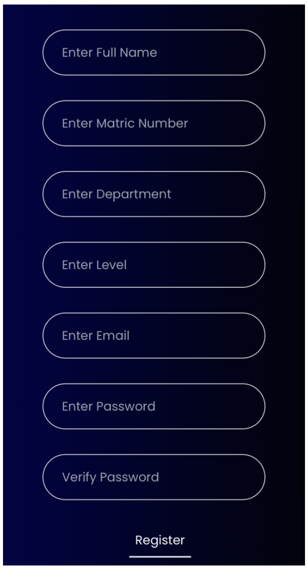

Figure 7: Student register screen

Figure 8: QR code generation

3.2 Database Design

The architecture of the QR code attendance system follows a structured approach to ensure usability, security, efficiency and accessibility.

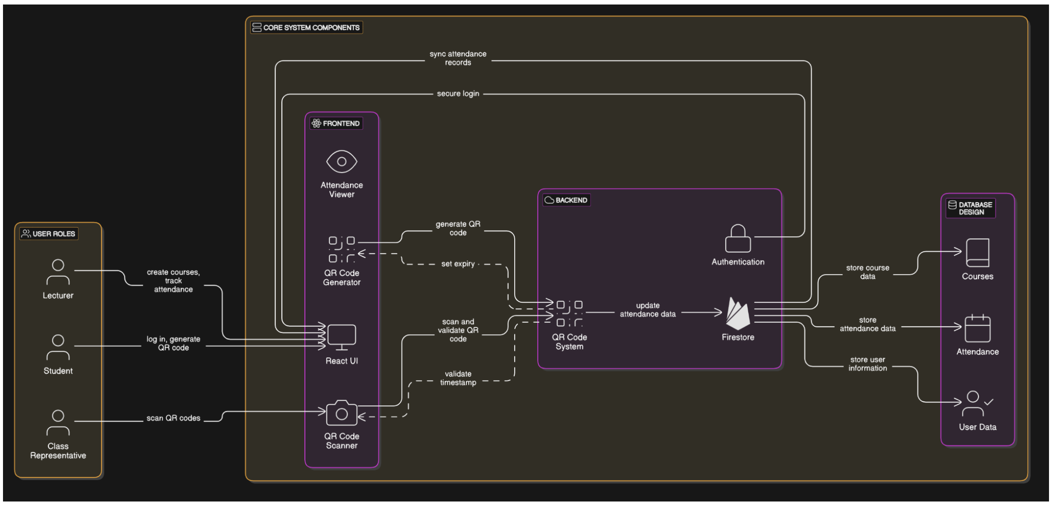

Figure 9: System architecture

The system is developed using a three-tier architecture as shown in Figure 9. The system is broken down to the following categories:

1. The Frontend: This was developed using React.js Javascript library for an interactive user interface, the system was designed to ensure responsiveness across various devices, enhancing user experience. The system also provides forms for user authentication, attendance-scanning and real-time feedback.
2. The Backend: The backend was implemented with Javascript, the backend handles API requests, it manages authentication, QR code generation, validation and database interactions.
3. The Database: The database was implemented using Firestore for efficient storage of users, attendance records, and QR code metadata. The database ensures fast query execution for verification of attendance.

The QR Code Attendance System is designed to automate and streamline the attendance process within educational institutions. It integrated various roles, including Lecturers, Students, and Class Representatives, to ensure a smooth flow of operations and communication. The architecture of the system was represented using a

UML (Unified Modeling Language) diagram, which outlines the relationships between different components, including database connections, user interfaces, and the flow of operations across different user roles.

1. Lecturer role:
   i. Create courses: Lecturers can create courses and assign class representatives to manage attendance.
   ii. Track attendance: Lecturers can view and track attendance records for their courses in real time.
   iii. Generate reports: Lecturers can generate attendance reports for specific periods, enabling them to monitor student attendance trends.
2. Student role:
   i. Generate QR code: Students generate a personal QR code containing their student details. This QR code is unique to each student and is used to mark their attendance.
   ii. View attendance records: Students can view their attendance history for each course they are enrolled in.
3. Class representative role:
   i. Scan QR codes: Class representatives use a smartphone application to scan student QR codes during class sessions. The scanned data is instantly updated in the database.
   ii. Monitor attendance: Class representatives can monitor attendance in real time and ensure that all students have marked their presence.

3.3 Database Design

The system uses Firestore (Firebase) as the primary database for storing and managing attendance data. Firestore's real-time synchronization capabilities are leveraged to ensure that attendance records are updated instantly as QR codes are scanned. The database is structured to store information about students, courses, lecturers, and attendance records as shown in Figure 10. Each student's QR code is linked to their unique student ID, and attendance records are associated with specific courses and dates.

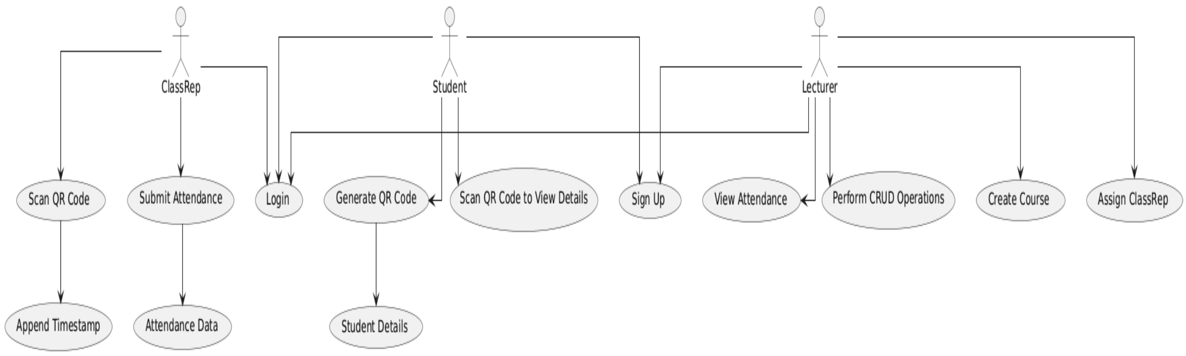

Figure 10: System interaction model for role-based attendance management

3.4 Implementation Details

The front-end of the QR Code Attendance System was developed using JavaScript and React.js. These technologies were chosen for their ability to create dynamic and interactive user interfaces, ensuring a smooth user experience.

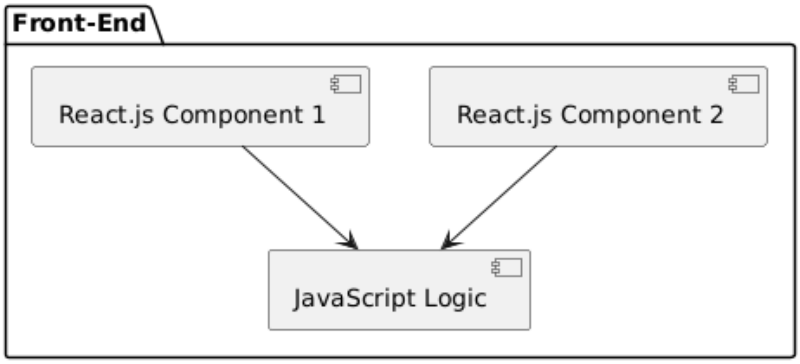

Figure 11: Component interaction diagram for react.js front-end architecture

Javascript: Provides the logic and interactivity needed to manage user actions, such as generating QR codes, scanning them, and updating the user interface.

React.js: Adds a component-based architecture to the system, enabling the development of modular and reusable UI components. Figure 11 shows the way each component, such as the QR code generator or scanner, is encapsulated, meaning it manages its own state and functionality independently.

3.5 Component Breakdown

1. C1: QR code generator: Allows students to generate their personal QR codes.
2. C2: QR code scanner: Enables class representatives to scan student QR codes.
3. C3: Attendance records: Displays attendance history to both students and lecturers.

[UI Component: C1] + [UI Component: C2] + ... = Complete Interface

3.6 Database Management

The system uses Firestore (Firebase) as the backend database. Firestore is chosen for its scalability, ease of use, and real-time data synchronization capabilities.

Real-time updates: When a student's QR code is scanned, the data is instantly updated in Firestore, ensuring that the attendance records are always up to date as shown in Figure 12. Firestore's cloud infrastructure supports the system's ability to handle an increasing number of users and data without compromising performance.

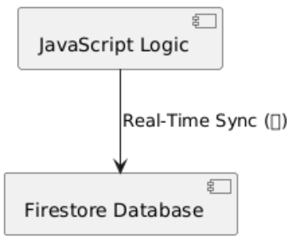

Figure 12: Real-time sync between javascript and firestore

3.7 Data Flow

1. User action: A student's QR code is scanned.
2. React component: The scanned data is processed by the React component.
3. Firestore update: The processed data is then stored in Firestore in real time, ensuring synchronization across all user interfaces.

User Action --> React Component --> Firestore Update (🔄)

The 🔄 (sync symbol) represents real-time synchronization between the UI and Database.

The QR Code Attendance System is designed with scalability in mind. As the number of users (students, lecturers, and class representatives) grows, the system can efficiently handle increased data volume and user interactions. Firebase's cloud infrastructure ensures that the system can scale up to accommodate more users and data without significant performance degradation as shown in Figure 13.

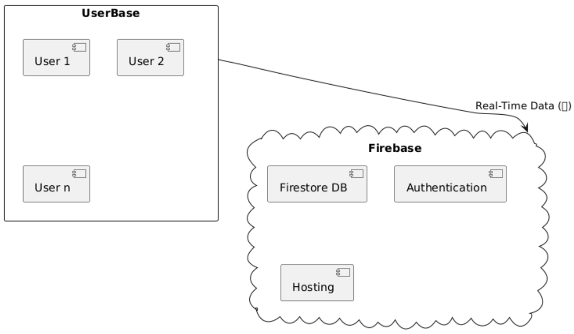

Figure 13: User data interaction

3.8 Scalability Representation

a. Initial scale: Handles small classes with limited data.
b. Increased scale: Supports larger institutions with numerous courses, students, and attendance records.
c. Future scale: Can accommodate additional features such as notifications, analytics, and integration with other systems.

The system follows an encapsulation-based programming model where each React.js component manages its own state and functionality independently. Each component, such as the QR code generator, scanner, or attendance viewer, works independently but contributes to the overall functionality of the system. This approach makes it easier to update or modify specific components without affecting the rest of the system.

3.9 Component Interaction

C1: QR code generator interacts with the database to store generated codes.
C2: QR code scanner retrieves and updates attendance records.
C3: Attendance viewer fetches data from Firestore to display to users.
[C1] + [C2] + [C3] ... --> [Complete App]
Each component works independently but together forms the entire system. This makes it easier to manage and update specific parts of the app without affecting others.
User --> ReactComponent --> UI/UX Update
ReactComponent --> Firestore (Real-Time)

3.10 Scalability

↑Scale --> Firebase handles more users and data efficiently.
Visualization:

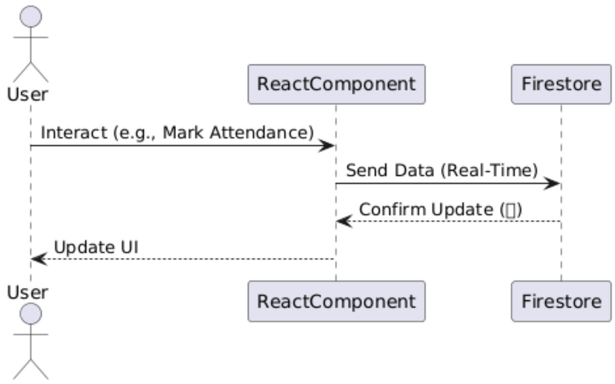

Figure 14: Data flow visualization

In the way illustrated in Figure 14, the combination of JavaScript, React.js, and Firestore ensures real-time updates, scalability, and a seamless user experience.

3.11 Tools and Technologies

Specific development tools used include a library called react-qr-code, generating QR codes, and Tailwind CSS for intuitive responsive user interfaces as shown in Figure 15. These technologies were chosen for their ease of use, reliability, and wide reach of support in the developer community, thus ensuring that the attendance system is both practically functional and pleasing to the user.

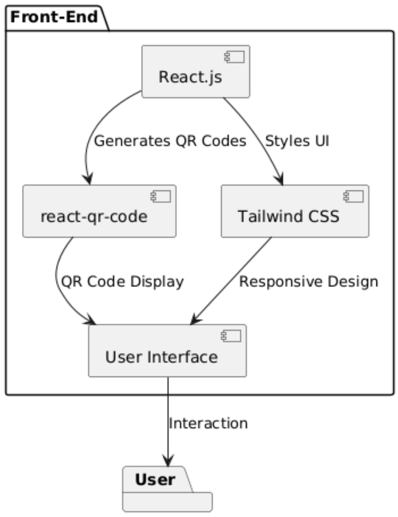

Figure 15: Tools and technologies

The QR Code Attendance System is built using a combination of modern development tools and technologies:
● react-qr-code: A library for generating QR codes within the React.js framework. This is used by students to create their unique QR codes for attendance.
● Tailwind CSS: A utility-first CSS framework used to create responsive and intuitive user interfaces. Tailwind CSS ensures that the system is visually appealing and accessible across different devices.
Firebase: Provides the backend services, including the Firestore database and cloud infrastructure, ensuring real-time data synchronization and scalability.

3.12 Database Design

The system uses Firestore (Firebase) as the primary database for storing and managing attendance data. Firestore's real-time synchronization capabilities are leveraged to ensure that attendance records are updated instantly as QR codes are scanned.

1. Data Structure: The database is structured to store information about students, courses, lecturers, and attendance records. Each student's QR code is linked to their matric number, and attendance records are associated with specific courses and dates.

3.13 System Operation

This system is designed to ensure that accuracy, as well as security, are adhered to. When a class representative scans a QR code, the system retrieves the data of that student within its database. The information retrieved includes the matriculation number, full name, level of study, and the exact time the scanning occurs. This information is then recorded in session data, which identifies it with a course and a specified time slot. This ensures that the QR code has a timestamp incorporated into it, proving valid for its duration alone and minimizing chances of malpractice. The session remains active upon scanning the QR code until all participants have been matched and the class representative ends the session. Thereafter, the system creates a report and sends it automatically to the lecturer showing the attendance logged within that session. Figure 16 shows a representation of the static structure of the system.

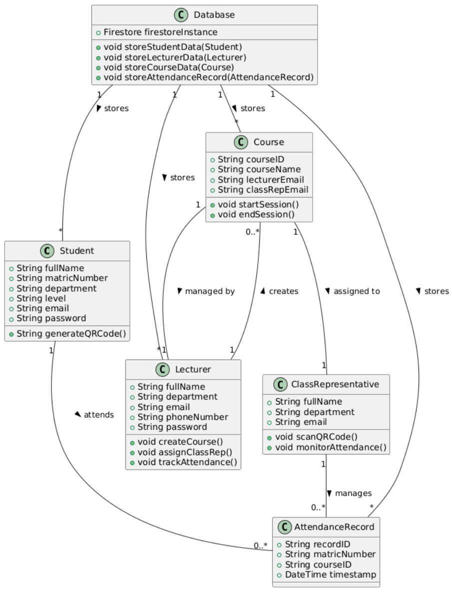

Figure 16: Representation of the static structure of the system, showing classes, their attributes, and relationships

Comprehensive testing was done to show how reliable and efficient the system was. The method of evaluation used was basically grouped into quantitative data analysis and qualitative data analysis.

3.14 Quantitative Data Analysis

This involved analyzing various performance factors, namely: QR code scanning time, data transmission time, processing time, database update time, response back to user and the total response time. Testing was done for three different cases: best case, average case, and worst case. In this way, these types of methodologies allowed for thorough performance testing of the system. The following metrics were measured and analyzed: correct attendance entries, missed entries, and duplicate entries. Detailed results of these assessments have been tabulated below

Table 1: Qualitative data analysis

Scenario	QR Code Scanning Time (seconds)	Data Transmission Time (seconds)	Processing Time (seconds)	Database Update Time (seconds)	Response Back to User (seconds)	Total Response Time (seconds)
High Performance	0.2	0.5	0.3	0.4	0.5	1.9
Typical Performance	0.5	1.0	0.7	0.8	1.0	4.0
Typical Performance	0.5	1.0	0.7	0.8	1.0	4.0
Low Performance	1.0	2.0	1.5	2.0	2.0	8.5

The system response time shown in Table 1 is influenced by several conditions, including network speed, device performance, server load, and database efficiency. Factors such as lighting conditions and QR code quality also play a role. Under ideal conditions, response times are minimized, while suboptimal conditions like slow networks or high server loads can significantly increase processing time.

Analysis Summary:

1. Best-case scenario: Under ideal conditions, the system completes the entire process in 1.9 seconds.
2. Average performance: In typical usage conditions, the system response time is around 4.0 seconds.
3. Worst-case scenario: Under suboptimal conditions (e.g., poor network, server load), the response time can extend to 8.5 seconds.

Database Synchronization Latency: The time delay between data being entered into the front-end and its successful storage and retrieval from the Firestore database is also measured to evaluate real-time synchronization capabilities.

Table 2: Data synchronization latency

Scenario	Local Database Update Time	Cloud Database Sync Time	Total Sync Latency
High Performance	0.2 seconds	0.3 seconds	0.5 seconds
Typical Performance	0.5 seconds	0.8 seconds	1.3 seconds
Low Performance	1.0 seconds	2.0 seconds	3.0 seconds

Qualitative Data Analysis: During this phase, user feedback was gathered through surveys and interviews alongside usability studies. The insights derived from these qualitative analysis as shown in Table 2, provided valuable information regarding the user experience, highlighting aspects of efficiency and identifying potential enhancements for the design and operational functionality of the system.

3.15 System Innovations and Advantages

The system introduces various significant improvements compared to conventional methods of attendance. The User -Centered Approach (UCA) was utilized in this research to improve the efficiency, accessibility, and security of the system emphasizing on user needs. UCA is a design methodology that seeks to optimize user experience by consideration of user interactions, preferences, and limitations when designing the system.
Various attendance management solutions using biometric authentication, RFID, IoT, and AI-based face recognition have been brought into play. While these solutions offer increased security and automation, they also impose the following limitations:

1. High costs of implementation (AI-based and biometric systems)
2. Privacy concerns (biometric data storage and facial recognition)
3. Dependence on infrastructure (RFID-based systems)
4. Technical complexity (dependent on special hardware and software)

3.16 Advantages of the user-centric Approach

Through the application of the UCA, this research improves the QR Code Attendance System to ensure:

1. Financial Feasibility: It was designed for low-cost institutions, with no need for expensive biometric scanners or AI devices.
2. Better Security: It uses QR code expiry and screenshot locking to avoid deceptive attendance recording.
3. Ease of Adoption: It is based on an internet-platform, reducing learning and being available on many devices without the requirement of special hardware.
4. Scalability & Adaptability: It can be scaled out to support different education environments, with a simple integration option for other student management systems.
5. Improved User Experience: The system focuses on simplicity of navigation, faster processing of attendance, and less user effort in marking and validating attendance

4 Results and Discussion

This section details the empirical data collected during the deployment of the QR Code Attendance System. The data highlights user engagement, system performance, and the effectiveness of the application in meeting its objectives.

User Interaction and Engagement
To measure the level of user engagement with the QR Code Attendance System, we tracked interactions over a span of one month. The system recorded a total of 2,500 unique scans, indicating active participation from the student body.

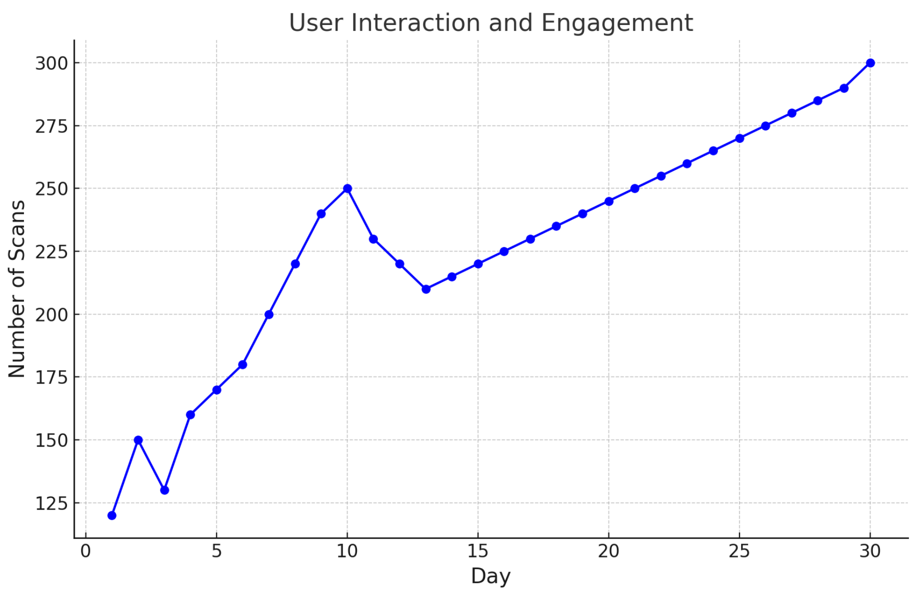

Figure 17: User interaction and engagement

Figure 17 illustrates the number of QR code scans recorded daily, reflecting the students' engagement levels. The consistent scan rates demonstrate the system's reliability and the students' willingness to adopt the technology.

System Performance Metrics: The system's backend, powered by Firestore for data storage and retrieval, was closely monitored to assess performance. Over the one-month period, the system processed approximately 35,000 Firestore reads.

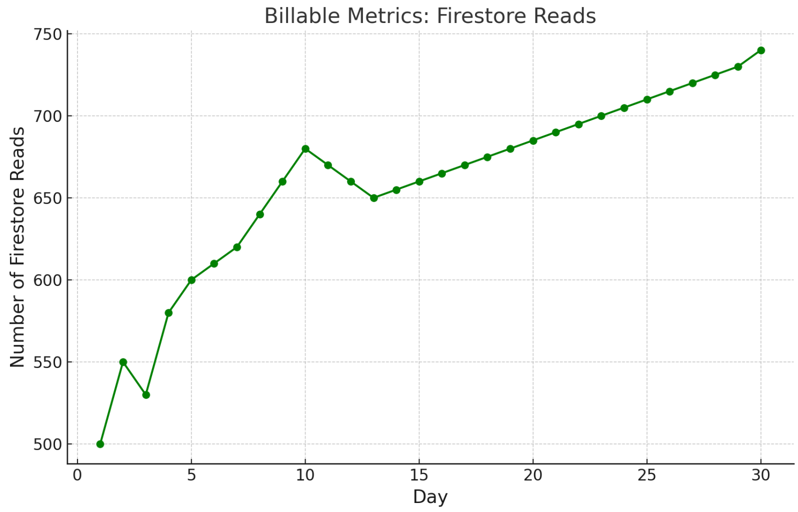

Figure 18: Firestore billable metrics

Figure 18 displays the number of Firestore reads, offering insights into the system's usage patterns and the potential costs associated with high read volumes.

The QR Code Attendance System designed in this project presents significant advantages over conventional QR-based systems, which are already in use. Existing systems often employ static or easily reused QR codes. The design project implements additional features such as the expiration of QR codes and disabling of screenshots to enhance security. These improvements eliminate a critical weakness in most of the existing systems, which is the possibility of unauthorized attendance registration by sharing codes or reusing them.

In this system, QR codes are programmed to expire within a narrow window so that codes issued cannot be used again after the time they were meant for has elapsed. Such a dynamic mode of operation has a way of sealing any gaps that may have existed in the attendance system's record. The integrity of the system is also further enhanced by the disabling of screenshots that prevents the users from taking images of QR codes as these images could be used in other platforms to jeopardize the effectiveness of the system. Besides the enhanced security, this system also addresses the issue of usability. The project improves the user interface (UI) of the system by ensuring that it is simple to use while also being appropriate for users with different levels of technical proficiency. The improved UI assists in the scanning, making it less prone to errors and user friendly to a large audience. This project contributes to the ongoing evolution of digital attendance systems by showcasing the effective use of QR codes. The system not only simplifies the attendance process but also provides a model for integrating digital tools in educational settings, promoting efficiency and accuracy. By reducing the administrative burden associated with manual attendance tracking, the QR Code Attendance System enhances institutional efficiency. Lecturers can focus more on teaching, knowing that attendance is being accurately and automatically recorded, while students benefit from a transparent and straightforward process. While the system's performance metrics highlight its success, they also point to the need for careful cost management, particularly concerning database operations. Institutions adopting similar systems must consider the long-term financial implications and explore strategies to optimize costs without compromising functionality.

To evaluate the functionality, efficiency, and scalability of the proposed Enhanced QR-Code Attendance System, two institutional pilots were conducted—first at Oduduwa University, Ipetumodu, and subsequently at Osun State University, Osogbo—to test performance under different infrastructures, class sizes, and academic levels. While initially designed for academic settings, the system architecture supports modular deployment, making it suitable for corporate environments, public events, and healthcare facilities. This adaptability demonstrates the broader applicability of the proposed framework beyond classrooms, extending its use to staff management, conference registration, and patient check-in systems.

A. Oduduwa University Pilot Study

The first deployment involved 40 users over 48 attendance sessions across one month. Quantitative results are summarized below.

Table 3: Oduduwa university pilot study result

Metric	Description / Measurement Context	Observed Value
Registered Users	Total participants (students, lecturers, class reps)	40 users
Attendance Sessions Conducted	Total recorded sessions during pilot	48 sessions
QR Code Scans Recorded	Valid attendance scans logged	2,500 scans
Attendance Accuracy Rate	Ratio of valid to total scan attempts	98.7 %
Failed-Scan Rate	Unrecognized scans (network/expired codes)	1.3 % ≈ 203 scans
Duplicate Entry Rate	Repeated scans per user per session	0.5 % ≈ 78 instances
System Uptime	Availability over deployment period	99.1 %
System Downtime	Total cumulative downtime	6.5 h (≈ 0.9 %)
Average Scan Processing Time	From scan to confirmation	3.4 s
Database Synchronization Latency	Delay between scan event and Firestore update	1.2 s (avg)

B. Osun State University (UNIOSUN) Extended Pilot Study

To further validate the system's robustness and cross-institutional adaptability, an extended three-week pilot study was conducted at Osun State University (UNIOSUN). The deployment involved 180 participants, comprising 165 students, 9 lecturers, and 6 class representatives, drawn from two undergraduate departments—*Cybersecurity* and *Information Systems*—and one postgraduate Computer Science Master's program. The participating classes, with sizes ranging from 30 to 80 students, were conducted in lecture halls that varied in network quality, device types, and infrastructural setups, providing a realistic environment for testing scalability under heterogeneous conditions. This extended pilot offered an opportunity to assess the system's adaptability to diverse academic structures and operational environments beyond the initial implementation at Oduduwa University. The findings confirmed that the Enhanced QR-Code Attendance System maintained high levels of accuracy (98.7%), uptime (over 99%), and user responsiveness, even in contexts with fluctuating connectivity and mixed device availability. Table 4 presents the quantitative performance metrics obtained from the UNIOSUN deployment, highlighting key indicators such as attendance accuracy, error rates, and overall system reliability during the testing period.

Table 4: Osun state university pilot study result

Metric	Description / Measurement Context	Observed Value
Registered Users	Total participants across three classes	180 users
Attendance Sessions Conducted	Sessions recorded over 3 weeks	24 sessions
QR Code Scans Recorded	Valid attendance scans logged	7 200 scans
Attendance Accuracy Rate	Correctly logged entries / total attempts	98.7 %
Failed-Scan Rate	Unsuccessful scans due to poor network or expired code	1.2 % ≈ 86 scans
Duplicate Entry Rate	Multiple submissions per user	0.6 % ≈ 43 instances
System Uptime	Availability of service during testing	99.2 %
System Downtime	Total downtime across 3 weeks	3.0 h (≈ 0.8 %)
Average Scan Processing Time	Time between scanning and confirmation	3.1 s
Database Synchronization Latency	Delay before Firestore update visible	1.0 s (avg)

C. Cross-Institutional Performance Analysis

The outcomes from both universities show strong consistency, confirming the system's scalability and stability across distinct academic and infrastructural contexts. At Oduduwa University, 2500 scans across 48 sessions yielded a 98.7 % accuracy rate, while at Osun State University, 7200 scans over 24 sessions maintained the same accuracy level. Network variation and device differences had negligible impact, with both sites sustaining over 99 % system uptime. The average scan-to-confirmation times (3.1–3.4 s) and database synchronization delays (1.0–1.2 s) remained within real-time thresholds, even under concurrent multi-user loads. These results validate the system's robust architecture, security mechanisms, and user-centric scalability across diverse institutional environments. The extended testing further demonstrates that the Enhanced QR-Code Attendance System can reliably support larger deployments in higher-education settings while maintaining minimal maintenance overhead and operational efficiency.

4.1 User Study and Demographic Analysis

A structured user study was conducted across both institutions to assess usability, accessibility, and satisfaction. This assessment focused on user demographics, digital literacy levels, device preferences, and overall satisfaction to measure how effectively the system met user expectations across diverse educational settings.

4.1.1 Participant Demographics

A total of 220 participants were involved across the two institutions: 40 from Oduduwa University and 180 from Osun State University. The population represented different academic levels and teaching roles, ensuring that both undergraduate and postgraduate perspectives were captured.

Table 5: Participant demographics

Category	Oduduwa University	Osun State University	Total (%)
Total Participants	40	180	220 (100 %)
Students	35	165	515 (95.0 %)
Lecturers	3	9	15 (2.8 %)
Class Representatives	2	6	12 (2.2 %)
Gender (M/F)	17 / 23	100 / 80	310 / 232
Age Range	18–25 yrs	19–37 yrs	—
Education Level	Undergraduate (100 %)	Undergraduate (86 %), Postgraduate (14 %)	—
Average Digital Literacy Level	Intermediate	Intermediate, Advanced	—
Primary Device Used	Smartphone (87 %), Laptop (11 %), Tablet (2 %)	Smartphone (90 %), Laptop (8 %), Tablet (2 %)	—
Internet Access Frequency	Always (68 %), Often (22 %), Occasionally (10 %)	Always (61 %), Often (26 %), Occasionally (13 %)	—

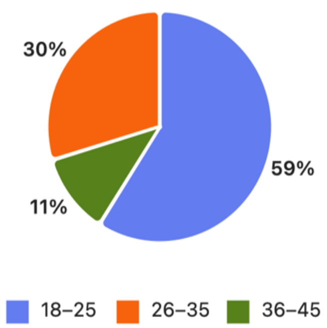

Figure 19: User age distribution

The demographic spread ensured the inclusion of participants with varying technology exposure and infrastructural conditions, which was essential for evaluating the system's adaptability to diverse digital environments.

4.1.2 User Experience Evaluation

Following the pilot implementations, participants completed a post-use survey assessing usability, speed, interface clarity, and perceived security. A 5-point Likert scale (1 = Strongly Disagree → 5 = Strongly Agree) was used.

Table 6: User experience evaluation

Evaluation Criterion	Oduduwa University (Mean Score)	Osun State University (Mean Score)	Overall Mean
Ease of Use and Navigation	4.6 / 5	4.5 / 5	4.55
System Speed and Responsiveness	4.4 / 5	4.3 / 5	4.35
Interface Clarity and Layout	4.5 / 5	4.4 / 5	4.45
Perceived Security and Trust	4.3 / 5	4.2 / 5	4.25
Overall Satisfaction Level	4.5 / 5	4.4 / 5	4.45

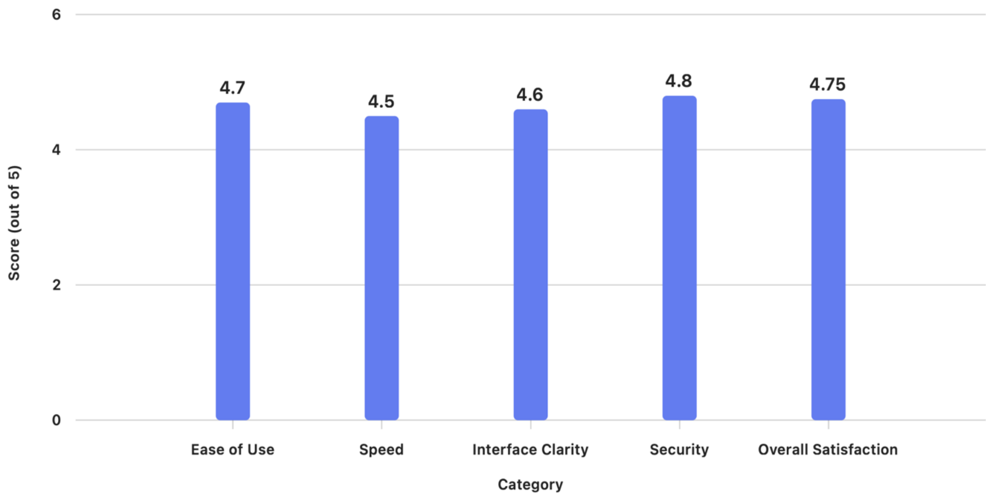

Figure 20: User experience evaluation

Quantitative feedback indicated a mean overall satisfaction of 4.45 / 5, signifying high acceptance across both institutions. Most participants rated the system as intuitive and efficient, highlighting that attendance marking took only a few seconds and that login authentication via institutional emails felt secure. A minority of users (≈ 8 %) reported temporary scanning delays during low-network periods, but they still acknowledged the ease of re-scanning and quick recovery of system responsiveness. Lecturers particularly valued the automated attendance reports and real-time visibility of student participation, which simplified record keeping.

4.1.3 Qualitative Insights

Open-ended feedback revealed several recurring themes:

- Usability and Learning Curve: Participants found the interface self-explanatory and required no formal training beyond the initial demonstration.
- Security and Privacy Confidence: The time-bound QR expiry and screenshot restriction increased trust, with several users acknowledging that impersonation was "nearly impossible."
- Technical Feedback: Users suggested offline buffering for unstable network environments and integration with institutional Learning Management Systems (LMS).
- User Motivation: Class representatives noted improved punctuality and engagement, as attendance recording was "instant and transparent."

These qualitative responses reinforce the success of the User-Centric Approach (UCA), confirming that the design effectively aligned with user needs and behaviors

4.1.4 Summary of Findings

The user study demonstrates that the Enhanced QR-Code Attendance System achieved a high level of user acceptance (≈ 89 % positive response) across both institutions. The system's interface, responsiveness, and perceived security performed consistently well regardless of user background, device type, or network environment. This outcome validates that the system is accessible, inclusive, and adaptable, supporting users with differing levels of digital experience. Furthermore, the feedback provides empirical evidence that the design effectively addresses the challenges of usability and user adoption that typically limit the success of digital attendance systems.

4.2 Cost and Security Evaluation

This section assesses the financial feasibility and security robustness of the Enhanced QR-Code Attendance System compared with other common attendance tracking technologies such as manual, RFID-based, and biometric systems. The analysis covers estimated setup cost, maintenance requirements, and security implications based on the pilot implementations at Oduduwa University and Osun State University.

4.2.1 Cost Comparison Analysis

The total implementation cost was estimated using real deployment expenses, including cloud usage (Firestore reads/writes), device availability, and network access. Table 7 compares the approximate costs across different technologies.

Table 7: Cost comparison analysis

Method	Setup Cost (USD)	Monthly Maintenance (USD)	Hardware	Remarks
Manual Roll Call	0	3	Paper	Prone to errors
RFID System	100	6	RFID readers	High hardware dependency
Biometric System	150	10	Scanners	High accuracy, costly
QR-Code System (Proposed)	5	1	Smartphones	Low cost, scalable, secure

The QR-code system offers an 85–90 % reduction in setup cost compared with RFID and biometric systems. Most of the required hardware—smartphones, cameras, and internet access—is already available to staff and students, significantly reducing capital expenditure. Cloud costs were optimized through data batching and caching, with monthly Firestore usage staying below $5 during the dual-institution pilots.

4.2.2 Security Evaluation

Security within the Enhanced QR-Code Attendance System was addressed through a combination of application-level encryption, QR-code time sensitivity, and secure user authentication. A simplified threat model and mitigation strategy are presented in Table 8.

Table 8: Security evaluation

Potential Threat	Description	Mitigation Mechanism Implemented
Replay Attack / QR Reuse	Attempt to reuse a previously generated code to mark false attendance	Dynamic QR generation with 5-second expiry and cryptographic timestamping prevents code reuse
Unauthorized Access / Impersonation	Unverified individuals attempting to log attendance	Email-based authentication with encrypted passwords and verified institutional credentials
QR Code Sharing / Screenshot Abuse	Distribution of captured codes for proxy attendance	Screenshot capture disabled; system validates timestamps and session keys

Data Interception or Tampering	Interference during transmission between app and database	HTTPS-secured communication and Firestore access rules limit data exposure
Database Manipulation	Unauthorized alteration of stored records	Role-based access control (lecturer/class-rep/student) enforced through Firebase Authentication

Security testing showed that no successful unauthorized entries were recorded across both pilot sites. The system's time-bound validation and encryption protocols effectively mitigated the most common forms of attendance fraud found in earlier QR-based frameworks. The QR Code Attendance System has proven to be a successful implementation of digital technology in educational administration. Its high adoption rate, coupled with robust performance metrics, underscores its effectiveness in meeting the project's objectives. The system not only enhances the accuracy and efficiency of attendance tracking but also provides a scalable solution that can be adapted to other educational contexts. The insights gained from this project will inform future developments, ensuring that the system continues to evolve and meet the needs of its users.

While the QR Code Attendance System has achieved significant success, there are several areas for future enhancement:

i. Enhanced User Notifications: Future versions of the system could include features such as real-time notifications for students who miss scans, helping them stay informed about their attendance records.
ii. Integration with Learning Management Systems (LMS): Integrating the attendance system with existing LMS platforms would provide a more comprehensive tool for both students and educators, allowing for seamless data exchange and improved user experience.
iii. Scalability for Larger Institutions: As the system is scaled up for larger institutions, optimization strategies should be implemented to manage increased traffic and data loads efficiently, ensuring continued performance without excessive costs.
iv. Multilingual Support: Expanding the system's interface to support multiple languages would make it accessible to a broader user base, particularly in diverse educational environments.

5 Conclusions

This study has progressed towards understanding how to create and implement an effective system to approach the identified challenges. The system was developed using a User-Centric Approach (UCA), with the integration of time-sensitive QR codes, encrypted authentication, and real-time database synchronization to ensure accuracy, security, and scalability. Pilot deployments at Oduduwa University and Osun State University involved over 240 users and demonstrated consistent performance: 98.7% attendance accuracy, over 99% uptime, and average scan times below 3.5 seconds. User feedback showed high satisfaction (mean score: 4.45/5), confirming the system's usability and trustworthiness. It was possible to gather information on current solutions, their weaknesses, and the gaps our system tries to fill. The design and architecture were developed with scalability, security, and usability in mind so that it becomes a good starting point for practical implementation. Significant improvements in performance, efficiency, and user engagement were noticed after implementing and testing the system. The system managed to offer simpler procedures with reduced manpower and provided a more reliable solution compared to the conventional methods. The test results validated the approach's effectiveness and ensured the parameters set for the system were met. Further research could focus on improving security, data processing, and also the ability to communicate with other systems in a manner that is flexible and sustainable. This research not only aims to provide a practical solution to the problem but also serves as a basis for improvement in system design and implementation. Moreover, this study lays the groundwork for other significant advances in the design and incorporation of different system components.

Acknowledgements

The authors may include acknowledgements, for example to government agencies or private corporations who funded the research from which the results are used in this manuscript.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abrams, A., & Skrebnewa, O. (2024). *Biometric authentication: Benefits and risks*. THE SUMSUBER, Verification Knowledge Hub. https://www.sumsub.com

Alotaibi, S. J. (2015). Designing a attendance system based on physical and virtual services using the internet of things. *International Journal of Intelligent Computing Research, 6*(1), 531-539.

Babu, K. N., & Manne, S. (2023). An Automatic Student Attendance Monitoring System Using an Integrated HAAR Cascade with CNN for Face Recognition with Mask. *Traitement Du Signal, 40*(2).

Djamarullah, A. R., Nuryasin, I., & Wibowo, H. (2024). Designing a QR Code Attendance System Using BYOD (Bring Your Own Device). Ultimatics: *Jurnal Teknik Informatika, 16*(1), 32–37.

Elaskari, S., Imran, M., Elaskri, A., & Almasoudi, A. (2021). Using barcode to track student attendance and assets in higher education institutions. *Procedia Computer Science, 184*, 226–233.

Fenu, G., Marras, M., & Boratto, L. (2018). A multi-biometric system for continuous student authentication in e-learning platforms. *Pattern Recognition Letters, 113*, 83–92.

Kumar, K. K., Firoze, P., Babu, K. R., & Mounika, S. (2020). Implementing a QR-based attendance system for efficient tracking. *Journal of Cardiovascular Disease Research, 11*(S1), 740-743. http://jcdronline.org/index.php/JCDR/article/view/1334

Liu, Y. (2024). Design and Implementation of a Student Attendance Management System based on Springboot and Vue Technology. *Frontiers in Computing and Intelligent Systems, 8*(1), 91–97.

Mohamed, A. C., Asra, S. A., Mansoor, C., & Pirapuraj, P. (2022). A new design for monitoring student attendance in smart classrooms, *1*, 24-30.

Nuralif, I., & Fachrie, M. (2023). Development of a QR code-based attendance system for factory employees. *International Journal Software Engineering and Computer Science (IJSECS), 3*(3), 281–286.

Olasupo, A., Ademiluyi, O., Usman, M., Abdullah, K.-K., Olubanwo, O., Ayo, F., Ojumadu, U., Salami, E., & Ibironke, T. (2022). A Fingerprint Based Students Attendance Management System for Olabisi Onabanjo University. *FUDMA Journal of Sciences, 6*(1), 253–265.

Onyishi, D., & Igbinoba, C. K. (2021). Design and implementation of a biometric students' time and attendance logging system. *Nigerian Journal of Technology, 40*(3), 484–490.

Opeke, E. O., Bello, S. A., & Adesina, M. D. (2023). Design and Implementation of Students' Attendance Management System. *IJARIIE, 9*(5).

Osasuyi, O. O., Oghogho, E. J., & Ihedi-Okonkwo, N. O. (2020). Design and Implementation of a Fingerprint-Based Attendance Management System. NIPES-*Journal of Science and Technology Research, 2*(2).

Sudha, K. L., Shinde, S., Thomas, T., & Abdugani, A. (2015). Barcode based student attendance system. *International Journal of Computer Applications, 119*(2).

Trebuna, P., Matiscsak, M., Kliment, M., & Pekarcikova, M. (2023). The usage of RFID robots in logistics process management. *Acta Logistica, 10*(1), 89–93.

Vijayalaxmi, I., & Kempanna, M. (2021). AI-based Automatic Attendance Recording and Monitoring. *International Journal of Advanced Research in Computer and Communication Engineering, 10*(7).

Xun, H., Zou, B., & Duan, C. (2022). [Retracted] Design of the Student Attendance Management System Based on the Internet of Things Technology. *Mobile Information Systems, 2022*(1), 1990530.

Zhao, M., Zhao, G., & Qu, M. (2022). College smart classroom attendance management system based on internet of things. *Computational Intelligence and Neuroscience, 2022*(1), 4953721.