الجمهورية الجزائرية الديمقراطية الشعبية

DEMOCRATIC AND POPULAR REPUBLIC OF ALGERIA

وزارة التعليم العالي والبحث العلمي

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

– جامعة أبي بكر بلقايد – تلمسان –

Abou bekr Belkaïd – Tlemcen University–
Faculty of Technology

MASTER THESIS

Presented for the award of the **MASTER degree**

In: Telecommunication

Specialty: Networks & Telecommunications

By: DIB Ayoub and SAHRAOUI Mohammed Amine

Theme

Modeling a wireless network of sensors for Human Activity Recognition

Presented publicly, the 13/06/2024, facing the jury composed of:

Mr. KAMECHE Samir	Pr.	Univ. Tlemcen	President
Ms. SLIMANE Zohra	MCA	Univ. Tlemcen	Supervisor
Ms. KORTI Djazila Souhila	Doctoral Student	Univ. Tlemcen	Co-Supervisor
Mr. ABDELMALEK Abdelhafid	MCB	Univ. Tlemcen	Examinator

Academic year: 2023-2024

Dedication

Completing this project with high intense of work and search I dedicate to

My parents that I've inspired from them wisdom, patience and ambition.

My Brother Sidi Mohamed his wonderful wife Ilhem and their adorable daughter Latefa for their moral support and encouragements.

My Brother Abderrahmane, his amazing wife Assia for their advices and Motivation.

My little brother Youcef filled with energy and courage a source of inspiration.

My beloved grandparents Abd el ouahab and Fatema Zohra your presence is a treasure.

May Allah Blessings be upon All

Dib Ayoub ✒

Dedication

I dedicate this work

*To my dear **parents**, may **Allah** prolong their lives, the source of love, tenderness and sacrifice, to whom I owe all the credit. I wish **my father** a speedy recovery,*

*To my beautiful brothers **Adel** and **Maissara***

*To all my **family**,*

*And all my **friends**,*

This project is the fruit of our collective efforts and shared determination. We hope it will live up to expectations and pave the way for new opportunities and achievements.

S. Med Amine

Thanking

*First of all, we should thank **Allah** who guide us to the light of knowledge and wisdom, and bless us with Health and strength.*

*All Gratitude to our supervisor **Slimane Zohra** for their help and support with advices and her presence among this work Helped us to Achieve this project with Success.*

*Our best words of Gratitude to our co-supervisor **Korti Djazila Souhila** for her presence to support, guidance and assistance throughout the entirety of our project.*

*Our sincere appreciation for **jury members** who accepted to review and evaluate our work.*

*Our deep thanks to all **professors** of **Telecommunication** whom participated on our learning process among all the previous years*

Finally, our special appreciations our friends and partners for their support and confidence among all the previous years.

Résumé

De nombreuses recherches ont été consacrées à l'étude de la reconnaissance de l'activité humaine. Cependant, la plupart de ces études utilisent des méthodes traditionnelles axées sur l'entraînement et l'optimisation des modèles. Notre recherche repose sur l'utilisation de bases de données adoptées avec des capteurs radar. Les données radars collectés sont d'abord prétraitées puis converties en images 2D, fournissant des informations sur la variation de fréquence au cours du temps, également appelée signature micro-Doppler. Ces images 2D sont ensuite utilisées pour entraîner des algorithmes d'apprentissage profond afin d'identifier et de classer différents types d'activités humaines. Nous avons choisi les algorithmes d'apprentissage profond pour leur capacité à traiter efficacement des données complexes et pour leur flexibilité dans la reconnaissance des formes. De plus, nous avons mis en place des techniques permettant de gérer efficacement les données provenant de plusieurs radars. Des modifications ont été apportées au niveau des bases de données ainsi que la nouvelle structure des modèles établis sur les différentes bases de données établies afin de cibler le traitement, le temps de calcul et la précision ainsi leur application dans la réalité que nous proposons sont adaptables et opérationnels dans des environnements réels.

Mots clés : IR-UWB, FMCW, reconnaissance de l'activité humaine, Deep Learning, CNN, LSTM, indication de cible mobile, confusion, concaténation, ConvLSTM

Abstract

Much research has been on the study of human activity recognition. However, most of these studies use traditional methods focus on training and models optimization. Our search based on using data base adopted with radar sensors The radar data collected is first preprocessed and then converted into 2D images, providing information on frequency variation over time, also called micro-Doppler signature. These 2D images are then used to train deep learning algorithms to identify and classify different types of human activities. We chose Deep Learning algorithms for their ability to efficiently process complex data and for their flexibility in pattern recognition. In addition, we have implemented techniques to effectively manage data from multiple radars. changes have been made at database level as well as the new structure of models been established on the different databases established in order to target the processing the time of calculation and the precision thus their application in reality, we offer are adaptable and operational in real environments.

Keywords: IR-UWB, FMCW, Human Activity Recognition, Deep Learning, CNN, LSTM, Moving Target Indication, confusion, concatenation, ConvLSTM

ملخص

تم إجراء الكثير من الأبحاث حول دراسة التعرف على النشاط البشري. ومع ذلك، فإن معظم هذه الدراسات تستخدم أساليب تقليدية تركز على التدريب وتحسين النماذج. يعتمد بحثنا على استخدام قاعدة البيانات المعتمدة مع أجهزة استشعار الرادار تتم معالجة بيانات الرادار التي تم جمعها أولاً ثم تحويلها إلى صور ثنائية الأبعاد، مما يوفر معلومات عن تباين التردد بمرور الوقت، وتسمى أيضًا بصمة دوبلر الدقيقة. ثم تُستخدم هذه الصور ثنائية الأبعاد لتدريب خوارزميات التعلم العميق لتحديد وتصنيف أنواع مختلفة من الأنشطة البشرية. وقد اخترنا خوارزميات التعلّم العميق لقدرتها على معالجة البيانات المعقدة بكفاءة ولمرونتها في التعرف على الأنماط. بالإضافة إلى ذلك، قمنا بتنفيذ تقنيات لإدارة البيانات من رادارات متعددة بفعالية، حيث تم إجراء تغييرات على مستوى قاعدة البيانات بالإضافة إلى البنية الجديدة للنماذج التي تم إنشاؤها على قواعد البيانات المختلفة التي تم إنشاؤها من أجل استهداف معالجة وقت الحساب والدقة وبالتالي فإن تطبيقها في الواقع الذي نقدمه قابل للتكيف والتشغيل في بيئات حقيقية.

الكلمات المفتاحية: التعرف على النشاط البشري، التعلم العميق، CNN، LSTM، إشارة الهدف المتحرك، الارتباك، التسلسل، ConvLSTM

Contents

Figures list

Tables List

Schemes List

Abbreviations

2

2D **Two**-**D**imensional

3

3D **Three**-**D**imensional

A

AI **A**rtificial **I**ntelligence

C

CNN **C**onvolutional **N**eural **N**etwork

CML **C**lassical **M**achine **L**earning

CW **C**ontinuous **W**ave

CV **C**omputer **V**ision

CTC **C**onnectionist **T**emporal **C**lassification

D

DT **D**ecision **T**rees

DL **D**eep **L**earning

DTW **D**ynamic **T**ime **W**arping

F

FMCW **F**requency **M**odulated **C**ontinuous **W**ave

FCC **F**ederal **C**ommunications **C**ommission

H

HAR **H**uman **A**ctivity **R**ecognition

HRRP **H**igh **R**esolution **R**ange **P**rofile

I

IoT **I**nternet **o**f **T**hings

IR-UWB **I**mpuls-**R**adio **U**ltra-**W**ide**B**and

L

LSTM **L**ong **S**hort-**T**erm **M**emory

M

ML **M**achine **L**earning

MTI **M**oving **T**arget **I**ndication

N

NLP **N**atural **L**anguage **P**rocessing

P

PRF **P**ulse **R**epetition **F**requency

R

RF **R**andom **F**orest

RNN **R**ecurrent **N**eural **N**etwork

RD **R**ange–**D**oppler

S

STFT **S**hort **T**ime **F**ourier **T**ransform

SVM **S**upport **V**ector **M**achine

T

TSN **T**emporal **S**egment **N**etwork

U

UWB **U**ltra-**W**ide**B**and

UBB **U**ltra-**B**road**B**and

X

XGBoost e**X**treme **G**radient **Boost**ing

General Introduction

Context

In this decade we are face to the huge amount of data and process where automation and precision plays significant roles, HAR one of the most used on our Generation decade that used for various Activities as health-care, surveillance systems, sports and education the development of fully automated HAR system capable of classifying person's activities with low error by developing a performed Artificial intelligence model that give a good result.

Recently, effective solutions have emerged for human activity recognition (HAR) systems that use a variety of sensors, such as radars. Their advanced technologies and accuracy make them suitable tools for detecting human activity.

With their many benefits, radars are positioned as a wise choice for successful integration in a variety of use cases. Among these benefits, one can highlight the lack of invasiveness, preservation of private life, low energy consumption, and insensitivity to environmental conditions.

Human activity recognition systems in general based on sensors recently the researches aim on radars functionality and marked as an effective sensor for this task

Using one sensor for HAR have important limitations on accuracy and precision for example physical obstacles even unfavorable weather conditions may influence on sensor accuracy during recognition of human activities.

The suggestion of using sensor network which is group of sensors where each sensor collect data on real-time then processing for treatment and analysis information, Sensor network play a crucial role in recognizing human activities by providing a more robust, accurate and scalable approach to using a single sensor. Their ability to collect data, provide increased redundancy and reliability, and improve detection accuracy make sensor networks a preferred choice for many HAR applications in various fields. Using a network of sensor reduce errors and ambiguity and ensures the validity of results data collected during process of HAR.

Problem Description

We use a network of sensors, so managing the data from the various radar sensors effectively presents challenges related to complexity level and data analysis.

Traditional DL algorithms are typically designed to process data from a single source or modality. However, our research focuses on a crucial challenge:

How to effectively manage data from multiple radars in order to develop a powerful classification algorithm? This complex issue requires a thorough exploration of various strategies to find suitable solutions.

The management of data from multiple radars for the training of classification algorithms requires a suitable approach by exploring solutions at both model and data levels, we can successfully address this complex challenge.

Project contributions

- Our work encompasses on treatment Data before feeding them into DL models this process called Fusion at level of Data, another treatment is at level of models where we apply (Early, Halfway and Late) Fusion Methods.
- We have extended our approach beyond simple model architecture changes to also include data-level fusion, which improves the management of data from multiple radars.
- It is essential to emphasize that the models we have developed have been entirely designed "from scratch". This approach allowed us to create architectures specifically adapted to our problem, without relying on pre-existing solutions.
- Unlike many previous works that focused solely on the performance of algorithms without considering their complexity, our project takes a balanced approach. We preferred simple and less complex algorithms while maintaining high performance. This approach reflects our commitment to finding a solution that is both effective and practical.

Chapter I: Initiations About Human Activity Recognition

Chapter I: Initiations About Human Activity Recognition

I.1 Introduction

To better understand the functioning of the systems of recognition of human activity we raise various issues in our search that occupied us during our project. We would like to know the approaches and tools used to identify these activities. Furthermore, what commonly challenges are encountered by HAR systems, even what is several applications of human activities Recognitions and their impact on future world? In this chapter we will answer these questions by giving explanations with examples that will evade doubts about the functioning of recognition of human activity.

I.2 Definition of human Activity

The World Health organization defines physical activity as any bodily movement produced by skeletal muscles that requires energy expenditure [1]. Physical activity refers to all the movements that we perform, particularly during leisure time, at the workplace or to move from one place to another [2].

Regular physical activity is proven to aid the prevention and management of noncommunicable diseases, such as heart disease, stroke, diabetes and several cancers. It also helps prevent hypertension, maintain healthy body weight and improve mental health, quality of life and well-being [1].

Human activity refers to any action, behavior, or task performed by individuals or groups of people that involves the use of physical or mental effort. This term encompasses a wide range of activities, including but not limited to work, leisure, communication, transportation, education, and social interactions [2-4].

I.3 Types of human activity

Human activities have been classified into two main Categories: namely static and dynamic [5].

Dynamic activities have a motion trend, which represents more intense environmental interactions. Walking, falling, and other short-lived behaviors are examples of how human body posture and behavior interact [6]. It typically refers to some of the most common human behaviors in everyday life. Activities with posture transition are combinatorial behaviors that accumulate through multiple basic activities. Multiple users may participate during the behaviors [7].

Static activities (such as Sit and Stand, etc.) are easier to recognize than regular activities (such as walking, picking up something, etc.). However, highly similar posture transitions, such as Sitting and Standing, will result in significant complexity during separation because the activity feature space is clearly overlapped. In the static activities of Sit and Stand, which means finished action without movements, and dynamic activities or posture transitions for Sitting down and standing [8].

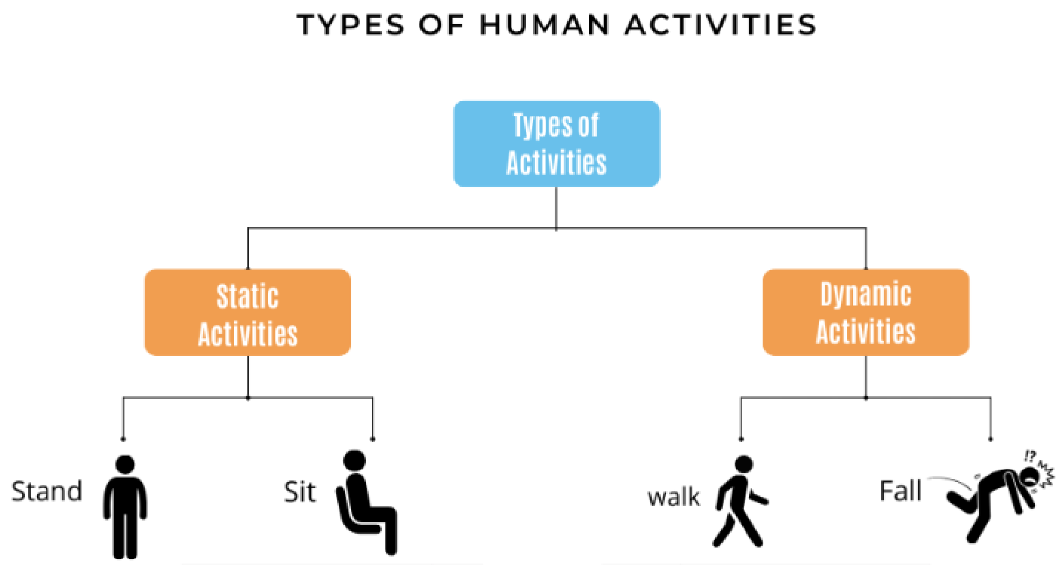


Figure I- 1: Different Types of Human Activities

I.4 Human activity Recognition

Human Activity Recognition (HAR) has a significant role in the everyday life of people enhancing health and fitness monitoring, ensuring safety by sensors. A substantial amount of research has been conducted on HAR and numerous approaches based on DL and ML have been exploited by the research community to classify human activities [9].

HAR researches focuses on automatically identifying and categorizing these Activities using data from various sensors such as accelerometers, gyroscopes, or video cameras, radars [8]. HAR has applications in healthcare monitoring, sports analysis, and smart environments, among others [10].

Sensors and wearable devices surround us in our daily life [11] HAR with sensors involves the acquisition of data by sensors, their processing to extract relevant characteristics, then the classification of detected activities according to the analyzed information.

HAR systems designed to recognize such Activities can provide valuable insights into human behavior and enable intelligent responses or interventions in the environment.

I.5 HAR Process

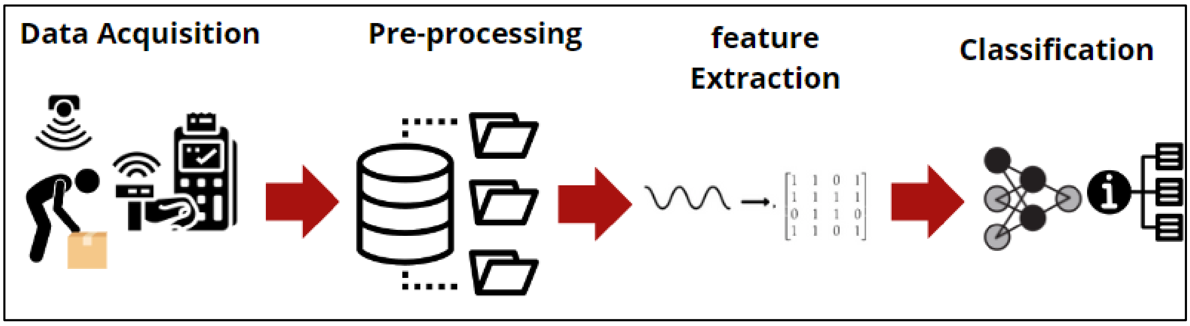


Figure I- 2: HAR Using Inertial, Physiological and Environmental Sensors

There are four main stages in the HAR process: data acquisition, pre-processing, feature extraction, Classification.

I.5.1 Data acquisition

The first step in recognizing human activity is data acquisition, a crucial process that involves selecting the appropriate types of sensors, the number of sensors to be used and their strategic placement [12].

These sensors can be accelerometers, gyroscopes, video cameras, radars or other environmental sensors [13].

This section by stating that volunteers require their assistance and their tasks and frequency of participation will be specified. This makes the phase the most difficult because it requires a lot of time and labor. The value of the data gathered determines the outcome of the entire process.

I.5.2 Pre-processing

Secondly the pre-processing step aims to prepare the raw data by filtering, sampling and standardizing them. the collected data contains useful information as well as noise and unwanted information This ensures data quality and facilitates the application of analytical techniques to extract meaningful information about human behavior [12].

I.5.3 Feature extraction

The feature extraction, used to identify actions from sensors data. This process of extractions divided in two methods requires experts to identify statical frequency metrics, the automatic extraction uses algorithms to discover patterns directly from raw data this method is more adaptable and more performed then manual extraction.

I.5.4 Classification

This part concerns the development of a classification algorithm that will be trained and tested on data. The training can be approached in various ways it can using:

I.5.4.1 Supervised learning

Supervised learning is characterized by the use of labeled training data. The concept of a 'supervisor' guiding the learning system by providing labels for the training examples. In classification tasks, these labels are usually class categories. Supervised learning algorithms generate models based on this labeled data, which can then be applied to classify new, unlabeled data. [16].

I.5.4.2 Unsupervised learning

Unsupervised learning can be used to identify similarities among a set of examples, a process known as clustering. It can also be employed to understand the data distribution, referred to as density estimation, or to reduce the dimensionality of the data, which is called subspace learning [17].

I.5.4.3 Semi-supervised learning

Semi-supervised learning is often used when labeled data is scarce or expensive to obtain, as it leverages both labeled and unlabeled data to improve the model's performance. By incorporating information from unlabeled data, semi-supervised learning algorithms aim to generalize better and make more accurate predictions compared to purely supervised learning approaches [18].

I.6 Sensors technology use on field of HAR

In the field of HAR various sensor technologies are utilized to capture data related to human movements and behaviors. These sensors are often embedded in wearable devices, smartphones, or deployed in the environment to monitor activities [5].

Here are some common sensor technologies used in HAR.

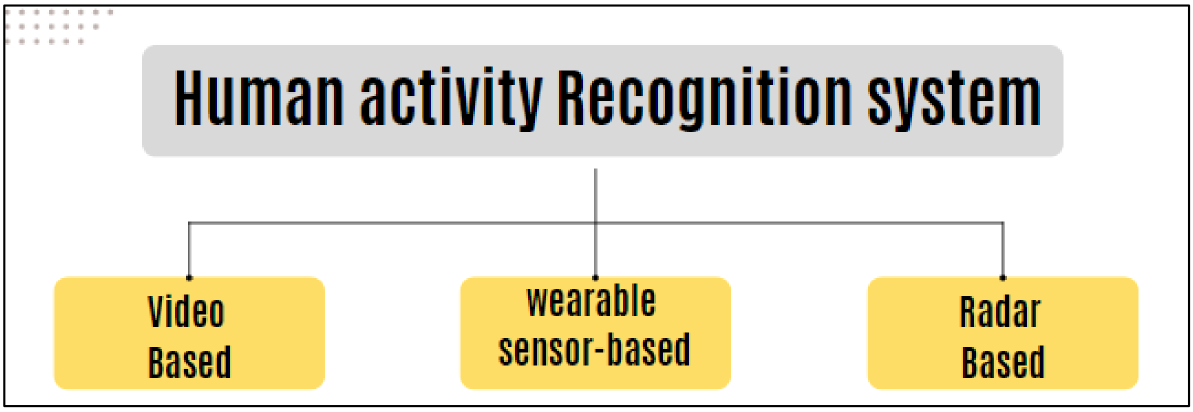


Figure I- 3: Approaches Employed for HAR

I.6.1 Vision Sensor (Camera)

A vision sensor dispersed in the environment, discrete and integrated into the environment, provides interference-free monitoring for the user, capturing visual data for HAR without interfering with the user experience.

Camera sensor ensures such as Supervision Track human activities, thus facilitating carrying. Visual behaviors, while remote control allows for rapid response also Human trafficking management optimizes flows and analyze risks.

Camera sensor Costs and maintenance must be managed to ensure sustainability. Data privacy and security are essential to protect the privacy of individuals. Dependence on electricity and the Internet can be a challenge, while the risk of piracy requires constant vigilance. Confidentiality limits must be clearly defined to ensure ethical use of the collected data.

I.6.2 Wearable Sensor

Wearable sensors have a high potential for improving current medical systems because they can monitor a wide range of physiological functions and provide highly efficient economic services in a variety of fields. However, wearable sensors require appropriate energy sources because their operation is dependent on a continuous, stable, and sufficient supply of power. Different types and applications of wearable sensors require different amounts of energy to function [19].

Wearable devices such as caps, smart watches and smart clothing are designed to monitor various aspects of health, fitness and sometimes even the environment around the user. Some examples of sensors commonly built into these devices include accelerometers for measuring body movements, heart rate sensors for monitoring heart activity [43] It is important to note that these sensors need to be worn, which can sometimes be forgotten by users. Finding the best location on the body is crucial to ensure accurate and comfortable measurement. In addition, sensors built into portable devices are generally designed to not cause intrusion problems [20] thus ensuring comfort and ease of use for users.

Connectivity and data sharing facilitate access and collaboration, while personalization and seamless integration enhance the user experience. However, data reliability, external interference and battery life are capacity challenges.

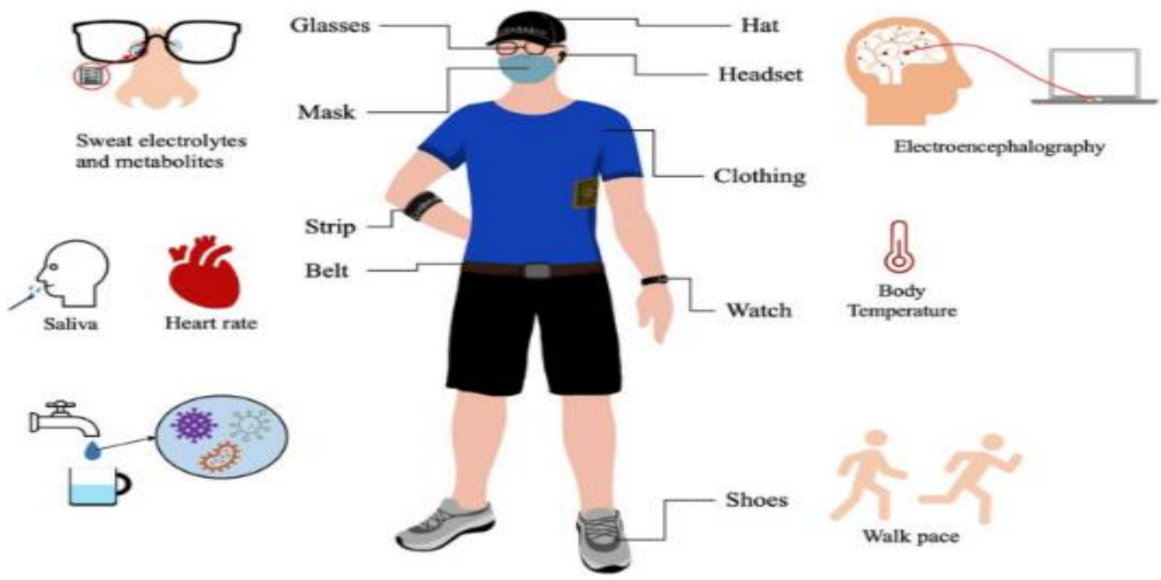


Figure I- 4: Wearable medical and healthcare devices for various regions of the body [19]

I.6.3 Radar Sensors

The context of research and technology, radar is a system that uses radio waves to determine the distance (ranging) [21], direction (azimuth and elevation angles), and radial velocity of objects relative to the site [22].

These sensors emit electromagnetic waves through space, which bounce off objects and return to the sensor. By measuring the time required for the return of waves and analyzing the frequency and phase changes of reflected waves, radar sensors can calculate various information about detected objects [23].

There are different types of radar sensors, including pulse radars, continuous wave radars, Doppler radars and synthetic aperture radars [24].

Each has its own advantages and is suitable for specific applications depending on range, resolution, speed and other factors.

Radar sensors have several distinct advantages over the other sensors, while filling some of their gaps. Some of these benefits' instant response to current events. Due to their penetration capacity, perceive various obstacles, such as smoke or even walls, ensuring complete and effective monitoring. Immunity to lighting conditions, privacy and confidentiality are preserved, Radar visible by size, its accuracy and reliability can be compromised in certain situations due to their sensitivity to electromagnetic interference [25].

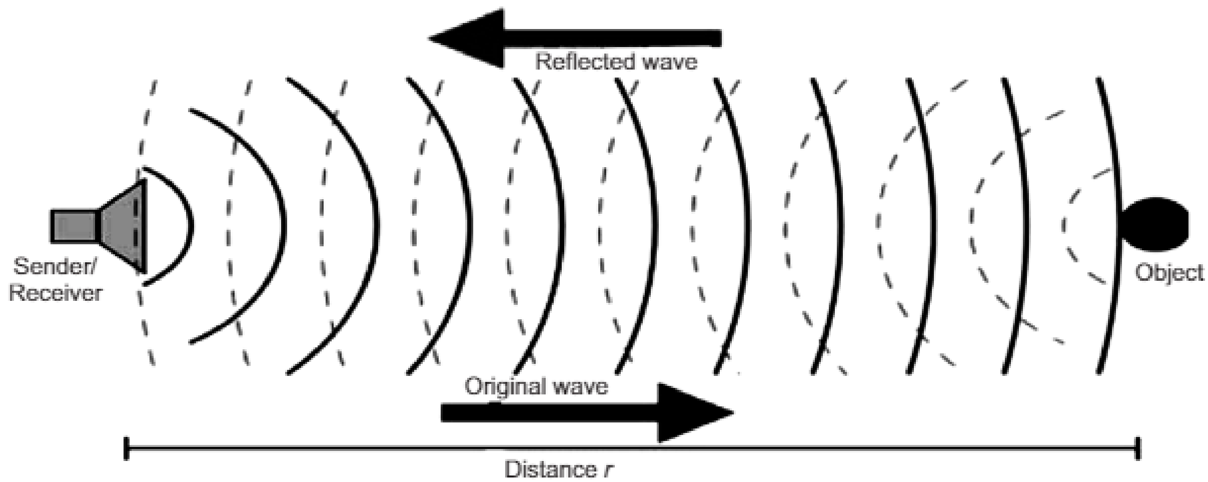


Figure I- 5: Continuous Wave Radar [12]

I.7 HAR Applications

HAR has a wide application prospect, and has great potential in healthcare, smart home, and other indoor monitoring applications [26-28].

HAR has advanced significantly over the last decade. Successful HAR applications include surveillance [29], smart homes [30], video analytics [31], autopilot [32], and human-computer interaction [33]. The goal of HAR is to identify a user's behavior so that computing systems can proactively assist the user [34].

Health and well-being: HAR systems can be used to monitor daily physical activities, sleep, dietary routines, etc. This information can be used to encourage healthy living and prevent chronic diseases.

Fitness and sport: wearable devices equipped with motion sensors can recognize and track physical activities, such as running, walking, cycling, etc. They can provide valuable data on sports performance and help improve workouts.

Security and surveillance: video surveillance systems equipped with human activity recognition software can detect and report suspicious behavior or dangerous situations in environments such as urban areas, the industrial facilities.

Assistance for the elderly and persons with disabilities: Technologies for the recognition of human activities can be integrated into assistive devices for the elderly or disabled. For example, they can be used to detect falls, monitor daily activity, and send emergency alerts.

Home automation: HAR systems can be used to automate certain household tasks according to the needs and habits of the occupants. For example, turn lights on or off, adjust heating, etc., depending on the activity detected.

For example, in elderly care, HAR systems can be applied to detect life-threatening activities like “falling” [33]. Commonly, HAR is achieved by exploiting the features of human motions embedded in the Doppler-time (DT) domain (or spectrogram) [34].

HAR can also be used to recognize entertainment activities, such as sport [35–36], dance [37,38] and gaming [39-41], in order to enrich lifestyles, Various systems activity recognition systems have been proposed [42].

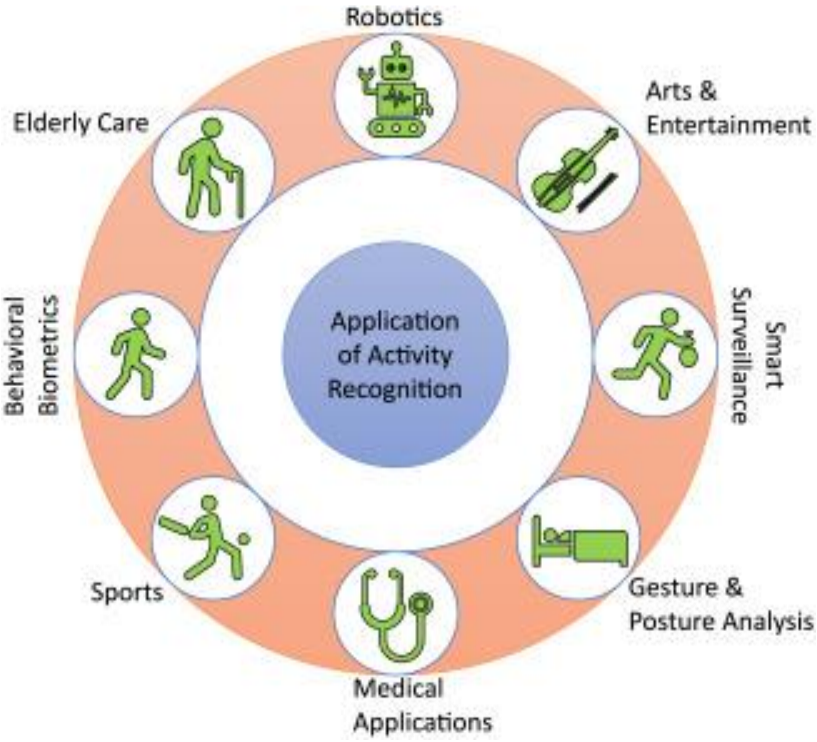


Figure I- 6: Applications of HAR [32]

I.8 Conclusion

This chapter concludes by offering a thorough overview of human activity recognition, covering everything from the definition of human activity as well as the discussion of different activity kinds and the recognition procedure. It explores the process of gathering data, preprocessing it, creating a model, and testing it. It also looks at the application of sensor technologies emphasizes how important this field is to comprehending. The next chapter we delve into UWB technology, ML, and DL, examining their significance in HAR via radar. We also review relevant literature, exploring multi-domain fusion strategies and the potential of employing multiple radars to address HAR challenges.

Chapter II: Literature review

Chapter II: Literature review

II.1. Introduction

The term UWB has appeared to refer to a number of synonymous terms for: pulses, carrierless, base band, time domain, no-consonantal, radio/radar signals of large bandwidth [45]. Although the foundations of technology have been known since the mid-1960s, the term UWB was first introduced in 1990 by the department of defence of the United States of America. UWB signals were mainly used for radar applications and not for radio communications [46].

This technique was developed in the mid-1980s. It's a spectrum-spaced technology. Its mode of operation is entirely different from other existing radio technologies. IT consists of sending and receiving signals without the use of carrying frequency [47,48], of extremely wide spectrum (of the order of GHz) and of fairly low power (close to the noise threshold). Thus, the spectral density of the radiated power is extremely low so that other systems perceive it as a background noise.

In this section, we'll present the history and principle of the UWB technique before addressing its main advantages and disadvantages as well as the different regulations imposed by European and American standardization bodies. Next, the different areas of application of UBB technology and the existing devices to date will be presented.

II.2. History of UWB Technology

Contributions to early electromagnetic studies based on non-sinusoidal signals in the temporal domain and work on UWB signals began in the late 1960s. This began with the contributions of Harmuth pioneers to the Catholic University of America [49], Ross and Rabbins to Sperry Rand Corporation [50,51], Paul van Etten to the USAF of Rome Air Development Center [52] and to Russia [45].

Many books and publications by Harmuth, (1969-1984), dealt with the basic concept of UWB transmitters/receivers. At about the same time and independently, many patents [53] were filed concerning the use of UWB signals in a number of application areas, including communications and radar. [54] is an example of UWB communication patent. 5year Etten experiments to test UWB radar systems resulted in the development of systems and antenna concepts [52].

Since the 1980s, several academic papers have made public information on the design of generators, antennas and impulsive receivers. In addition, these studies have demonstrated the interest of this technique for communications with low probability of detection and interception (LPI/D). In 1989, Time Domain Corporation loc. demonstrated the feasibility of a low-energy radio communication system based on very short pulses and time modulations [45]. Subsequently, it was the turn of Aether Wire & Location loc. to create localization systems based on this technique with a low-speed communication capability [55].

Before 1994, most of the work was aimed at improving certain subsystems of this new technique. At that time, it was given several names: technology "without a carrier", or "base band" or "impulsive" technology. It was in 1989 that the term "ultra-broadband" appeared in a publication by the Department of Defense in the United States. Most of the work on this technology is carried out within the framework of confidential U.S. programs (military). In 1994, the confidentiality of UBB work was lifted, and since that date, research has grown significantly both in industry and in universities.

Since the 1998s, civil industries have become increasingly interested in this technology and have pushed the U.S. government to take steps to regulate UWB technology. Subsequently, the FCC (Federal Communications Commission) issued a public notice [56] to evaluate the possibility of allowing the use of systems using the UWB. As a result of this publication, she received 150 responses and comments, from various organizations and companies involved in close or remote use of the UWB. In 2000, it proposed to include UWB systems within the framework of regulation [57]. This proposal has been updated following the comments of various industrial partners on the document. In February 2002, the FCC published an initial report on this technology, granted permission for commercial deployment, established suitable technical guidelines and imposed usage limitations [58].

Today, the great flexibility of UWB technology in terms of variable flow rate, different possible modulations, low consumption and low cost predicts its use for different applications: communications (reconfigurable networks, high-speed), localization (better precision) and radar (high detection precision).

II.3. UWB Technology

II.3.1 Definition

The emission of ultra-wideband (UWB) trains into very broad bands is the foundation of UWB technology. Can be used to measure travel times for location estimates, or to transmit data by modulation (in position or in polarity or other types of modulation). They can also be used to examine reflections on obstacles in radar and imaging applications, which will be discussed in more detail in the second part of this chapter.

In addition, the shape of UWB and their spectrum are affected by the antenna that radiates the signal. The antenna acting like a pass-down filter. This phenomenon will be explained in detail in the following paragraphs.

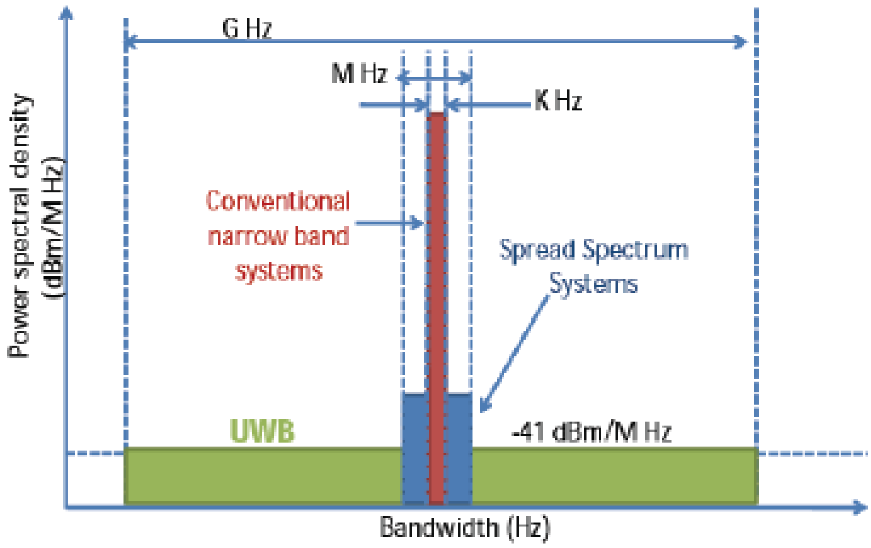


Figure II- 1: Comparison of spectrum occupancy between UWB and other radio systems [59]

II.3.1 According to Taylor

The first definition is given by Taylor [47]. According to the latter, the term UWB refers to systems that emit and receive waves whose relative bandwidth LB_r is greater or equal to 25% of the central frequency f_e. The bandwidth is given by the following equation:

$$LBr = \frac{f_h - f_l}{f_c} \quad \textbf{II.1}$$

Where f_e refers to the central frequency and its expression is given by:

$$f_c = \frac{(f_h - f_l)}{2} \quad \textbf{II.2}$$

f_h and f_l, show the upper and lower edges of the signal frequency band, respectively.

II.3.2 According to the FCC (Federal Communication Commission)

The FCC or Federal Communications Commission suffices in Taylor's sense for its regulatory proposal opinion, but by making two clarifications to the given definition [57]:

- The relative bandwidth must be that measured at -10dB below the maximum power.
- A signal that occupies more than 1.5GHz of band width is also considered a UWB signal.

The mean powers associated with these types of signals are generally very low because the cyclical ratio, described as the ratio of the pulse duration to the impulses repetition period, is also very low.

However, the FCC proposes to limit the increasing power emitted in an ultra-wide-band signal. This power should never exceed a certain P level above the average emission limit allowed by the "Part 15" of the Commission Rules [57]. The formula used to calculate this P level in dB over the entire bandwidth is as follows:

$$P=20+20log_{10}[\frac{(La\ largeur\ de\ la\ bande\ \text{à}-10dB\ du\ signal\ en\ Hz)}{50MHz}] \qquad \textbf{II.3}$$

For a bandwidth of 1GHz to -10dB, for example, the increasing power emitted should not exceed the average emission limit of 46dBm. Furthermore, the FCC proposes that the maximum *P* value be set at 60dB.

II.4. Advantages and disadvantages of UWB

Like all other radio communication techniques, UWB has its advantages and disadvantages.

II.4.1 The advances

Compared to other conventional transmission systems, the ultra-wide band has a number of notable advantages.

- UWB signals have a low spectral power density, since the signal power is distributed over a large bandwidth. This particularity gives systems using the UWB a low probability of detection and interception [60].

- UWB signals interfere little or no with other signals, such as narrow-band signals because their power is very low on the small part of the spectrum concerned. Thus, their use guarantees greater safety [61].
- UWB signals have a good penetration capacity due to their large bandwidth. They can thus cross surfaces, such as walls, unlike other technologies such as infrared [62].
- UWB signals allow for high precision in distance measurement, because the resolution achieved is inversely proportional to the duration of the pulse. It is thus possible to obtain accurate information on the position of the issuer. UWB signals are therefore of great interest to the radar detection functionality.
- UWB signals suffer very little from the influence of multiple trajectories, because, thanks to the shortness of UWB impulses, the direct signal arrives well ahead of those corresponding to the secondary pathways, without any recovery occurring [63].
- UWB-based systems have better capacity. Indeed, Shannon's Theorem teaches us that the capacity of a system is determined by the following formula:

$$C=B.\log2(1+SNR) \qquad \textbf{II.4}$$

Where B is the bandwidth of the system and SNR is the signal to noise ratio [64].

From this formula, we deduce that a system's capacity increases logarithmically with the signal to noise ratio, but grows linearly with bandwidth. Thus, the best way to increase a system's capacity is to increase its bandwidth, which the UWB technique accomplishes.

- UWB systems are a good compromise due to their simplicity and low cost of components [63].

II.4.2 Disadvantages

UWB communication systems are limited to solving a few technical challenges:

- The signals various frequency components propagate at different speeds.
- Frequency dispersion phenomena can then appear. For example, waves do not propagate at the same speed in the 0 to 3GHz band, particularly in a heterogeneous propagation environment [60].
- The unequal mitigation of the different frequency components can lead to phenomena of distortion in frequencies.
- UWB signals are filtered by the antennas. The challenge is to design small and efficient antennas across the entire frequency band used (Ultra-broadband antenna).

- The UWB system will probably never be able to surpass the performance of high-speed optical systems (of the order of several Gbits/s), but these are generally much more expensive [60].

II.5. HAR Radar System

Over the past few decades, radar has mainly been used in remote sensing systems, such as geophysical monitoring, air and land traffic control, and satellite remote sensing [65]. Additionally, short-range radar has recently been expanded for HAR duties.

Because radar is not affected by light or weather, HAR methods based on radar have higher reliability than those based on vision [65]. Without a tag affixed to the human body, they are able to directly detect human presence and activity. The Doppler frequencies and speeds of an individual's body components fluctuate in relation to their motion when they are in motion. As a result, the time ranges of these parts are not linear. By utilizing the targets' range, speed, and angle information, radar may be able to detect human activity [66].

Radar is a suitable technology for measuring human motion because of its inherent advantages, which include penetration capability, easy system integration, low cost, and simple architecture [67]. HAR uses a variety of radars, including IR-UWB and CW radars. Although there are numerous sophisticated studies on noise radar HAR systems [68-73], they are outside the purview of this work because they do not use machine learning approaches. We then go over a few types of radar that are intended for HAR applications. Table 1 lists those radars and their fundamental attributes in brief.

Table II- 1: Radar system and basic characteristics [68-73]

CW Radar	Doppler radar	Emitting single-tone radio waves with the ability to gather target doppler/radial velocity data
	FMCW radar	Giving simultaneous target range and speed information appropriate for situations where multiple targets are present
UWB Radar		Delivering fine range resolution that can identify the main dispersion centers

II.5.1 UWB techniques

A UWB spectrum can be obtained using a variety of waveforms (Figure II–2).

Two techniques can be distinguished:

- The temporal technique. The signals emitted are pulses, random noise or frequency-modulated signals. Measurement is performed in the time domain.
- The harmonic technique, which uses sinusoidal signals whose frequency is such as Step frequency or Frequency Modulated Continuous Wave (FMCW). Measurement is performed in the frequency domain.

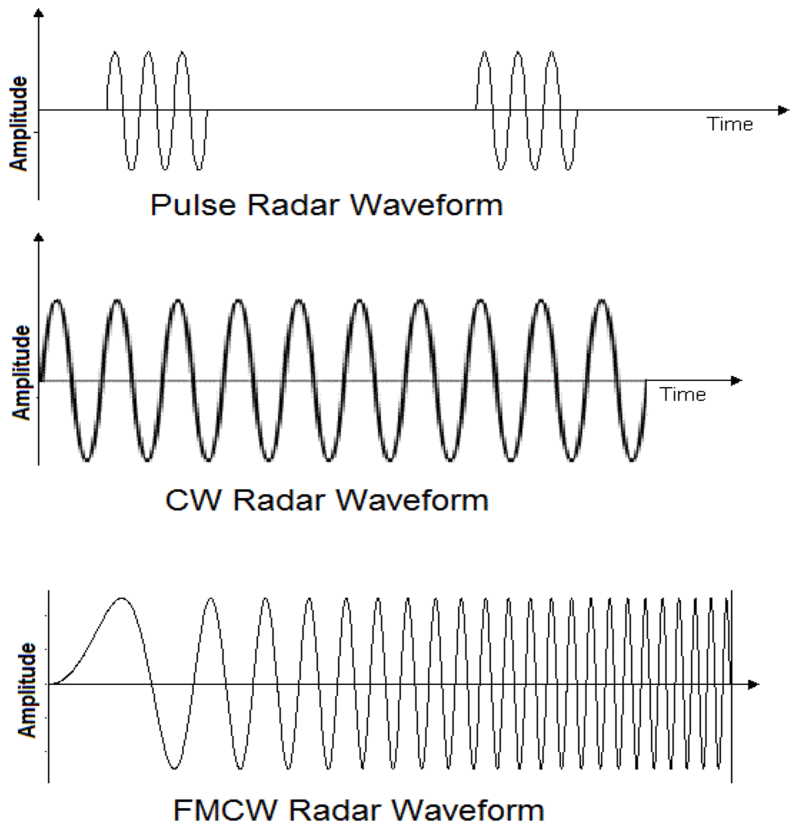


Figure II- 2: Different Types of UWB Waveforms [74-76]

II.5.2 IR-UWB

The principle of IR-UWB consists of very brief energy switching in an UWB transmission chain. The UWB signal emitted is an ultra-short, nanosecond, carrierless pulse. Its instantaneous spectral content ranges from hundreds of MHz to several GHz.

Typically, the system consists of a generator associated with a UWB transmitting antenna. The temporal response of a target is received by a UWB receiving antenna. Time-domain acquisition is performed by a fast sampler.

The transient response of a target illuminated by an impulse signal is very rich. This response can be used to locate the target by calculating the pulse's travel time.

The shape of the received signal may differ from that of the transmitted signal. It is made up of different contributions:

- Direct scattering from the target, known as specular reflection (from the Latin speculum = mirror).
- The sustained regime, the target's behavior in the presence of illumination.
- The free response when the pulse has passed. It is at the origin of the analysis of the poles for identification purposes.

II.5.3 Continuous-Wave (CW) Radar

A known stable-frequency CW ratio signal is transmitted by CW radar, which then receives the reflected signal modulated by objects on the ratio signal path [65]. It can function in both unmodulated and modulated modes. Because of its straightforward architecture, ease of system integration, and low power consumption, CW radar is a good choice for portable and mobile applications. For HAR applications, a variety of commercial CW radar chips and systems are available [66]. These include the 24GHz BGT24MTR11 of Infineon (Neubiberg, Germany), the 77GHz TEF8181EN and TEF8102EN of NXP (Eindhoven, The Netherlands), and the 77GHz AWR1642 and AWR1443 of Texas Instruments (Dallas, TX, USA).

II.5.3.1 Doppler Radar

One of the most widely used radars in HAR is the doppler radar [77,78]. There is no modulation involved in the single-tone radio waves that a Doppler radar emits. The Doppler effect causes the frequency of the received signals to be shifted away from the transmitted ones when the target is moving. The backscattered signal's frequency $f r$ is displayed as follows [67].

$$f_r = f_t(1+v/c)\,/\,(1-v/c) \qquad \textbf{II.5}$$

where c is the speed of light, f_t is the frequency of the Doppler radar single-tone radio signals, and c is the target's radial speed. Therefore, f_d or Doppler frequency shift is

$$f_d = f_r - f_t = 2v f_t\,/\,(c-v) \qquad \textbf{II.6}$$

Doppler radar capacity to record Doppler shifts makes it useful for detecting time-varying radial speeds of human motion. Doppler radar can achieve notable performance in motion and displacement measurement because of its comparatively simple signal processing [79, 80].

II.5.3.2 Frequency Modulation Radar (FMCW)

The signal in the waveform is sinusoidal, with a frequency variation that is linear in time. The transmitting system consists of a voltage-controlled oscillator (VCO) driven by a system responsible for controlling the frequency variation and making it linear as a function of time. To achieve this, the control system takes a tiny fraction of the transmitted signal [81].

After that, the signal is sent to a transmitting antenna. The target responds to a receiving antenna. A portion of the transmitted signal is combined with the received signal and detected by a probe.

This type of radar is often used for distance measurements (altimeters, fluid level measurement, snow layer surveys, etc.).

For example, Thales is developing a radar of this type operating in the 800 MHz - 1 GHz band, designed to detect targets in buildings. This system is made up of 3 radar modules that detect targets and triangulate to locate them [82].

II.6. Machine Learning (ML)

ML, by definition, is a sub-domain of AI (Artificial Intelligence), which is to teach a computer to do a task without explicitly programming it by teaching it from examples and data. A central idea is to replace writing scripts or abstract instructions with describing algorithms and models that learn from experiments and data. Machines then become able to adapt to new information and make choices based on stored knowledge when performing past tasks. The computer system then analyzes the information, detects patterns, and makes predictions, without the system user intervening.

The most common algorithms:

II.6.1 Support vector machines

Support vector machine (SVM) family of machine learning techniques was first introduced to address the classification problem and subsequently extended to a number of other contexts. Based on the concepts of convex optimization and statistical learning theory, they are currently applied in computer vision, bioinformatics, and text classification, among other fields [83].

II.6.2 Random forest

The random forest prediction algorithm was created by Ho in 1995 and formally unveiled by researchers Adele Cutler and Leo Breiman in 2001. As we shall see, it integrates the ideas of bagging and random subspaces.

Random forest decision trees are trained independently using subsets of the data learning set (bagging method). Once all estimates have been produced, the data are combined to determine the final prediction, which is shown by a lower variance. Put another way, the idea is to gain a deeper understanding of a problem by drawing inspiration from multiple approaches to it. Random subsets of decision trees are chosen from every model.

II.6.2.1 Working of random regression forest

You can use a random forest to make a regression. The random forest of regression, based on a bagging system, is to compute schematically the average of the predictions obtained from the set of estimates of the decision trees of the random forest.

II.6.2.2 Working of classification random forest

In a random classification, the final estimate is to select the most common response category. Instead of exploiting all the results obtained, a selection is made by looking for the prediction that returns most often [84].

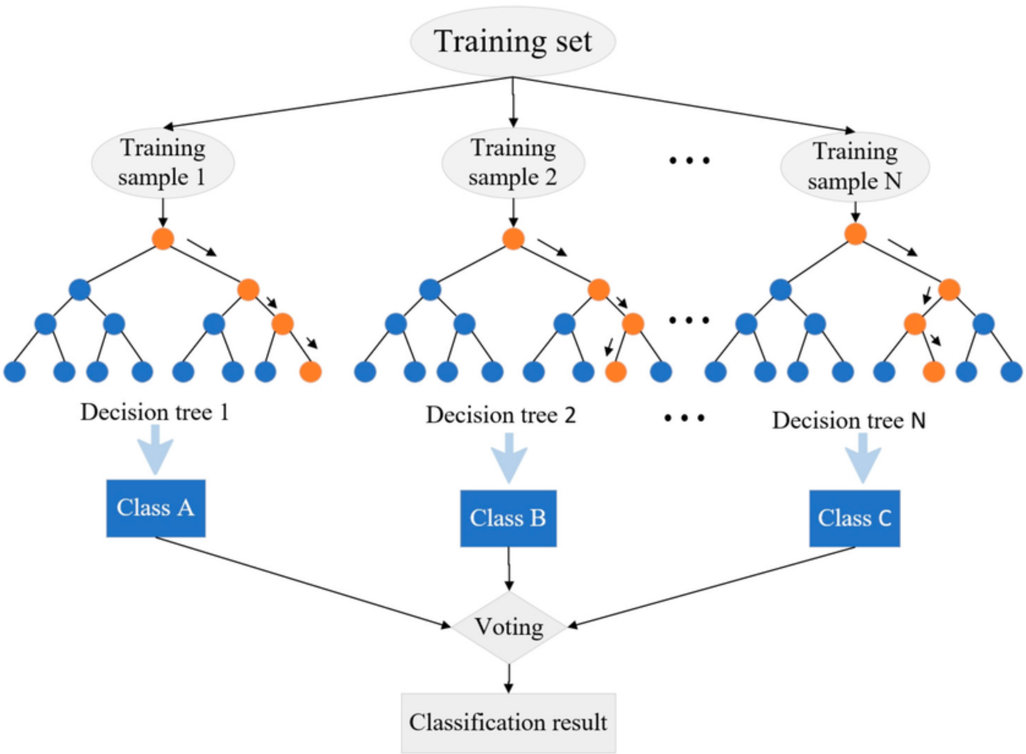


Figure II- 3: Random Forest Architecture

II.6.3 Decision Tree

A decision tree consists of an organizational chart built around a main idea, then structured according to the repercussions of your choices. This tool is named because it looks very much like a tree and has many branches.

Trees of this kind visually illustrate the results, expenses and possible consequences of a complex decision, which you can then study. This tool is perfect for evaluating the value of each expected outcome based on the decision taken, as well as clarifying the implications of your decisions. Then compare the various results to quickly determine the most relevant action plan to implement. A decision tree can also be used to manage expenses, find cost sources, and solve problems [85].

Symbols of a Decision Tree

The symbols of a decision tree are as follows:

- Alternative branches: lines that start from a decision represented on your tree. These branches indicate the results (options) or decisions that arise from the initial decision.
- Decision node: A square that represents a decision to be made. Any decision tree starts with this symbol.
- Opportunity node: A circle that represents a result among all possible ones.
- Final node: a triangle that indicates the final result (the end of the tree).

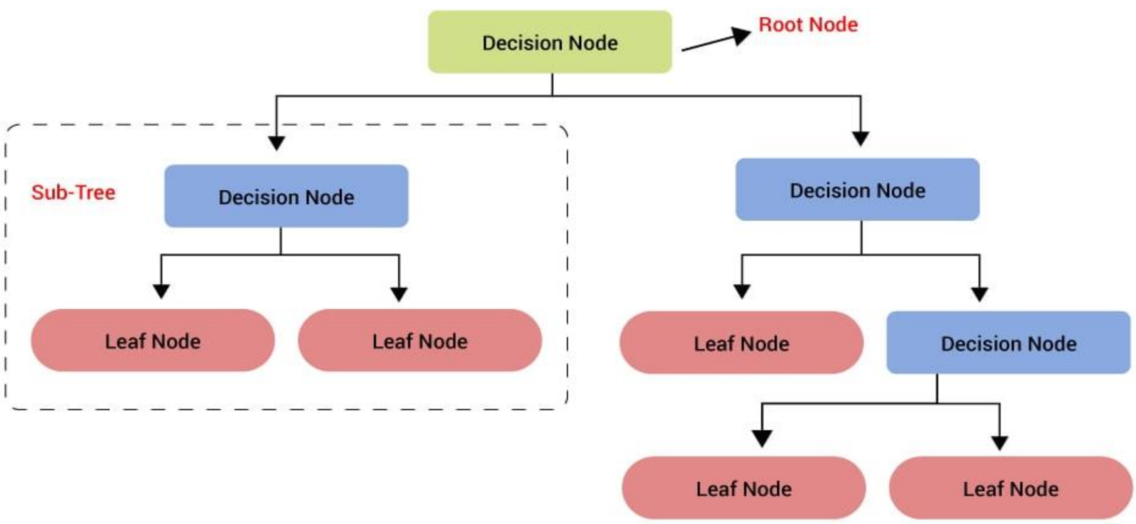


Figure II- 4: Decision Tree Architecture

II.6.4 Extra-Trees models

Extra-Trees is an ensemble machine learning technique that trains multiple decision trees and combines their output to produce a forecast. It trains decision trees using the complete dataset, RANDOMLY SELECTING the values at which to split a feature and build child nodes to guarantee significant variances across individual decision trees. Unlike Random Forest, Extra Trees uses a greedy search algorithm to choose the value at which to split a feature [86].

Extra-Trees can lessen the bias of the model by using the full dataset, but it increases bias and variation by randomly selecting the feature value at which to split. A bias-variance study of several tree-based models found that Extra Trees exhibit greater bias and lower variance compared to Random Forest on most classification and regression tasks.

Extreme Gradient Boosting (XGBoost) is an approach to learning derived from Boosting Tree models. It uses the CPU multithreading capability and applies a second-order Taylor expansion to the loss function to achieve parallel computing. XGBoost employs techniques to prevent overfitting, such as a splitting threshold and random attributes when creating each tree. These techniques improve XGBoost performance in real-world situations and make it more broadly applicable.

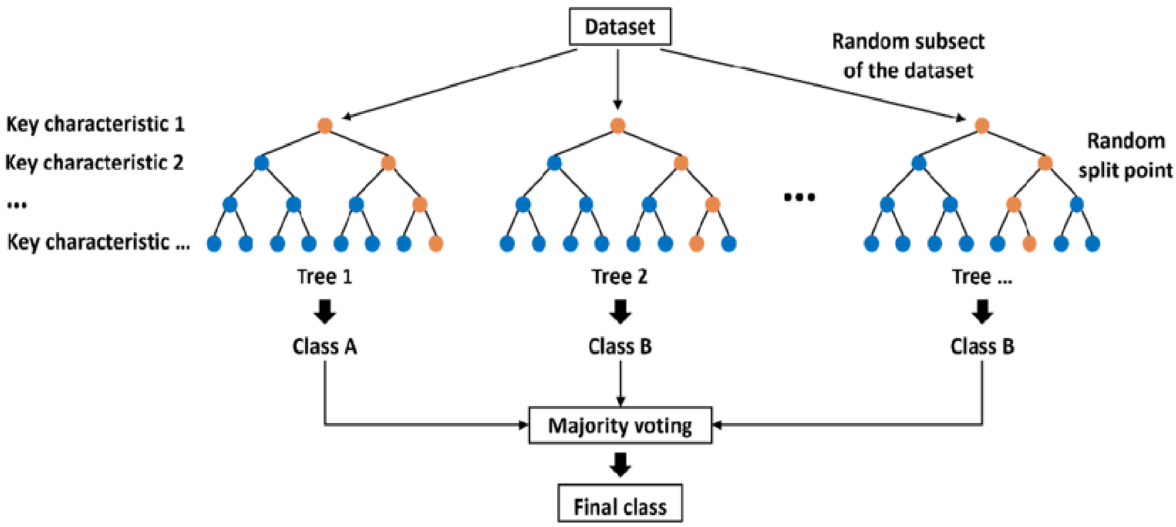


Figure II- 5: Extra Tree Architecture

II.7. Deep Learning (DL)

Teaching computers to process data similarly to the human brain is the goal of deep learning. DL models can produce precise information and predictions by identifying intricate patterns in text, images, sounds, and other types of data. DL techniques can be used to automate tasks that are typically performed by humans, like text transcription from audio files or image description [87].

II.7.1 The importance of DL

AI aims to teach machines to think and learn like people. DL technology powers numerous AI applications found in commonplace goods, including:

- Electronic assistants.
- Voice-activated remote controls for television.
- Fraud identification.
- Automated facial identification.

It is also a crucial component of modern technologies like virtual reality, driverless vehicles, etc.

Data scientists create DL models from computer data to carry out tasks based on pre-established steps or algorithms. Businesses evaluate data and make predictions using deep learning models for various purposes.

II.7.2 Working of DL

Neural networks consist of interconnected nodes or neurons in a layered structure that relate the inputs to the desired outputs. The neurons between the input and output layers are referred to as hidden layers. Consider the human brain as an example. It is made up of millions of interconnected neurons that collaborate to process and learn information.

DL models can have hundreds or even thousands of hidden layers, which is why they are called "deep". These models are trained using large sets of labeled data and can learn features directly from the data.

Nodes, sometimes referred to as artificial neurons, are computer programs that process data through mathematical operations. These nodes are used by artificial neural networks, which are DL algorithms, to solve complex cases [87].

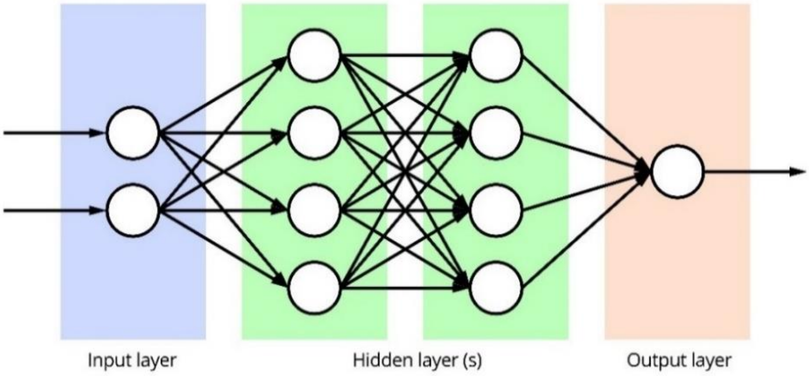


Figure II- 6: Working of DL

II.7.3 The advantages of DL over ML

The following are the advantages of a DL network over traditional ML.

II.7.3.1 Efficient processing of unstructured data

ML techniques discover that because the learning data set can have an infinite number of variations, processing unstructured data, like text documents, is complex. Moreover, deep learning models have the ability to enter unstructured data and make general observations without having to extract the characteristics manually.

To navigate serenely between ML and DL, discover this summary table 2 specially designed by DataScientest:

Table II- 2: ML and DL summary [88]

Aspect	ML	DL
Data Dependency	Typically requires structured data.	Well-suited for unstructured data, e.g., images, audio, text.
Feature Engineering	Requires manual feature engineering.	Automatically learns features from data.
Network Depth	Shallower network architectures.	Involves deep neural networks with multiple hidden layers.
Training Data Size	Works well with smaller datasets.	Benefits from large datasets for better performance.
Interpretability	Easier to interpret and explain results.	Complex models can be less interpretable.
Computation Resources	Requires fewer computational resources.	Demands significant computational power (GPUs/TPUs).
Use Cases	Common for tasks like regression, classification.	Ideal for tasks like image recognition. NLP, and more complex problems.
Training Time	Faster training with smaller models.	Longer training due to deeper architectures

II.7.3.2 Hidden relationships and pattern discovery

A DL application has the ability to analyze larger amounts of data in greater depth and to uncover new information for which it may not have been competent [87].

Take an example of a deep learning model specifically designed to study human activity. Consider utilizing a fitness tracker software that gains knowledge from your regular workout routine. At first, it just records the exercises you have recorded, such as jogging, yoga, or walking. But it goes farther than that. The app compares your workout habits with those of people who have similar fitness objectives and lives by utilizing artificial intelligence. It makes recommendations for new activities based on the aggregate data, taking into account your needs and interests. For variation, it can suggest taking up dancing lessons or HIIT workouts if you enjoy cardio. If you enjoy yoga, it might suggest trying out tai chi or meditation classes for a more all-encompassing approach to wellness. Essentially, this customized method makes it easier for you to stay motivated, vary your exercise routine, and accomplish your fitness goals.

II.7.4 Radar data processing for HAR

We have covered a number of popular DL models and their benefits for HAR. For radar, it is preferable to create special DL algorithms for radar echoes because a radar echo includes time, range, and Doppler information. With this in mind, in this section, we describe slow time - fast time matrix processing, how a CW gets a matrix for processing, and deep learning methods based on the dimensionality of radar echoes for detecting human activity in radar.

To process a CW in the context of signal processing, it typically involves transforming the wave into a form that is amenable to matrix operations. This often entails discretizing the signal, followed by applying various mathematical techniques to analyze, filter, or manipulate the signal.

These data samples arranged in a two-dimensional (2D) Matrix, where each row is a range profile associated to a transmitted pulse. Thus, the row index corresponds to pulse number n (associated with its receiver turn-on time nTr), while the column index corresponds to a sample number within a range profile. The time corresponding to range on the horizontal axis is called “fast time”, since it is measured on a scale comparable to the individual pulse duration. The time on the vertical axis is known as “slow-time,” because it is measured on a scale comparable to the pulse repetition interval Tr.

Range-Doppler (RD) processing [89] uses the Doppler effect to determine the radial component of the target's velocity and then transforms radar signals into a 3D time-range-Doppler data cube. This makes it possible to resolve the different components of a target in terms of both Doppler and range. The 3D RD 'movie' depicts the slow-time evolution of the target's activity, as can be observed in Figure II-7a. Furthermore, 2D representations of radar signals can be created, like the rang-Doppler map (Figure II-7d), time-Doppler map (Figure II-7b), and time-range map (Figure II-7c). More careful customization of deep learning techniques to the diverse types of echoes is necessary to effectively leverage the information contained in them.

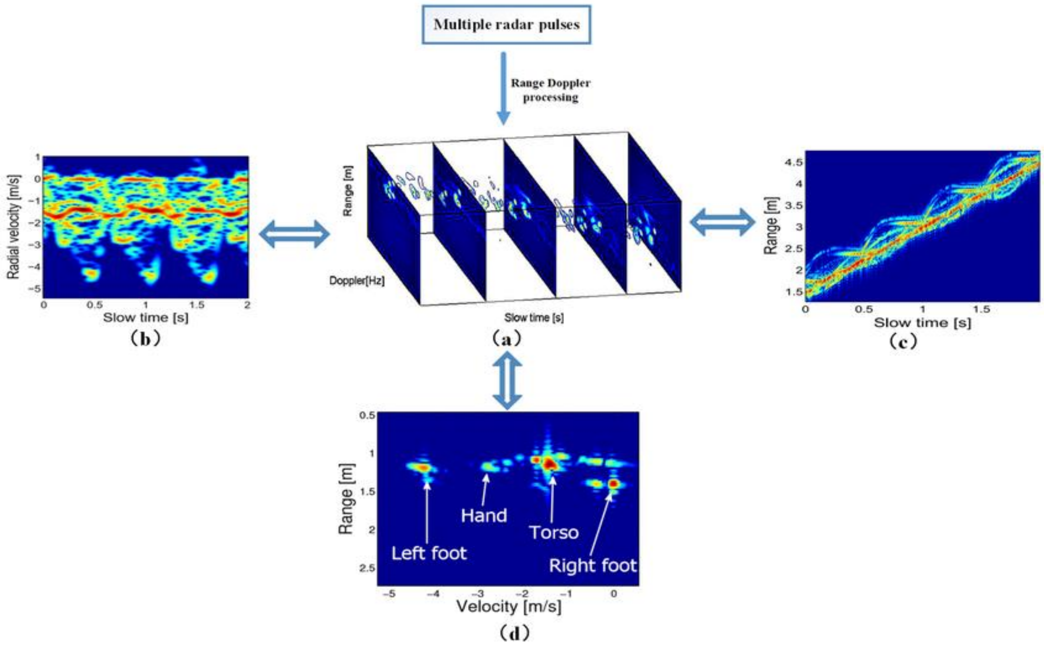


Figure II- 7: Summation of existing works on DL based HAR in radar

II.7.4.1 DL Approaches in 3D Radar Echo

Targets micro-Doppler characteristics and movement characteristics can be seen in range-Doppler frames [66]. The 3D RD video sequence, which is made up of N time-sampled 2D range-Doppler frames, exhibits both temporal and spatial properties. Every RD frame contains range and Doppler information, and time information is present in between frames. When compared to 2D echoes, nearly all of the activity information received by radar is contained in the joint time-range-Doppler echoes. It is necessary to have models that can extract both spatial and temporal information.

For 3D echo-based HAR, DL approaches are more practical and preferred than manual feature creation, which is difficult to achieve from 3D echoes, because they can automatically extract deep features. Furthermore, DL models can now efficiently process 3D data thanks to the introduction of GPUs. DL techniques for 3D echoes show promise for HAR, despite the fact that few DL algorithms for 3D radar echoes have been proposed thus far.

3D CNN is one of the most widely used models for processing 3D data nowadays [90,91]. Spatial-temporal features are automatically learned in the spatiotemporal model it builds, which is an expansion of the spatial CNN. Zhang et al. [92] introduced a recurrent 3D CNN model for continuous dynamic gesture recognition using an FMCW radar. Following the extraction of brief temporal-spatial features from continuous time-range maps using 3D CNN, an LSTM was employed for global temporal feature learning. When 3D CNN was utilized in place of a traditional 2D CNN, the experiment found that the recognition rate dropped by roughly 5%, suggesting that 3D CNN was better at learning hand gesture representations than 2D CNN. This method works well with a 3D data cube even though the 3D CNN requires time-range maps as its input because the cube contains almost all of the activity data in continuous time-range maps.

Using 3D radar echoes for HAR, HASD is a representative example of $G\,o\,o\,g\,l\,e\,S\,o\,l\,i$. Based on short-range FMCW radar, $G\,o\,o\,g\,l\,e\,S\,o\,l\,i$ is the first gesture recognition system that can identify a wide range of dynamic gestures [93,94]. The model is based on an end-to-end trained combination of deep convolutional and recurrent neural networks, and the dataset comprises 3D radar echoes. The ability to recognize different activities with a range of temporal and spatial distributions may be enhanced by the combination of CNN and LSTM. It was shown that the 3D range-Doppler video approach outperformed the frame-level classification approaches, and that the end-to-end "CNN + LSTM" method was able to explore the gesture information more fully than the single CNN or LSTM models. Following the release of $G\,o\,o\,g\,l\,e\,S\,o\,l\,i$ [91,92].

II.7.4.2 DL Approaches in 2D Radar Echo

Even though 3D human backscattering echoes contain a wealth of information about human activity, processing them is still challenging. Enough information about human activity is also carried by 2D radar echoes, which are primarily called time-Doppler maps, time-range maps, and range-Doppler maps. Since 2D echoes are usually treated as images, CNN has become the most popular model for 2D echoes, along with the computer vision line. As such, 2D echo-based HAR is often used to create an image classification task.

(1) Time-Doppler map, also known as micro-Doppler signatures, are crucial for radar-based HAR [95]. They provide time-varying Doppler information, affecting movement and vibration of body parts. Every body part has a varied range of motion and speed. The time-Doppler maps corresponding to these activities vary when the target behaves in different ways. Using STFT [96] and other joint time frequency analysis techniques, these maps are easily obtained. Compared to other 2D radar echoes, they are easily understood and are most frequently utilized for radar-based HAR [97–102].

Human gaits on time-Doppler maps were classified using a 14-layer deep CNN (DCNN), which outperformed SVM and artificial neural networks [100]. With a recognition performance of 94.2%, indoor human activities were distinguished using a CAE architecture [103]. It was suggested to use a DCNN-based hand gesture recognition system, with micro-Doppler signatures that changed according to aspect angle and radar distance. It was also suggested to use a DCNN architecture with cascading convolutional network layers, which outperformed the current feature-based techniques [104].

(2) Time-range map is made up of several pulses spaced out over time (Figure II-7c). It includes information about the target's and the radar's time-varying range. When a person is moving, different body parts are at different relative distances from the radar. Consequently, the time-varying range information of the human body can be utilized to identify human activities even though time-range maps ignore Doppler information [92].

(3) Range-Doppler map: see Figure II-7d to view the range and Doppler data of a moving target at a particular moment. It is capable of precisely locating the target and disentangling various moving human body parts. Furthermore, it is promising that range-Doppler maps can track multiple targets at once for the recognition of multiple human activities. To detect dynamic hand gestures, P. Molchanov et al. [105] used a short-range monopulse FMCW radar with one Tx and three Rx.

(4) Hybrid 2D maps, most HAR systems based on 2D radar echoes have only used one of the three types of maps mentioned above. Sometimes, though, it is noticed that certain activities that were clearly identifiable on one map might not be on another. This encourages the use of multiple maps in an effort to lower false alarms [107]. used range-Doppler, time-Doppler, and time-range maps to detect falls. The three maps' range and Doppler information were extracted, which decreased the false alarm rate associated with fall detection. Three stack AEs and three SoftMax classifiers were used in [106] to categorize the four human motions—

walking, bending, sitting, and falling. To fully explore the motion information contained in radar echoes, this method applied time-Doppler maps, time-range maps, and range-Doppler maps. Subsequently, the final result was determined by voting strategy by combining the three classification results. Tests revealed that the performance outperformed the one using a single type of map.

II.7.5 DL Techniques

With the introduction of DL algorithms, a number of domains have experienced rapid growth, including drug discovery, speech recognition, and visual object recognition [108]. A DL model can automatically learn high-level representations by utilizing multiple processing layers. Comprehensive feature engineering and domain knowledge are not required for DL. Furthermore, a great deal of deep transformations made it possible to solve difficult classification and recognition problems and learn incredibly complex functions [109, 110]. DL has thereby aided in the advancement of numerous fields, including HAR. This section examines a number of DL models and evaluates each one's special benefits for HAR tasks. Table II-3 provides a brief description of each model.

Table II- 4: DL models and advantages for HAR.

Models	Descriptions and Advantages
CNN	Capturing the spatial relationship by multiple convolutional layers, often utilized as an excellent localized feature extractor.
RNN	Exploring the temporal relationship in data, variants are often utilized, such as LSTM .
Hybrid deep models	The combination of some deep models, built on each model's own strength to obtain better performance.

II.7.5.1 Convolutional Neural Network

The arrangement of simple and complex cells that make up the visual cortex served as the model for the CNN. Four fundamental concepts are implemented: multi-layering, parameter sharing, pooling, and local connections. It is noteworthy that the first DL architecture with hierarchical layers was CNN [111]. In contrast to traditional ML algorithms, which require manual feature extraction, CNN can extract higher-level spatial features from lower-level ones hierarchically thanks to its multiple convolutional layers. Pooling and fully connected layers are typically used for classification or regression tasks after convolutional layers. Furthermore, CNN outstanding deep feature learning capabilities make it a popular choice for automatic feature extraction in a range of applications [112,113].

CNN offers two primary benefits when used for HAR [114]: it accounts for surrounding signals and does not scale differently for different rhythms or frequencies. The first advantage allows CNN to extract localized characteristics from positions connected to space, instead of extracting features from a single position. The second benefit is that the features that are retrieved maintain the information related to rhythm or frequency. As shown in (Figure II-8a), CNN classifiers with different architectures and convolution kernel sizes were employed to detect human activities using time-Doppler maps [97-102,104,115,116] . As demonstrated in Figure II-8b, CNN extracts high-level representations of human activities for further identification in [48-50,59] by acting as a spatial feature extractor.

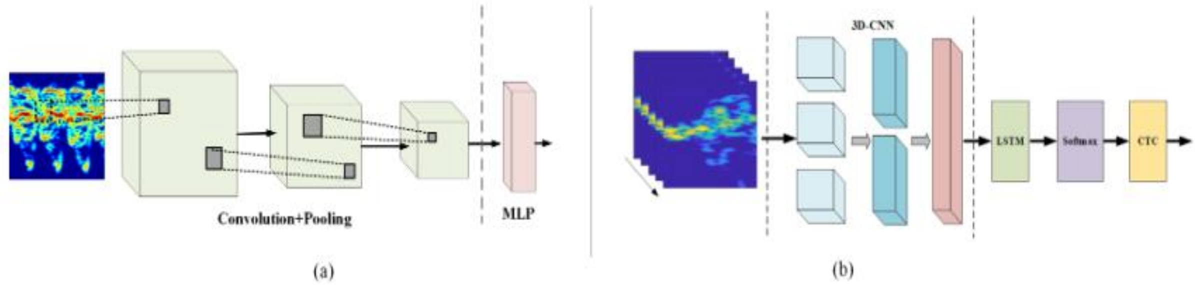


Figure II- 8: (a) CNN is used in radar-based HAR as a classifier [56], (b) High-level features are extracted from the input time-range maps using CNN as a feature extractor [54]

II.7.5.2 Recurrent Neural Network

Recurrent neural networks (RNN) have drawn interest from researchers in HAR due to their successful applications in speech recognition and natural language processing. Because RNN can mine both temporal and semantic information, it has shed light on modeling temporal sequences. From a network structure standpoint, RNN retains the prior data and applies it to shape the output of subsequent nodes. Long-term dependencies, however, are a limitation of traditional RNN. In order to address this weakness, LSTM which functions better in many tasks was developed [117], refer to Figure II-9. The input, output, and forget gates are the three unique gates that an LSTM is equipped with. The forget gate in particular allows LSTM to access a long-range context of sequential data. It is assumed that the prediction accuracy of RNN and its variants increases with the amount of available data, in contrast to CNN, which is limited to processing data of a certain size. With time, the prediction's outcome is evolving. As a result, RNN is more susceptible than CNN to changes in the input data.

RNN and its variants are superior to HAR in that they can better take advantage of temporal correlations within an activity, which is an important problem for human activity recognition. all used RNN and its variations to simulate the temporal features of human behavior [92–94, 118]. LSTM was used in [93] to learn complex dependencies across time in the features that a CNN extracted from range-Doppler maps because those features are time-correlated. In this way, CNN and LSTM worked together to explore both spatial and temporal information.

An LSTM network calculates a mapping from an input sequence $x = (x_1, \ldots, x_T)$ to an output sequence $y = (y_1, \ldots, y_T)$ using the following iteratively calculated equations from $t = 1$ to T:

$$i_t = \sigma\,(W_{ix}xt + W_{im}mt\text{-}1 + W_{ic}c_{t-1} + b_i) \qquad \textbf{II.8}$$

$$f_t = \sigma\,(W_{fx}xt + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \qquad \textbf{II.9}$$

$$c_t = ft \odot c_{t-1} + i_t \odot g\,(W_{cx}x_t + W_{cm}m_{t-1} + bc) \qquad \textbf{II.10}$$

$$o_t = \sigma\,(W_{fx}x_t + W_{om}m_{t-1} + W_{oc}c_{to} + b_o) \qquad \textbf{II.11}$$

$$m_t = o_t \odot h(c_t) \qquad \textbf{II.12}$$

$$y_t = \varphi\,(W_{ym}m_t + b_y) \qquad \textbf{II.13}$$

The weight matrices denoted by the W terms (e.g. W_{ix} is the matrix of weights from the input gate to the input), the diagonal weight matrices for peephole connections denoted by W_{ic}, W_{fc}, W_{oc} the bias vectors denoted by the b terms (b_i is the input gate bias vector), σ is the logistic sigmoid function; and the input, forget, output, and cell activation vectors, respectively, are *i, f, o*, and *c*, which are all of the same size as the cell output activation vector m; ⊙ is the element-wise product of the vectors g and h are the cell input and cell output activation functions, generally and in this paper tanh; and φ is the network output activation function, SoftMax in this paper.

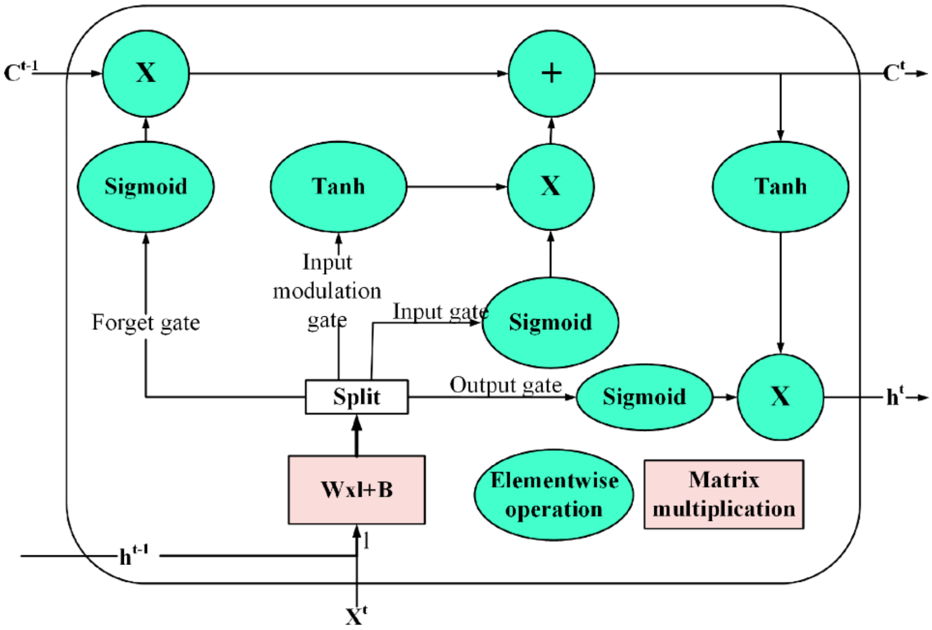


Figure II- 9: A schematic overview of LSTM cell.

II.7.5.3 Hybrid Deep Model

Every model has drawbacks of its own and isn't capable of doing everything. Hybrid deep models combine multiple networks and leverage the capabilities of each network. Such collaboration is based on the individual strengths of each model in order to achieve higher performance. Since CNN and RNN are effective at abstracting various domain features, they are currently frequently combined in HAR, as seen in Figure 8.b: RNN records temporal relationships, whereas CNN records spatial relationships [114]. gave useful illustrations of how to integrate CNN and RNN [92–94]. Research has shown that integrating CNN and RNN tends to strengthen the capacity to identify activities that differ in terms of time and space.

II.7.5.4 Convolutional Long-Short Term Memory (ConvLSTM):

LSTM cells have convolution operations, making ConvLSTM a 2D (two-dimensional) convolutional model [119]. Long-term dependencies can therefore be processed by this model. This process will be added with the convolution operation when the input matrix multiplication is calculated using LSTM cells. Two inputs the input matrix and the kernel matrix are required for the convolution process. Each kernel element is multiplied by the matching input component before being added together by the kernel matrix as it scans the input matrix [120, 121]. The prediction is optimized by iteratively adjusting the kernel weights during the training process [90]. The architecture of ConvLSTM cell is identical to that of the LSTM, with input, forget, and output gates as well as candidate values. Convolution operations on the hidden state and memory cells from the previous timestep h_{t-1} and the input at the current timestep x_t are used in the ConvLSTM cell to calculate the input, forget, and output gate information.

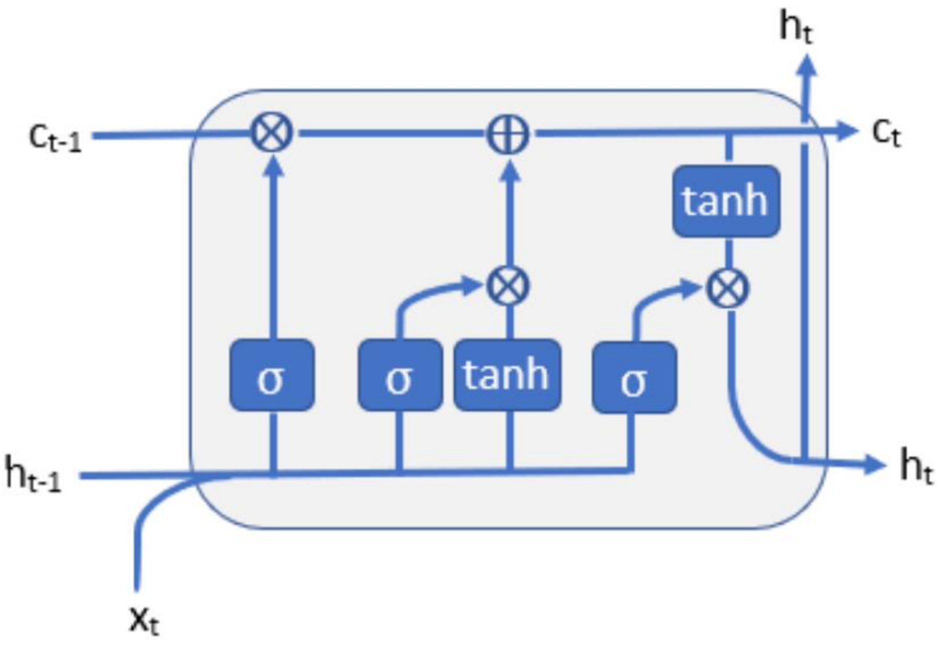
$$f_t = \sigma\left(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} \circ c_{t-1} + b_f\right) \quad \textbf{II.14}$$

$$i_t = \sigma\left(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} \circ c_{t-1} + b_i\right) \quad \textbf{II.15}$$

$$C_t = f_t \circ C_{t-1} + i_t \circ R\left(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c\right) \quad \textbf{II.16}$$

$$O_t = \sigma\left(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} \circ c_{t-1} + b_o\right) \quad \textbf{II.17}$$

The model's convolutional kernels are represented by W_{cf}, W_{ci}, W_{co}, W_{hi}, W_{xi}, W_{ho}, W_{xo}, W_{xf}, W_{xc}, W_{hf} represent convolutional kernels [119,122]. The additional connections found in the ConvLSTM cell above the LSTM cell, which are derived from the current and previous cell states (Figure II-10), which depicts the ConvLSTM architecture. The red line indicates which input, output, and forget gates for example, forget gates have a kernel matrix multiplication operation with the previous cell states $W_{cf} \circ c_{t-1}$. In addition, the LSTM's convolution process allows it to simultaneously represent fine grained local information in the feature data and capture temporal correlation. ConvLSTM can assist in reducing the size of the



model, particularly for large input sizes. The advantage of the ConvLSTM approach is that convolution is more flexible to represent more specific local information, while the LSTM prior performs well in terms of overall information interaction in weight calculation.

Figure II- 10: ConvLSTM [120,121]

II.7.5.5 Convolutional neural networks Long-Short Term Memory (CNN-LSTM)

LSTM detects both short- and long-term dependencies, CNN are frequently used in image and video recognition, image classification, and other domains requiring the processing of grid-structured data [123]. The local perception feature and weight sharing, which can drastically lower the number of parameters and thereby increase training efficiency, are two of CNN key benefits.

The CNN-LSTM architecture model is shown in Figure II-11. The two primary parts of the CNN process are the convolutional layer (2D Conv), which takes input from time steps in a 2D array. After that, it applies mathematical operations to extract features from the input data by taking unique features from the concentration data, such as patterns, trends, or specific variations, and adding non-linearity to the output by applying activation and convolution operations.

A feature representation matrix made up of multiple layers during this process represents the feature extraction outcomes from various filters. In the CNN's second layer, the Max-Pooling layer is used to pool input dimensions from the convolutional layer process, resulting in smaller segments (refer to the red color). To expedite the training process, dimensional reduction is achieved by choosing the maximum value of each piece. Next, there is a flattened layer that converts the CNN layer's matrix output into a 2D vector in order to process the data into the format required by the long short-term memory.

The LSTM procedure comes next, and the dense layer, or Fully-connected Layer (FC), assists in converting the LSTM output into a predictive value [123]. It follows that the work produced by the CNN layer or primary component will be gathered into a smaller dimension before being directed into the LSTM layer, resulting in predictions at the output layer [124].

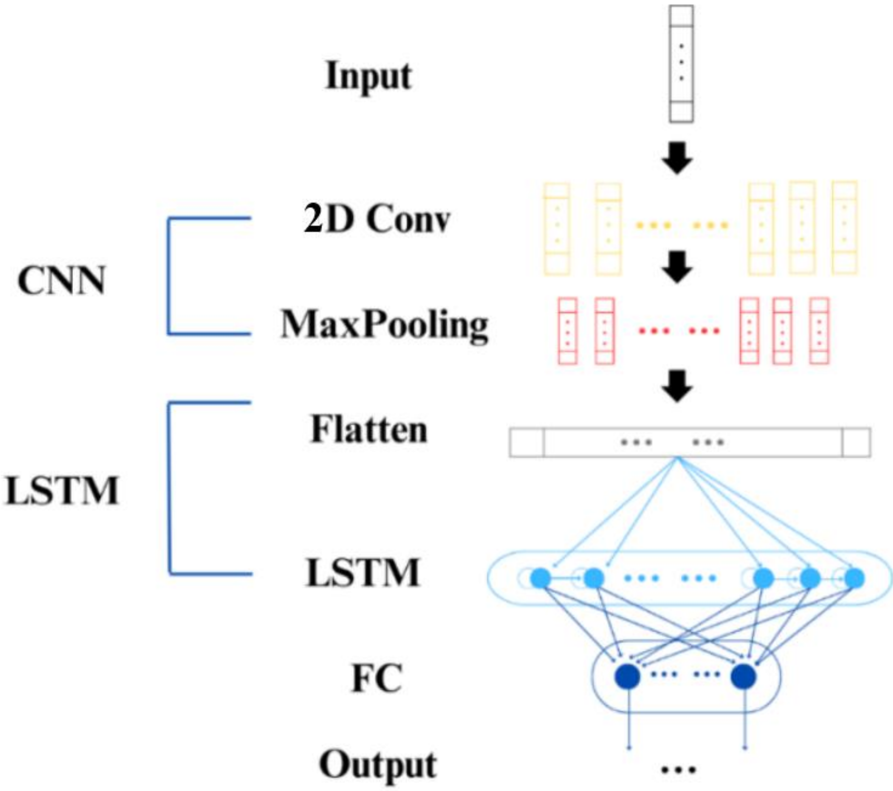


Figure II- 11: An Approach to CNN-LSTM for Time Series Applications [79]

II.8. Related work on ultra-wideband radar HAR

In the field of UBB human radar detection and recognition, the merger offers promising perspectives for improved perception and understanding of the environment. The aim of this method is to combine data from different radars on the same representation, such as micro-Doppler signatures, range-time or frequency-range images, to acquire a more thorough and precise depiction of the scenes under observation [125].

Various studies have examined different multi-domain fusion approaches for the HAR-UWB. For example, fusion algorithms have been developed to integrate the spatial and temporal data provided by range-time images with the micro-Doppler characteristics extracted from radar signals. This method makes it easier to distinguish human targets from other objects present in the environment, which improves the accuracy of detection and classification [126].

Nevertheless, despite its benefits, the multi-domain merger also has significant constraints. Dependency on a single radar is one of the main concerns. If it fails, the entire system may become disabled, jeopardizing the reliability and availability of HAR-UWB monitoring. Furthermore, a single radar may not be adequate to adequately account for all activity in a large monitored area, especially in complex or populated environments [127,128].

To overcome these constraints, studies were conducted on the use of a radar set for the HAR-UWB. Multi-radar surveillance is a method of coordinated deployment of multiple radars to improve system spatial coverage, redundancy and resilience [129]. In this area, research focuses on developing strategies for collaboration and data exploitation between radars, with the aim of improving the detection, localization and classification of human movements [130].

In conclusion, the combination of merger and the use of a radar set are promising methods for improving human detection and recognition capabilities by ultra-wide-band radar. While combining different data modalities to create a more complete representation of the environment is the goal of the multi-domain merger, using a radar set enhances the capture system's robustness and spatial coverage. These studies promote the improvement of detection and recognition skills in complex and ever-changing situations.

II.9. Conclusion

In the fields of smart surveillance and human-computer interaction, HAR is a developing field of study. Because radar systems have distinct advantages, HAR uses radar systems like UWB and interferometry. In order to achieve desirable classification performance, DL models have been developed to automatically extract deep hierarchical features. Radar echoes must be divided into 2D and 3D forms in order to be classified for the development of DL based HAR. Since 2D radar echoes provide enough activity information and are intuitive, they are more frequently used for HAR. Time-Doppler maps in particular are a good choice. Even though they are harder to interpret, 3D echoes provide more details. The feature learning capability of DL techniques makes them promising for radar-based HAR. Radar has distinct benefits, such as greater privacy protection and insensitivity to the environment, which make it a promising technology for HAR, even though adoption is still lagging behind that of vision-based technologies.

Chapter III: Multi sensor system for Human Activity Recognition

Chapter III: Multi sensor system for human activity recognition

III.1 Introduction

The technology progress further and massively on the different fields implemented the new technology tools AI to overcome the different challenges of this generations on HealthCare, work, writing, data analysis, gaming, Studies, data optimization and security that gives an importance to the tool for learn and use it and give a touch of creativity and smart touch on this world.

HAR one of future systems qualified to be implemented on the field by recognize the several human activities and gives another perspective into this world, recognize is a very important operation to have an general analysis and predictions for human activities.

The implementation of a system of multi sensors guide us for important challenges to overcomes in term of complexity data, complexity of models and precision Accuracy, Classification Results of different Classes also occurred from different inputs.

On this chapter we aim to surpass the challenges and suggest a suitable models for HAR system used in the different fields ensure the flexibility and precision of the system built using DL of several training and executions.

We introduce to the experimental research to find a solution for HAR systems starting with implementation with all resources and tools needed to progress in our mission also the different basics parameters that need to know and work with them to use them in DL models.

After building a sufficient knowledge about DL and their different parameters also vision about the existing models and methods implemented previously and observe their limitations by comparing them and evaluate their performances we've been able to find multiple solution includes, After observation of their performances we remarked an limitations on this approach motivated us to find other solutions to enhance the performances of classification , we introduced modifications on datasets (concatenation, Fusion and 3D) and solutions on models with (Early fusion, halfway fusion Late fusion) methods implemented on (CNN, CNN-LSTM, ConvLSTM) models.

Then we process to observe and analyze even compare models performances the different Results found on our project and the standard solutions existed previously to make a general conclusion to the suggested solution for our HAR system by comparing with previous study on this field implemented other methods.

III.2 Description of database studied for network of sensors

The deployment of a network of radar sensors for the HAR in a complex environment involves the use of different radars, these radars can be of the same frequencies or different frequencies for sensor network system, for our project we used a public database referred to [131] shared for researchers.

Three different Radar Frequencies (RF) sensors were used in this study:

- **77GHz FMCW radar:** The Texas Instruments IWR1443 Frequency Modulated Continuous Wave (FMCW) transceiver was set at a 77GHz center frequency and 750MHz bandwidth.
- **24 GHz FMCW radar:** An Ancortek SDR-Kit was set to transmit an FMCW waveform at 24GHz center frequency and 1500MHz bandwidth.
- **10 GHZ IR-UWB:** The XeThru X4 sensor transmits across a band of roughly 7GHz - 10GHz.

Such as 10GHz, 24GHz and 77GHz. Sensors at 10GHz are preferred to detect general movements over large areas, despite their reduced sensitivity to fine details in difficult weather conditions.

In contrast, integrating these sensors into a robust data processing system, including fusion and filtering algorithms as well as regular calibration (well adjusted), is essential to ensure optimal performance in a complex environment.

All sensors were placed side-by-side on top of a table of 1- meter height from the ground, with the test subject moving about 0.5-3 meters from the sensors. (Figure III- 2) setup of the sensors positioning. The Xethru and 24GHz sensors were operated from one laptop, whereas a separate, dedicated laptop was used to operate 77GHz sensor. Each sensor has its own graphical user interface and a common graphical user interface has been used to enable synchronization across the sensors and simultaneous data collection [131].

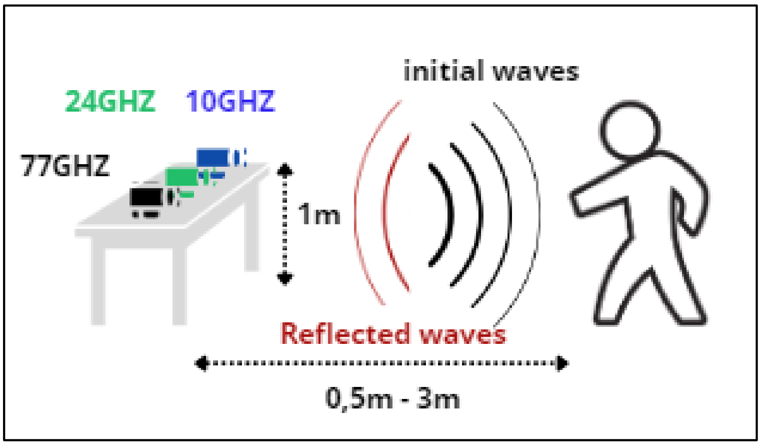


Figure III- 1: Illustrates the sensors positioning

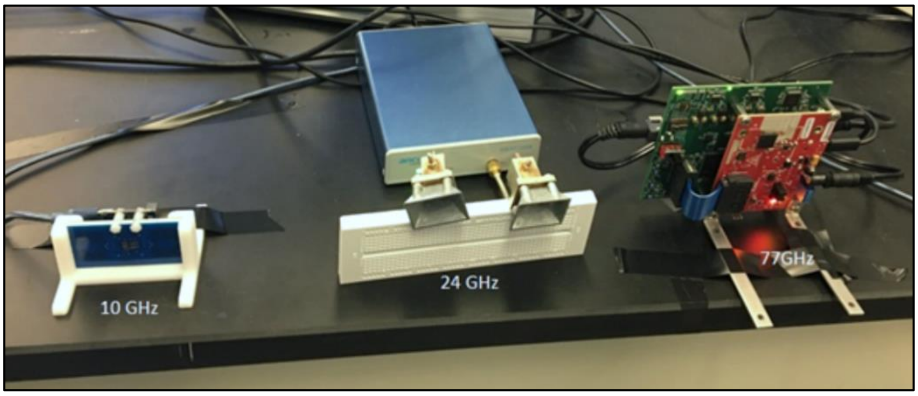


Figure III- 2: Setup of the sensors positioning [131]

III.3 Database

III.3.1 Activity Datasets Acquired

Six participants of various ages, heights and weights were involved in this study. A total of 11 different activities and ambulatory gaits were considered, as listed in Figure III-3.

The choice of these activities were motivated by smart environment applications, where monitoring of activities of daily living are required to support health monitoring and gesture recognition. Each participant conducted 10 repetitions of each activity, resulting in a total of 60 samples per class per sensor. Additionally, data previously acquired at an earlier date from the 24 GHz sensor was used to enrich the dataset with an additional 180 samples per class. All activities were conducted along the radar line-of-sight sensor.

Additionally, data previously acquired at an earlier date from the 24 GHz sensor was used to enrich the dataset with an additional 180 samples per class. All activities were conducted along the radar line-of-sight.

The activities are simple identify as:

1. Walking toward the sensor “WTOWARDS”
2. Walking away from the sensor “WAWAY”
3. Limping with right leg stiff “LIMP”
4. Scissors gait walking “SCSSR”
5. Walking short step “SHSTEP”
6. Walking on both toes “WTOES”
7. Crawling on the ground “CRWL”
8. Picking up an object “PICK”
9. Bending “BEND”
10. Sitting on a chair “SIT”
11. Kneeling “KNEEL”

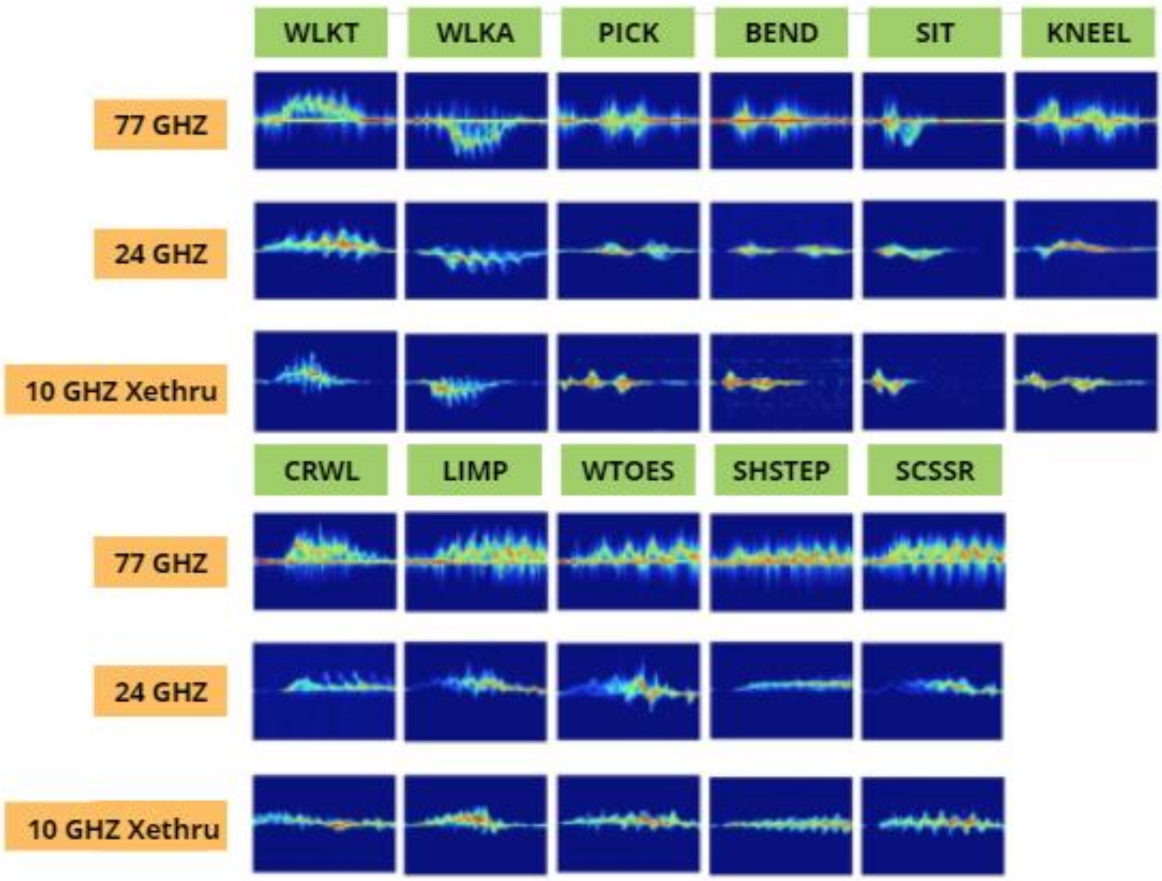


Figure III- 3: Activity Datasets Acquired [131]

III.3.2 Data Processing

Using UWB radars installed on a specified environment on the right angles we have several steps illustrated on (Figure III-1) to obtain micro doppler signatures almost ready for use and analyse.

First, we have UWB signal emission the radar emits signals that UWB emission are short radio frequency these signal spreads through the environment the doppler effect occurs while an object moves relative to the radar, it induces a change of frequency these signals detected named the echo of transmitted signals received by the radar mark a frequency change in the reflected signal.

Processing by correction In Phase/ Quadrature (IQ) this step is important to correct the variations signals occurred by the object which allows to determine radial speed of object.

These data samples arranged (mentioned in section II.7.4).

After organizing these data, Moving Target Indication (MTI) filtering is implemented for clutter suppression, and range gating may be used to mitigate noise and clutter by isolating those data samples that correspond to the person's location. Figure below depicts the effect of MTI filtering and range gating on the micro-Doppler signature. It's been reported. [132] A typical MTI filter employs a high-pass filter to eliminate energy at low Doppler frequencies. Because the frequency response of a high-pass filter is periodic, some energy at high Doppler frequencies is removed. Targets at high Doppler frequencies will thus be undetectable by radar. This is known as the "blind speed problem".

The MTI filter is a digital filter that processes signals by performing digital computations. It includes a shift register with multiple stages, each stage delaying the input pulse by one time unit (denoted as Z^$^{-1}$). Each stage of the shift registers outputs (figure III-4) a delayed signal to a multiplier. The multipliers are also fed with filter weight coefficients (h□, h□, ..., h□). The summed output from these multipliers is then processed by a Doppler FFT module to further analyze and filter the signal [133].

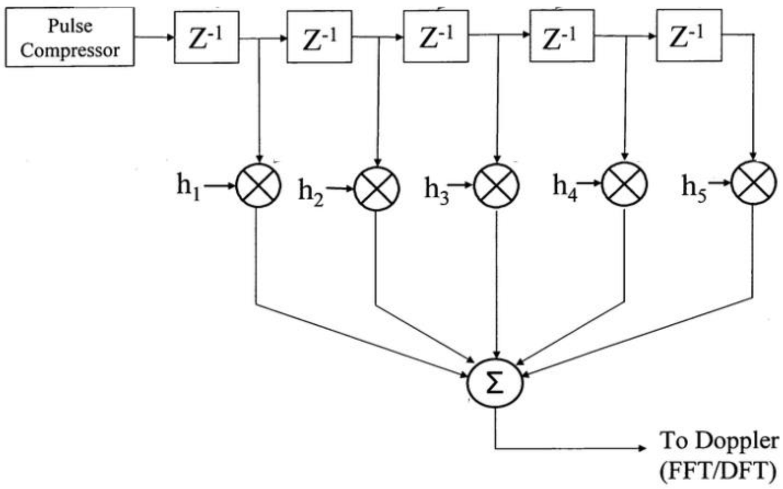


Figure III- 4: MTI Filter block [133]

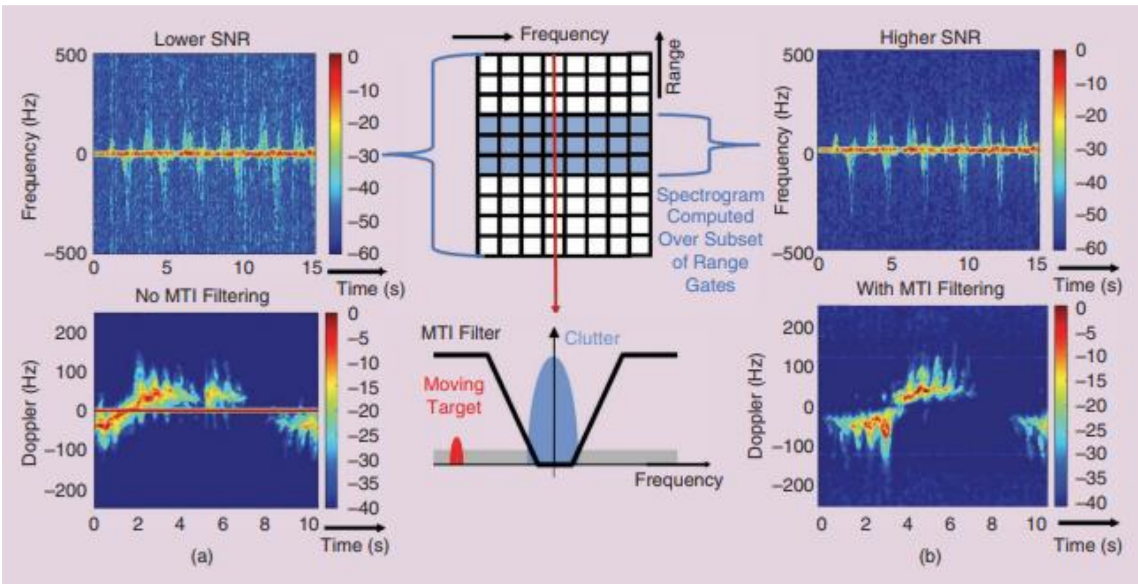


Figure III- 5: Four spectrograms and two illustrations showing the effect of clutter-suppression and noise-reduction techniques on spectrograms of radar data. Clutter-suppression and noise-reduction techniques were not applied to spectrograms in (a). Clutter-suppress

The data is then used to generate a micro-Doppler spectrogram. A spectrogram is a visual representation of the frequency content of a signal over time [134]. a micro-Doppler spectrogram used to visualize the Doppler shift of radar echoes from small, moving targets. The Doppler shift a change in frequency of a wave [135] caused by the relative motion. By analyzing the micro-Doppler spectrogram, it extracts information about the size, velocity, and direction of moving targets.

Compute the Doppler spectrum of the received pulses as in conventional pulsed Doppler radar by processing algorithms such as short-term Fourier transform (STFT) are applied the most commonly used time–frequency representation is the spectrogram, which is defined as the modulus of the short-time Fourier transform [136].

$$STFT(\tau, \omega) = x(t)w(t-\tau)e^{-j\omega t}dt \qquad \textbf{III.1}$$

- STFT (τ, ω) is the STFT of the signal $x(t)$.
- $x(t)$ is the original time-domain signal.
- $w(t{-}\tau)$ is the window function centered at time τ.
- τ is the time shift (or the time index where the window is centered).
- ω is the angular frequency variable.
- j is the imaginary unit.

From time domain to the frequency domain is an important passage in our search used to simplify the computational operations also facilitate the analysis of signal properties.

STFT provides a time-frequency representation of the signal allow us to see the frequencies variations over the time it's applied by dividing signal into windows where the choice of the length of window is important then we apply Fourier transform on each window segment where the signal will be decomposed into frequencies components then each window is combined to form a time-frequency representation visualized as a spectrogram.

The advantages of using STFT visualize non stationary signals different than traditional FT allows to see how the frequencies change over time specially used for analyzing signals speech, images where the frequencies evolve by time, also used to extract features from the signal helps for classification applications.

To resume, the data collected by the sensor is a time-stream of complex I/Q data. After reshaping the data matrix using the Pulse Repetition Frequency (PRF) mentioned in section..., phase corrections on I and Q channels are applied. After that, a (MTI) filter is used to suppress stationary returns, such as those from ground clutter. The magnitude of the signal's Short-Time Fourier Transform (STFT) is then applied to generate a 2D time-frequency representation of the data [137]. The process is resumed on (Figure III-6).

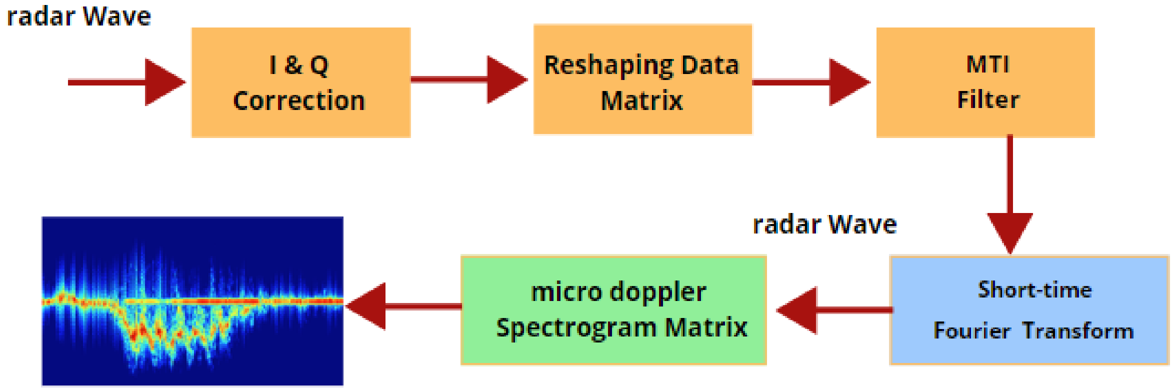


Figure III- 6: Illustration of data processing [131]

III.4 Implementation

III.4.1 Machines Specifications

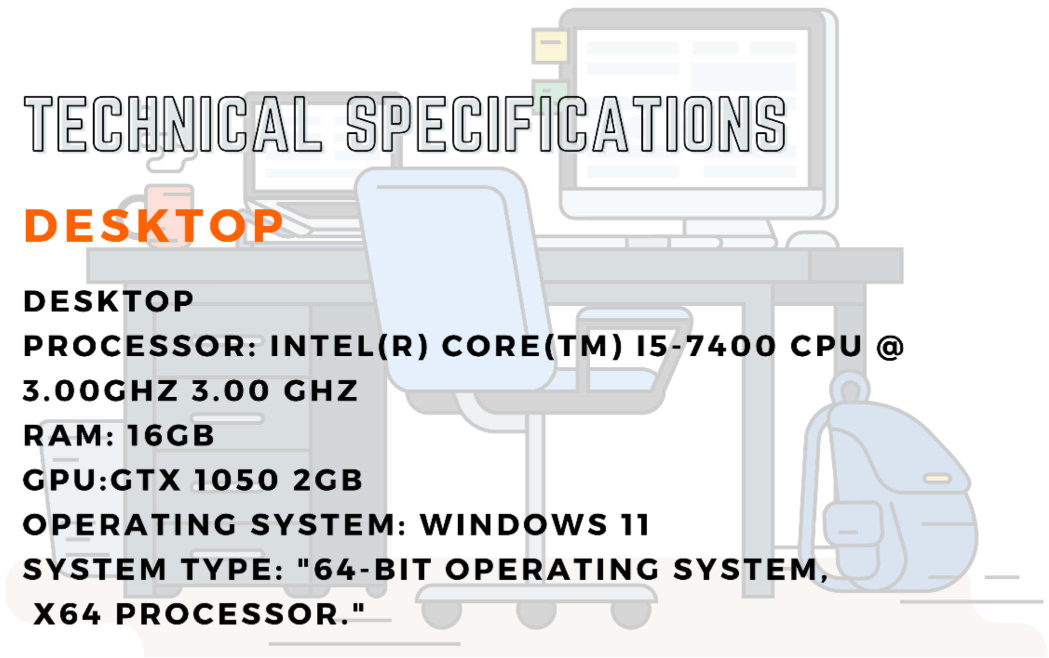


Figure III- 7: Machine Specifications

III.4.2 Platforms

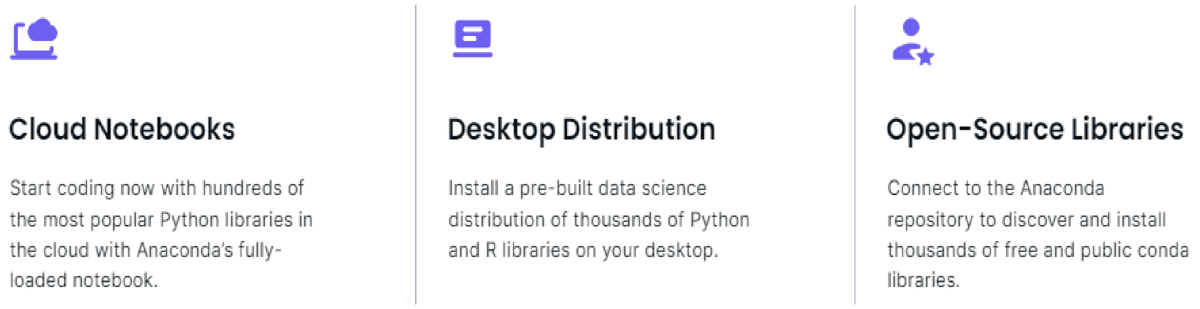
III.4.2.1 Anaconda

Anaconda is a powerful tool for data scientists, researchers, and developers, Open-source package and environment management system that runs on Windows, macOS, and Linux. Install, run, and update packages and their dependencies [138].

Anaconda info:

CONDA VERSION: 23.7.4

CONDA-BUILD VERSION: 3.26.1



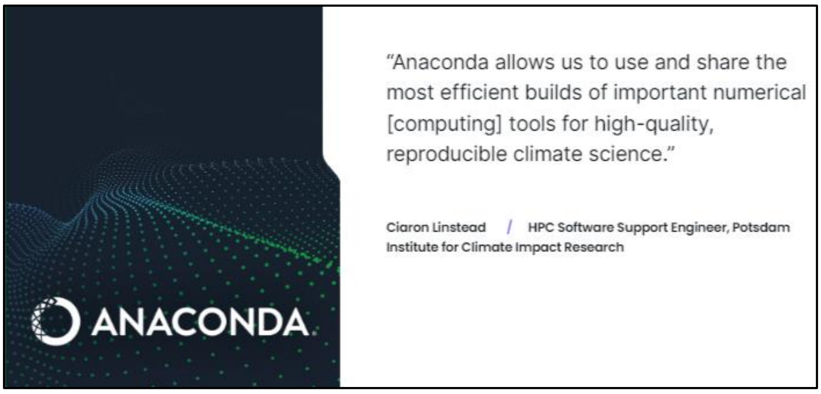


Figure III- 8: Software support describing anaconda Tool

Figure III- 9: Jupyter Application

Applications Installed.

Jupyter notebook for Writing and editing compilation execution scripts.

III.4.2.2 Platform remark

Anaconda is used to executed locally has much more advantages such as free access using local data and easier modifications and updatable environment the important thing is an offline Software that what make powerful and choice to work within it.

III.4.2.3 Programation Language

Is a dynamic and transformative field, and when it comes to learning AI, Python programming language emerges as a standout choice. Python's extensive library support and user-friendly features make it a compelling language for AI development [139].

Python 3 used on our research with necessary packages to assure the Working Environment of our Program.

III.4.2.4 Libraries & versions

The list bellow displays the necessary libraries installed for our environment.

- Pandas ==1.5.2
- numpy == 1.23.5
- tensorflow_addons == 0.19.0
- tensorflow == 2.10.0
- Ipywidgets == 7.6.5
- scikitplot == 0.3.7
- Seaborn == 0.12.1
- PIL.Image == 9.3.0
- PIL == 9.3.0
- Pydot == 1.4.2
- Graphviz == 0.20.1
- Pip == 22.3.1

III.4.3 Data preparation

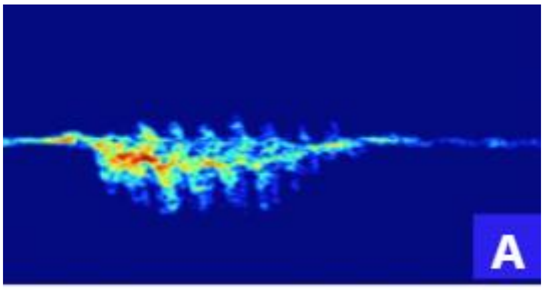
III.4.3.1 Image resize

Each image is resized to a size of 128,128 pixels. This step is set for make all images on the same size and perform the model less pixels and fast treatment.

- in_channel =3 Number of channels of input image
- img_rows, img_cols = 128, 128 image size
- num_classes = 11 number of classes

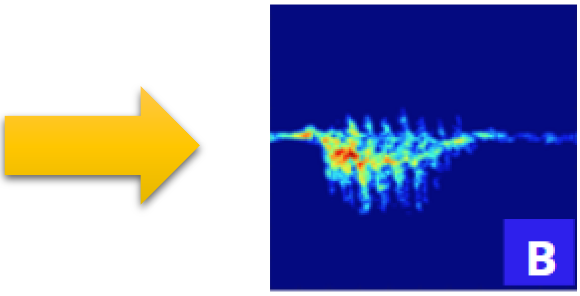
This process affects the learning model in term of:

- Reduce overfitting where the model memorizes specified details while the there is a less number of pixels
- Upgrade efficacity learning while using the machine performances, reduced images require less efforts of computations.



DIMENSION:1178 X 628
WIDTH: 1178 PIXELS
LENGTH: 628 PIXELS

Figure III- 10: (A)Before resize



DIMENSION:128 X 128
WIDTH: 128 PIXELS
LENGTH: 128 PIXELS

Figure III- 11:(B) After resize

Image loading loop: the code iterates through each file path in the file_paths list it uses the imread function to load the image from the file and add it to the image list.

III.4.3.2 Nump array conversion

Np.asarray() function used to convert an image into array creates a new array from an input object(on our search is images) the type of parameters specified the data type of the element, np.float 32 indicates pixels of the images convert them to 32bits after this conversion each value on the array represents an pixel value in the images.

Reshape() method to change the shape of the image array. The resulting shape is (**number of images, img_rows**, **img_cols**, **channels**), where img_rows and img_cols represent the spatial dimensions of each image, and channels represent the number of color channels of each image (channels=3 for RGB images).

III.4.3.3 Normalization

Normalizing a set of data transforms the set of data to be on a similar scale. For machine learning models, our goal is usually to recenter and rescale our data we used

Images = images/ np.max(images)

That Each pixel value in the images array is divided by the maximum pixel value found in the array, depending on the data itself. One common way to accomplish this is to calculate the mean and standard deviation of the set of data and transform each sample by subtracting the mean and dividing by the standard deviation, which is good if we assume that the data follows a normal distribution because this method helps us standardize the data and achieve a standard normal distribution [140].

Help with training our neural networks because the different features are on a similar scale, which helps to stabilize the gradient descent step, allowing us to use larger learning rates or help models converge faster for a given learning rate [140].

III.4.3.4 Train, validation and test split

This step is essential in deep learning models where we process cimport the function from sklearn.model_selection import train_test_split ,splitting data on three portions explained below:

a) Training set

First, we have primary datasets should be between 0.0 and 1.0 and represent the proportion of the dataset to include in the train split.

Used to train the model it contains labeled examples where each input match to a desired output then the model analyzes these examples to identify patterns and relationships features between labels.

An iterative process, the model adjusts its internal parameters to reduce errors occurred on predictions and labels in the training data [141].

b) Validation set:

Validation used to fine tune the model during the training process and identify potential problems early detection of overfitting, creating reliable and robust machine learning models. The validation data helps assess the model's performance.

The validation process practice on a smaller set of unseen data and adjust the model approach and prevent overfitting ensures the model can effectively learn effectively from training data [142].

c) Test Set:

Asses how well the trained model performs also evaluate model accuracy, testing the model ability to identify/classify correctly the new data given for test process, allow us to evaluate models performances.

III.4.3.5 Random seed

This seed is generally determined as a function of the current time. However, the developer may decide to give a value to this random seed to control the process. If the same seed is used in an algorithm [143], A random seed is used to ensure that results are reproducible.

III.4.3.6 Stratify

"Stratify" typically refers to a technique used in data splitting or sampling to ensure that the distribution of classes or labels remains similar across different subsets of the data. This is particularly important in tasks like classification, where the goal is to learn a model that can generalize well to unseen data with similar class distributions [144].

III.4.4 Models approach

III.4.4.1 Using single frames each database separately

The process of training each database separately, training databases of 10GHZ, 24GHZ, 77GHZ separately to understand the influence of each frequency, in order to optimize our integrated sensor network. This approach helps us identify the strengths and weaknesses specific to each frequency to improve the entire system (network of sensors).

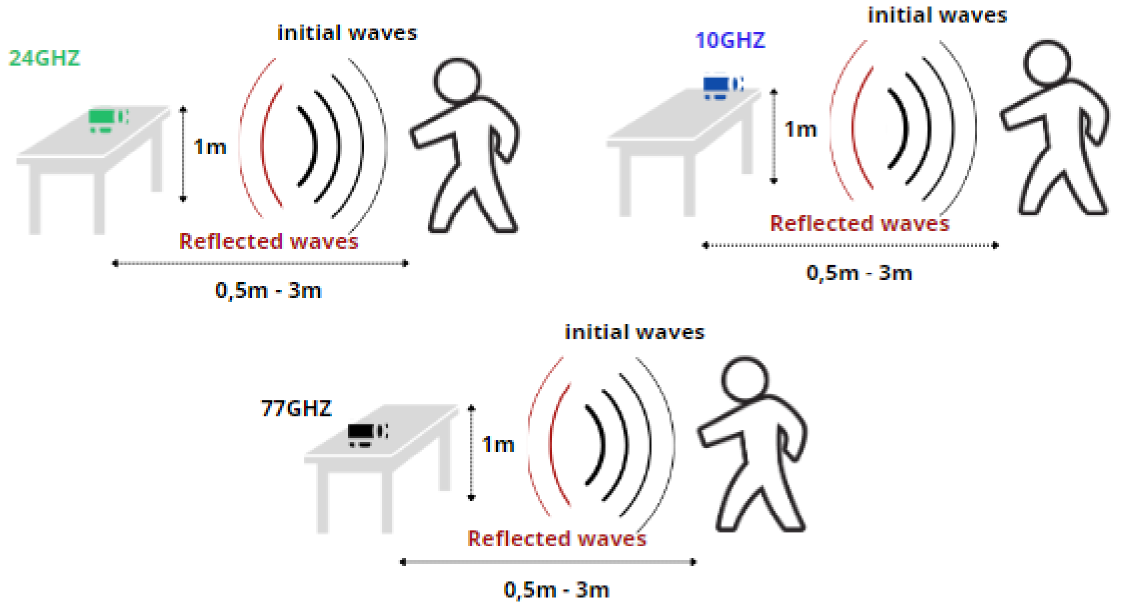


Figure III- 12: Illustration collecting data of each Radar database separately

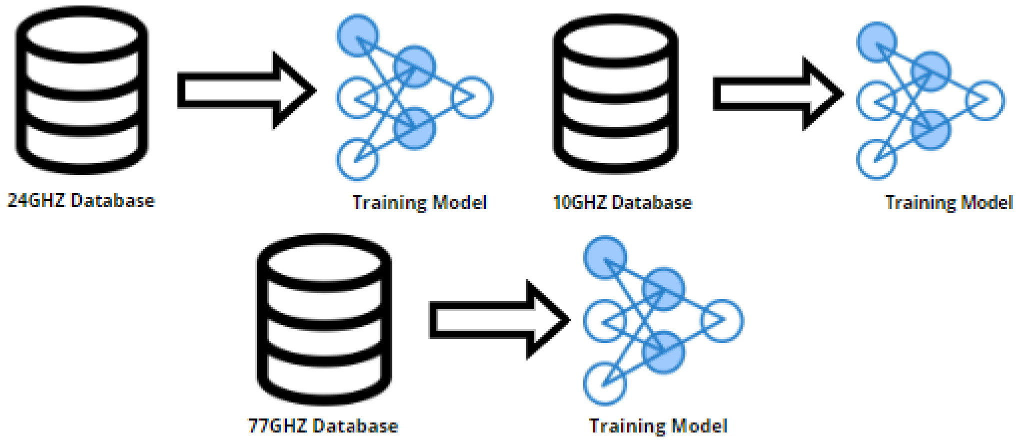


Figure III- 13: Illustration training each database separately

Combining data from different radars

Combining data from three radars into a single dataset to make things easier as using one database contains the different activities classes of the different radars this Method also will make our training More complex for the classification a multimodal data what's make it difficult is the variety of data of the different radars also the huge amount of data treated in one time

Another goal set from this combination of data to observe the robustness of our model and the ability classify successfully the activities

Each radar database has 642 signals of 11 classes of human activity.

- 653 signals of 77GHZ
- 653 signals of 24GHZ
- 653 signals of 10GHZ

This operation executed by moving datasets to one database and make a change on the labels Automatically for correspond to our new database named Combined Data.

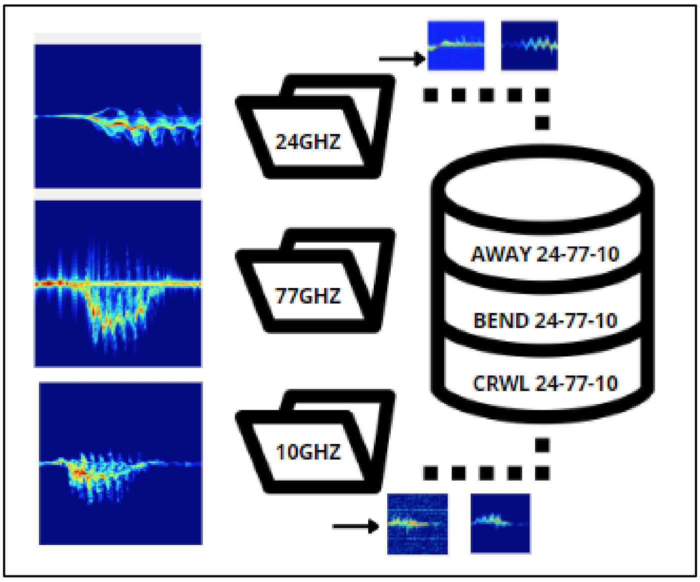


Figure III- 14: Process of combining Data

III.4.4.2 Solutions Proposed

In the context of our project we faced into a problematic the use of standard methods used previously and codes existing before also challenges of our sensor network system influences on results, treatment even analyze of results also extraction features, we suggested a solutions of different approaches models and methods divided into two solutions Database level and models level surpass this challenges these solutions aims to maximize the rate of Success classification the different class of activities, here we mention our learning methods implemented:

a) Database level solutions

➢ **Fusion data of different Radars**

This operation is achieved using a script allow us to fuse data from separated databases in our search we have data of three radars fused into one image as the figure illustrates the process of fusion.

The database established contain 653 image results of fusion.

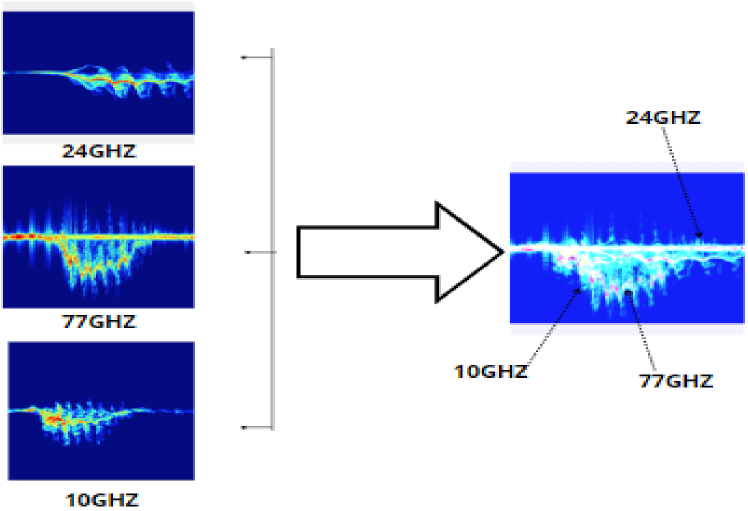


Figure III- 15: Process of Data fusion

➢ Concatenate data of three radars in one frame

The process of concatenate data acquired from the different radars executed automatically using the code which allow us to Concatenate horizontally the frames cv2.hconcat() is used to concatenate images of same height horizontally [145] the image obtained is established by importing an image from each radar database with same Class.

- ❖ The frame 1000 of 77GHZ class AWAY
- ❖ The frame 1000 of 24GHZ class AWAY
- ❖ The frame 1000 of 10GHZ class AWAY

Resuming this process it to make the frames of the same class side by side by concatenating horizontally

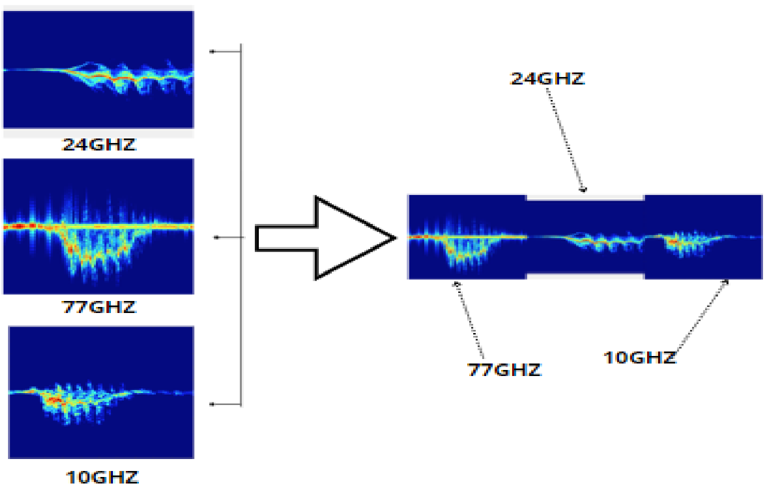


Figure III- 16: Illustration of concatenation process

b) Cube of Data

The process of treatment of manipulation of data at shape of 3D consists of the change of the input model by processing the stack function to generate a Cube of data of different sets the figure III-6 below show the process of generating a cube of Data.

3D preprocessing with input (batch_size, time_steps, height, width, depth, channels)

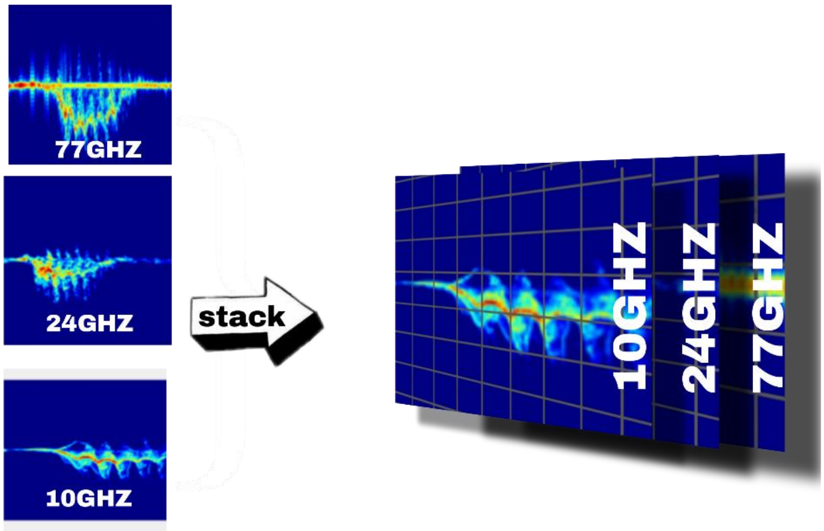


Figure III- 17: Illustration of Concatenation Doppler signatures to a cube of Data

c) Models level

Models level solution consist of mange the training performances providing changes on models architectures includes fusion of results before training of data and fusion of results after training process, we propose this solution to manage efficiently the data from multiple sources and have a better HAR system performances.

➢ **Early fusion**

Before moving on to the learning/feature extraction process, the various raw data modalities must be integrated into a unified representation [146].

Early fusion, in deep learning, is a method by which characteristics from different sources or modalities are combined at the beginning of the processing process, before being introduced into a model for learning. This approach aims to simultaneously integrate information from multiple sources.

- **Halfway fusion**

Halfway fusion is a method used in DL models; data are combined at an intermediate stage of processing aims to perform result.

- **Late fusion**

Late fusion, where the classification results are combined by applying stacking or averaging in our project, we applied majority vote [147] (FigureIII-18).

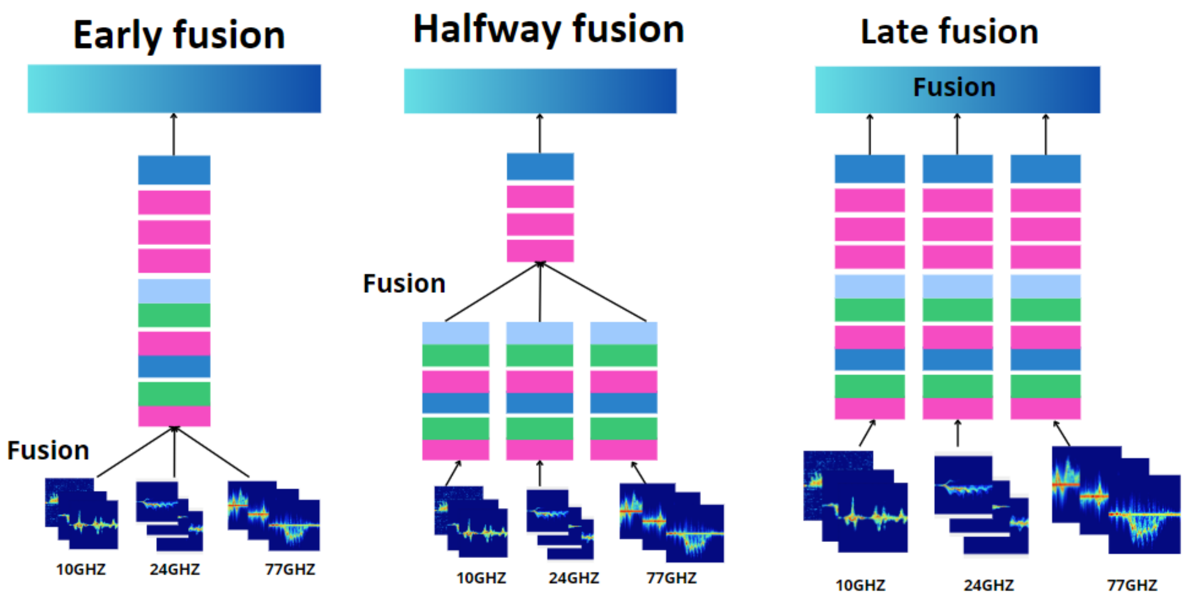


Figure III- 18: Illustration of early fusion, halfway fusion and late fusion

III.4.5 Implemented models

III.4.5.1 Architectures

a) Model layers

- **MaxPooling**

MaxPooling is a type of operation in convolutional neural networks that selects the maximum value within a specified region, effectively reducing the dimensions of the feature map using a filter. This process results in an output feature map that emphasizes the most significant features [148].

MaxPooling2D, tailored for 2-dimensional spatial data like images, operates on data with dimensions including width, height, and channels (batch_size, height, width, channels). Its main objective is to condense the spatial dimensions of feature maps while retaining important features. This reduction in dimensionality not only simplifies computational complexity but also serves as a regularization technique, aiding in the prevention of overfitting.

- pool_size: Size of the pooling window (e.g., 2x2x2).
- strides: Steps to move the pooling window across the input (defaults to the same as pool_size).

MaxPooling3D is a technique used in convolutional neural networks (CNNs) for reducing the dimensions of 3D spatial data where the input has three spatial dimensions plus a depth the computational burden and memory usage while retaining the most critical aspects of 3D data.

- **Maxout**

Designed of activation functions for deep learning models maxout function is to improve the performance and flexibility of neural networks addressing some of the limitations of traditional activation functions mitigate issues [148].

Maxout function then takes the element-wise maximum of these transformed vectors. Mathematically, if:

$$\max_{i=1}^{k}(w_i^T x + b_i) \qquad \textbf{III.5}$$

- T: transposition into scalar
- k: number of linear functions
- b: biases of linear function
- Wi: weights

The output: $\text{Maxout(x)} = \max(\mathbf{z}1, \mathbf{z}2, \ldots, \mathbf{z}k)$ for i=1, 2, …, k

$$z_i = W_i\text{x} + \mathbf{b}_i$$

Example with k=2 we have computation of two linear transformations

$$z_1 = W_1\text{x} + \mathbf{b_1}$$

$$z_2 = W_2\text{x} + \mathbf{b_2}$$

$$\text{Maxout(x)} = \max(z_1, z_2)$$

➢ **Dense layer**

One of the fundamental building blocks in a neural network architecture the input to a Dense layer is a vector, matrix, or tensor representing the features or activations from the previous layer [148].

The most common situation would be a 2D input with shape (batch_size, input_dim).

➢ **ReLU**

ReLU outputs zero for negative inputs and linearly scales positive inputs efficient and easy to compute compared to activation functions like sigmoid or tanh helps in mitigating the vanishing gradient problem commonly used in hidden layers of deep neural networks for various tasks such as image classification, object detection specialized by its simplicity, effectiveness, and computational efficiency [148].

➢ **Dropout**

Prevent overfitting and improve the generalization of the model During each training iteration, a certain proportion of neurons in the network are randomly selected and "dropped out" with a probability p p. The dropout rate p p is typically a hyperparameter set by the user, commonly ranging from 0.1 to 0.5 Dropout acts as a form of regularization by adding noise to the network during training, similar to adding noise to the input data This improves the model's ability to generalize to unseen data Increased Training Time Training Speed [148].

➢ Flatten

converts multi-dimensional data into a single-dimensional array which is essential before feeding the data into fully connected (dense) layers.

b) CNN architecture

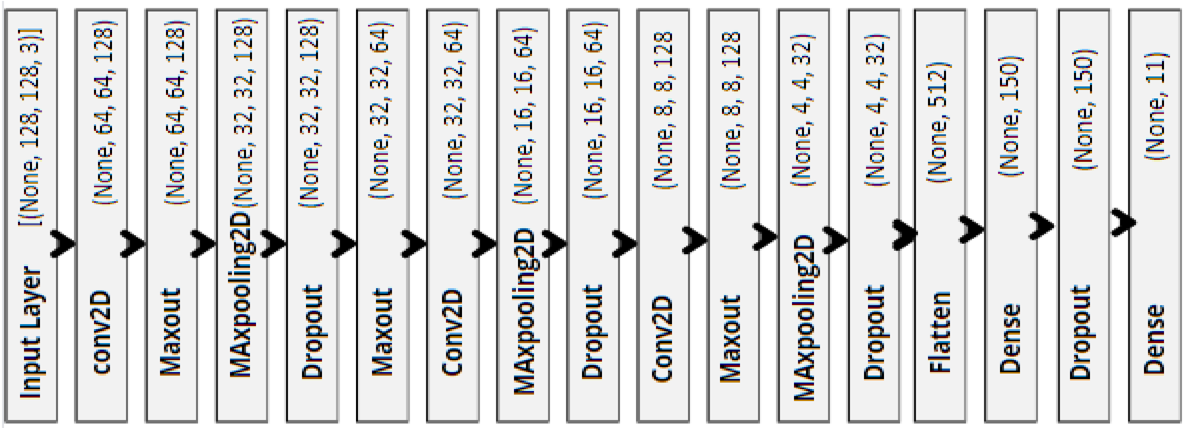
We realized a model with multiple layers of convolutional networks the table below resume the implementation of the different model parameters.

Table III- 1: Illustrate the different hyperparameters on CNN architecture

Filters Size	Learning Rate	Function Activations	Optimizer	Evaluation metrics	Epochs	Batch Size	Maxout	Pool Size
(4.4)	0.001	ReLU Softmax	Adam	F1 score Accuracy Recall Confusion Matrix	100	32	128 64 32	2.2

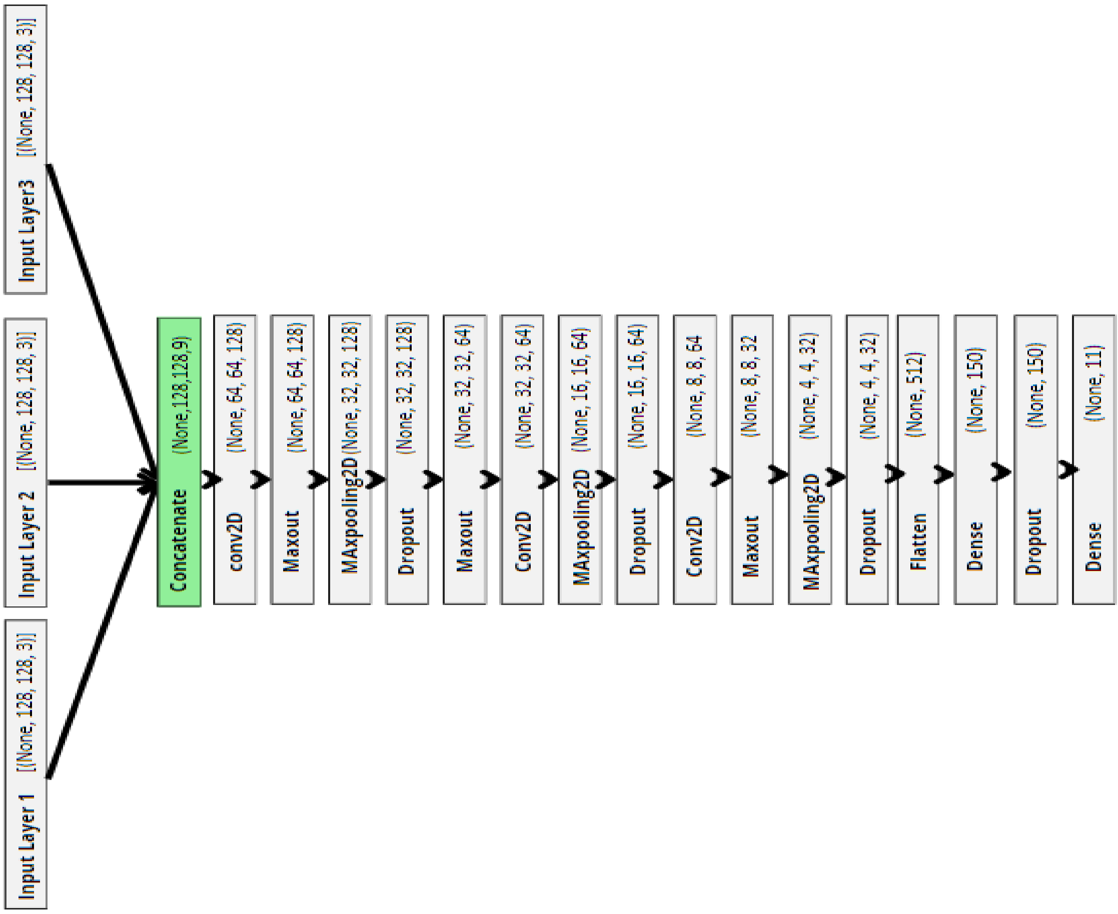
- **Scheme CNN used for single input**

Single training (mentioned in section 4.4.1) separately data of the different radars used also for combined datasets and fused database also concatenated datasets.



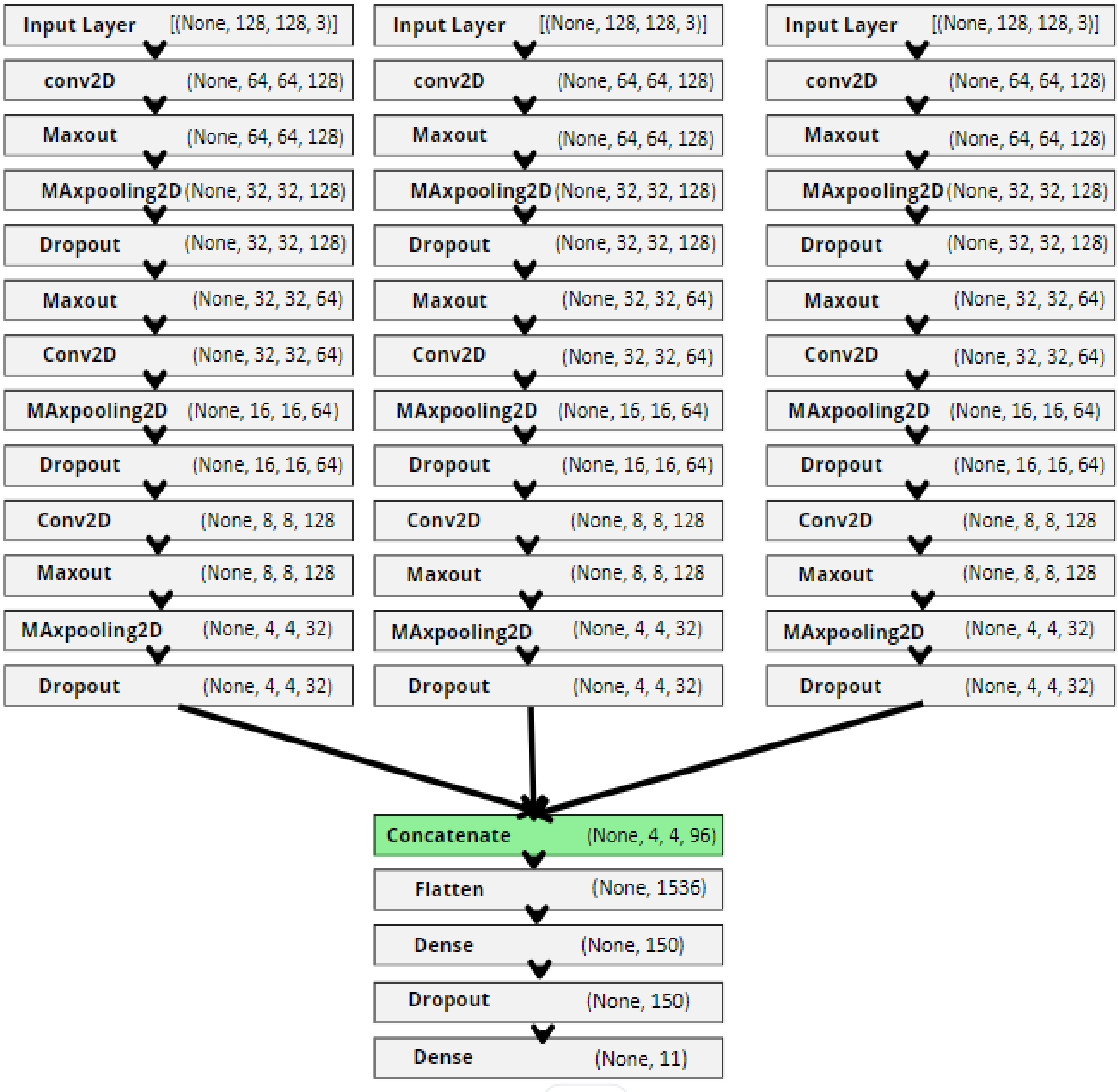
Scheme III- 1: CNN used for single input

- **Scheme CNN used for three inputs Early fusion model**



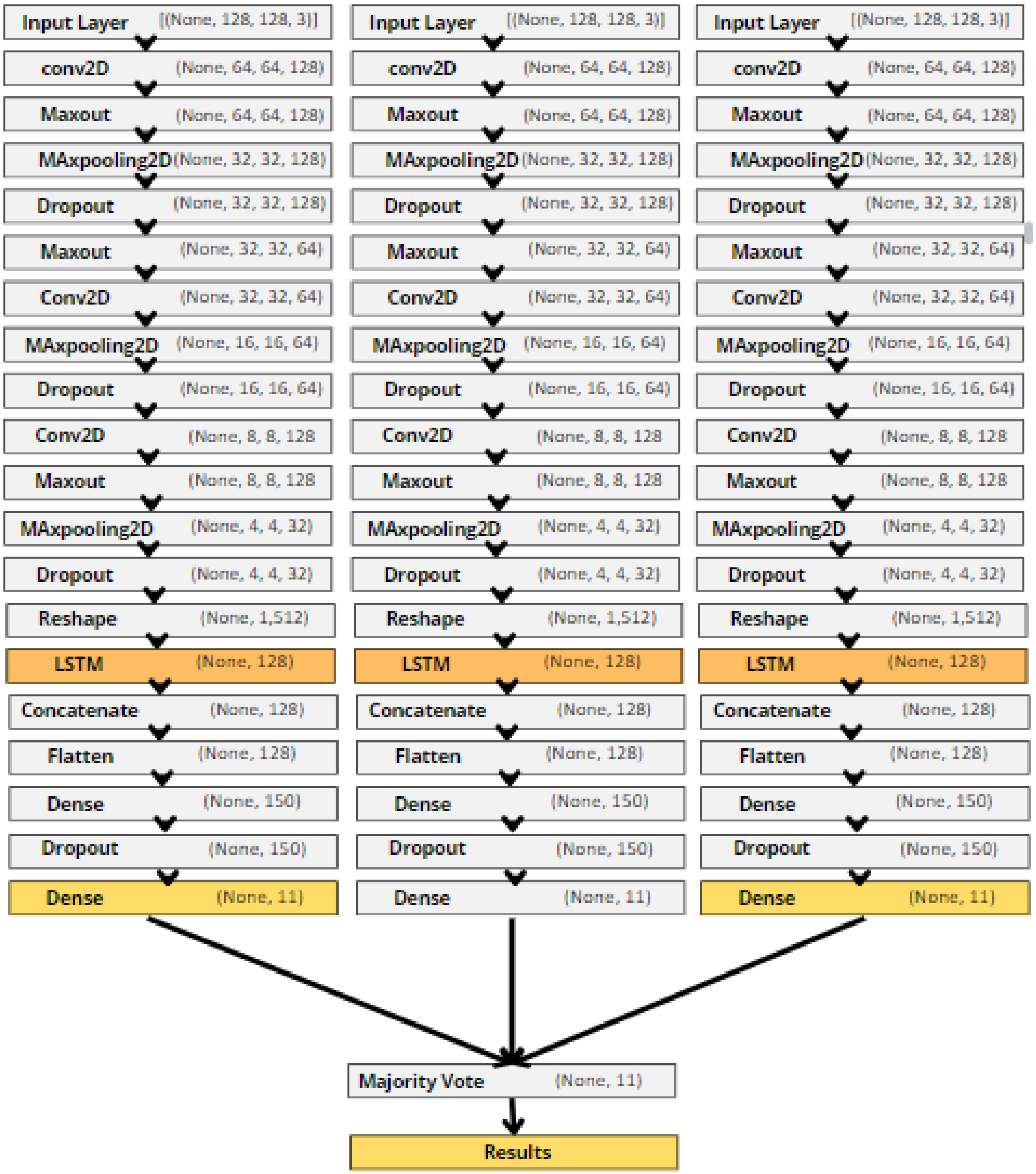
Scheme III- 2: CNN used for early fusion

- **Scheme CNN used for three inputs Halfway fusion model**



Scheme III- 3: CNN used for halfway fusion

Late fusion scheme is the combination of the three training single inputs using CNN model and fuse the data on last stage (view **Scheme III- 4**).



Scheme III- 4: CNN used for Late fusion

c) CNN-LSTM architecture

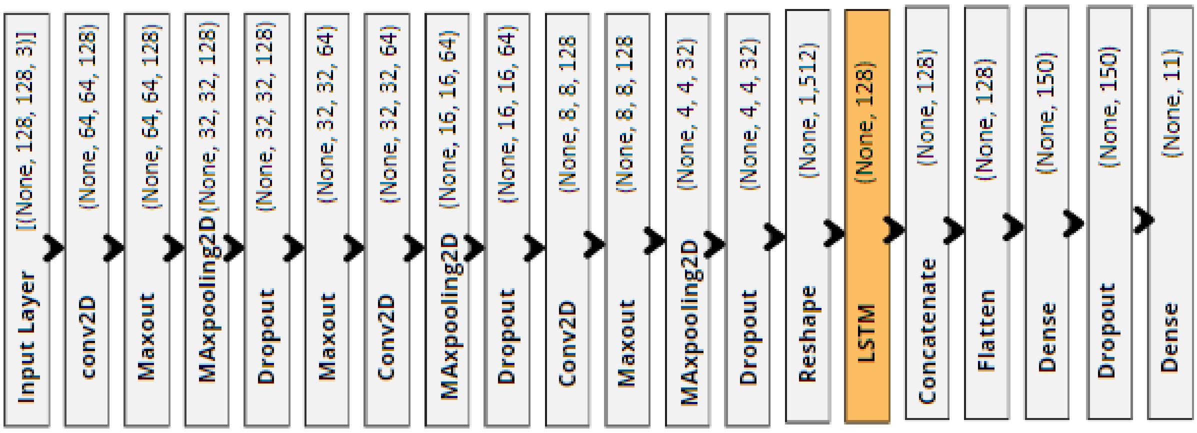
Table III- 2: Proposed experimental hyperparameters of the CNN-LSTM model

Filters Size	**Learning Rate**	**Function Activations**	**Optimizer**	**Evaluation metrics**	**Epochs**	**Batch Size**	**Maxout**	**Pool Size**	**LSTM Layer**
(4.4)	0.001	Relu Softmax	Adam	F1 score Accuracy Recall Confusion Matrix	100	32	128 64 32	2.2	128

CNN architecture (batch_size, height, width, channels)

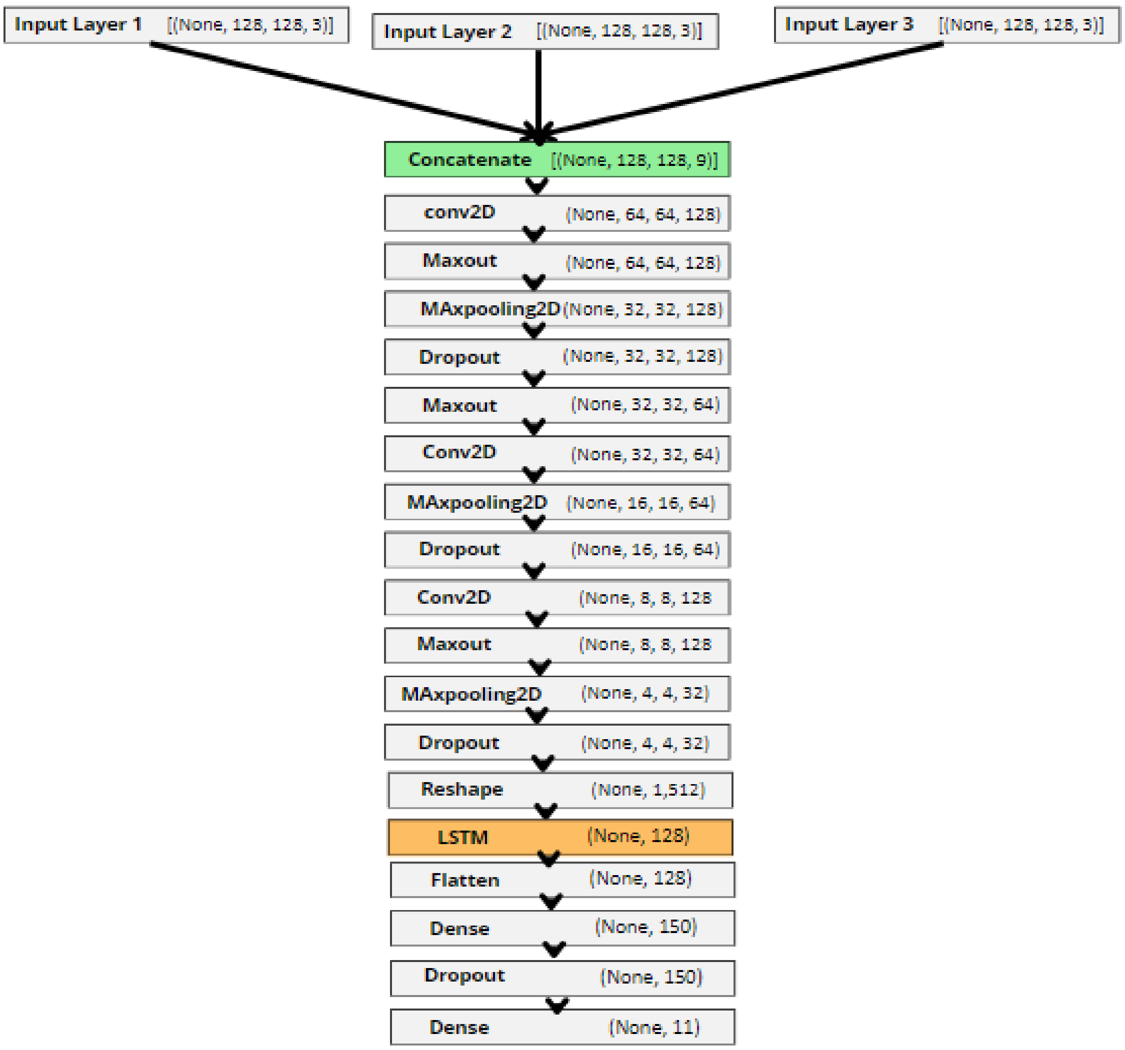
➢ **Scheme LSTM used for single input**

Single training separately data of the different radars used also for combined database and fused database also concatenated database.



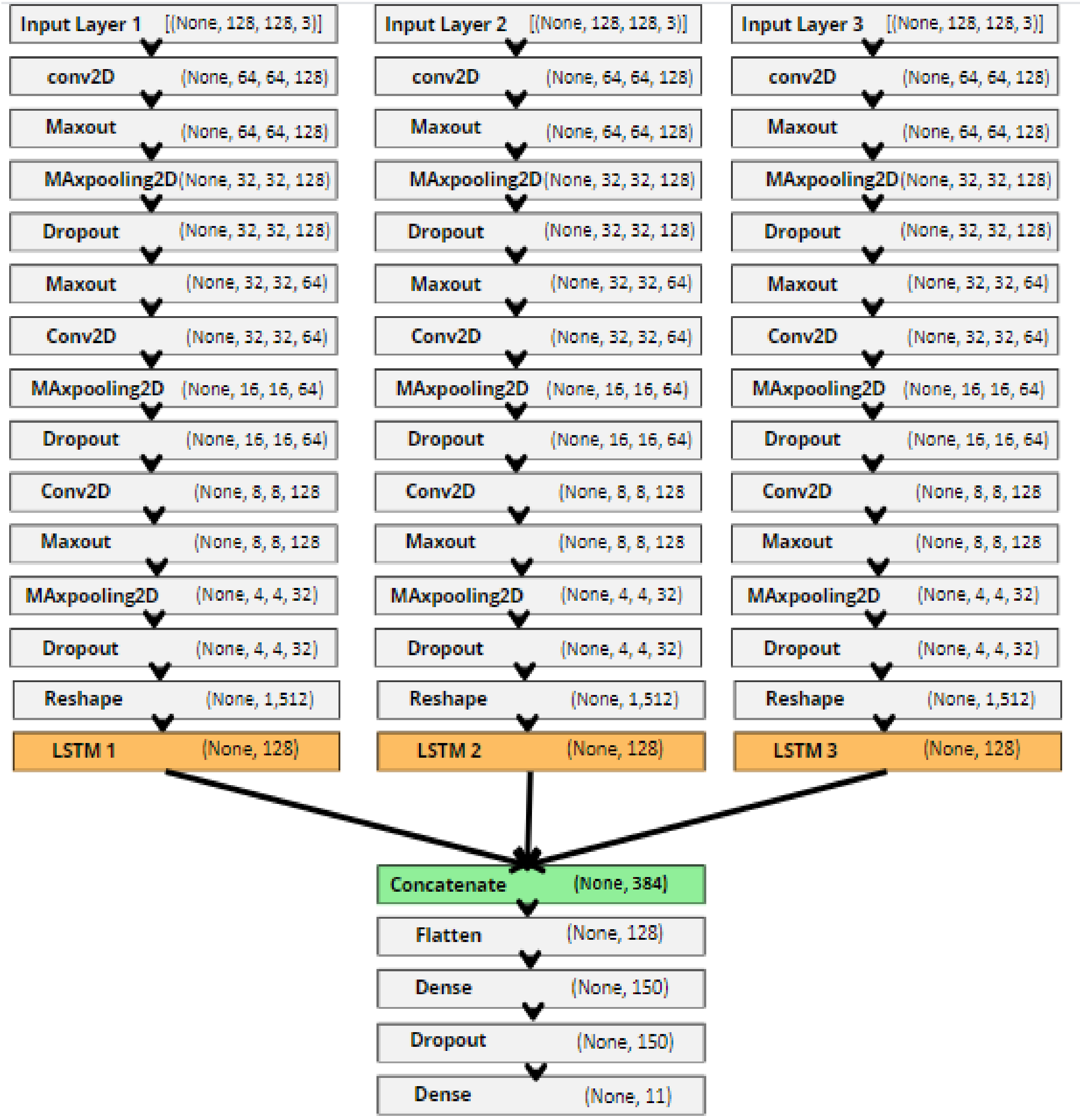
Scheme III- 5: CNN-LSTM used for single input

Scheme LSTM used for three inputs early fusion model



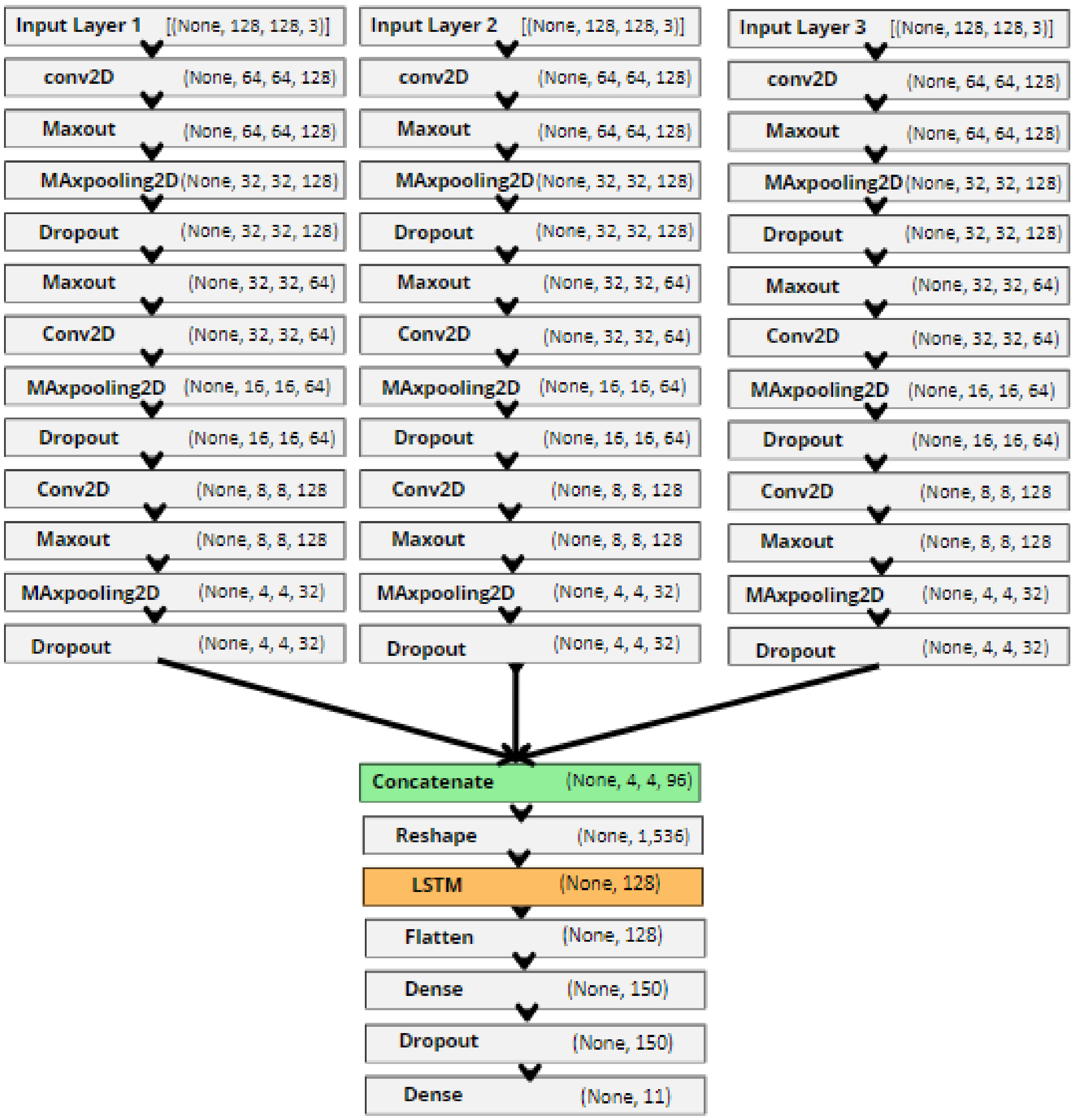
Scheme III- 6: CNN-LSTM used for three inputs early fusion model

➢ **Scheme LSTM used for three inputs halfway fusion model before concatenation**



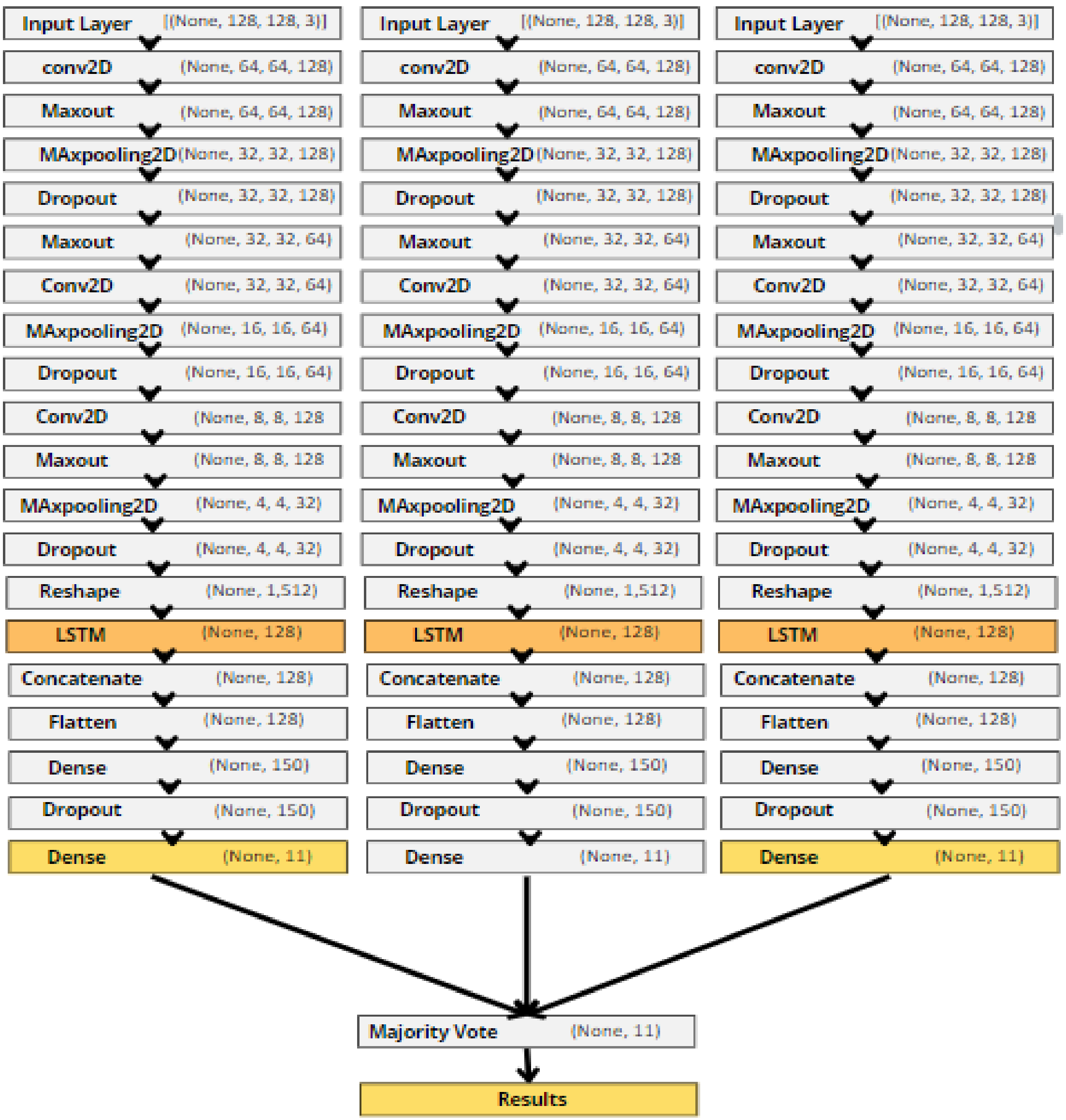
Scheme III- 7: CNN-LSTM used for three inputs halfway fusion model before concatenation

➢ **Scheme CNN-LSTM used for three inputs halfway fusion model after**



Scheme III- 8: CNN-LSTM used for three inputs halfway fusion model after concatenation

Late fusion scheme is the combination of the three training single inputs using LSTM and fuse the data on last stage (view **Scheme III- 9**).



Scheme III- 9: CNN-LSTM used for Late fusion

d) ConvLSTM architecture

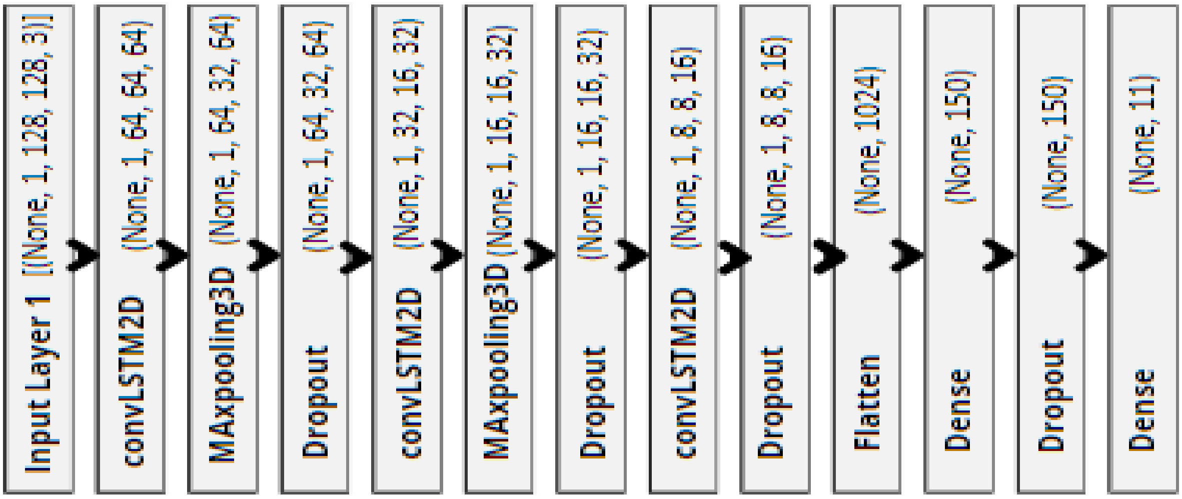
Table III- 3: Proposed experimental hyperparameters of the ConvLSTM model

Filters Size	Learning Rate	Function Activations	Optimizer	Evaluation metrics	Epochs	Batch Size	Maxout	Pool Size	Conv LSTM Layer
(4.4)	0.001	ReLU Softmax	Adam	F1 score Accuracy Recall Confusion Matrix	100	32	128 64 32	2.2	128

ConvLSTM (batch_size, time_steps, height, width, channels).

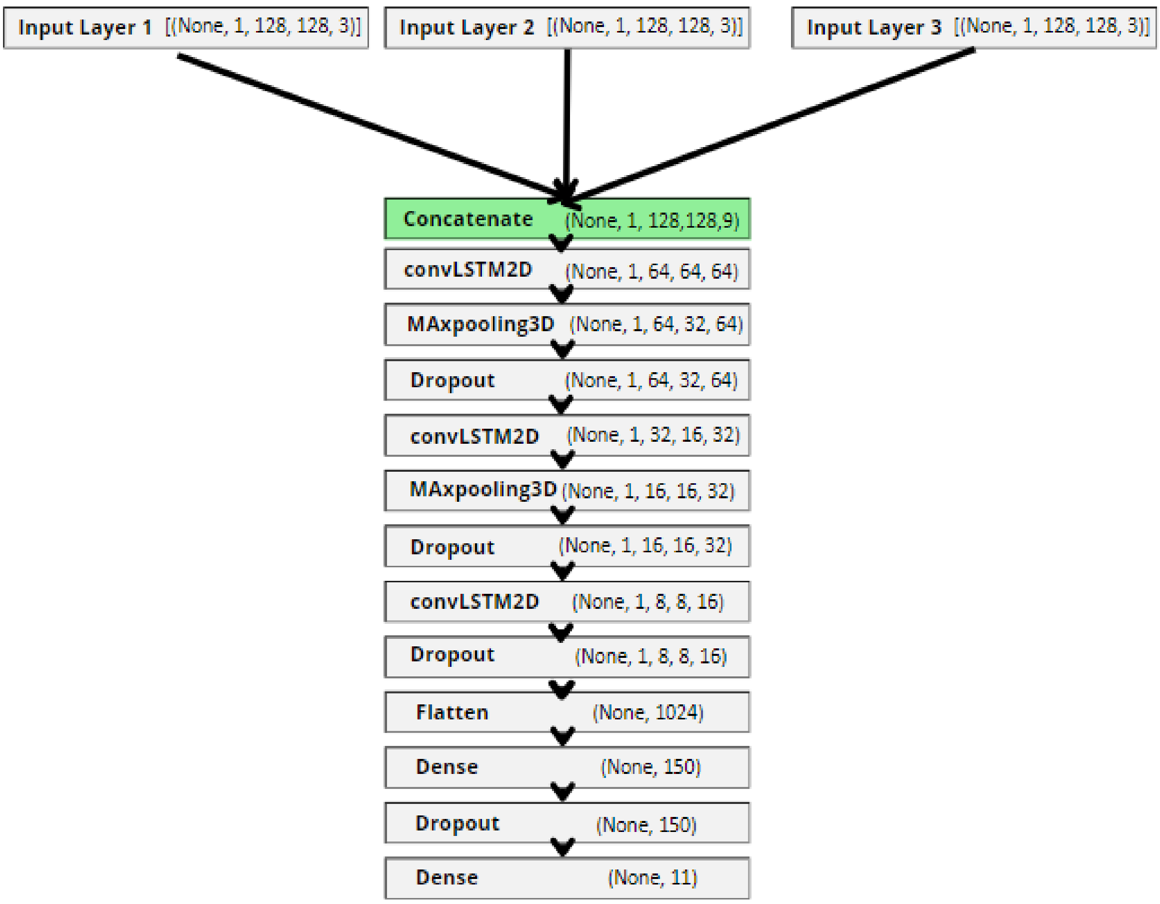
➢ Scheme ConvLSTM used for single input

Single training separately data of the different radars used also for combined database and fused database also concatenated database.



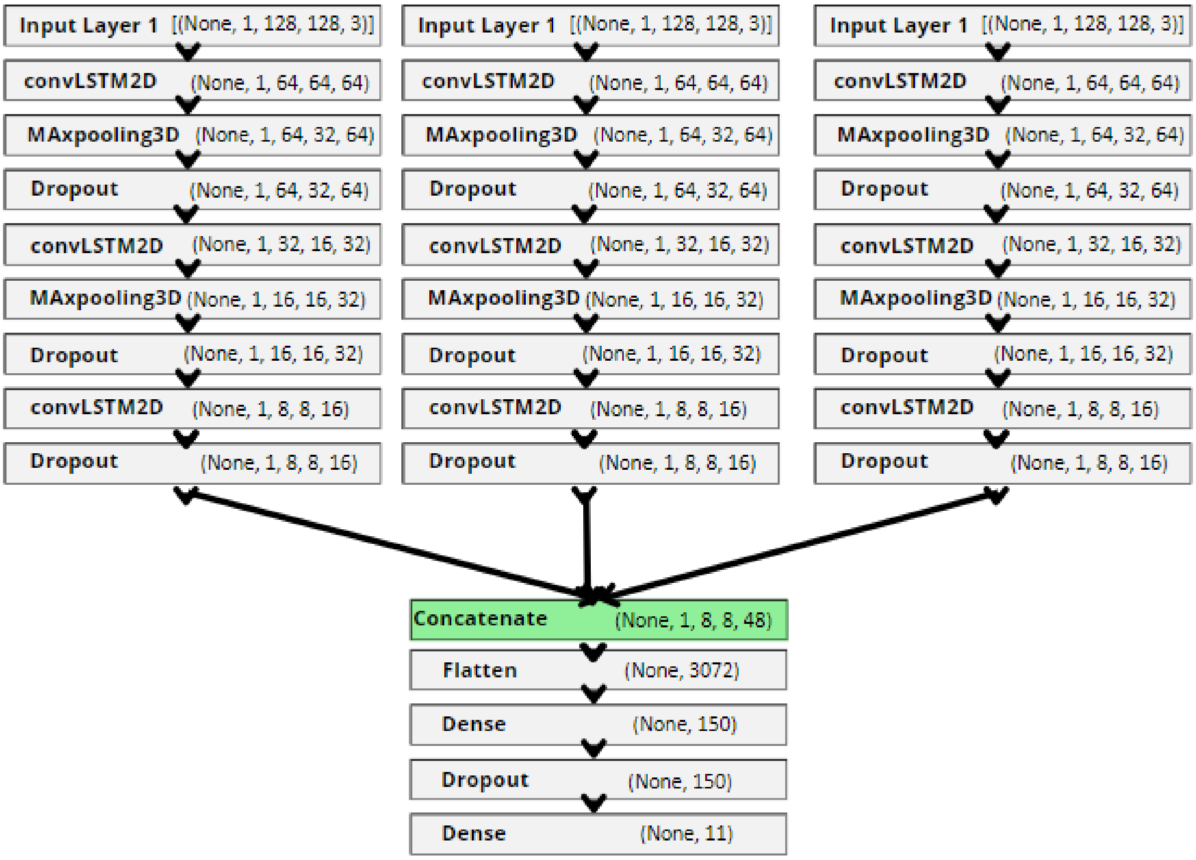
Scheme III- 10: ConvLSTM used for single input

➢ Scheme ConvLSTM used for three inputs early fusion model



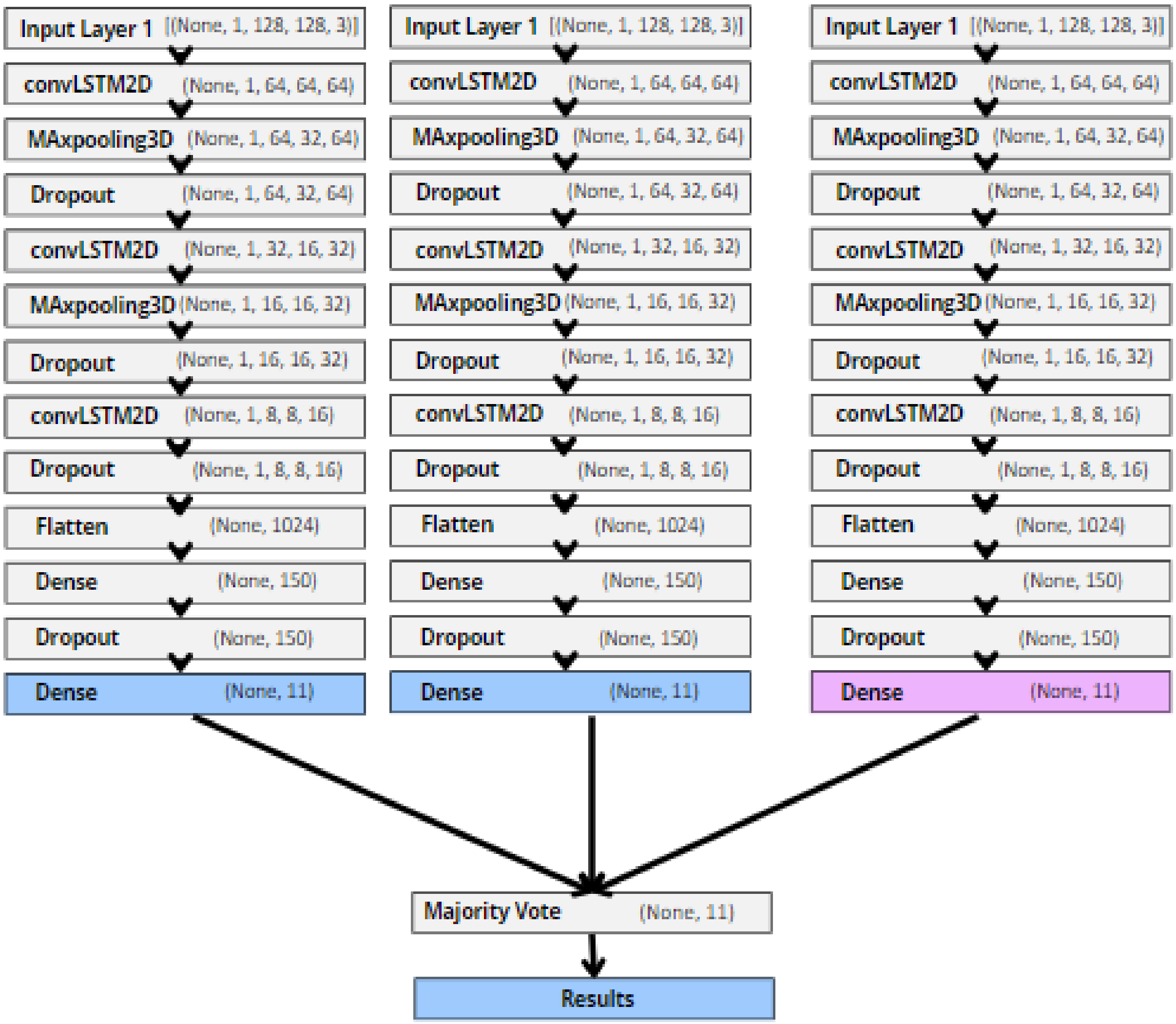
Scheme III- 5: ConvLSTM used for three inputs early fusion model

➢ **Scheme ConvLSTM used for three inputs Halfway fusion model**



Scheme III- 6: ConvLSTM used for three inputs Halfway fusion model

Late fusion scheme is the combination of the three training single inputs using ConvLSTM model and fuse the data on last stage (view **Scheme III- 13**).



Scheme III- 7: ConvLSTM using Late-fusion model

III.4.5.2 Hyperparameters

Hyperparameters is important to maximize the performance of machine learning models. Hyperparameters directly influencing its ability to generalize on unviewed data. Poor configuration can lead to under optimal models, either by overfitting or by under learning (underfitting). In this project, we have adopted a systematic approach to selecting and adjusting hyperparameters, using (Gridsearch). This allows us to explore a wide and complex range of parameters, ensuring that we identify the configurations that provide the best performance for our models.

Gridsearch

Gridsearch is an hyperparameter function that automates the process exist on scikit learn library allow an exhaustive search over a specified parameters grid to find the combination of hyper parameters that give the best performance for model training, this hyperparameter used for large parameter grids Figures bellow illustrate an example about Gridsearch processing.

Figure III- 19: Hyperparameters example

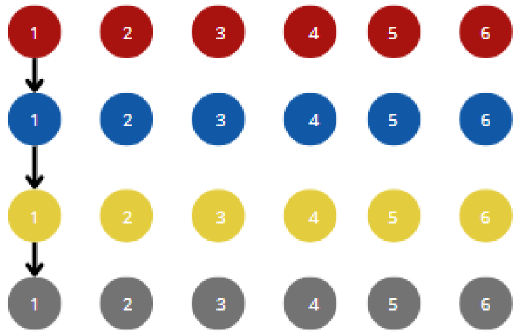


Figure III- 20: First step of processing

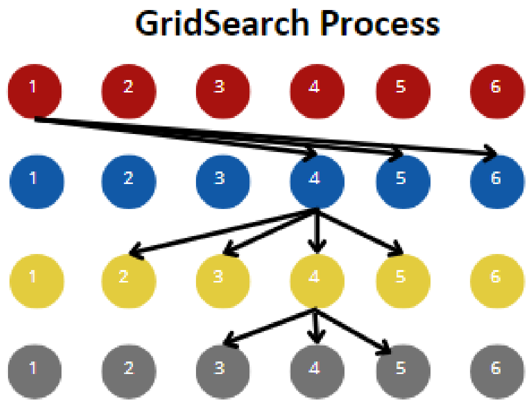


Figure III- 21: Illustration about Gridsearch processing

Figure III- 22: The combination of best hyper parameters (result final)

At the end of process this line will appear

Exemple:

Best Parameters: {'batch_size': 32, 'dropout': 0.35, 'filters': 64, 'lr': 0.001} Best Accuracy: 84.23%

Batch Size refers to the number of training examples used in one iteration common batch sizes ranges from 8 to 256 choosing an appropriate batch size is important for efficient training [149].

Epochs: The number of epochs is a hyperparameter that defines the number times that the learning algorithm will work through the entire training dataset [150].

Learning rate is a hyper parameter that determines the size of steps taken during the optimization process the optimal learning rate depends on factors as architecture of neural network also data characteristics learning rate range {0-1}.

Dropout hyperparameter effectively reduces the co-adaptation between neurons, forcing the network to learn more robust features dropout range {0.1-0.5}.

III.4.6 Evaluation Methods

In this section, we will define the main components of the measures we will use to evaluate classification models.

III.4.6.1 Confusion Matrix

Confusion matrix for multiple classes similar to the binary class matrix where the columns represent the original or the excepted class distribution, and row represent the predicted or output distribution by the classifier, The diagonal elements represent the correct predictions for the classifier, using this concept we can calculate the class-wise Accuracy, precision, recall and f1-scores evaluations parameters [151].

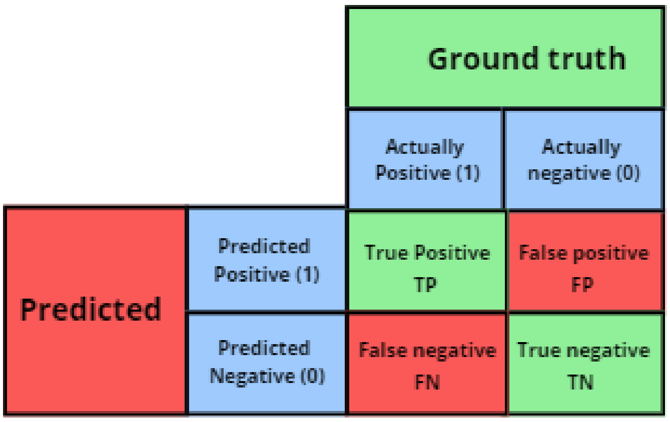


Figure III- 23: Confusion Matrix

III.4.6.2 Accuracy

Accuracy criteria are concerned with more general aspects of classification performance. Note that the term "classification accuracy" is occasionally used in the literature to denote simply the proportion correctly classified in percentage [152].

III.4.6.3 Precision (P)

P measures the percentage of samples correctly classified.

from sklearn.metrics import precision_score

$$P = \frac{TP}{TP+FP} \qquad \text{III.3}$$

III.4.6.4 Recall

It tells you that out of all the positive classes, how many we predicted correctly. Recall should be as high as possible.

Note that it is also called as sensitivity.

from sklearn.metrics import recall_score

$$\text{Recall} = \frac{TP}{TP+FN} \qquad \text{III.4}$$

III.4.6.5 F1-Score

The F1 score reaches its maximum value of 1 when accuracy and recall are perfect, and its minimum value of 0 when precision and recall are zero.

#from sklearn.metrics import f1_score

$$\text{Score F1} = 2 \times \frac{Recall \times Precision}{\text{Recall+ Precision}} \qquad \text{III.5}$$

III.5 Results

In this section, we introduce the results obtained from the different architectures experienced on the Anaconda platform. The data presented are the result of training research in the field of deep learning, aimed at evaluating the effectiveness of various approaches on HAR Systems. We start by training the basic methods training of each radar separately and evaluate their performances then we process to the solutions suggested.

III.5.1 Single radars training Results

We present the results of a basic method using one radar and observe their results using CNN and multimodal methods to have a vision and perspectives to compare and discuss with the other methods suggested in our project these radars presented 10GHZ, 24GHZ, 77GHZ.

Table III- 4: Results of single inputs (single radar training).

Results of training separated radars on the three models implemented						
MODEL	RADARS (GHz)	ACCURACY %	PRECISION %	FI-SCORE %	RECALL %	TRAINABLE PARAMETERS
CNN	10	87	88	87	87	281.683
	24	80	84	79	80	281.683
	77	93	94	93	93	281.683
CNN-LSTM	10	87	89	88	87	555.275
	24	81	85	81	81	555.275
	77	90	92	90	90	555.275
ConvLSTM	10	90	92	91	90	448.467
	24	88	90	88	88	448.467
	77	92	93	92	92	448.467

Models approach suggested as solution For HAR system

III.5.2 Combined data Results

Results of combined data using CNN Model, CNN-LSTM, Conv-LSTM model

On this approach we process to a database level solution suggested defined previously in models approach and architecture used for one input for our HAR system here we observe the training results of the different models implemented.

Table III- 5: Results Combination of datasets

Results of combination datasets on the three models implemented						
MODEL	DATASETS	ACCURACY %	PRECISION %	FI-SCORE %	RECALL %	TRAINABLE PARAMETERS
CNN	COMBINED	83	84	89	89	281.683
CNN-LSTM		80	84	81	80	552.275
ConvLSTM		82	82	82	82	448.467

III.5.3 Fused data Results

Results of fused data using CNN Model, CNN-LSTM, Conv-LSTM model

Fused data is a database level solution, we implement the models and architectures defined previously for our HAR system, here we present the training results of various models implemented.

Table III- 6: Results Fused data

Results of fused datasets on the three models implemented						
MODEL	DATASETS	ACCURACY %	PRECISION %	FI-SCORE %	RECALL %	TRAINABLE PARAMETERS
CNN	FUSED	89	91	89	89	281.683
CNN-LSTM		80	84	81	80	552.275
ConvLSTM		93	94	93	93	448.467

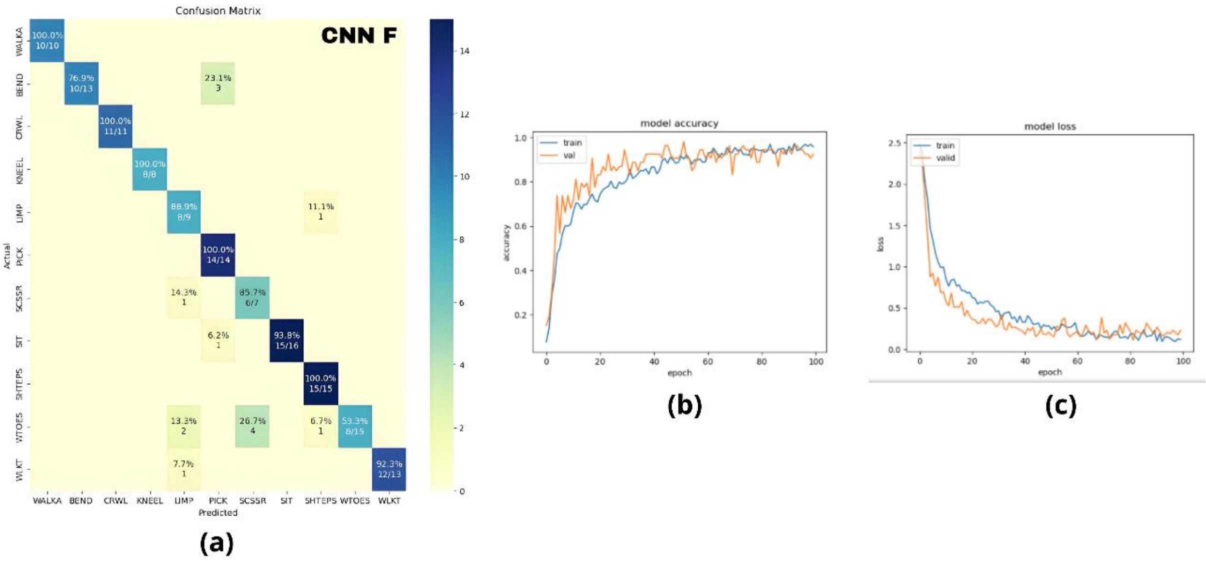


Figure III- 24: Fused data (CNN model): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

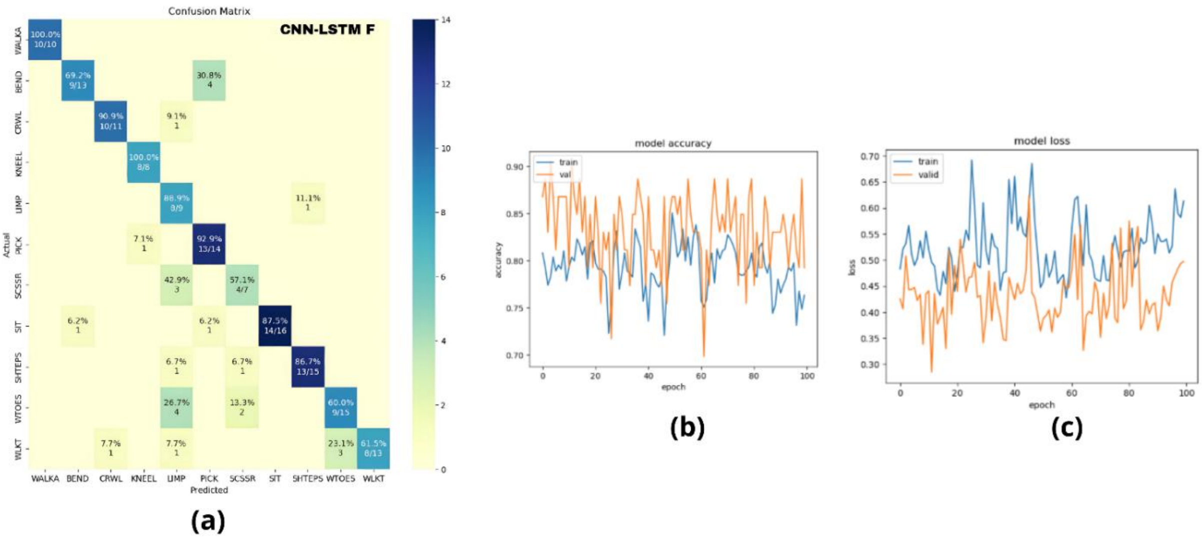


Figure III- 25: Fused data (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

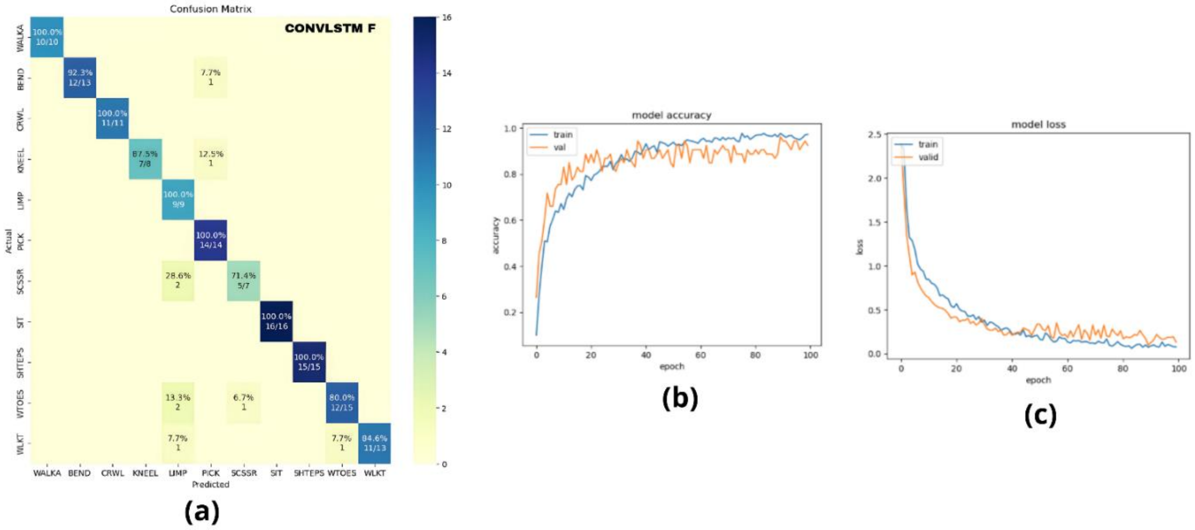


Figure III- 26: Fused data (ConvLSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

III.5.4 Concatenated data Results

Results of concatenated data using CNN Model, CNN-LSTM, Conv-LSTM model

We are employed a solution by making a change at database using architectures defined previously four our HAR system here we present the training results which can allow us to observe and analyze the performances of our system.

Table III- 7: Results Concatenated data in one frame

Results of concatenated datasets (3 in 1) on the three models implemented						
MODEL	DATASETS	ACCURACY %	PRECISION %	FI-SCORE %	RECALL %	TRAINABLE PARAMETERS
CNN	Concatenated	94	95	94	94	435.283
CNN-LSTM		90	90	90	90	1.076.563
ConvLSTM		90	92	90	90	755.667

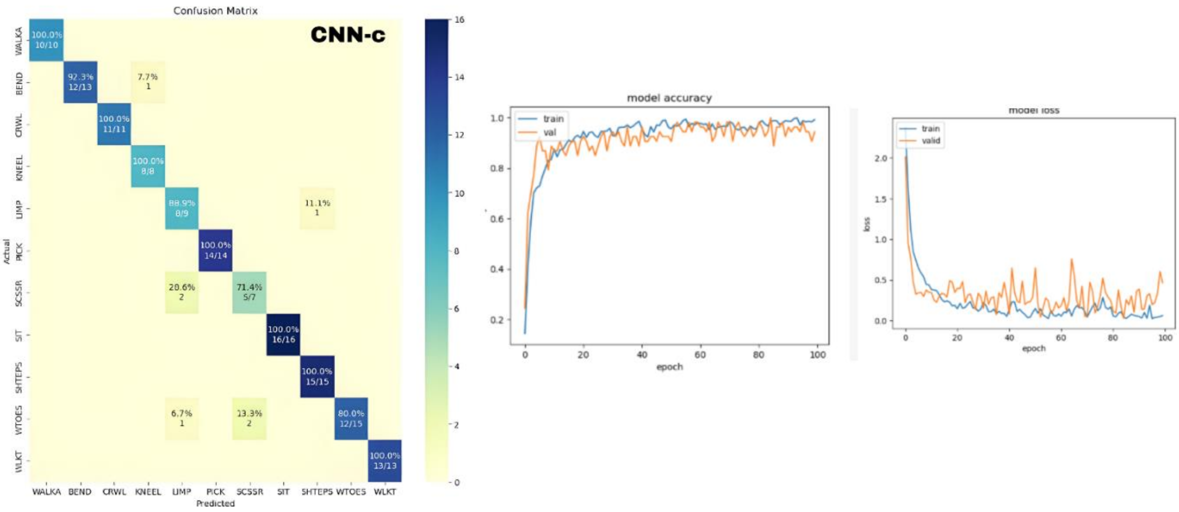


Figure III- 27: Concatenated data (CNN model): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

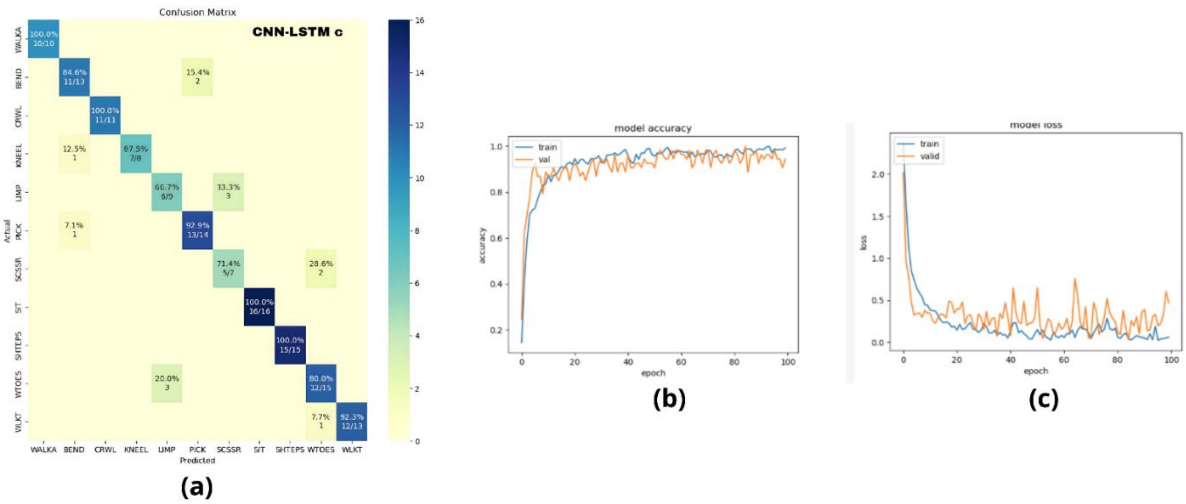


Figure III- 28: Concatenated data (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

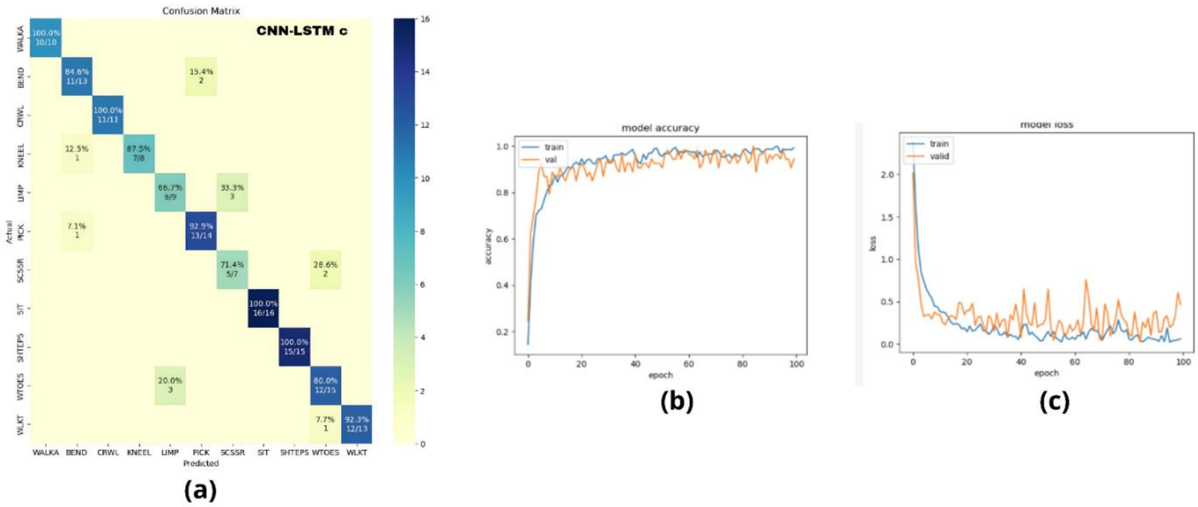


Figure III- 29: Concatenated data (ConvLSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

III.5.5 3D results

We are employed a solution by making changes at the input database Using 3D architecture defined previously in our HAR system, which uses a cube of data resulting from the use of stack function, here we present the training results, which can allow us to observe and analyze the performance of our system.

Table III- 8: 3D Model results

Results of 3D model cube of data on the three models implemented						
MODEL	METHOD	ACCURACY %	PRECISION %	FI-SCORE %	RECALL %	TRAINABLE PARAMETERS
CNN	Cube of data	78	79	78	78	327.027
CNN-LSTM		73	75	74	73	570.355
ConvLSTM		82	82	82	83	1.466.515

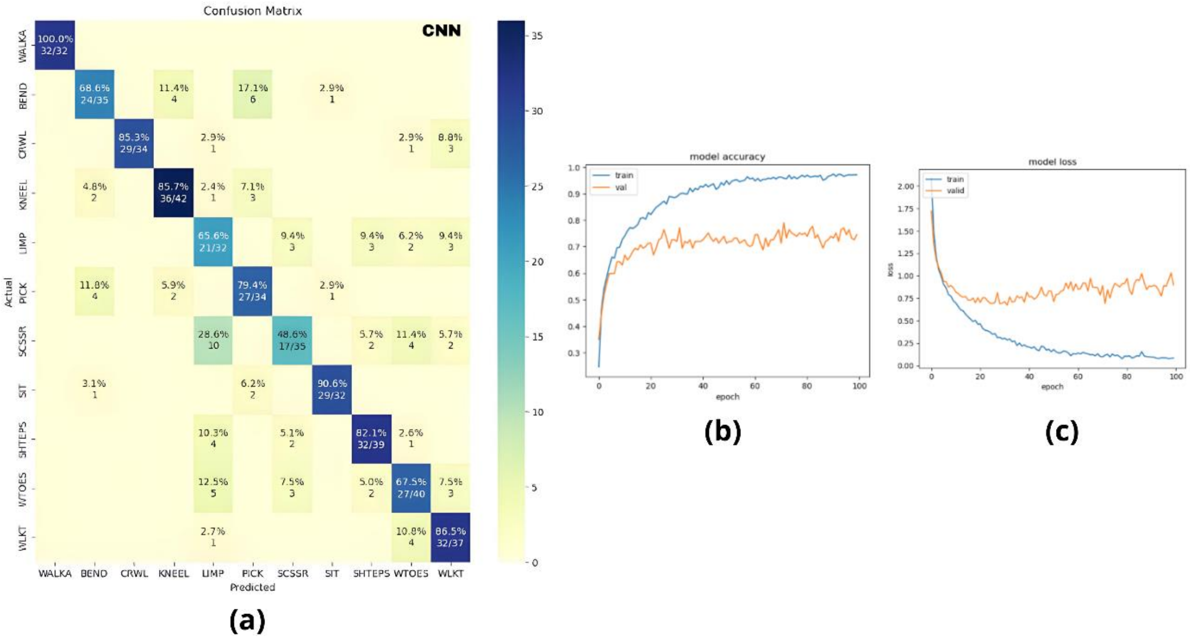


Figure III- 30: 3D Model (CNN): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

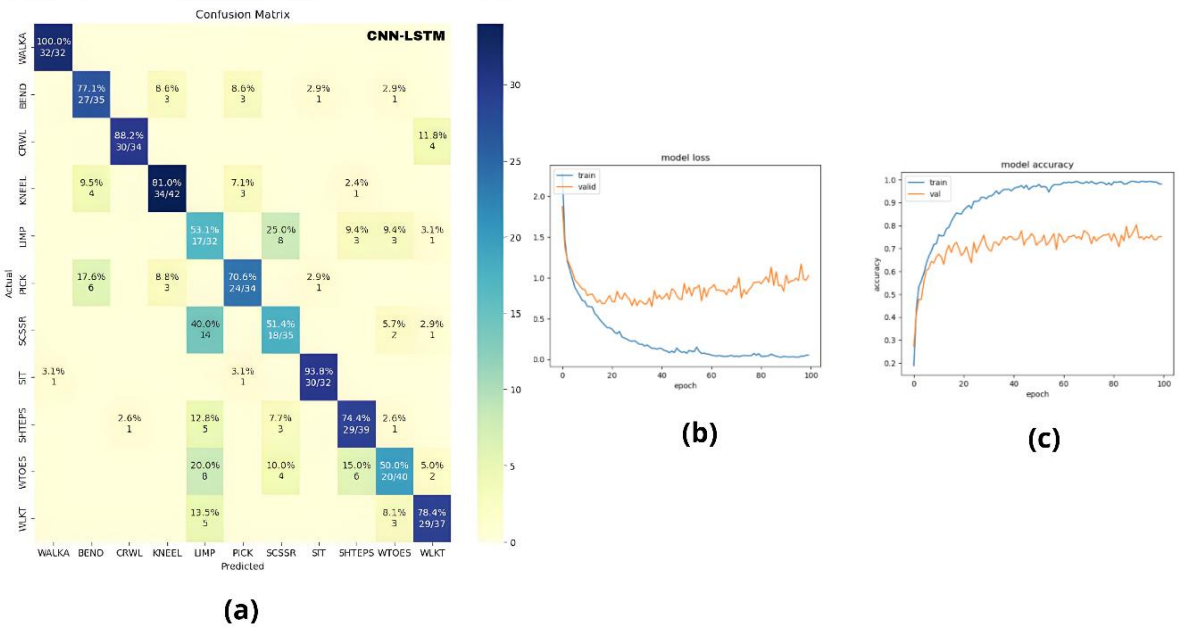


Figure III- 31: 3D Model (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

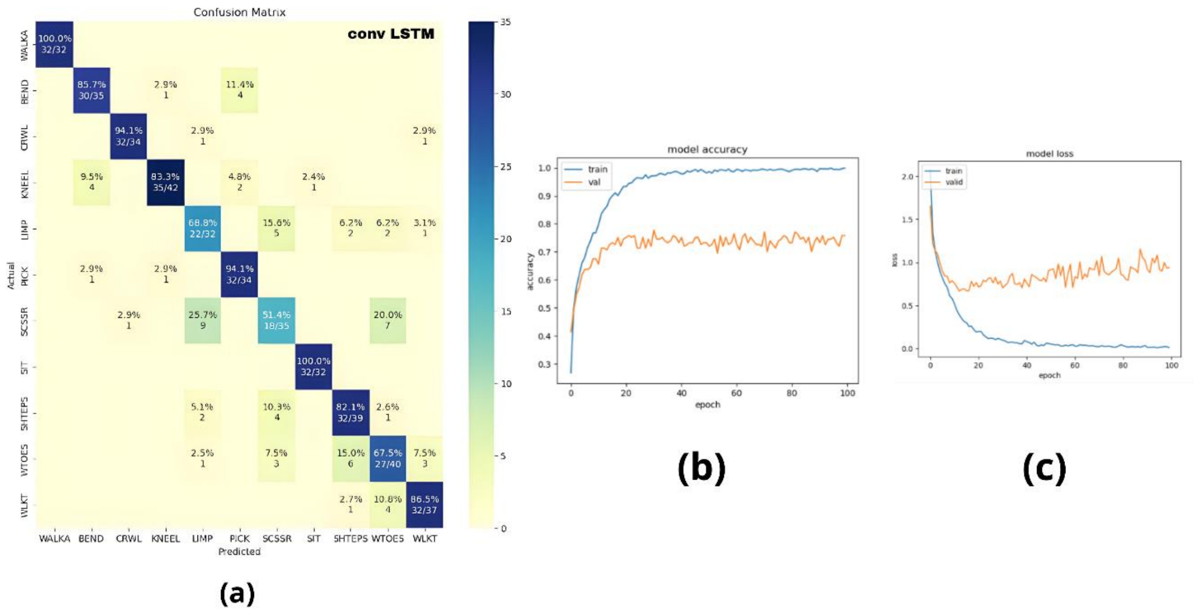


Figure III- 32: 3D Model (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

III.5.6 Models fusion results

Results of fusion methods integrated on CNN Model, CNN-LSTM, Conv-LSTM model

The current approach process on models levels as defined previously for the fusion methods early fusion, halfway fusion and late fusion we are above to implement the architectures mentioned for three inputs data which allows us to observe and analyze the performances of our solutions implemented to the HAR system.

III.5.6.1 Early fusion Results

Early Fusion

Table III- 9: Results Early fusion

Results of Early fusion on the three models implemented						
MODEL	FUSION	ACCURACY %	PRECISION %	FI-SCORE %	RECALL %	TRAINABLE PARAMETERS
CNN	Early Fusion	85	92	85	85	228.371
CNN-LSTM		94	95	92	94	498.963
ConvLSTM		90	93	90	91	462.291

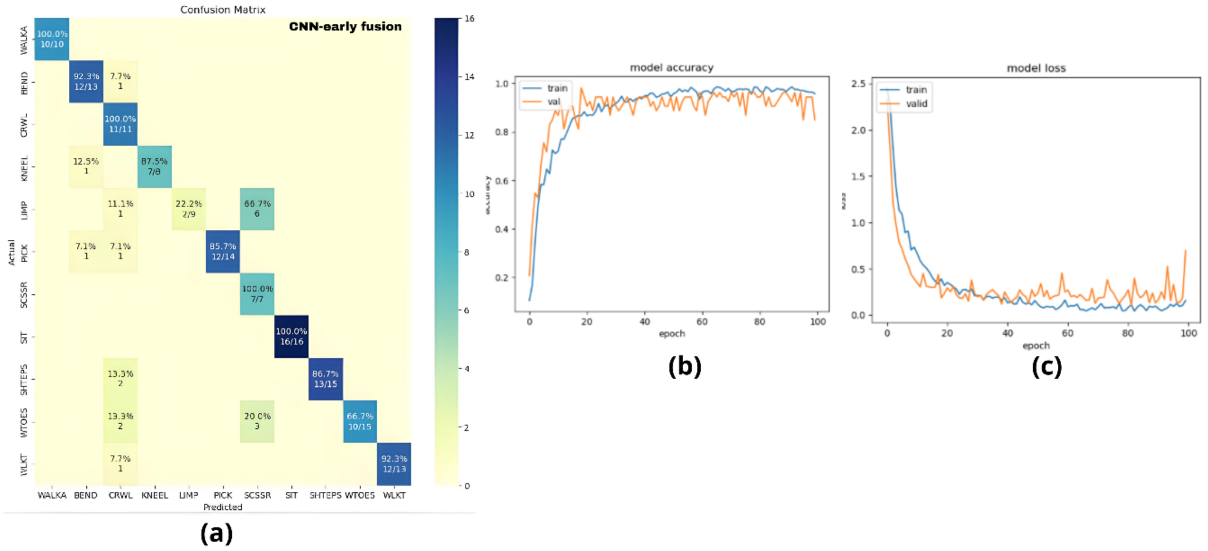


Figure III- 33: Early Fusion (CNN model): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

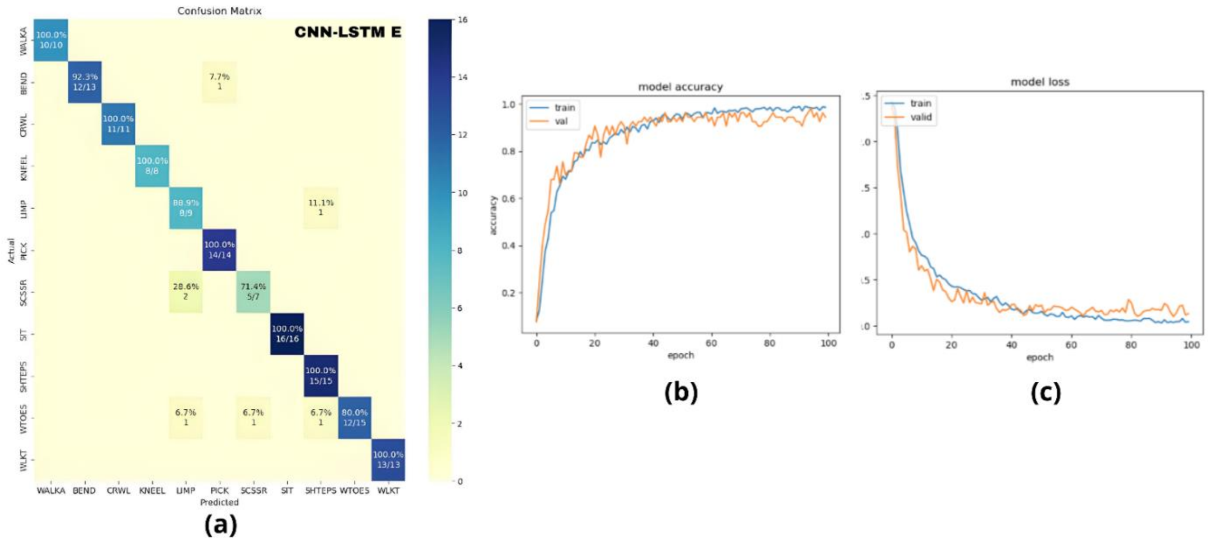


Figure III- 34: Early Fusion (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

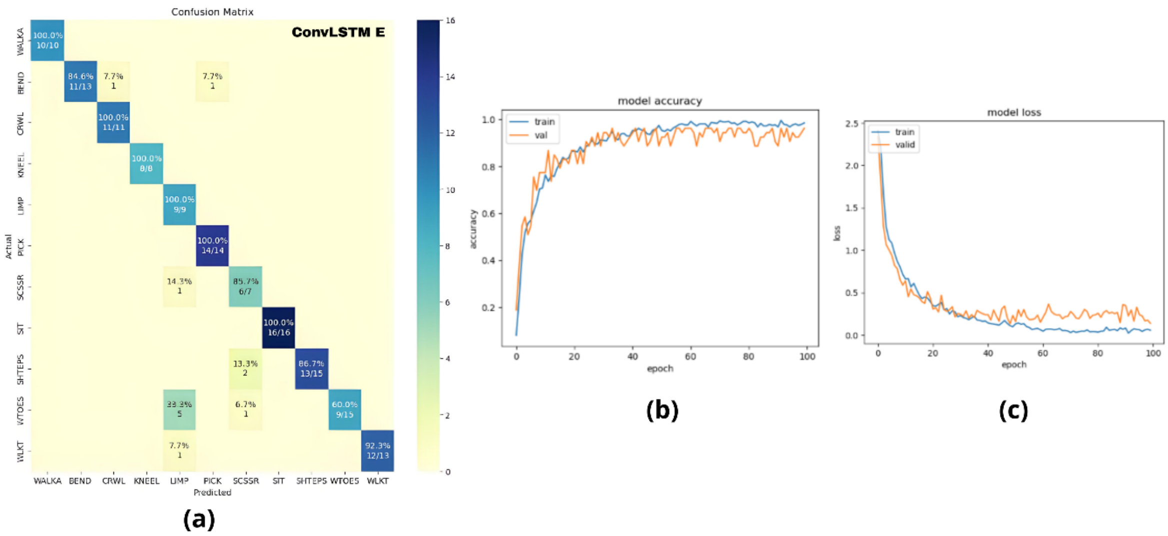


Figure III- 35: Early Fusion (ConvLSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

III.5.6.2 Halfway Fusion Results

Table III- 10: Results Halfway Fusion

Results halfway fusion on the on the three models implemented						
MODEL	Fusion	ACCURACY %	PRECISION %	FI-SCORE %	RECALL %	TRAINABLE PARAMETERS
CNN	Halfway Fusion	92	94	92	92	841.427
CNN-LSTM Before Concatenate		96	96	96	96	1.653.203
CNN-LSTM After Concatenate		96	96	96	96	1.285.907
ConvLSTM		93	95	94	93	1.341.779

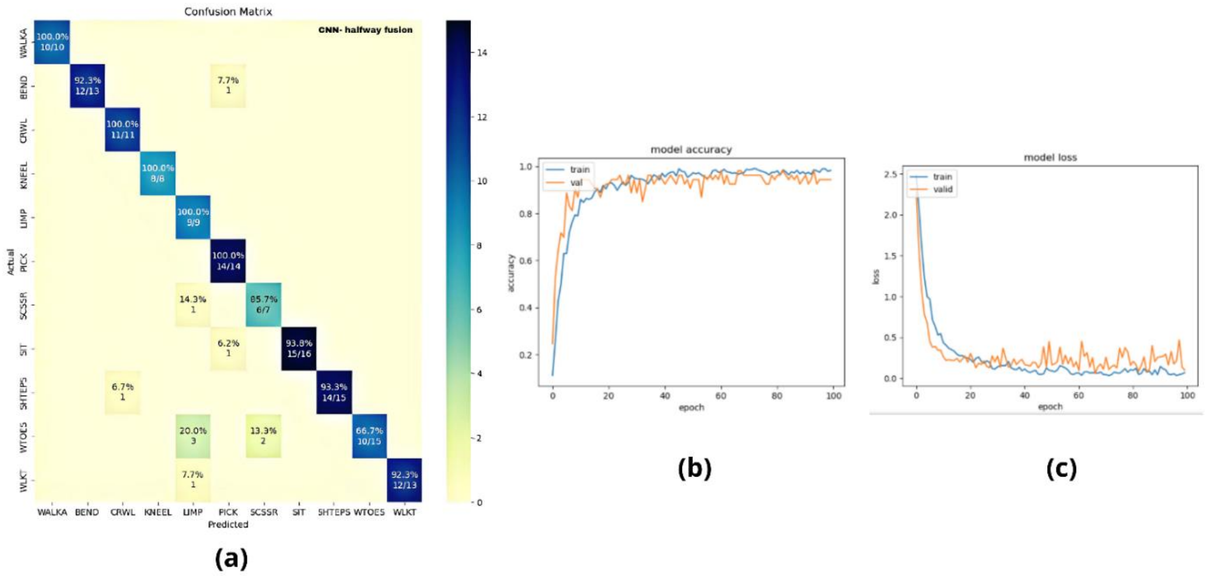


Figure III- 36: Halfway Fusion (CNN model): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

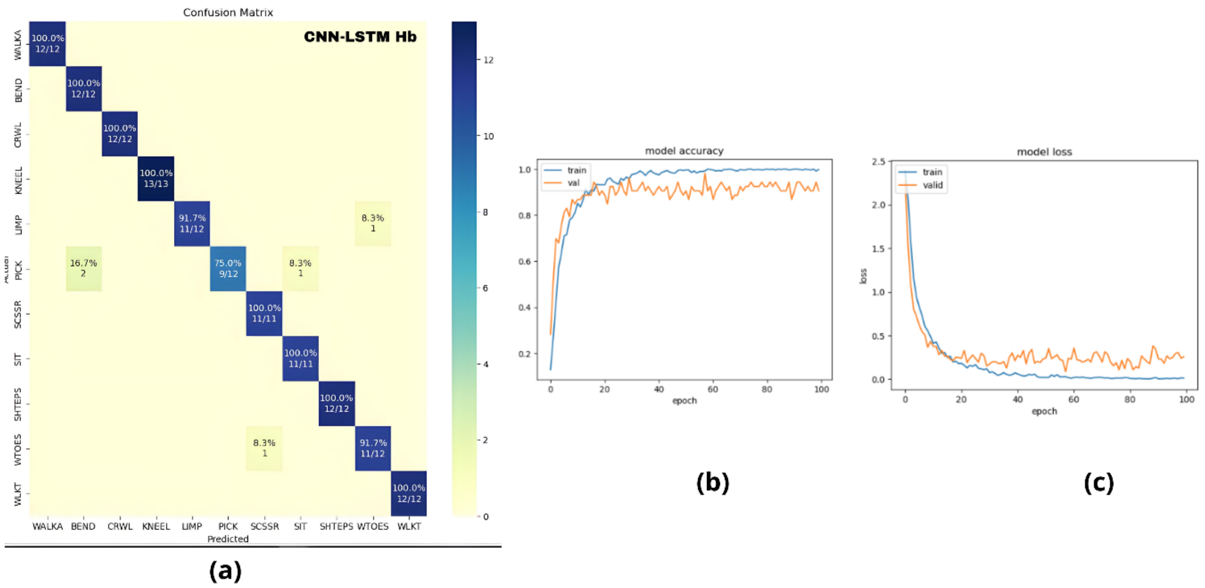


Figure III- 37: Halfway Fusion before concatenation (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

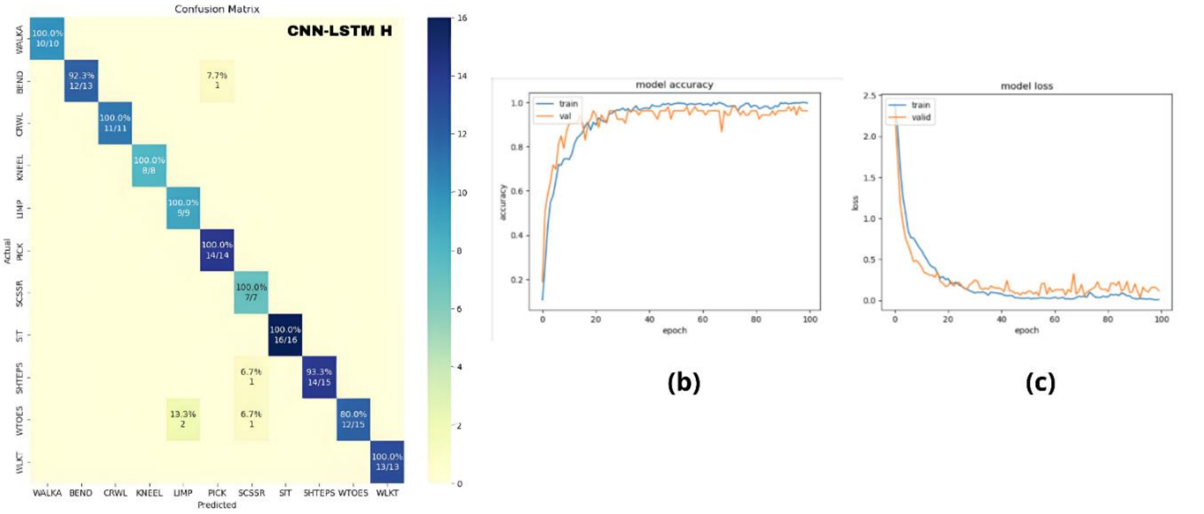


Figure III- 38: Halfway Fusion after concatenation (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

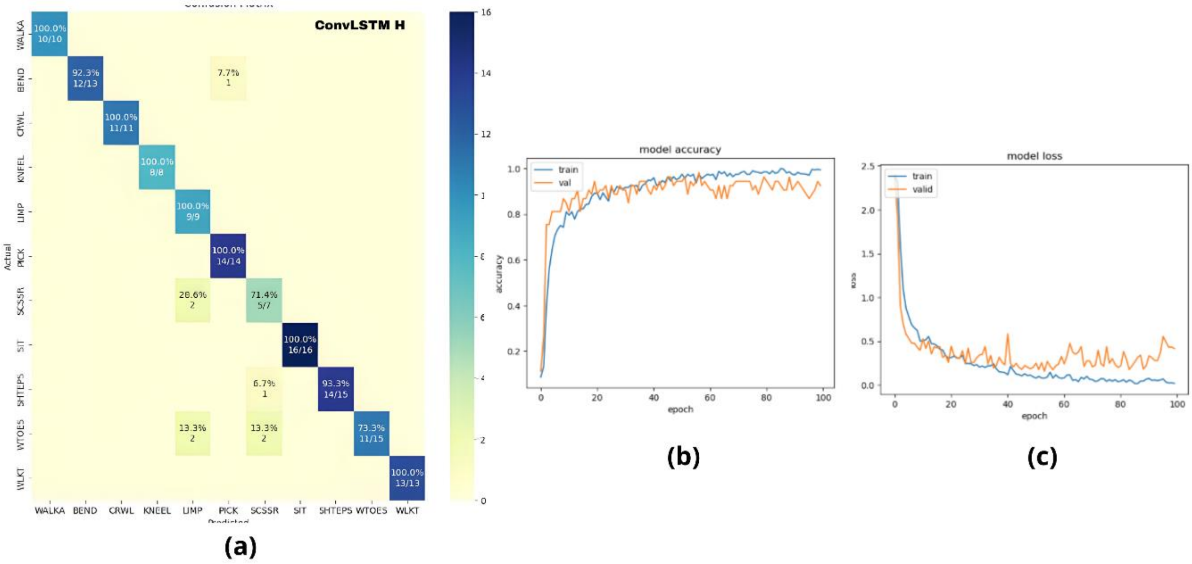


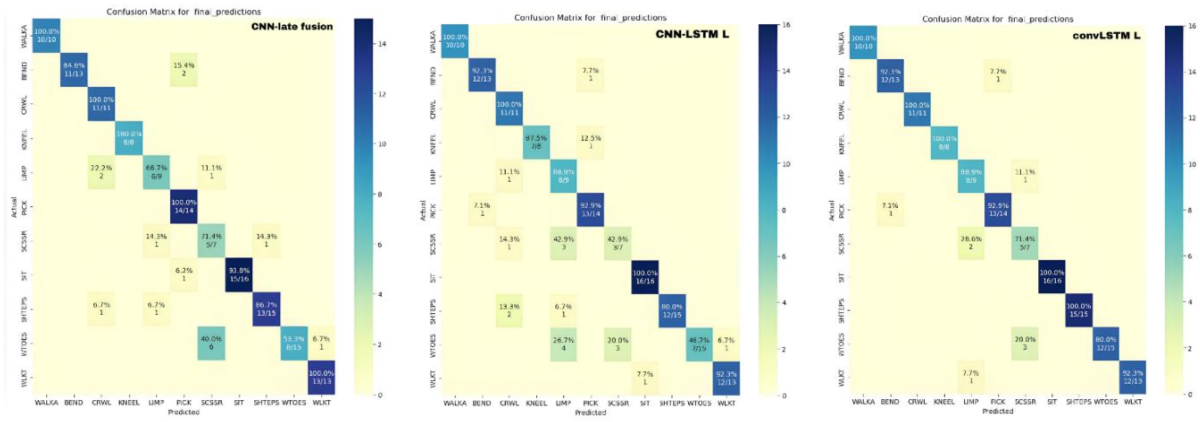
Figure III- 39: Halfway Fusion (ConvLSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

III.5.6.3 Late fusion Results

Late-fusion:

Table III- 11: Results Late fusion

Results Of late fusion on the three models implemented					
MODEL	DATASETS	ACCURACY %	PRECISION %	FI-SCORE %	TRAINABLE PARAMETERS
CNN	Late Fusion	87	89	87	845.049
CNN-LSTM		84	87	84	1.656.825
ConvLSTM		93	94	93	1.345.401



(a) (b) (c)

Figure III- 40: Confusion Matrix of late fusion (a.CNN model) (b.CNN-LSTM model) (c.ConvLSTM model)

III.6 Discussion

On this section we present analysis about architectures and results of the different approaches and comparing between the different solutions and standard ones.

Table III-4: results of single inputs (single radar training)

The **Table III-4** shows the accuracy, precision and F1-score also Recall of the different radars we can observe from those results.

Accuracy: overall, the accuracy in range of 0.80 to 0.93 the system can classify correctly the radars frames, Precision is also good rated from 0.84 to 0.94 there is a high proportion through the targets that the radar identify true positives , F1-Score show a range of 0.79 to 0.93 conclude there is a balance in our HAR system, The recall shows an interval from 0.80 to 0.93 the system have an suitable performances to the capability to ensure achieving true classification of the activity classes.

The number of trainable parameters influences the model's training in term of complexity and capacity CNN-LSTM parameters is higher than CNN and ConvLSTM models, on training results it appears that ConvLSTM outperform CNN and CNN-LSTM models in term of the different evaluations methods.

To conclude we see a good result using Conv-LSTM model which is adapted to HAR offer a better result with higher precision comparing the other models, The performance between the different models is significant for the HAR system but still have a Challenges on the recognition system using a single radar to complete this mission.

Combined datasets results

Table III-5: results Combined data

The **Table III-5** shows the accuracy, precision and F1-score also Recall of the models implemented to the combined database we can observe from those results

Accuracy: all the models achieved relatively high average accuracy with combined CNN a highest value 0.83 comparing to the other implemented models, Precision the three models achieved a balanced results between them there is an ability to identify true positive ones, F1-score and recall we observe a significant results highest scores using CNN model.

The models training influenced by training parameters CNN-LSTM parameters is higher than CNN and ConvLSTM models, For the models implemented at the combined database it appears that CNN outperform CNN-LSTM and ConvLSTM models in term of the different evaluations methods.

Fused data results

Table III-6: results Fused data

The **Table III-6** Shows the accuracy, precision and F1-score also Recall of the models implemented to the fused database we can observe from those results.

Accuracy: Fused ConvLSTM with the highest accuracy (0.93) followed by Fused CNN (0.89) and Fused LSTM (0.80), **Precision:** Fused ConvLSTM achieves the highest precision (0.94) followed by Fused CNN (0.91) and Fused LSTM (0.84), F1-score and recall.

There is an important difference between the two models implemented CNN and ConvLSTM, where the ConvLSTM surpass the CNN model on performances, where on the trainable parameters 448,467 which surpass the CNN model on training parameters where both of architectures used 3 convolutions layers (scheme III-1 & scheme III-10) but the ConvLSTM specification of recognize give an advantage on evaluation methods.

The figure III-26 that show's results of fused data with implementation of conv-LSTM, the x-axis represents the epochs, which are iterations over the training data. where the y-axis represents accuracy, a metric indicates how often the model makes correct predictions. We observe on the graph accuracy and loss there is an overfitting where the model performs on training with low performances on the unseen data.

The confusion matrix showed presents the performance of our HAR system classified some classes correctly in the other side our system has some difficulty to classify some Activities.

Concatenated Data results

Table III-7: results Concatenated data

The Table III-7 Shows the accuracy, precision and F1-score also Recall of the models implemented to the concatenated database we can observe from those results.

Accuracy the training models achieved a good performance a high accuracy 0.94 using CNN model, for the precision a significant result observed using the three models on Concatenated database a huge accuracy marked for the CNN model, the recall and f1-Scores a similar result for the implementation of CNN-LSTM and ConvLSTM models with 0.90 of rate where the CNN model surpass with higher performances with score and recall marked 0.94

On trainable parameters we observe a significant difference between the models implemented which CNN model with reduced amount of parameters indicates low complexity of model while comparing with CNN-LSTM model and ConvLSTM over 750,000 trainable parameters which we have an pre-vision about the complexity of the models implemented.

The concatenated data refers to combine data of different radars into a single frame gives an advantage for learning process specially for CNN model implemented on the concatenated data.

In second we see the (figure III-27) of implementation CNN model on the concatenated data that contains the confusion matrix and accuracy /loss graphs, The x-axis represents the epochs, which are iterations over the training data. where the y-axis represents accuracy, a metric indicates how often the model makes correct predictions.

The plot show’s model accuracy on training and validations datasets over iterations(epochs) indicated that we have a training higher than validation which is a sign of overfitting model accuracy reaches its maximum of 93% at epoch 71.

The confusion matrix shows that the object been well classified into each predicted class all classes placed on the diagonal are correct classifications we observe that our model performs on some classes as (WLKT/WTOES) classes and fails to classify on others with low-rate errors.

To conclude the CNN shows a suitable result for classifying classes that helps further in our HAR system.

Table III-8: 3D Model result

The **Table III-8 shows** Precision, Recall and F1-Score: ConvLSTM has the highest values in precision (82%), recall (83%) and F1-Score (82%) compared to CNN and CNN-LSTM.

Trainable Parameters ConvLSTM (1,466,515) has a significantly higher number of trainable parameters compared with others models This implies that ConvLSTM is a more complex model.

On evaluation metrics (figure III-32) ConvLSTM shown for better performances let's Visualize and discuss the graph which presents accuracy curve, loss curve and confusion matrix.

The curve displays lower accuracy and higher loss indicates a poorer fit to the training data, while the confusion matrix shows the performance of classification.

Conclusion of the methods suggested at database level.

We conclude that the concatenated data in one frame using CNN model gives such a step-in advance to our project goal to build a performed and suitable solution in term of less complexity and precision accuracy for the mission of recognition human activities.

Models Fusion results

Fusion methods with implementation of CNN, ConvLSTM and CNN-LSTM

Here we discuss the results of the fusion methods integrated on the models defined Early fusion, Halfway fusion and late fusion.

Table III-9: results Early fusion

The table displays **Table III-9** the models performances using early fusion method.

The accuracy shows above the 0.85 a sign that performs well on classifications of classes highest accuracy marked for using LSTM layer, the precision there is a results balance with recall F1-score mark an important value using CNN-LSTM that surpass the other models implemented, however the trainable parameters is medium with 498,963 higher than the models implemented also using for LSTM layer gives a pre-vision of complexity of the model implemented.

On evaluation metrics we aim on the figure that presents the accuracy/loss and confusion matrix of early fusion method using LSTM.

The (figure III-34) that show's results of early fusion methods with implementation of conv-LSTM we observe on the graph accuracy and loss there is an overfitting where the model performs on training with low performances on the unseen data.

The confusion matrix showed presents the performance of our HAR system classified some classes correctly in the other side our system has a less difficulty to classify some activities classes.

To conclude by adding LSTM layer Show’s the early fusion the most effective model comparing with the other models implemented.

Table III-10: results Halfway fusion

The table **Table III-10** displays the models performances using Halfway fusion method

The halfway fusion using CNN model achieves an accuracy of 0.85 and precision 0.92 with f1-score and recall 0.85 and using LSTM an important Results presented with Accuracy and precision also f1-score and recall with 0.96 sign that is the performed model by comparing with CNN model implemented and conv-LSTM model.

Both CNN-LSTM models (before and after concatenation) have a significantly higher number of trainable parameters compared to the CNN model. suggests that the CNN-LSTM model have a more complex architecture.

The CNN-LSTM after concatenation model has the highest number of trainable parameters among all three. This could be due to the additional parameters introduced by the concatenation operation, which combines the outputs of the CNN and LSTM layers before feeding them into the final classification layer (scheme III-7 & scheme III-8).

On evaluation metrics we aim on (the figures III-36 & III-37) that presents the accuracy/loss and confusion matrix of Halfway fusion method using CNN-LSTM model, The x-axis represents the epochs, which are iterations over the training data. where the y-axis represents accuracy, a metric indicates how often the model makes correct predictions.

The training accuracy (figureIII-36b & III-37b) reaching nearly 1.0 suggests the model is effectively fitting the training data, we observe in both of 2 suggested solutions before concatenation and after concatenation layer, the graphs which presents a great display and more efficient is the solution implemented after concatenation layer.

the confusion matrix shows that the object been well classified into each predicted class all classes placed on the diagonal are correct classifications we observe that our model performs on some classes as (WLKT/SIT/LIMP) classes and fails to classify on others with very low-rate errors.

To conclude by adding LSTM layer after concatenation Show’s the Halfway fusion the most effective model comparing with the other models implemented.

Table III-11: results Late fusion

CNN has low accuracy (0.87) and F1 score (0.87), LSTM: Has the lowest accuracy of 0.84 same as F1-score (0.84) than Conv-LSTM. The precision (0.94) is greater than the recall (0.93). ConvLSTM has the highest accuracy (0.93) and F1-score of the three. It also has a balance of precision (0.94) and recall (0.93), observing late fusion architecture more complex architecture compared to other methods where each datasets trained separately (scheme III-4,9,13) one by one where we observe the trainable parameters is the result of addition of training separated data then applying majority vote for final result.

The confusion matrix (figure III-40) of late fusion method showed presents the performance of our HAR system classified some classes correctly in the other side our system has a less difficulty to classify some activities classes.

To conclude using ConvLSTM Show's the Late fusion the most effective model comparing with the other models implemented on Late-fusion method.

Conclusion of using several solutions as methods of fusion integrated on the models.

We're above to have a vision about the different HAR systems performances and observed that the halfway fusion method using CNN-LSTM model gives such a step-in advance to our project goal to build a performed and suitable solution in term of less complexity and precision accuracy for the mission of recognition human activities.

General Comparison between the solutions

We have a solution chosen by changing on data base level with implementation of CNN model and a method of fusion integrated LSTM after concatenation model

We have the CNN model proposed model been a choice the data base build with concatenated data in one frame we have an operation of concatenation and use a simplified architecture one input and using LSTM layer repose on three inputs of data caused by different radars.

At level of architecture, we have the CNN model typically less complex than LSTM complexity for their specifications store and process data requires more memory and computations also LSTM often take longer to train.

Finally, we can build our recognition of human activities system based on modification of Database much easier implementation also for the model of convolutional neural network less complexity and ensure a fast training of the model.

III.7 Comparative study

In our project we proposed a different solutions model to confirm their quality and validity a comparison of our models with another classification model for the same concept modeling a system based on wireless network of sensors for human activity recognition was carried that use the same datasets Siamese-based architecture which is designed for multiple inputs of data

A Siamese network is a neural network architecture that processes two inputs to determine their similarity. It is commonly used in tasks such as image similarity. It consists of twin networks with shared weights, ensuring identical transformations of the inputs. The outputs are compared using a distance function, such as Euclidean distance. The workflow involves converting input pairs into feature vectors, which are then compared to determine similarity. To train the network, a loss function, such as contrastive or triplet loss, is used, with distances decreased for similar pairs and optimized for dissimilar pairs. This model is designed for multiple-entry situations and can be tested only in multi-input cases, with all weights tied across parallel paths.

The proposed solution is specialized by its simplicity and effectiveness., using models with optimized processing instead of standard data processing, a minimized architecture was achieved that reduces the complexity and the number of parameters required, by applying only the necessary functions. This approach ensures high precision and high accuracy(Table III-12), facilitating database and model modifications and ensuring that the model handles data from different sources and versions and frequencies. In comparison, the standard solution is based on a more complex architecture, combining multiple layers and traditional data processing. While these solutions can provide robust performance, they may be more difficult to maintain and optimize, complicating future adaptations.

These twin networks apply the loss function to infer the relationship between input images A and B (figure III-41). Traditional CNNs handle single inputs aim for tasks like classification or detection and use a class label, while Siamese networks handle pairs of inputs are designed for similarity measures and tasks like verification. Use similarity score or distance metric.

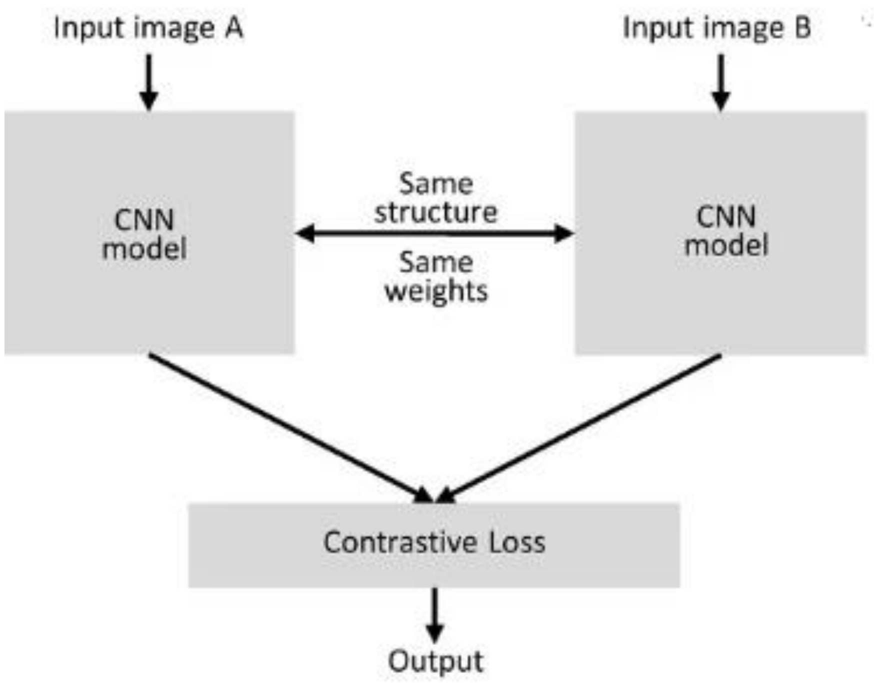


Figure III- 41: Siamese based Architecture

Table III- 12: Performances presented our approach and Mert Ege approach [153]

RESULTS performances of different methods proposed compared with previous researches			
MODEL	METHOD	TRAINABLE PARAMETRES	ACCURACY %
CNN	CONCATENATED DATASETS	435.283	94.6
ConvLSTM	FUSED DATASETS	448.67	93
CNN-LSTM	EARLY FUSION	498.963	94
CNN-LSTM AFTER CONCATENATION	HALFWAY FUSION	1.285.907	96
SIAMESE-BASED [BY MERT EGE]	/	1.283.637	94.3

III.8 Conclusion

On this chapter we employed different approaches using different architectures, models as (CNN, CNN-LSTM, ConvLSTM) and advanced solutions includes data fusion and models fusion with simple architecture, these methods allowed us to overcome the challenges and limitation on our project, a multiple execution and multiple architectures been build and didn't validated because of their performances and complexity levels, the models validity processed by evaluation of the different parameters also comparison to build an robust system for human activity recognition mission.

Where we reached to an interesting system approach using CNN model on Concatenated datasets in one frame surpassed previous searches established allow us to confirm the efficacity of our proposed model on the different terms accuracy and reduced complexity with advantages.

General Conclusion

HAR systems are integrated on various fields which consist to a real time respond with environment with precision, on this purpose a various searchers and developers started to create new systems based on deep learning and machine learning models for recognition of human activities task.

The development of artificial intelligence gives new opportunities on technology field using certain tools and parameters to build a robust system with efficacity and precision, using basic methods is still efficiently on specified data includes shapes dimension and treatment.

Understanding human activities purposes with their different types also the integration of recognitions systems for detection and classification of human activities allows us to have a knowledge about the different systems and tools to build an HAR system.

We used a database of spectrogram images built by UWB radars generated by micro doppler signals treated and processed by STFT an advanced tool for treatment of signals used to decompose with window parameters

The study of basic methods single radar systems with programs from scratch and personal designed architectures allowed us to know the negative points these existing systems then used to optimize standards methods by changing the parameters with implementation of different models (CNN, CNN-LSTM, ConvLSTM) and evaluate their performances also observe their limitations classifications

We proposed solution of combination of datasets of different radars with implementation of different models (CNN, CNN-LSTM, ConvLSTM) and evaluate their performances and their limitation allowed to explore further solutions and opened the gate to new solution as Fused datasets and Concatenation of datasets in one frame

The fused results given an important results but with a high complexity, Concatenated datasets offers best performances with reduced complexity built an prevision about the final model approach, the solutions proposed didn't limited at level of database fusion operation also used on Advanced fusion on models used to implement (early-fusion, halfway-fusion and late-fusion) those methods also implemented on the different models approaches CNN and hybrid models by observation and evaluation of their performances

Finally, to confirm our model system we used to compare the previous researches carried out to check the validity of the proposed solution we have been able to validate our suggested solution for Huaman activity recognition Mission overcoming all Challenges encountered During our project.

Future perspectives

Future perspectives in field of HAR with AI and DL includes improving models for greater accuracy and efficiency as well adaptability of systems to users, even motivated to implement HAR systems on the real world by integrate on the different fields health monitoring reduces risks and costs and offer more detailed analysis of athletes, offering modern solutions for home automation .the implementation of a network of sensors provide adaptivity and accuracy their implementation beneficial on a large scale.

Bibliography

[1] https://www.who.int/news-room/fact-sheets/detail/physical-activity

[2] https://www.physio-pedia.com/Physical_activity

[3] Blöbaum, B., & Hunecke, M. (2005). Motives, environmental knowledge, and environmental behavior: A study of tourists. Environment and Behavior, 37(3), 436-460.

[4] Van Vliet, W., Bregt, A. K., & Hagen-Zanker, A. (2013). Revisiting KAB (knowledge, attitudes, behavior) theory for environmental policy and management. Environmental Education Research, 19(1), 1-20

[5] Y. Zheqi et al., "A Radar-Based Human Activity Recognition Using a Novel 3-D Point Cloud Classifier," IEEE Sensors Journal, vol. PP, pp. 1-1, 10/01 2022, doi: 10.1109/JSEN.2022.3198395.

[6] Augusto, Juan C; Callaghan, Vic (2013). "Intelligent Environments: a manifesto" (PDF). Human-centric Computing and Information Sciences. 3: 12

[7] Steventon, A., and Wright, S. (eds) (2006) Intelligent Spaces: The Application of Pervasive ICT, Springer-Verlag

[8] Wang, L., Gu, T., Tao, X., Lu, J., & Bahsoon, R. (2019). Human activity recognition based on deep convolutional neural networks. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 50(2), 726-735.

[9] M. M. Islam, S. Nooruddin, F. Karray, and G. Muhammad, "Human activity recognition using tools of convolutional neural networks: A state of the art review, data sets, challenges, and future prospects," Computers in Biology and Medicine, vol. 149, p. 106060, 2022/10/01/ 2022, doi: https://doi.org/10.1016/j.compbiomed.2022.106060.

[10] Bao, L., & Intille, S. S. (2004). Activity recognition from user-annotated acceleration data. In Pervasive computing (pp. 1-17). Springer, Berlin, Heidelberg.

[11] F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, "Human Activity Recognition Using Inertial, Physiological and Environmental Sensors: A Comprehensive Survey," IEEE Access, vol. 8, pp. 210816-210836, 2020, doi: 10.1109/ACCESS.2020.3037715

[12] CHUNG, Seungeun, LIM, Jiyoun, NOH, Kyoung Ju, et al. Sensor data acquisition and multimodal sensor fusion for human activity recognition using deep learning. Sensors, 2019, vol. 19, no 7, p. 1716.

[13] Diraco, G.; Rescio, G.; Caroppo, A.; Manni, A; Leone, A. Human Action Recognition in Smart Living Services and Applications: Context Awareness, Data Availability, Personalization, and Privacy. Sensors 2023, 23, 6040. https://doi.org/ 10.3390/s23136040

[14] S. Vijayarani, M. J. Ilamathi, and M. Nithya, "Preprocessing techniques for text mining-an overview," International Journal of Computer Science & Communication Networks, vol. 5, no. 1, pp. 7-16, 2015.

[15] Y. Z. Lin, Z. H. Nie, and H. W. Ma, "Structural damage detection with automatic feature-extraction through deep learning," Computer-Aided Civil and Infrastructure Engineering, vol. 32, no. 12, pp. 1025-1046, 2017.

[16] K. Tyagi, C. Rane, R. Sriram, and M. Manry, "Unsupervised learning," in Artificial intelligence and machine learning for edge computing, pp. 33-52, Academic Press, 2022.

[17] T. Hastie, R. Tibshirani, and J. Friedman, "The elements of statistical learning: data mining, inference, and prediction," Springer Science & Business Media, 2009.

[18] R. Guoguang, Y. Zheng, and M. Sawan, "Energy Solutions for Wearable Sensors: A Review," Sensors, vol. 21, no. 11, p. 3806, 2021, doi: 10.3390/s21113806.

[19] Çağlar Seçkin, B. Ateş, and M. Seçkin, "Review on Wearable Technology in Sports: Concepts, Challenges and Opportunities," Appl. Sci., vol. 13, p. 10399, 2023. [Online].

[20] Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. ACM Computing Surveys (CSUR), 46(3), 33.

[21] Wikipedia contributors, "Radar," Wikipedia, The Free Encyclopedia, Wikimedia Foundation, 25 Apr. 2024. [Online]. Available: https://en.wikipedia.org/wiki/Radar.

[22] V. Jain and P. Heydari, "Radar Fundamentals," in Automotive Radar Sensors in Silicon Technologies, V. Jain and P. Heydari Eds. New York, NY: Springer New York, 2013, pp. 5-11.

[23] M. Parker, "Chapter 16 - Radar Basics," in Digital Signal Processing 101, M. Parker Ed. Boston: Newnes, 2010, pp. 191-200

[24] N. Bhatta and G. P. M, "RADAR and its applications," vol. 10, pp. 1-9, 01/01 2017.

[25] Nanzer, J.A. A Review of Microwave Wireless Techniques for Human Presence Detection and Classification. IEEE Trans. Microw. Theory Tech. 2017, 65, 1780–1794. [CrossRef]

[26] Lukin, K.; Konovalov, V. Through wall detection and recognition of human beings using noise radar sensors. In Proceedings of the NATO RTO SET Symposium on Target Identification and Recognition Using RF Systems, Oslo, Norway, 11–13 October 2004; pp. 15-1–15-11

[27] Susek, W.; Stec, B. Through-the-wall detection of human activities using a noise radar with microwave quadrature correlator. IEEE Trans. Aerosp. Electron. Syst. 2015, 51, 759–764.

[28] J. Maitre, S. Bouchard, K. Gaboury, “Data filtering and deep learning for enhanced human activity recognition from UWB radars”, Journal of Ambient Intelligence and Humanized Computing, Vol. 14, No. 6, 2023, pp. 7845-7856.

[29] Z. Wang, A. Ren, Q. Zhang, A. Zahid, Q. H. Abbasi, “Recognition of Approximate Motions of Human Based on Micro-Doppler Features”, IEEE Sensors Journal, Vol. 23, No. 11, 2023, pp. 12388-12397.

[30] S. Hassan, X. Wang, S. Ishtiaq, N. Ullah, A. Mohammad, A. Noorwali, “Human Activity Classification Based on Dual Micro-Motion Signatures Using Interferometric Radar”, Remote Sensing, Vol. 15, No. 7, 2023, pp. 1–21.

[31] L. Jiang, M. Wu, L. Che, X. Xu, Y. Mu, Y. Wu, “Continuous Human Motion Recognition Based on FMCW Radar and Transformer”, Journal of Sensors, Vol. 2023, 2023

[32] Z. Slimane and D. Korti, "Advanced Human Activity Recognition through Data Augmentation and Feature Concatenation of Micro-Doppler Signatures," International journal of electrical and computer engineering systems, vol. 14, pp. 893-902, 10/24 2023, doi: 10.32985/ijeces.14.8.7.

[33] S. K. Yadav, K. Tiwari, H. M. Pandey, and S. A. Akbar, "A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and

[34] Jian Kang Wu, Liang Dong and Wendong Xiao, "Real-time Physical Activity classification and tracking using wearble sensors," 2007 6th International Conference on Information, Communications & Signal Processing, Singapore, 2007, pp. 1-6, doi: 10.1109/ICICS.2007.4449890.-

[35] O.Chin, "Human Activity Recognition: A Review," In Proc. IEEE International Conference on Control System, Computing and Engineering, 28 - 30 November 2014, pp.389-393.

[36] Ke, Y.; Sukthankar, R.; Hebert, M. Spatio-temporal Shape and Flow Correlation for Action Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8

[37] Lu, W.; Little, J.J. Simultaneous tracking and action recognition using the PCA-HOG descriptor. In Proceedings of the 3rd Canadian Conference on Computer and Robot Vision, Quebec, PQ, Canada, 7–9 June 2006; p. 6.

[38] Luo, Y.; Wu, T.; Hwang, J. Object-based analysis and interpretation of human motion in sports video sequences by dynamic Bayesian networks. Comput. Vis. Image Underst. 2003, 92, 196–216.

[39] Blank, M.; Gorelick, L.; Shechtman, E.; Irani, M.; Basri, R. Actions as Space-time Shapes. In Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV), Beijing, China, 17–21 October 2005; Volume 2, pp. 1395–1402.

[40] Shechtman, E.; Irani, M. Space-time Behavior Based Correlation. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 20–26 June 2005; Volume 1, pp. 405–412.

[41] Bodor, R.; Jackson, B.; Papanikolopoulos, N. Vision-based Human Tracking and Activity Recognition. In Proceedings of the 11th Mediterranean Conference on Control and Automation, Rhodes, Greece, 18–20 June 2003; Volume 1, pp. 18–20

[42] Huo, F.; Hendriks, E.; Paclik, P.; Oomes, A.H.J. Markerless Human Motion Capture and Pose Recognition. In Proceedings of the 10th IEEE Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), London, UK, 6–8 May 2009; pp. 13–16.

[43] Wren, C.R.; Azarbayejani, A.; Darrell, T.; Pentland, A.P. Pfinder: Real-time tracking of the human body. IEEE Trans. Pattern Anal. Mach. Intell. 1997, 19, 780–785.)

[44] https://www.techtarget.com/searchmobilecomputing/definition/wearable-technology

[45] T. W., « History of Ultra-Wideband (UWB) radar & communications: pioneers and inventors », Progress in Electromagnetics Symposium 2000, Cambridge, MA, July 2000

[46] C. Fowler, J. Entzminger, et J. O. Corum -« Assessment of ultra-wideband (UWB) technology » - IEEE Aerospace and Electronic Systems Magazine, vol. 5, n° 11, pages 45-49, novembre 1990.

[47] Taylor, James D. t al « Introduction to Ultra-Wideband radar system », James D. Taylor Editor, CRC Press, Boca Raton, 1995, 670 p.

[48] F. Elbahhar, A. Rivenq-Menhaj, J.M. Rouvaen and M. Heddebaut, « lnter-Vehicle communication based on Ultra-Wide Band techniques », Proc. Telecomm 2001, Casablanca, Maroc, pp. 58-61, March 2001.

[49] Harmuth, H.F., « Transmission of Information by Orthogonal Functions », First Edition, Springer, NewYork, 1969.

[50] Ross, G.F., «A Time Domain criterion for the design of wideband radiating elements ». IEEE Trans. On Antennas & Propagation, 16, 355, 1968.

[51] Robbins, K. w. « short baseband pulse receiver. » U.S. Patent 3,662,316 dated May 9, 1972.

[52] Van Etten, P., « The present technology of impulse radars. », Int. Radar Conf. Proc., Oct, 535- 539, 1977.

[53] Ross et Robbins (R & R) de brevets, 1972-1987.

[54] Ross' US Patent 3728632 du 17 avril1973.

[55] R. J. Fontana et al., « An Ultra-Wideband Communications Link for Unmanned Vehicle Applications » Proc. Assn. for Unmanned Vehic. Syst. 1997 International Conference, Baltimore, 3-6 June 1997.

[56] FCC- «Revision of the Rules Regarding Ultra-Wideband Transmission Systems, Notice of Inquiry »-Rapport, Federal Communications Commission, 1998.

[57] Federal Communication Commission. «Notice of proposed rulemaking», FCC 00-163, Et Docket 98-153, In the matter of revision of part 15 of the commission's rules regarding Ultra-wideband transmission system, Washington, 11 may 2000.

[58] FCC - « First report and order, revision of Part 15 of the Commission's rules regarding ultra-wideband transmission systems»- Document technique Et Docket 98-153, FCC, Avril2002.

[59] Manzi, Giuliano, et al. "Coexistence between ultra-wideband radio and narrow-band wireless LAN communication systems—Part I: modeling and measurement of UWB radio signals in frequency and time." *IEEE transactions on electromagnetic compatibility* 51.2 (2009): 372-381.

[60] F.Elbahhar. « Conception d'un system de communication ultra large band appliqué aux transports » thèse de doctorat de l'université de Valenciennes et du Hainaut Cambrésis, November 2003.

[61] Lamari A. « Conception et modélisation d'un système de communication MultiUtilisateurs basé sur la technique Ultra Large Bande » thèse de doctorat de l'université de Valenciennes et du Hainaut Cambrésis, janvier 2007.

[62] L. Pecastaing, "Conception et réalistaion d'un système de génération d'impulsions haute tension ultra brèves, application aux radars large bande », thèse de doctorat de 1 'université de Pau et des pays de l'Adour, Décembre 2001.

[63] L. Sakkila « Détection du type de modulation et synchronisation », Mémoire de DEA Juillet 2005.

[64] L. Babour « Etude et conception d'antenne Ultra Large Bande miniaturisées en impulsionnel » thèse de doctorat de l'institut polytechnique de Grenoble, mai 2009.

[65] Li, C.; Peng, Z.; Huang, T.Y.; Fan, T.; Wang, F.K.; Horng, T.S.; Muñoz-Ferreras, J.M.; Gómez-García, R.; Ran, L.; Lin, J. A Review on Recent Progress of Portable Short-Range Noncontact Microwave Radar Systems. IEEE Trans. Microw. Theory Tech. 2017, 65, 1692–1706. [Google Scholar] [CrossRef]

[66] Peng, Z.; Li, C. Portable Microwave Radar Systems for Short-Range Localization and Life Tracking: A Review. Sensors 2019, 19, 1136. [Google Scholar] [CrossRef]

[67] Nanzer, J.A. A Review of Microwave Wireless Techniques for Human Presence Detection and Classification. IEEE Trans. Microw. Theory Tech. 2017, 65, 1780–1794. [Google Scholar] [CrossRef]

[68] Lukin, K.; Konovalov, V. Through wall detection and recognition of human beings using noise radar sensors. In Proceedings of the NATO RTO SET Symposium on Target Identification and Recognition Using RF Systems, Oslo, Norway, 11–13 October 2004; pp. 15-1–15-11. [Google Scholar]

[69] Lai, C.P.; Ruan, Q.; Narayanan, R.M. Hilbert-Huang transform (HHT) analysis of human activities using through-wall noise radar. In Proceedings of the International Symposium on Signals, Systems and Electronics, Montreal, QC, Canada, 30 July–2 August 2007. [Google Scholar]

[70] Narayanan, R.M. Through-wall radar imaging using UWB noise waveforms. J. Frankl. Inst. 2008, 345, 659–678. [Google Scholar] [CrossRef]

[71] Lai, C.P.; Narayanan, R.; Ruan, Q.; Davydov, A. Hilbert–Huang transform analysis of human activities using through-wall noise and noise-like radar. IET Radar Sonar Navig. 2008, 2, 244–255. [Google Scholar] [CrossRef]

[72] Lai, C.P.; Narayanan, R.M. Ultrawideband random noise radar design for through-wall surveillance. IEEE Trans. Aerosp. Electron. Syst. 2010, 46, 1716–1730. [Google Scholar] [CrossRef]

[73] Susek, W.; Stec, B. Through-the-wall detection of human activities using a noise radar with microwave quadrature correlator. IEEE Trans. Aerosp. Electron. Syst. 2015, 51, 759–764. [Google Scholar] [CrossRef]

[74] Alabaster, C. M., & Hughes, E. J. (2012). *Arbitrary pulsed radar waveform*. https://doi.org/10.1109/wdd.2012.7311278

[75] Wei-Ming, Y. (2007). Study on the Design Method of a Novel Quasi-CW Radar Waveform. *Modern Radar*. https://en.cnki.com.cn/Article_en/CJFDTOTAL-XDLD200709006.htm

[76] Abousetta, M., & Cooper, D. (1998). Noise analysis of digitised FMCW radar waveforms. *IEE Proceedings. Radar, Sonar and Navigation*, *145*(4), 209. https://doi.org/10.1049/ip-rsn:19981310

[77] Le, H.T.; Phung, S.L.; Bouzerdoum, A. Human Gait Recognition with Micro-Doppler Radar and Deep Autoencoder. In Proceedings of the IEEE 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 3347–3352. [Google Scholar]

[78] Wang, M.; Zhang, Y.D.; Cui, G. Human motion recognition exploiting radar with stacked recurrent neural network. Digit. Signal Process. 2019, 87, 125–131. [Google Scholar] [CrossRef]

[79] Mercuri, M.; Liu, Y.; Lorato, I.; Torfs, T.; Wieringa, F.P.; Bourdoux, A.; Van Hoof, C. A Direct Phase-Tracking Doppler Radar Using Wavelet Independent Component Analysis for Non-Contact Respiratory and Heart Rate Monitoring. IEEE Trans. Biomed. Circuits Syst. 2018, 12, 632–643. [Google Scholar] [CrossRef] [PubMed]

[80] Rahman, A.; Lubecke, V.; Boriclubecke, O.; Prins, J.; Sakamoto, T. Doppler Radar Techniques for Accurate Respiration Characterization and Subject Identification. IEEE J. Emerg. Sel. Top. Circuits Syst. 2018, 8, 350–359. [Google Scholar] [CrossRef]

[81] D.J. DANIELS "SURFACE PENETRATING RADAR" Livre édité par IEE, 1996.

[82] J.M. FERRIER « Présentation Thalès sur les systèmes DEMISTER et CLEO » Réunion DGA, Décembre 2005.

[83] 2009 John Wiley & Sons, Inc. WIREs Comp Stat 2009 1 283–289

[84] https://www.journaldunet.fr/intelligence-artificielle/guide-de-l-intelligence-artificielle/1501905-random-forest-ou-foret-aleatoire/

[85] https://asana.com/fr/resources/decision-tree-analysis

[86] https://towardsdatascience.com/what-when-how-extratrees-classifier-c939f905851c

[87] https://aws.amazon.com/fr/what-is/deep-learning/

[88] Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., & Kumar, P. (2021). Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Molecular Diversity*, *25*(3), 1315–1360. https://doi.org/10.1007/s11030-021-10217-3

[89] He, Y.; Le Chevalier, F.; Yarovoy, A.G. Range-Doppler processing for indoor human tracking by multistatic ultra-wideband radar. In Proceedings of the 13th International Radar Symposium (IRS), Warsaw, Poland, 23–25 May 2012; pp. 250–253. [Google Scholar]

[90] Ji, S.; Xu, W.; Yang, M.; Yu, K. 3D convolutional neural networks for human action recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2013, 35, 221–231. [Google Scholar] [CrossRef] [PubMed]

[91] Sang, Y.; Shi, L.; Liu, Y. Micro Hand Gesture Recognition System Using Ultrasonic Active Sensing. IEEE Access 2017, 6, 49339–49347. [Google Scholar] [CrossRef]

[92] Zhang, Z.; Tian, Z.; Zhou, M. Latern: Dynamic Continuous Hand Gesture Recognition Using FMCW Radar Sensor. IEEE Sens. J. 2018, 18, 3278–3289. [Google Scholar] [CrossRef]

[93] Wang, S.; Song, J.; Lien, J.; Poupyrev, I.; Hilliges, O. Interacting with Soli: Exploring Fine-Grained Dynamic Gesture Recognition in the Radio-Frequency Spectrum. In Proceedings of the ACM Symposium on User Interface Software and Technology, Tokyo, Japan, 16–19 October 2016; pp. 851–860. [Google Scholar]

[94] Lien, J.; Gillian, N.; Karagozler, M.E.; Amihood, P.; Schwesig, C.; Olson, E.; Raja, H.; Poupyrev, I. Soli: Ubiquitous gesture sensing with millimeter wave radar. ACM Trans. Graph. 2016, 35, 142. [Google Scholar] [CrossRef]

[95] Tahmoush, D. Review of micro-Doppler signatures. IET Radar Sonar Navig. 2015, 9, 1140–1146. [Google Scholar] [CrossRef]

[96] Chen, V.C.; Qian, S. Joint time-frequency transform for radar range-Doppler imaging. IEEE Trans. Aerosp. Electron. Syst. 1998, 34, 486–499. [Google Scholar] [CrossRef]

[97] Kim, Y.; Moon, T. Human Detection and Activity Classification Based on Micro-Doppler Signatures Using Deep Convolutional Neural Networks. IEEE Geosci. Remote Sens. Lett. 2016, 13, 8–12. [Google Scholar] [CrossRef]

[98] Kim, Y.; Toomajian, B. Hand Gesture Recognition Using Micro-Doppler Signatures with Convolutional Neural Network. IEEE Access 2016, 4, 7125–7130. [Google Scholar] [CrossRef]

[99] Kim, Y.; Toomajian, B. Application of Doppler radar for the recognition of hand gestures using optimized deep convolutional neural networks. In Proceedings of the European Conference on Antennas and Propagation, Paris, France, 19–24 March 2017; pp. 1258–1260. [Google Scholar]

[100] Trommel, R.P.; Harmanny, R.I.A.; Cifola, L.; Driessen, J.N. Multi-target human gait classification using deep convolutional neural networks on micro-doppler spectrograms. In Proceedings of the European Radar Conference, London, UK, 5–7 October 2016; pp. 81–84. [Google Scholar]

[101] Shao, Y.; Dai, Y.; Yuan, L.; Chen, W. Deep Learning Methods for Personnel Recognition based on Micro-Doppler Features. In Proceedings of the 9th International Conference on Signal Processing Systems, AUT, Auckland, New Zealand, 27–30 November 2017; pp. 94–98. [Google Scholar]

[102] Zhang, J.; Tao, J.; Shi, Z. Doppler-Radar Based Hand Gesture Recognition System Using Convolutional Neural Networks. In Proceedings of the IEEE International Conference in Communications, Signal Processing, and Systems, Harbin, China, 14–16 July 2017; pp. 1096–1113. [Google Scholar]

[103] Seyfioğlu, M.S.; Özbayğglu, A.M.; Gurbuz, S.Z. Deep Convolutional Autoencoder for Radar-Based Classification of Similar Aided and Unaided Human Activities. IEEE Trans. Aerosp. Electron. Syst. 2018, 54, 1709–1723. [Google Scholar] [CrossRef]

[104] Le, H.T.; Phung, S.L.; Bouzerdoum, A.; Tivive, F.H.C. Human Motion Classification with Micro-Doppler Radar and Bayesian-Optimized Convolutional Neural Networks. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 2961–2965. [Google Scholar]

[105] Molchanov, P.; Gupta, S.; Kim, K.; Pulli, K. Short-range FMCW monopulse radar for hand-gesture sensing. In Proceedings of the IEEE Radar Conference, Arlington, VA, USA, 10–15 May 2015; pp. 1491–1496. [Google Scholar]

[106] Jokanovic, B.; Amin, M.; Erol, B. Multiple joint-variable domains recognition of human motion. In Proceedings of the IEEE Radar Conference, Seattle, WA, USA, 8–12 May 2017; pp. 0948–0952. [Google Scholar]

[107] Erol, B.; Amin, M.G. Fall motion detection using combined range and Doppler features. In Proceedings of the 24th European Signal Processing Conference (EUSIPCO), Budapest, Hungary, 29 August–2 September 2016; pp. 2075–2080. [Google Scholar]

[108] Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436. [Google Scholar] [CrossRef]

[109] Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the National Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; pp. 4278–4284. [Google Scholar]

[110] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [Google Scholar]

[111] Schmidhuber, J. Deep learning in neural networks: An overview. Neural Netw. 2015, 61, 85–117. [Google Scholar] [CrossRef] [PubMed] [Green Version]

[112] Kang, M.; Ji, K.; Leng, X.; Lin, Z. Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection. Remote Sens. 2017, 9, 860. [Google Scholar] [CrossRef]

[113] Lin, Z.; Ji, K.; Leng, X.; Kuang, G. Squeeze and Excitation Rank Faster R-CNN for Ship Detection in SAR Images. IEEE Geosci. Remote Sens. Lett. 2019, 16, 751–755. [Google Scholar] [CrossRef]

[114] Wang, J.; Chen, Y.; Hao, S.; Peng, X.; Hu, L. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognit. Lett.* 2018, *119*, 3–11. [Google Scholar] [CrossRef]

[115] Chen, Z.; Li, G.; Fioranelli, F.; Griffiths, H. Personnel Recognition and Gait Classification Based on Multistatic Micro-Doppler Signatures Using Deep Convolutional Neural Networks. IEEE Geosci. Remote Sens. Lett. 2018, 15, 669–673. [Google Scholar] [CrossRef] [Green Version]

[116] Lin, Z.; Ji, K.; Kang, M.; Leng, X.; Zou, H. Deep convolutional highway unit network for sar target classification with limited labeled training data. IEEE Geosci. Remote Sens. Lett. 2017, 14, 1091–1095. [Google Scholar] [CrossRef]

[117] Graves, A. Long Short-Term Memory. Neural Comput. 1997, 9, 1735–1780. [Google Scholar]

[118] Klarenbeek, G.; Harmanny, R.I.A.; Cifola, L. Multi-target human gait classification using LSTM recurrent neural networks applied to micro-Doppler. In Proceedings of the European Radar Conference, Nuremberg, Germany, 11–13 October 2017; pp. 167–170. [Google Scholar]

[119] S. Jung, J. Park, and S. Lee, “Polyphonic Sound Event Detection Using Convolutional Bidirectional Lstm and Synthetic Data-based Transfer Learning,” in ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, May 2019, pp. 885–889. doi: 10.1109/ICASSP.2019.8682909.

[120] W. Zhao, Y. Zhou, and W. Tang, “Novel convolution and LSTM model for forecasting PM2.5 concentration,” International Journal of Performability Engineering, vol. 15, no. 6, pp. 1528–1537, 2019, doi: 10.23940/ijpe.19.06. p4.15281537.

[121] C. Guo, W. Guo, C.-H. Chen, X. Wang, and G. Liu, The Air Quality Prediction Based on a Convolutional LSTM Network, vol. 11817 LNCS. 2019. doi: 10.1007/978-3-030-30952-7_12.

[122] J. Liu, T. Zhang, Y. Gou, X. Wang, B. Li, and W. Guan, “Convolutional LSTM networks for seawater temperature prediction,” in 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), IEEE, Dec. 2019, pp. 1–5. doi: 10.1109/ICSIDP47821.2019.9173301.

[123] Calzada, “Democratising Smart Cities? Penta-Helix Multistakeholder Social Innovation Framework,” Smart Cities, vol. 3, no. 4, pp. 1145–1172, 2020, doi: 10.3390/smartcities3040057.

[124] M. Yang and M. C. Chen, “Composite Neural Network: Theory and Application to PM2.5 Prediction,” IEEE Trans Knowl Data Eng, 2021, doi: 10.1109/TKDE.2021.3099135.

[125] Zhang, H., & Fu, H. (2020). "Multimodal Fusion for Human Activity Recognition Based on UWB Radar and Wearable Inertial Sensors." Sensors, 20(11), 3060.

[126] Li, J., Xu, W., & Stoica, P. (2017). "MIMO Radar Waveform Design with Sparse Bayesian Learning for Human Activity Classification." IEEE Transactions on Signale Processing, 65(15), 3953-3967.

[127] Aubry, A., Grelier, M., & Pastina, D. (2018). UWB Radar-Based Human Detection and Identification: Fusion of Doppler Signatures and Range Profiles. IEEE Transactions on Aerospace and Electronic Systems, 54(2), 947-961.

[128] Skolnik, M. I. (2008). Introduction to Radar Systems (3rd ed.). New York : McGraw-Hill.

[129] Wang, P., Braunstein, M. L., & Cetin, M. (2014). A Review of Ultra-Wideband Radar Techniques for Concealed Weapon Detection. Signal Processing, 97, 282-298.

[130] Mahalanobis, C., Muise, R. R., & Brennan, R. W. (1997). Multisensor Data Fusion. Proceedings of the IEEE, 85(1), 6-23.

[131] Rahman, M., Gurbuz, S. (2021): Multi-Frequency RF Sensor Data Adaptation for Motion

[132] Recognition with Multi-Modal Deep Learning. Paper presented at 2021 IEEE Radar Conference: Radar on the Move.

[133] Harasawa et al. "An Adaptive Moving Target Indicator Using Median Filters". Electronics and Communications in Japan, Part 1. 1998. vol. 81, No. 5. pp. 41-50.

[134] https://www.gaussianwaves.com/author/mathuranathan/

[135] https://sciencenotes.org/doppler-effect-definition-formula-and-examples/

[136] S. Z. Gurbuz and M. G. Amin, "Radar-Based Human-Motion Recognition with Deep Learning: Promising Applications for Indoor Monitoring," IEEE Signal Processing Magazine, vol. 36, no. 4, pp. 16-28, 2019, doi: 10.1109/MSP.2018.2890128.

[137] DOGARU, Traian. Doppler processing with ultra-wideband (UWB) impulse radar. Adelphi (MD): Army Research Laboratory (US), 2013.

[138] https://scikitlearn.org/stable/modules/generated/sklearn.model_selection.train_test_split.

[139] https://www.projectpro.io/article/ai-with-python/988#mcetoc_1hnt00j9t1h

[140] Using Normalization Layers to Improve Deep Learning Models - MachineLearningMastery.com

[141] Train-Test Split for Evaluating Machine Learning Algorithms - MachineLearningMastery.com

[142] https://encord.com/blog/train-val-test-split/

[143] How to Use Random Seeds Effectively | by Jai Bansal | Towards Data Science

[144] Random Seed TensorFlow - How to obtain stable results with a model (inside-machinelearning.com)

[145] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R. Springer Science & Business Media.

[146] https://www.geeksforgeeks.org/concatenate-images-using-opencv-in-python/

[147] S. Boulahia, A. Amamra, M. Madi, and S. Daikh, "Early, intermediate and late fusion strategies for robust deep learning-based multimodal action recognition," Machine Vision and Applications, vol. 32, 11/01 2021, doi: 10.1007/s00138-021-01249-8.

[148] https://www.geeksforgeeks.org/

[149] “Difference Between a Batch and an Epoch in a Neural Network - MachineLearningMastery.com.” [Online]. Available: https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/

[150] https://scikit-learn.org/stable/modules/cross_validation.html

[151] https://www.v7labs.com/

[152] Hand, D. J. (2012). Assessing the performance of classification methods. International Statistical Review, 80(3), 400-414.

[153] Ege, M. (2022). *Human activity classification with deep learning using FMCW radar* (Doctoral dissertation, Bilkent Universitesi (Turkey)).

APPENDICES

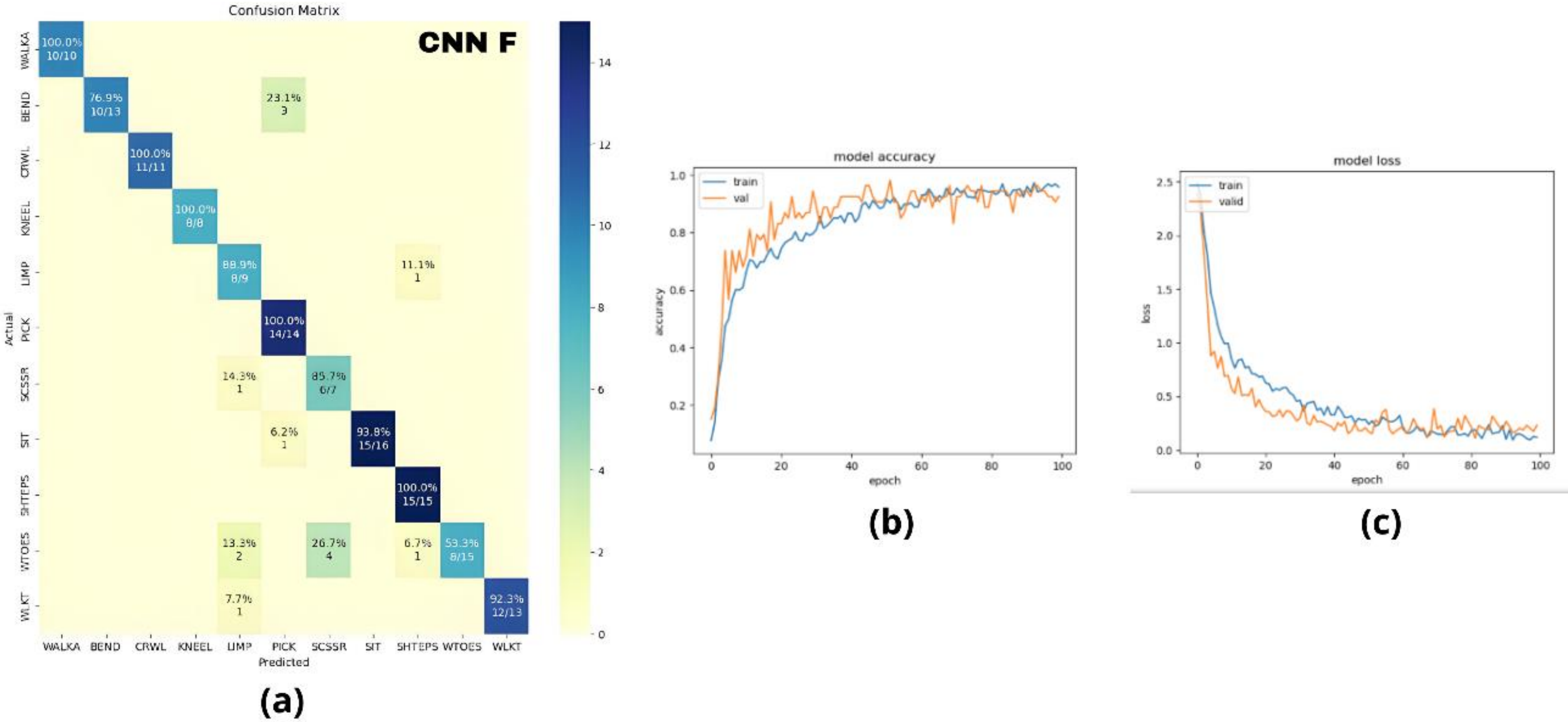


Figure III- 42: Fused data (CNN model): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

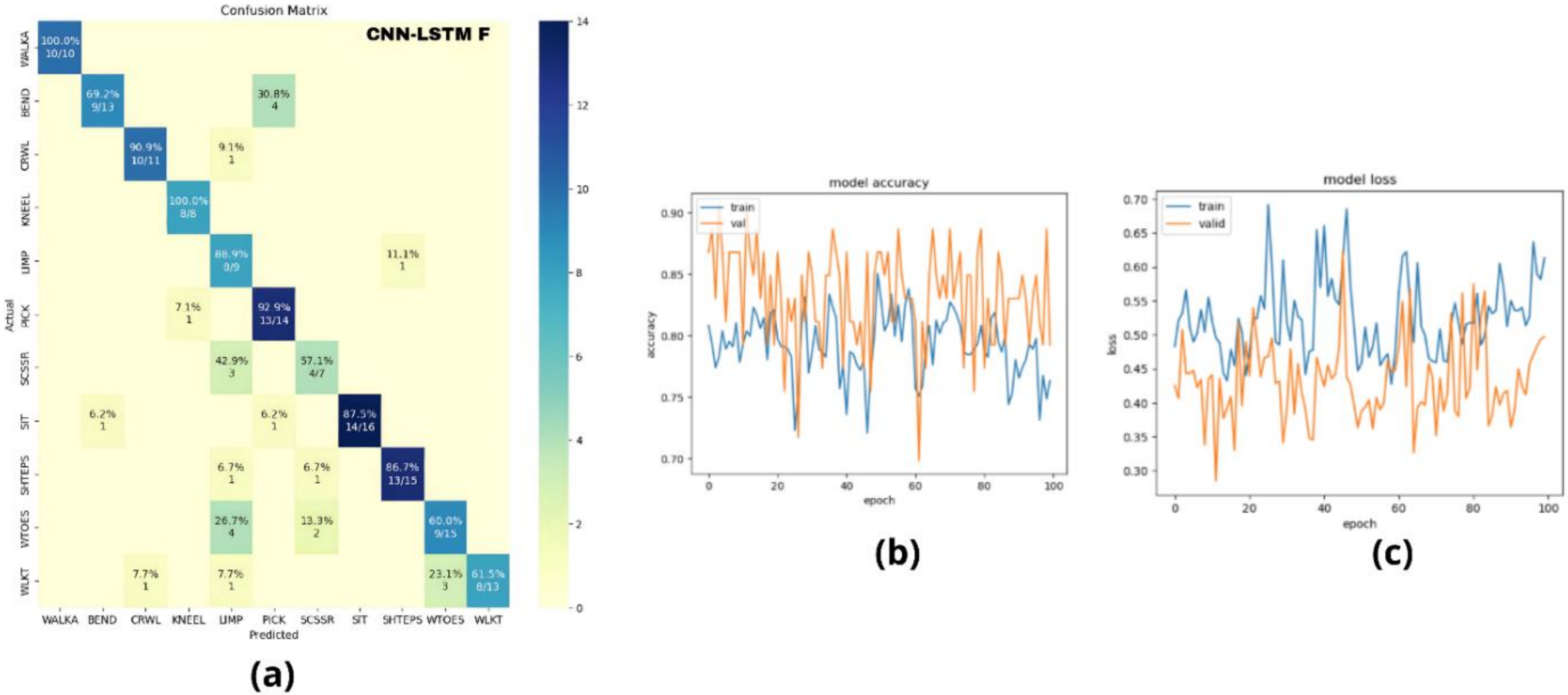


Figure III- 43: Fused data (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

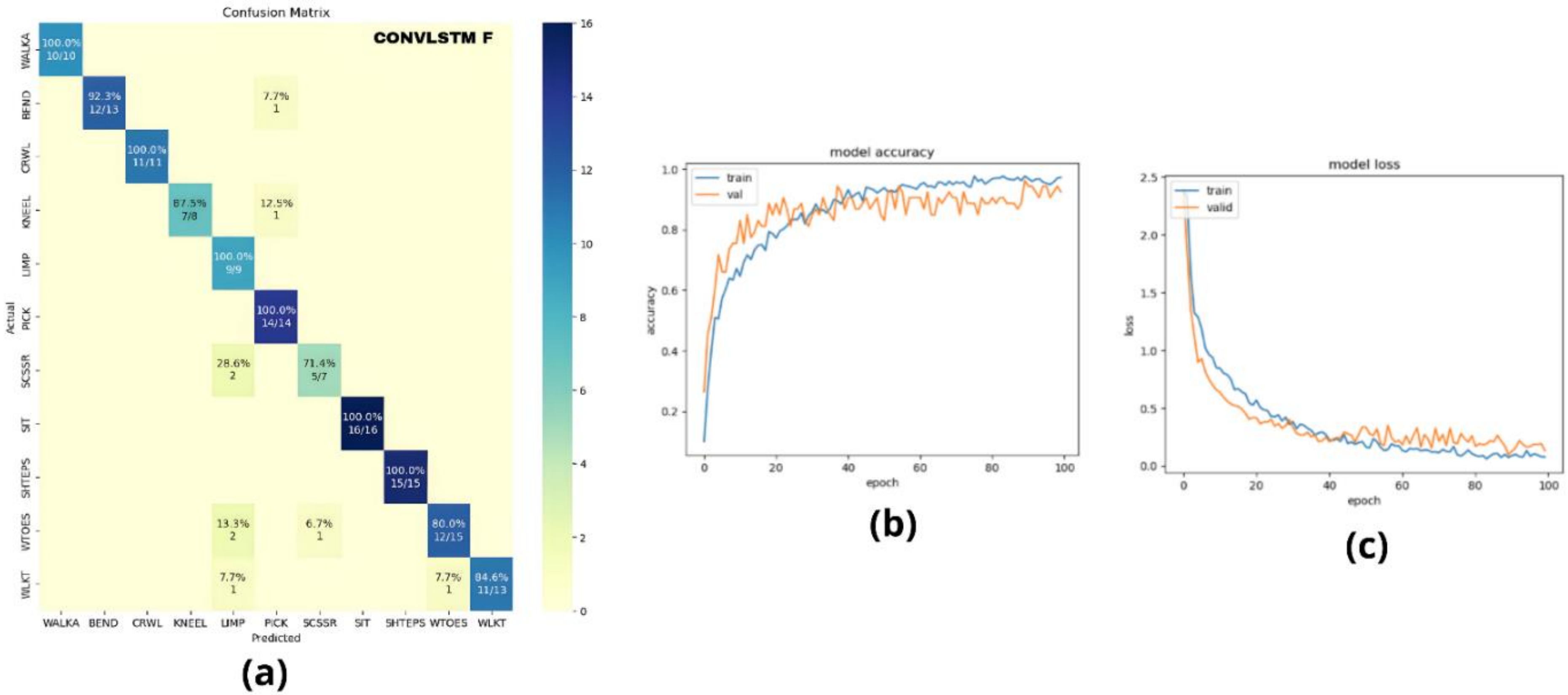


Figure III- 44: Fused data (ConvLSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

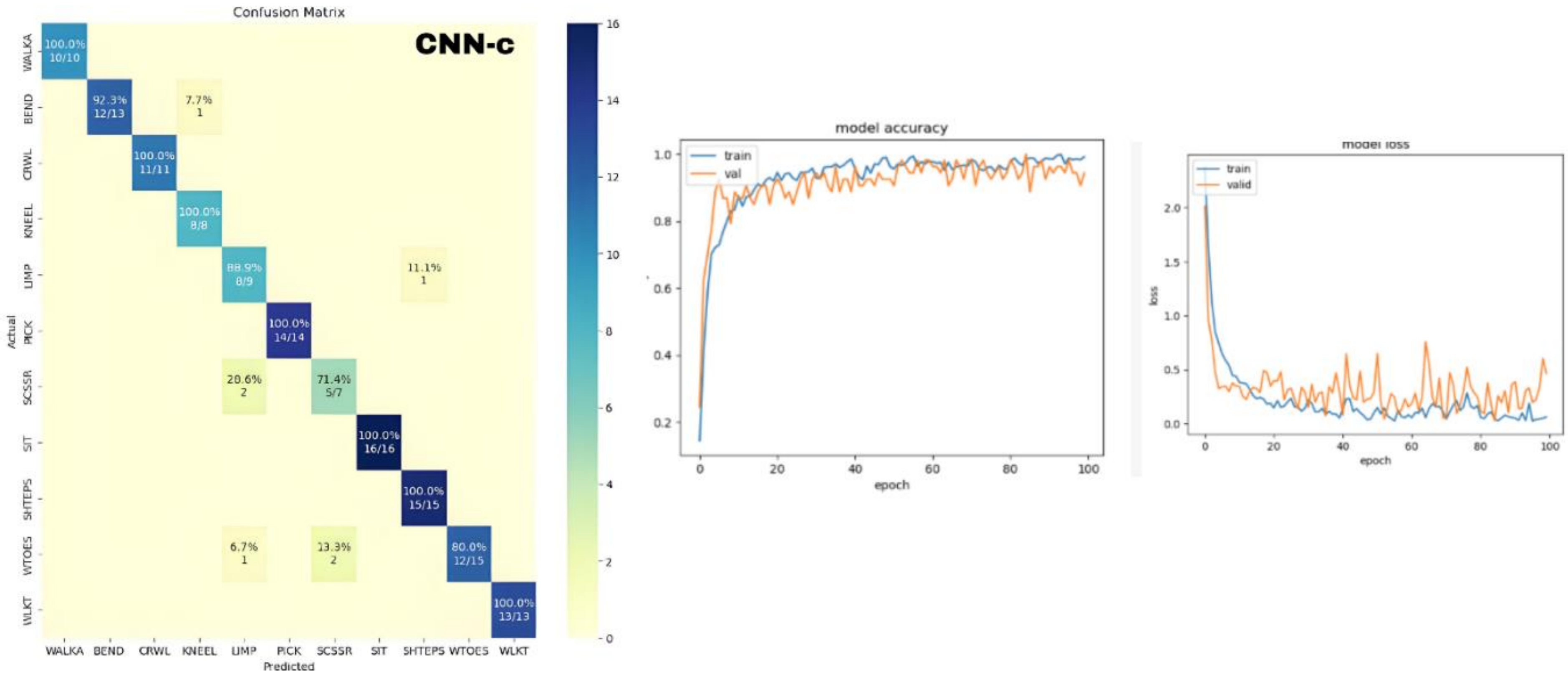


Figure III- 45: Concatenated data (CNN model): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

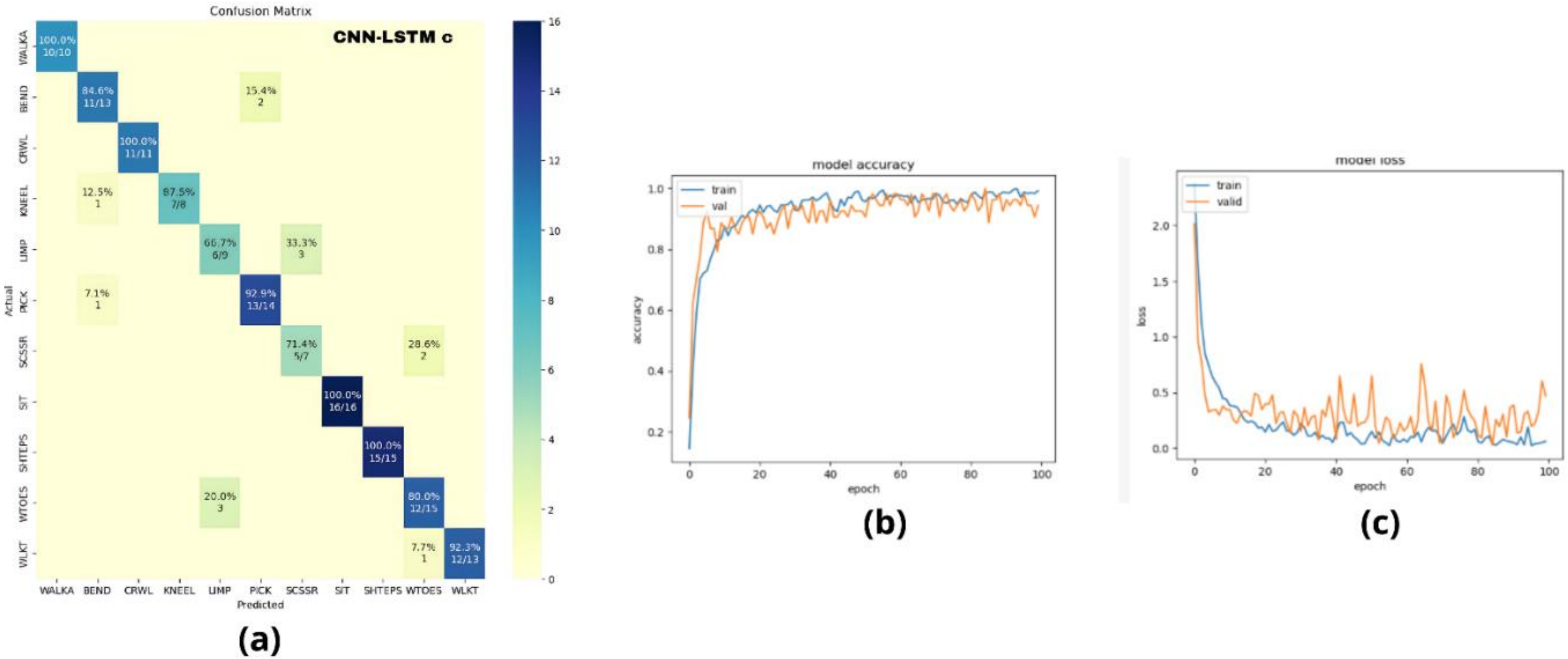


Figure III- 46: Concatenated data (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

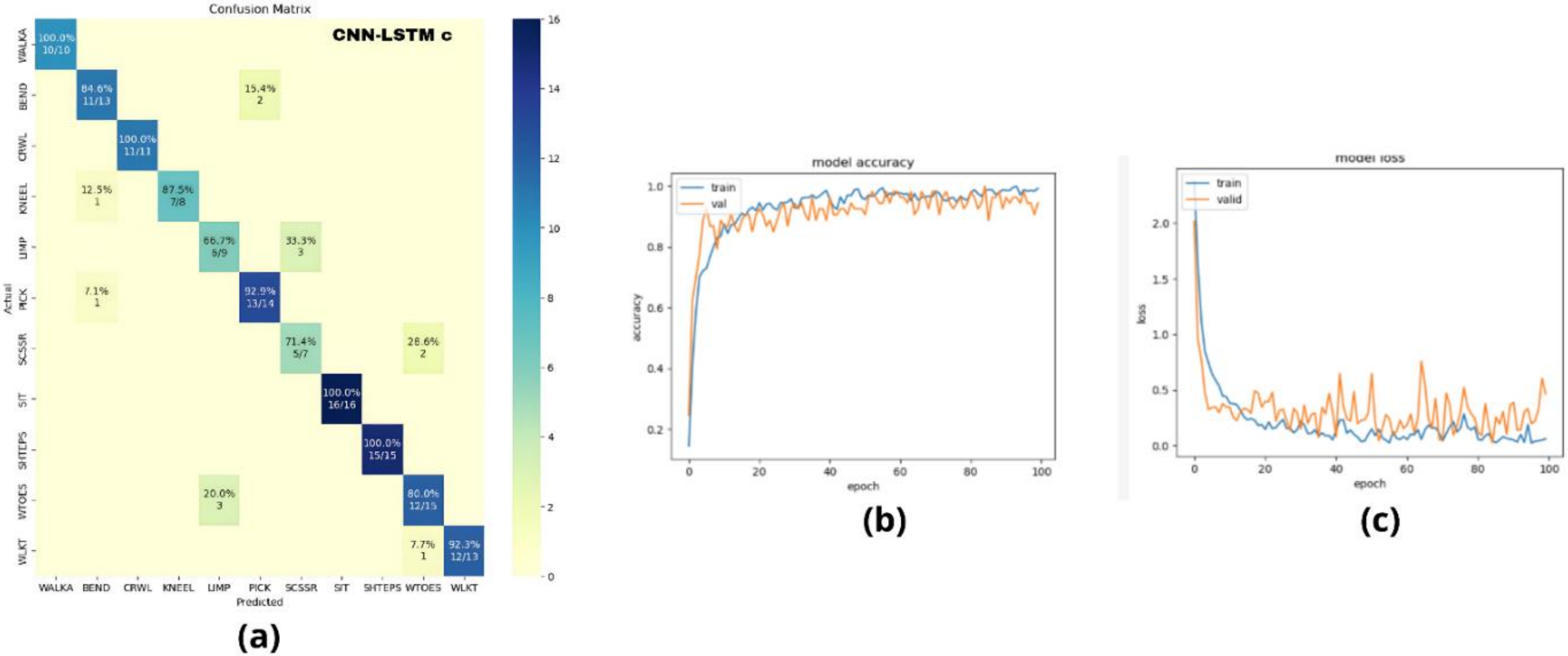


Figure III- 47: Concatenated data (ConvLSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

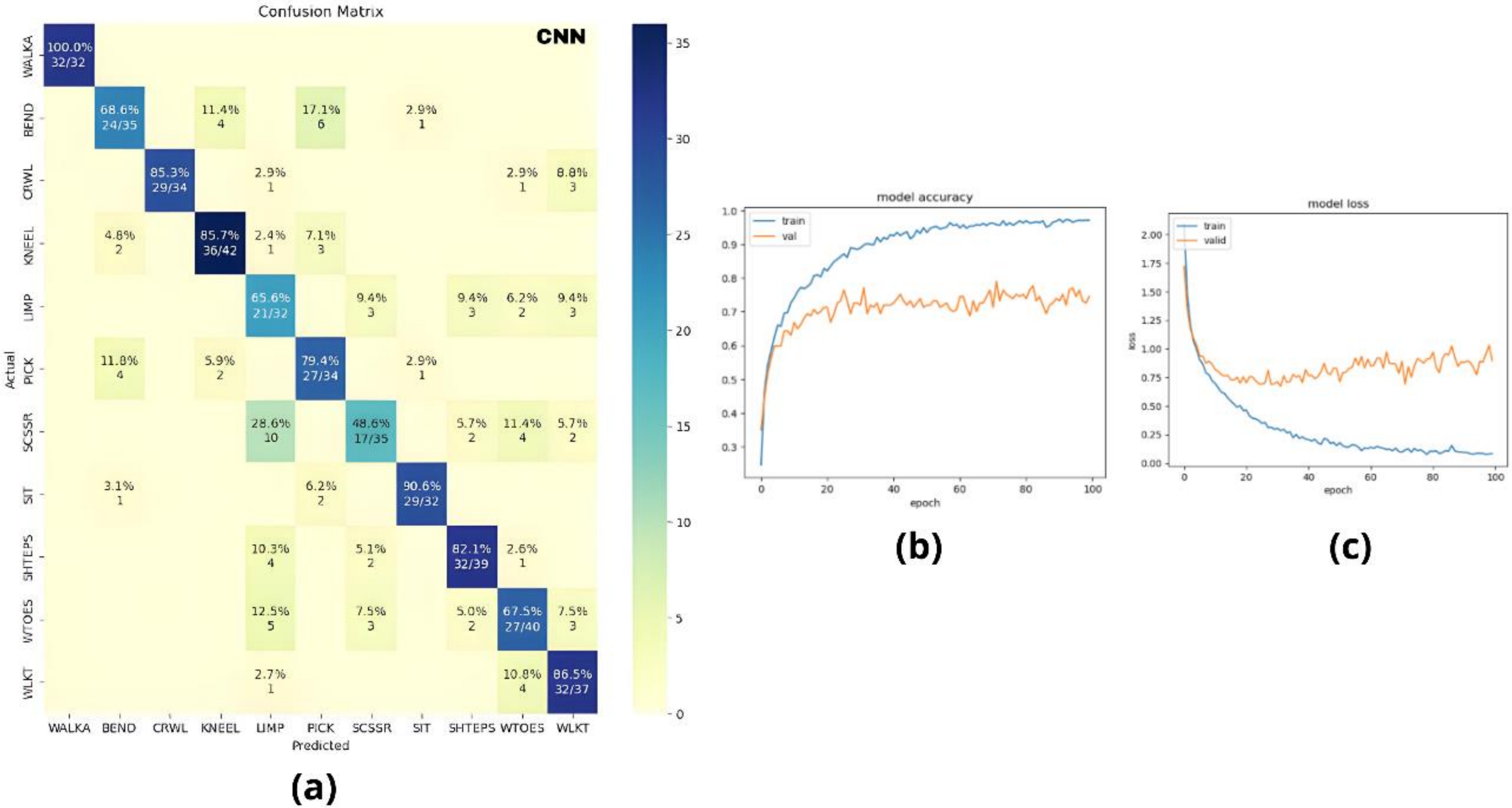


Figure III- 48: 3D Model (CNN): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

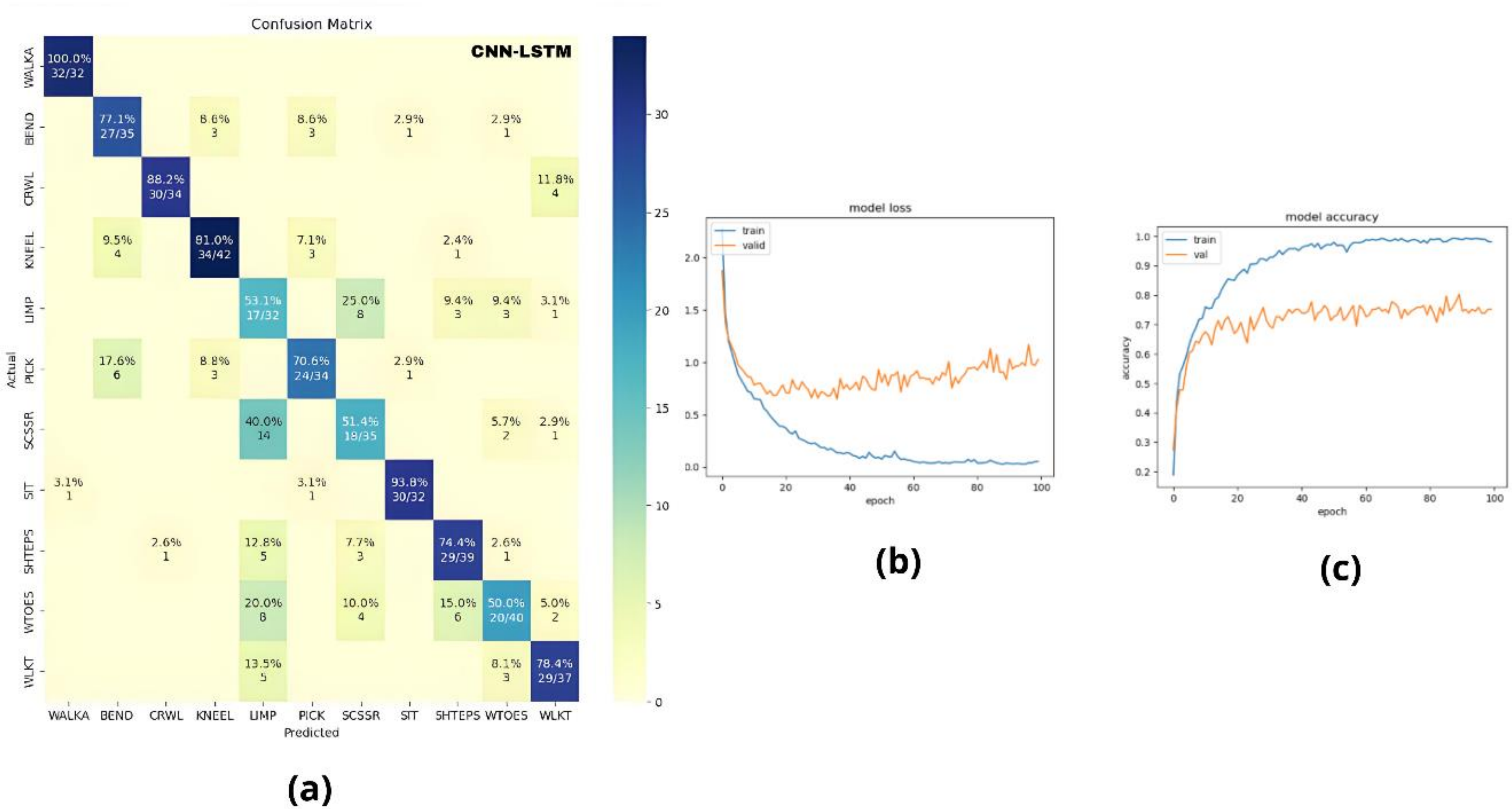


Figure III- 49: 3D Model (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

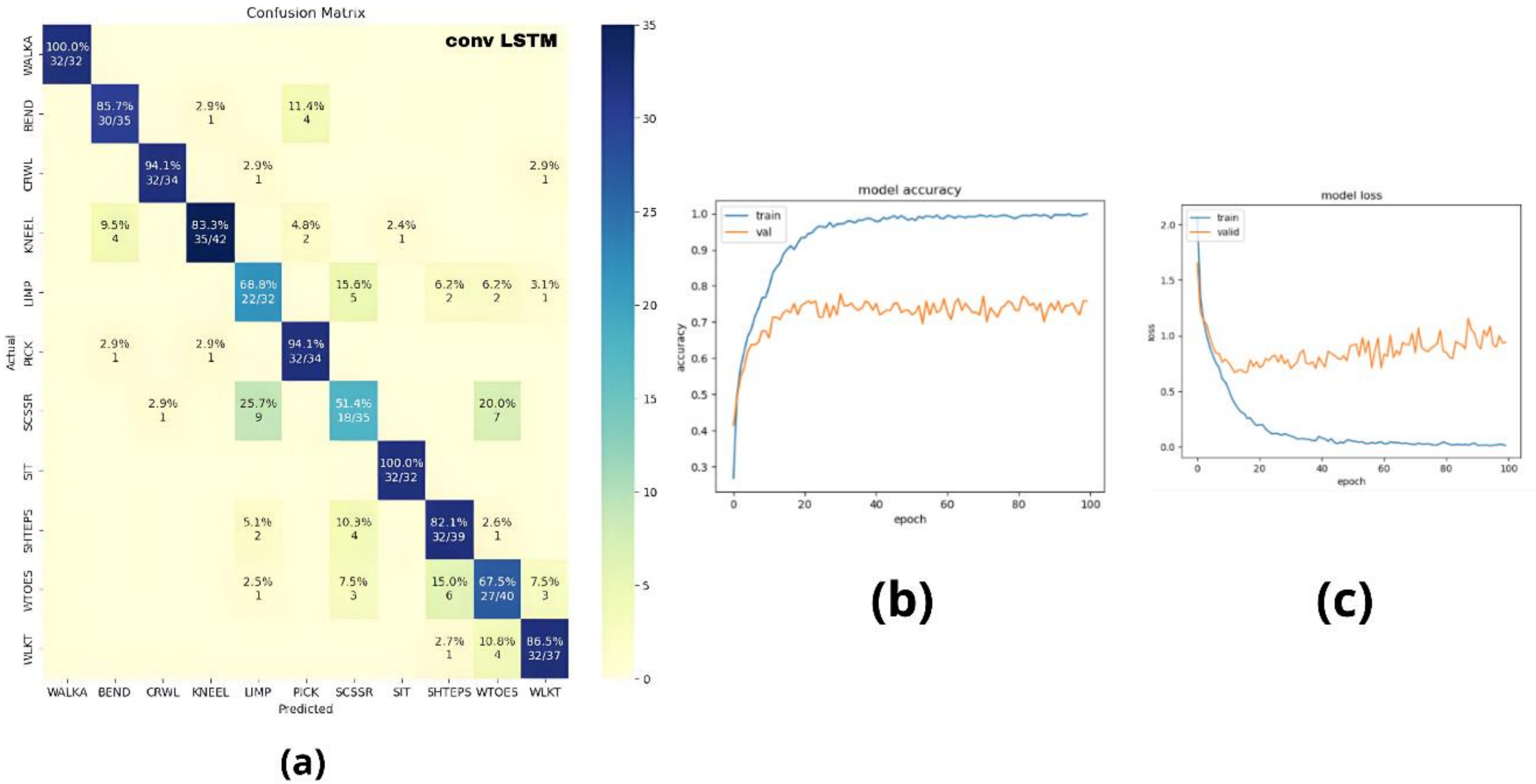


Figure III- 50: 3D Model (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

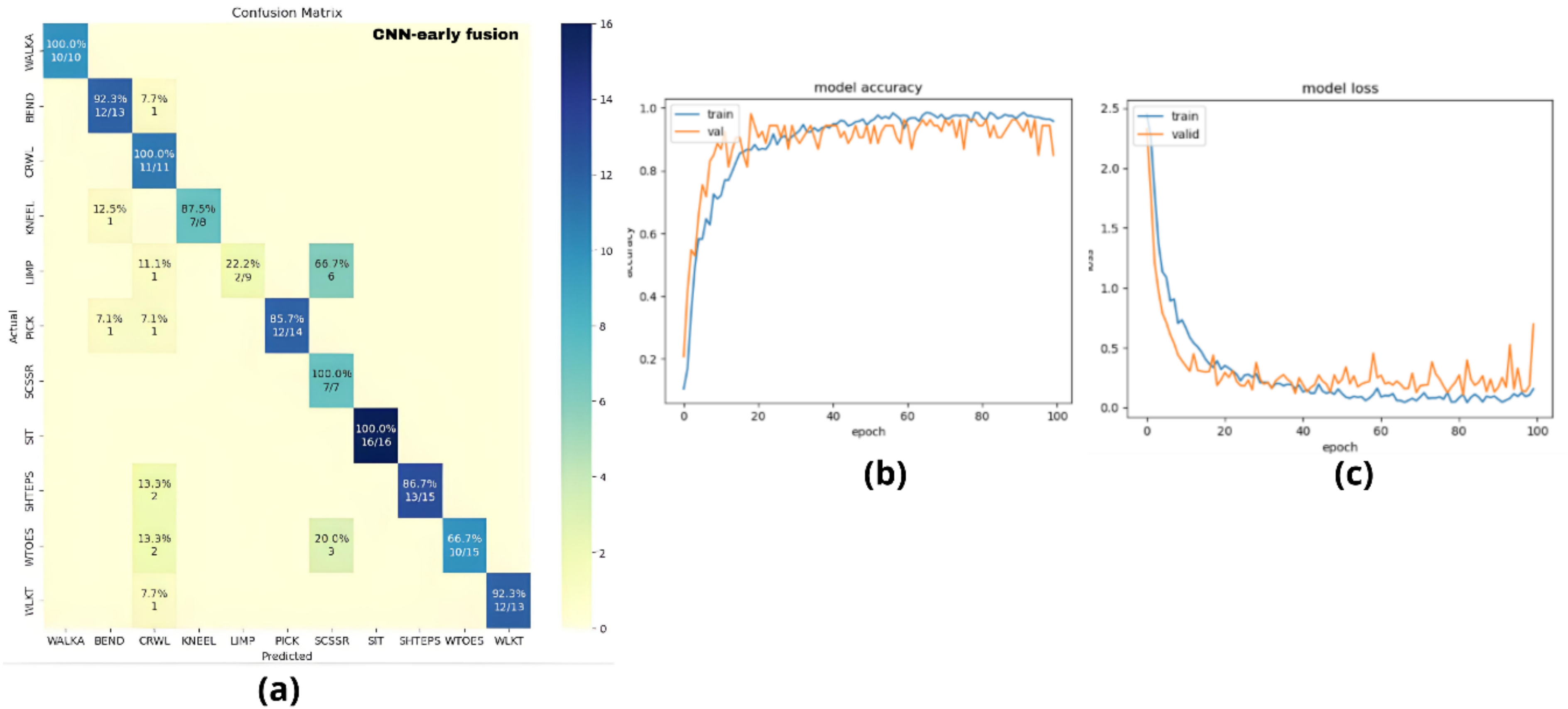


Figure III- 51: Early Fusion (CNN model): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

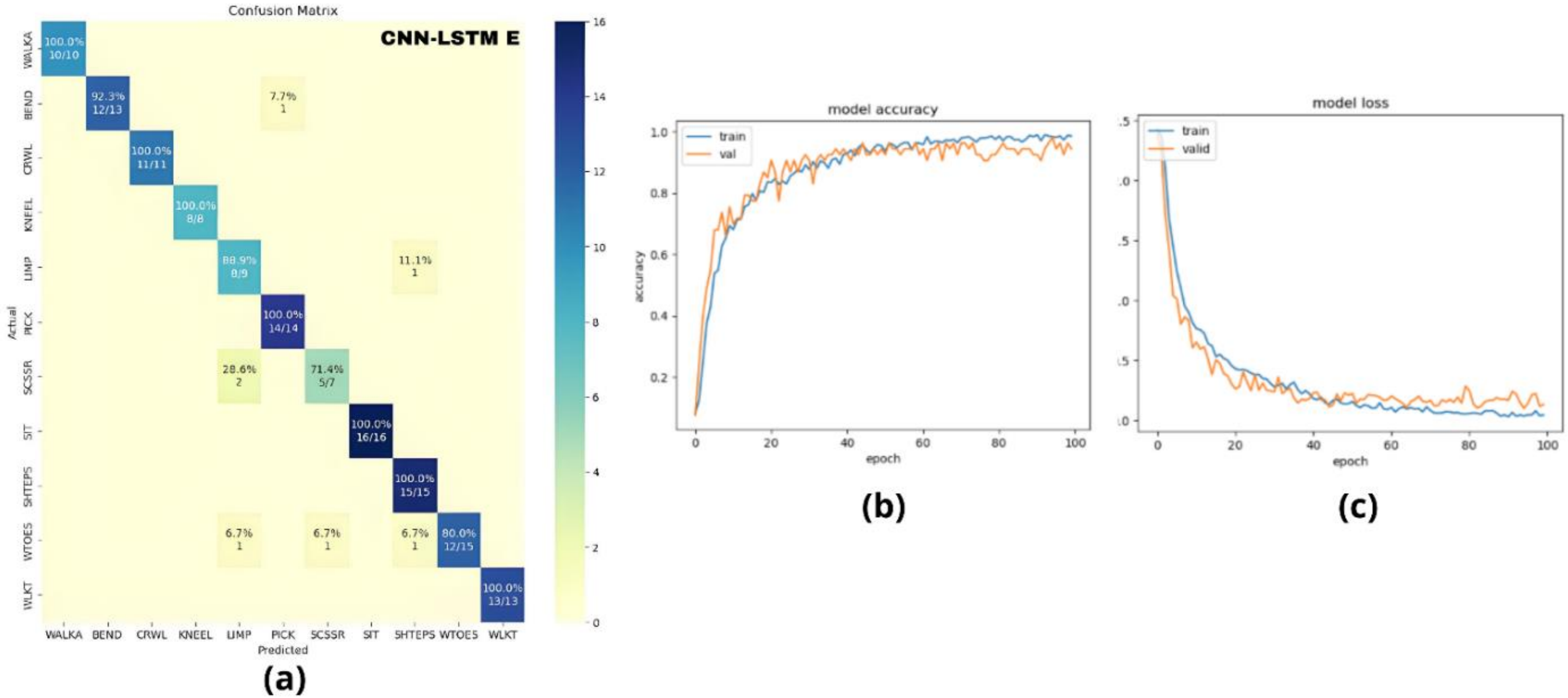


Figure III- 52: Early Fusion (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

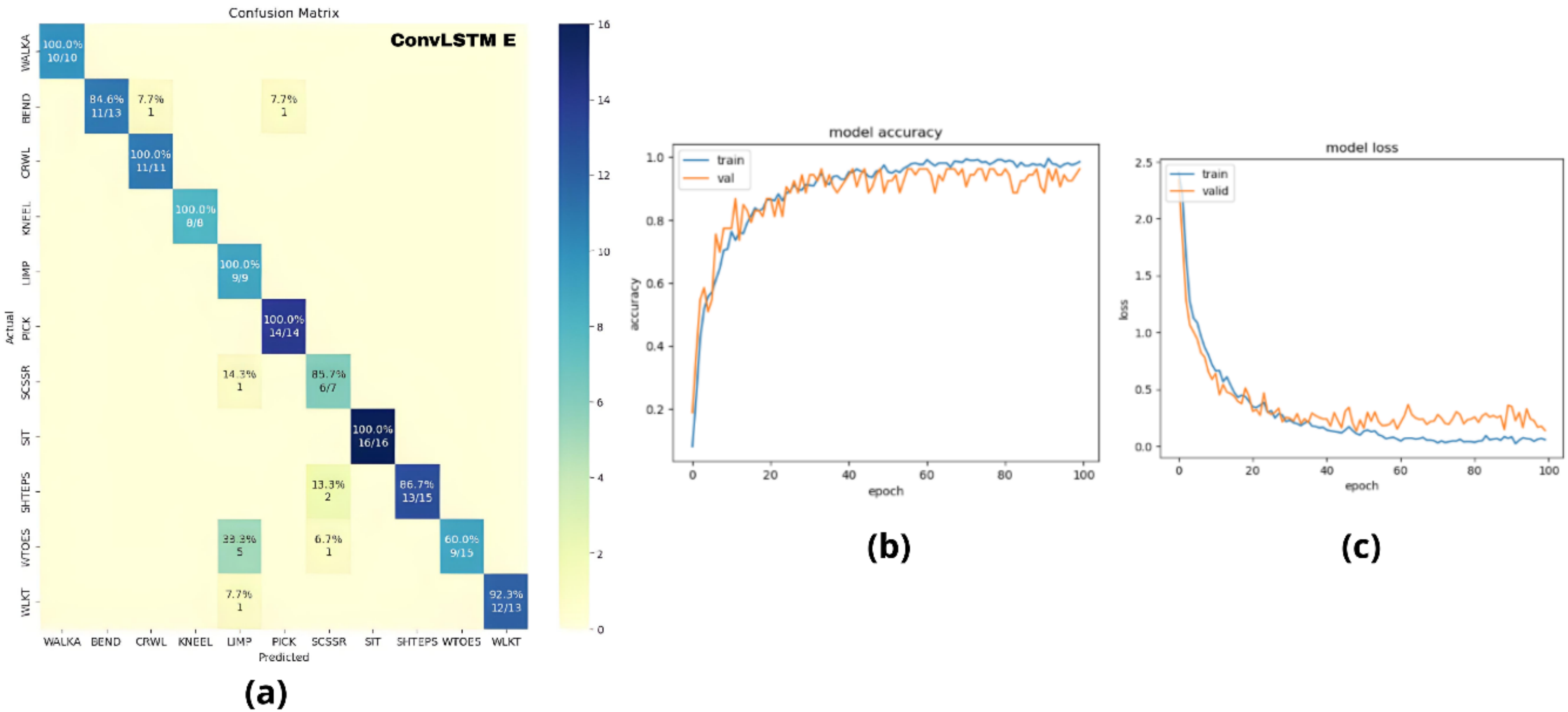


Figure III- 53: Early Fusion (ConvLSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

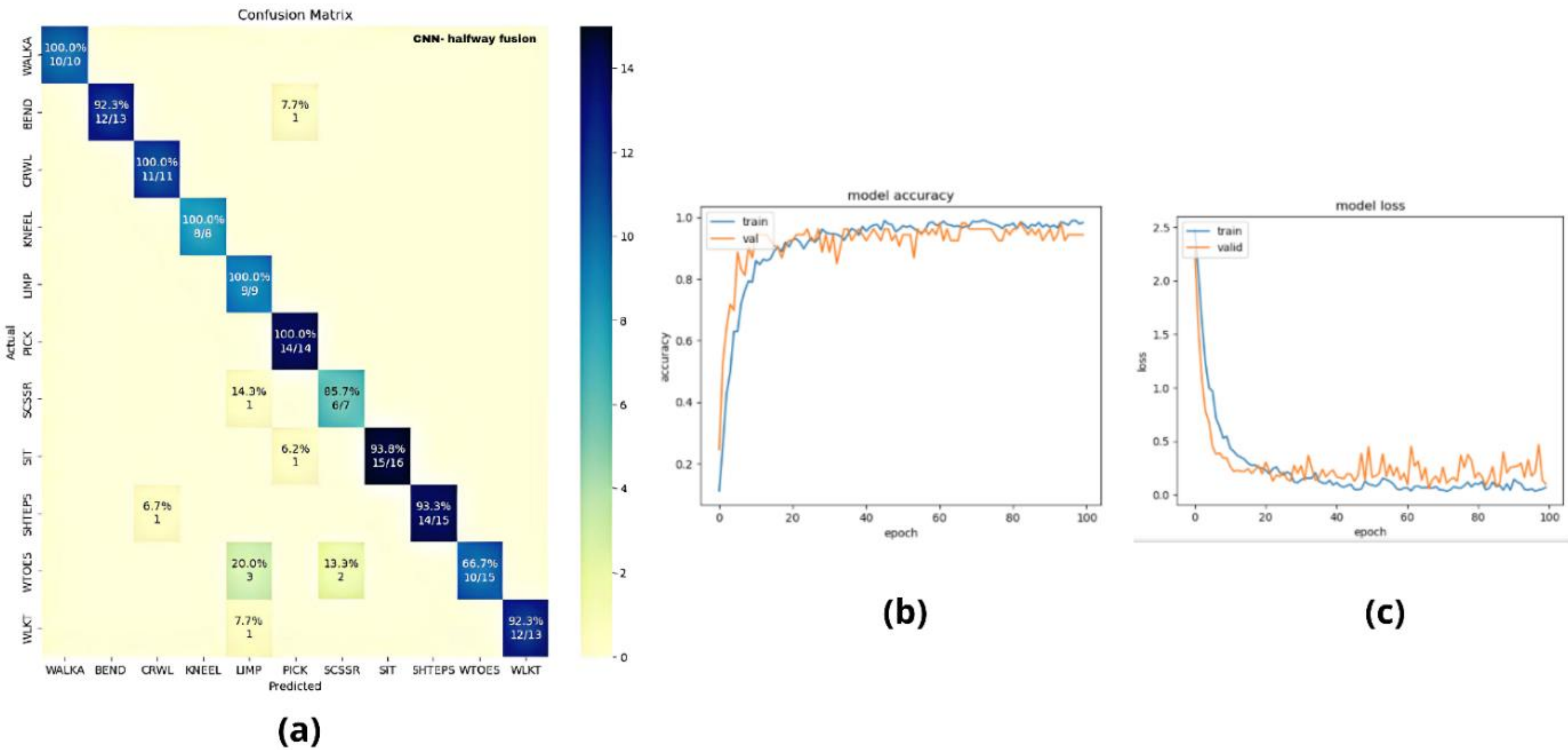


Figure III- 54: Halfway Fusion (CNN model): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

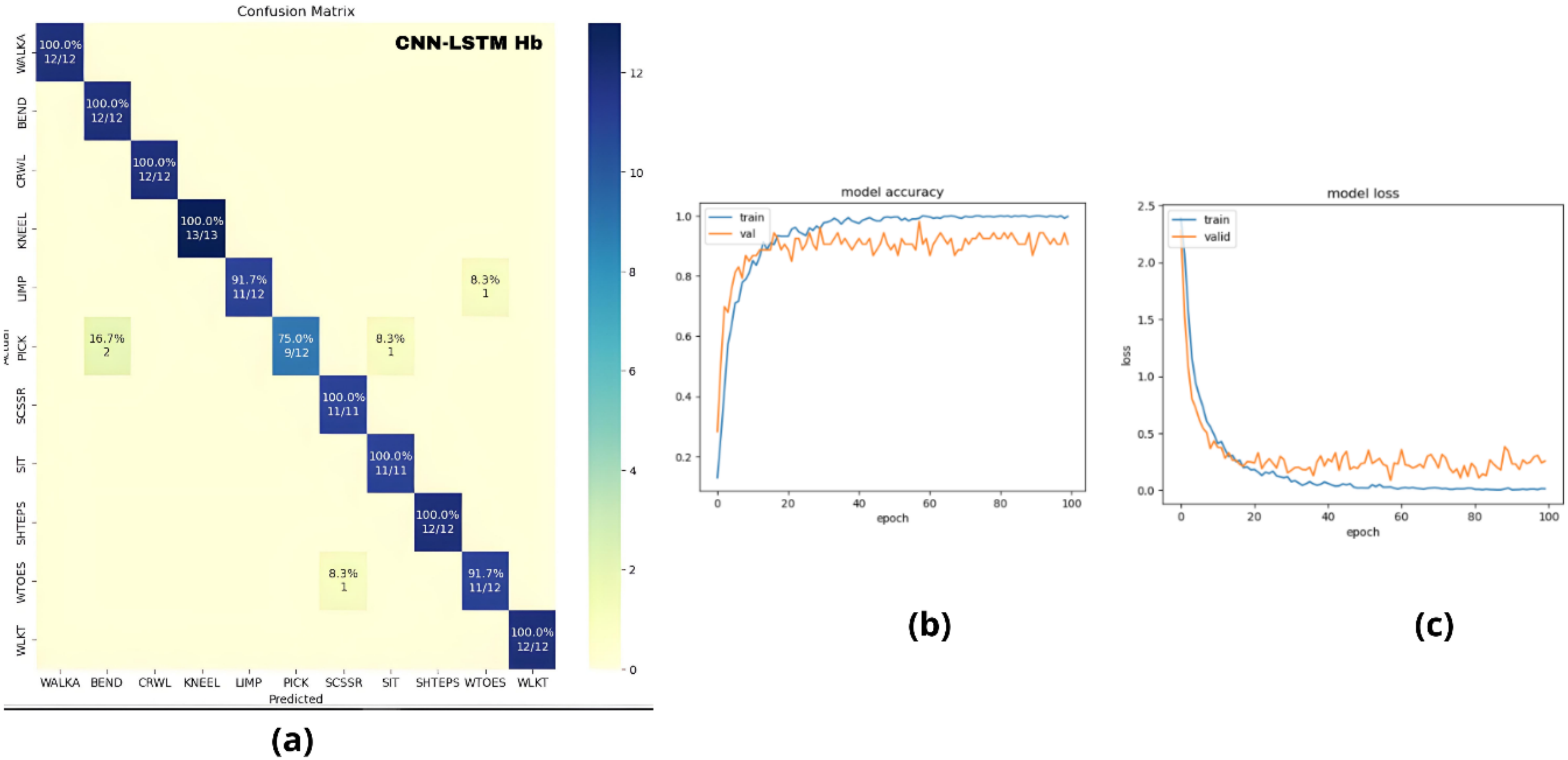


Figure III- 55: Halfway Fusion before concatenation (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

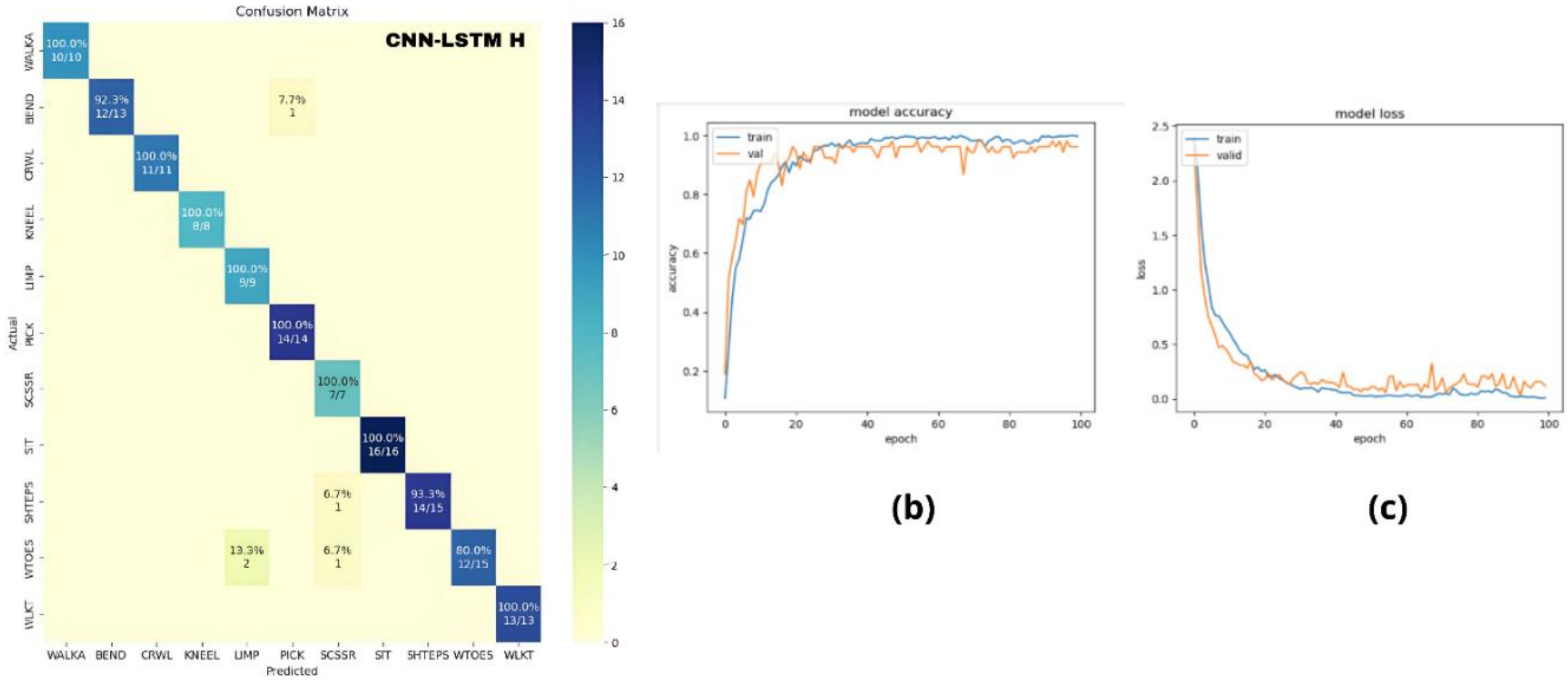


Figure III- 56: Halfway Fusion after concatenation (CNN-LSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

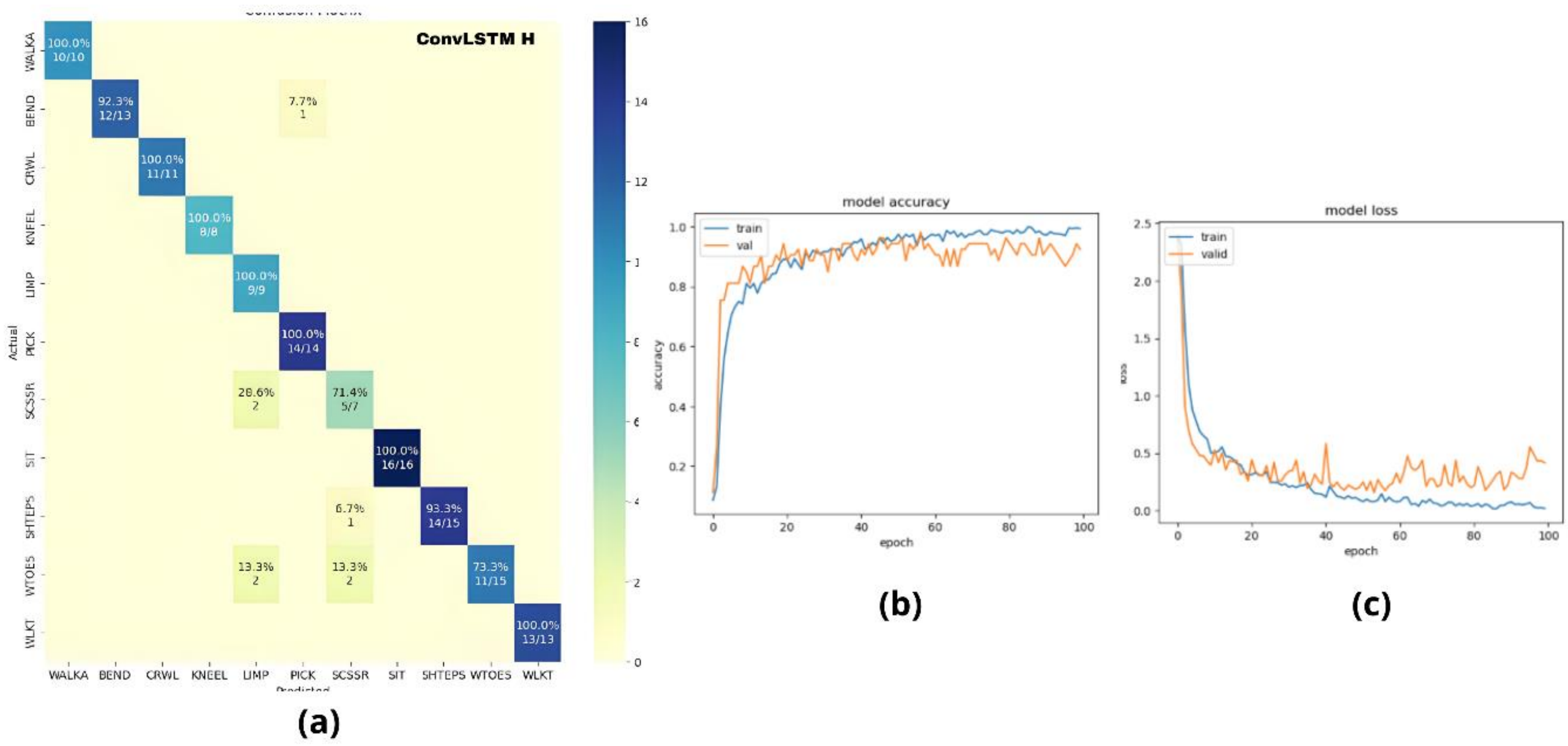


Figure III- 57: Halfway Fusion (ConvLSTM): (a) Confusion Matrix, (b) Accuracy curve, (c) loss Curve

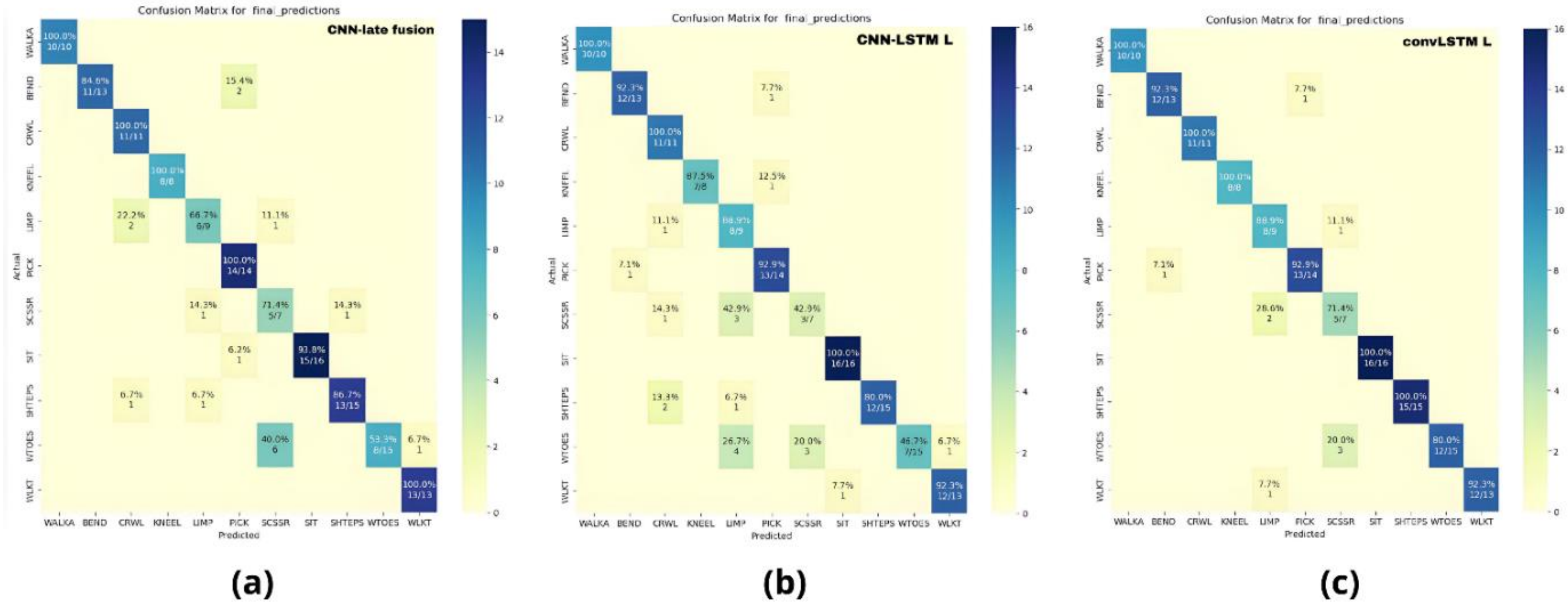


Figure III- 58: Confusion Matrix of late fusion (a.CNN model) (b.CNN-LSTM model) (c.ConvLSTM model)