Deep learning approaches for fingerprint localization in low-power wide area networks

by

Albert Selebea Lutakamale

Submitted in partial fulfillment of the requirements for the degree

Doctor of Philosophy (Electronic Engineering)

in the

Department of Electrical, Electronic and Computer Engineering

Faculty of Engineering, Built Environment and Information Technology

UNIVERSITY OF PRETORIA

October 2024

SUMMARY

Deep learning approaches for fingerprint localization in low-power wide area networks

by

Albert Selebea Lutakamale

Supervisor:	Prof. Herman C. Myburgh
Co-supervisor:	Dr Allan de Freitas
Department:	Electrical, Electronic and Computer Engineering
University:	University of Pretoria
Degree:	Doctor of Philosophy (Electronic Engineering)
Keywords:	Internet of things, convolutional neural networks, deep learning, fingerprint localization, long-range wide area network, low-power wide area network, machine learning, squeeze and excitation, transformer encoder, wireless communications.

In recent years, low-power wide area networks (LPWANs), particularly long-range wide area networks (LoRaWAN), have been increasingly adopted into large-scale internet of things (IoT) applications due to their ability to offer energy-efficient and cost-effective long-range wireless communication. The need to provide location-stamped communications to IoT applications for meaningful interpretation of physical measurements from IoT devices has increased the demand to incorporate location estimation capabilities into LPWAN networks.

Factors such as high-power consumption, high implementation costs, and poor localization performance in urban canyons or environments with many obstructions render outdoor localization solutions based on standalone GPS technology unfit for deployment in large-scale IoT applications, where the emphasis is on energy efficiency and cost-effectiveness. Implementing localization methods in short-range wireless communication networks, such as Bluetooth and ZigBee networks, to estimate locations of target

nodes in large outdoor environments is also not economically feasible due to their short-range nature, as there will be a requirement for dense deployment of wireless nodes, leading to high implementation costs.

In LoRaWAN (one of the key LPWAN technologies operating in unlicensed frequency bands), fingerprint-based localization methods are known to be robust in challenging environments with multipath and non-line-of-sight phenomena, making them relatively more accurate than range-based methods. However, most currently available fingerprint-based localization methods in LoRaWAN networks rely on conventional 'shallow' machine learning models. While such models may yield satisfactory results under specific conditions, their complexity tends to increase as the size of training datasets increases, ultimately resulting in a decline in localization accuracy. In this thesis, driven by the goal of improving the performance and efficiency of fingerprint-based localization methods in LoRaWAN networks, two deep learning-based fingerprint-based methods to estimate the locations of target nodes in LoRaWAN networks are proposed. The first proposed method is a branched convolutional neural network (CNN) localization method enhanced with squeeze and excitation (SE) blocks (referred to as the CNN-SE method). The second proposed method is a hybrid CNN-transformer fingerprint-based localization method (referred to as the CNN-transformer method). The main contribution of the first method is the joint use of CNN (proven to be very efficient in learning useful positional information in structured data) and SE blocks, which improves channel-wise interdependencies. The novel contribution of the second method is the development of a hybrid CNN-transformer fingerprinting-based localization model by leveraging the strengths of both CNNs and transformers. CNNs capture features from the input data at the local level, while the attention mechanism of the transformer captures features from the input data at the global level.

Adopting a 0.7/0.15/0.15 data split scheme for the training, validation, and test set, respectively, and using the entire LoRaWAN dataset, the CNN-SE method achieved localization accuracies of 291.51 m and 147.55 m mean and median localization errors, respectively, on the test set, using the powed data representation scheme. With the CNN-transformer method, the localization accuracy of 288.1 m and 143.7 m mean and median localization errors, respectively, were achieved, using the same experimental settings. The localization accuracies achieved by these two methods have outperformed the localization accuracies of the currently available state-of-the-art fingerprint-based localization methods in the literature, evaluated using the same publicly available LoRaWAN dataset. An R^2 score of 0.93 obtained by both methods further indicates the high degree to which the proposed methods

have been able to fit data in their respective regressors, enabling them to localize target nodes with satisfactory localization accuracies.

DEDICATION

This thesis is dedicated to my wife, Stella B. Msoffe; my siblings, Edward S. Mgaya, Eunice M. Lutakamale, and Julie M. Lutakamale; my late parents, Michael N. Lutakamale and Jane D. Mgaya and my aunt, Laurencia N. Lutakamale. Each one of them has, in one way or another, contributed to making me flourish in my PhD pursuit. I thank them all for their unwavering love, support, prayers, and confidence in me, which have given me much-needed strength and encouragement to overcome countless challenges that came my way as I have been navigating through this very arduous academic journey.

ACKNOWLEDGEMENT

I thank God Almighty for granting me good health, grace, knowledge, intellectual capacity, and mental strength that enabled me to pursue my PhD studies.

I am taking this opportunity to thank my employer, the University of Dodoma, for granting me permission to pursue my PhD studies and for helping me in securing funding for my studies through the Ministry of Education, Science, and Technology of the United Republic of Tanzania under the MoEST-HEET Project fund. Without this funding, I would not have been able to successfully pursue my PhD studies at the University of Pretoria.

My heartfelt appreciation and gratitude go to two very important people in my PhD academic journey: Professor Herman C. Myburgh and Dr Allan de Freitas. I recall when I was in the process of applying for a PhD study at the University of Pretoria (UP), I searched academic profiles of professors working at the Department of Electrical, Electronic, and Computer Engineering to familiarize myself with their research interests. I identified a few of them whose research interests align with mine. I then sent emails to each one of them, asking for their willingness to supervise me in my PhD study. These two people were the only ones who not only responded to my email but also agreed to be my supervisors. Their goodheartedness was also evident when they willingly spent their precious time offering me online supervision for the whole year, during which I could not register for my PhD study while I was struggling to secure funding. I thank them very much for believing that I could successfully pursue my PhD studies. When I finally secured funding, registered for my studies, and arrived at the University of Pretoria to continue with my studies in person, they never stopped in offering me their tireless efforts and supervision to ensure that I would succeed in my PhD studies. The invaluable guidance, support, advice, comments, suggestions, patience, and encouragement that they offered me, not forgetting the countless meetings we had to share the research progress, were vital in ensuring that I successfully carried out my research and completed my PhD studies on time. I pray to God to continue granting them good health, wisdom, grace, kindness, and goodheartedness so that they can continue offering their academic expertise and experience to more and more people in need thereof.

I would also like to extend my sincere gratitude to the University of Pretoria through the Center for Connected Intelligence (CCI) lab, headed by Prof. Herman C. Myburgh for providing me with a

conducive environment and generous financial support to enable me to pursue my studies.

My acknowledgment also goes to all academic and administrative staff of the Department of Electrical, Electronic, and Computer Engineering for ensuring that all my academic and administration issues at UP have been dealt with on time, so as not to affect my studies. I extend my appreciation to my lab mates at the CCI, notably Dr Rachel Kufakunesu, Kelvin Afachao, and Marchant Fourie, for all the wonderful memories we shared during the course of my studies. The invaluable academic advice and support I received from them were key to helping me to achieve my goals.

My sincere thanks also go to my aunt, Laurencia N. Lutakamale, for taking on the role of my late parents by providing financial support to fund my ordinary and advanced secondary school education, which set a good foundation for what I have achieved so far. I will forever be grateful for this help, and I pray to God to continue blessing her and granting her good health and happiness.

Finally, I would like to express my gratitude to my wife, Stella B. Msoffe, my siblings, Edward S. Mgaya, Eunice M. Lutakamale, and Julie M. Lutakamale, my mother-in-law, Lucy Mchau, my father in law, Benard R. Msoffe, as well as my brothers and sisters-in-law for their support and prayers. My special thanks go to my beloved parents, Michael N. Lutakamale and Jane D. Mgaya. Though I lost them while I was still very young and in need of their parenthood, I believe that they are continuing to pray for my well-being and success. I thank them for bringing me into this world and for their role in raising me to who I am today.

With a grateful heart, I humbly attribute all the glory to the divine grace of Almighty God.

LIST OF ABBREVIATIONS

2D	two-dimensional
3D	three-dimensional
3GPP	3rd Generation Partnership Project
AES	advanced encryption standard
ANN	artificial neural network
AoA	angle of arrival
AoD	angle of departure
AR	association rule
BLE	Bluetooth low energy
CDF	cumulative distribution function
CNN	convolutional neural network
CSA	Connectivity Standard Alliance
CSI	channel state information
CSS	chirp spread spectrum
DBPSK	differential binary phase shift keying
D-GPS	differential GPS
DL	downlink
DSSS	direct sequence spread spectrum
DT	decision trees
EC-GSM-IoT	extended coverage-global system for mobile internet of things
ED	end devices
EPC	evolved packet core
ERSS	extreme received signal strength
ESPRIT	estimation of signal parameters via rotational invariant techniques
ETSI	European Telecommunications Standards Institute
FC	fully connected
FFN	feedforward networks
FHSS	frequency hopping spread spectrum
GB	gradient boosting
GFSK	Gaussian frequency shift keying

GMSK	Gaussian filtered minimum shift keying
GNSS	global navigation satellite system
GPRS	general packet radio service
GPS	global positioning system
GSM	global system for mobile communications
GW	gateway
HDOP	horizontal dilution of precision
HF	high-frequency
IDW	radius-inverse distance weighting
IIoT	industrial internet of things
IoT	internet of things
ISM	industrial scientific and medical
KNN	k-nearest neighbor
LASSO	least absolute shrinkage and selection operator
LF	low-frequency
LoRaWAN	long-range wide area networks
LOS	line of sight
LPWAN	low-power wide area network
LSTM	long-short-term memory
LTE	long-term evolution
LTE-M	long-term evolution machine type communication
M2M	machine-to-machine
MAC	medium access control
MAE	mean absolute error
MHCA	multi-head cross attention
MHSA	multi-head self-attention block
ML	maximum likelihood
MLP	multi-layer perceptron
MSE	mean squared error
NB-IoT	narrow-band internet of things
NLOS	non-line of sight
NLP	natural language processing
NS	network simulator
OFDM	orthogonal frequency division multiplexing

OFDMA	orthogonal frequency division multiple access
OLS	ordinary least square
OTDoA	observed time difference of arrival
PCA	principal component analysis
PHY	physical layer
QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
RBF	radial basis function
ReLU	rectified linear unit
RF	random forest
RFID	radio-frequency identification
RFR	random forest regressor
RNN	recurrent neural network
RPMA	random phase multiple access
RSSI	received signal strength indicator
RTK	real-time kinetic
SC-FDMA	single carrier frequency division multiple access
SE	squeeze and excitation
SF	spreading factor
SGD	stochastic gradient descent
SHF	super-high frequency
SIG	special interest group
SIR	sequential importance resampling
SVM	support vector machines
SVR	support vector regression
TDoA	time difference of arrival
ToA	time of arrival
TSMA	telegram splitting multiple access
UHF	ultra-high frequency
UL	uplink
UNB	ultra-narrowband
WLAN	wireless local area network

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1 BACKGROUND

Low-power wide area networks (LPWANs) comprise different types of wireless communication technologies which, depending on their respective operating frequencies, are categorized into two main groups, i.e., those which operate in licensed frequency bands, e.g., long-term evolution machine type communication (LTE-M), narrowband-IoT (NB-IoT) and extended coverage-global systems for mobile communications-IoT (EC-GSM-IoT), and those which operate in unlicensed frequency bands, e.g., Mioty, Random phase multiple access (RPMA), Weightless, Wize, Sigfox, and LoRaWAN [1–3]. These technologies are characterized by their abilities to offer energy-efficient long-range wireless communication of up to 10-40 km and 1-5 km in rural and urban areas, respectively [4].

The high scalability, sustainability, and low power nature of LPWAN technologies have contributed to their rapid adoption in large-scale internet of things (IoT) applications. Additionally, researchers are attracted to adopting these technologies into IoT applications on account of their ability to offer long-range wireless communications at relatively lower costs by using inexpensive batteries that can last for years. These attributes make LPWAN technologies highly suitable for long-range IoT applications [5]. Unlicensed LPWAN technologies, particularly LoRaWAN, are leading technologies in large-scale IoT applications as a result of their utilization of unlicensed industrial, scientific and medical (ISM) bands and being relatively cheap and easy to deploy, compared to licensed LPWAN technologies [4, 6–8].

The explosive increase in the adoption of IoT-related technologies in fields such as smart cities, precision livestock farming, agriculture, asset tracking, emergency operations, environmental monitoring, and logistics has created a strong demand to incorporate location estimation capabilities into large-scale IoT applications in order to meaningfully interpret physical measurements being collected from the

IoT devices. For instance, context awareness, which is one of the essential aspects of IoT applications that refer to the ability of an IoT device to alter its behaviour based on the measurement that it records, is only possible if a device's location can be established with minimal localization error [9]. Likewise, precise localization of workers/equipment is also crucial in industrial internet of things (IIoT) applications where smart objects, such as sensors, wearables, and actuators, are connected to facilitate information management and provide a real-time progress of work processes to increase operational efficiency and level of work safety [10]. Being able to localize objects in IIoT settings is also helpful in preventing collisions, as well as unauthorized access to machinery and dangerous worksites, through proper planning and distribution of workforce [10].

With this increased adoption of unlicensed LPWAN networks into large-scale IoT applications and the importance of equipping them with location-stamped communication, more effort is needed from research communities to develop effective and efficient localization solutions, which are crucial to the implementation of smart solutions in different spheres of human life.

1.2 PROBLEM STATEMENT

For many years, the global positioning system (GPS) has been the most adopted technology for localization purposes in outdoor environments [11], while Bluetooth low energy (BLE), WiFi, ZigBee, and radio-frequency identification (RFID) are examples of short-range communication technologies that can be applied to both indoor and outdoor localization solutions. However, the inability to give reliable and accurate localization results in rich-scattering environments and urban canyons where there is no clear and unobstructed reception of satellite signals, as well as the high cost and power consumption of GPS receivers, make GPS-based localization solutions not ideal for large-scale IoT applications.

On the other hand, because of their short-range nature, localization solutions based on short-range wireless technologies such as WiFi, Zigbee, and Bluetooth, have been predominately adopted in indoor localization scenarios, as in [12–17]. If they are to be adopted in large-scale outdoor localization solutions, they will need dense deployment of access points, which is not economically feasible.

These notable limitations of GPS-based and short-range-based localization solutions in large-scale IoT applications necessitate more and more research dedicated to proposing and developing efficient and effective localization solutions to be implemented into LPWAN networks, particularly LoRaWAN,

as they are more equipped to offer scalable, reliable, robust, and low power localization and tracking solutions [5].

1.2.1 Preference of fingerprint-based localization methods over range-based localization methods

In wireless communication systems, a target node can be localized using either range or fingerprint-based approaches. In range-based localization approaches, a path loss model is usually applied to convert a received signal strength indicator (RSSI) value from a specific gateway into an estimated distance to that gateway. Any available geometrical (e.g., trilateration and min-max) or statistical (e.g., Bayesian and maximum likelihood) techniques can then be utilized to estimate the final location of a node [18]. Fingerprint-based approaches, on the other hand, rely on a database of features (fingerprints) at pre-determined physical locations to estimate the location of a target node through feature matching.

Due to their robustness, fingerprint-based localization approaches are increasingly being applied in different types of wireless communication networks to estimate the locations of target nodes. For instance, in [19–23], the authors adopt a fingerprint-based approach to estimate location of target nodes in massive multiple-input multiple-output networks, where in [19–21], angle of arrival vectors were extracted from the uplink signal and used as fingerprints to aid in the localization process. In [22], the angle of arrival and received signal strength parameters were extracted and used jointly as fingerprints to assist in the location estimation of target nodes, whereas in [23], an angle delay power matrix is extracted from channel state information and used in pinpointing locations of target nodes. Likewise, in [24–26], signals emitted by WiFi access points were extracted and used to build radio maps, which were then used as fingerprints to assist in location estimation of target nodes using machine learning-based algorithms. Furthermore, in [27–30], fingerprint-based methods are proposed for localization purposes using Bluetooth signals. In [27], a passive fingerprinting approach is proposed using received signal strength, while in [28], a bidirectional long short-term memory (Bi-LSTM) network is used to implement the fingerprint-based method. In [29], an inverse fingerprinting approach implemented on the server side of a Bluetooth network using signals from Bluetooth beacons is proposed, while in [30], a fingerprinting localization approach using multivariate polynomial interpolation methods is proposed. Another notable example of a research work where a fingerprint-based approach has been applied in estimating locations of target nodes is as reported in [31]. In this work, authors used a ZigBee interface to collect WiFi signals and build a radio map and developed a machine learning-based localization

method using a k-nearest neighbour algorithm for target node location estimation.

Like in other wireless communication networks, in unlicensed LPWAN networks (LoRaWAN in particular), the use of fingerprint-based methods in estimating the locations of target nodes is gaining popularity among researchers due to challenges facing the range-based localization methods. Frequency hopping is one of the challenges facing range-based localization methods in LoRaWAN networks, which in most cases is the main cause of poor performance of range-based methods, leading to significant degradation in localization accuracies [18]. Additionally, the localization performances of range-based localization approaches can be negatively affected by shadowing and fading phenomena due to multipath propagation. Moreover, the need to install dedicated hardware like expensive antenna arrays and the requirement for accurate clock synchronization among anchor nodes to implement range-based localization methods effectively limits their adoption in LPWAN networks.

1.2.2 Research gap

According to [32], fingerprint-based localization approaches in unlicensed LPWAN networks can achieve high localization accuracies, compared to the range-based approaches in a large urban area. By leveraging the advancement in machine learning algorithms, as well as increasing computational power of computers, fingerprint-based localization approaches can lead to relatively high localization accuracies, compared to range-based localization approaches due to their ability to effectively learn useful locational information even from noisy data being collected in non-line of sight (NLOS) environment [11].

Most currently available fingerprint-based localization methods in LoRaWAN rely on conventional ‘shallow’ machine learning models. While such models may yield satisfactory results under specific conditions, their complexity tends to increase as the size of training datasets increases, ultimately resulting in a decline in localization accuracy [11].

Therefore, the focus of this thesis is on researching effective ways of addressing the limitations of currently available fingerprint-based localization methods in LoRaWAN networks, which are predominantly based on ‘shallow’ machine learning models. The objective is to propose and develop efficient, robust, and effective fingerprint-based localization methods to estimate the locations of target nodes in LoRaWAN networks.

1.3 RESEARCH OBJECTIVE AND QUESTIONS

The main objective of this research is to propose methods which can improve the performance of machine learning-based fingerprint localization approaches in LoRaWAN networks. In order to meet this objective, the following specific objectives have to be accomplished:

1. To conduct a comprehensive literature review of machine learning-based fingerprint localization approaches in unlicensed LPWAN networks in order to analyse their strengths and weaknesses.
2. To explore currently available machine learning-based fingerprint localization techniques deployed in other wireless networks in order to analyze the possibility of adopting them in localizing target nodes in LoRaWAN networks.
3. To explore recent advancement in machine learning algorithms, especially deep learning techniques with a view to deploy them in enhancing the localization performance of fingerprint-based localization methods in LoRaWAN networks.
4. To explore available data preprocessing techniques and find out how they can be adopted to improve data representation in LoRaWAN-based fingerprint databases to enhance the generalization ability of machine learning-based fingerprint localization models.
5. To propose and develop efficient, robust, and effective fingerprint localization models to localize target nodes in LoRaWAN networks by using knowledge acquired after completion of specific objectives 1,2,3, and 4 above, and to validate their performance by using simulated or publicly available LoRaWAN datasets.

The following research questions will assist in the accomplishment of the stated main objective of this research:

1. What are the strengths and limitations of currently available machine learning-based fingerprint localization approaches in LoRaWAN networks?
2. How can the currently available machine learning-based fingerprint localization techniques being used in other wireless networks, be adopted to localize target nodes in LoRaWAN networks?
3. What are the recent advancements in machine learning algorithms, especially deep learning techniques, and how can they be adopted to enhance the performance of machine learning-based fingerprint localization models in LoRaWAN networks?

4. What are the available data preprocessing techniques, and how can they be used to improve LoRaWAN-based fingerprint databases to enhance the generalization ability of machine learning-based fingerprint localization models?
5. How can the knowledge obtained in research questions 1, 2, 3, and 4 above be used to develop enhanced machine learning-based fingerprint localization methods to estimate locations of target nodes in LoRaWAN networks?

1.4 APPROACH

The first step in conducting this research is to comprehensively review the machine learning-based fingerprint localization approaches in unlicensed LPWAN networks being proposed by different researchers in the literature. The aim is to analyze their strengths and weaknesses. Completing the first step will pave the way for the second step, in exploring currently available machine learning-based fingerprint localization techniques deployed in other wireless networks, in order to analyze the possibility of adopting them in LoRaWAN networks. The third step will be dedicated to exploring recent advancements in machine learning algorithms, especially deep learning techniques, and finding a way to deploy them to develop enhanced fingerprint-based localization methods to estimate the locations of target nodes in LoRaWAN networks. The fourth step will be to explore available data preprocessing techniques and examine how they can improve data representation in LoRaWAN-based fingerprint databases to enhance the generalization ability of machine learning-based fingerprint localization models. Finally, the last stage will be dedicated to proposing and developing enhanced machine learning-based fingerprint localization models, using the knowledge obtained in the preceding steps.

1.5 RESEARCH CONTRIBUTION

The contribution of this research to the existing body of knowledge can be summarized as follows:

1. The comprehensive literature review on LPWAN networks, short-range wireless communication technologies, localization solutions in wireless networks, currently available machine learning-based fingerprint localization approaches in unlicensed LPWAN networks, as well as the exploration of recent advancements in machine learning algorithms (deep learning in particular) and data preprocessing techniques, have provided valuable knowledge and tools crucial to assisting researchers in developing effective and accurate localization solutions in unlicensed LPWAN networks, particularly fingerprinting-based approaches.

2. An efficient, robust and effective deep learning-based fingerprint localization method based on convolutional neural network and squeeze and excitation blocks (CNN-SE) has been proposed and developed to estimate the locations of target nodes in LoRaWAN networks with acceptable levels of localization accuracies. The use of SE blocks is to complement the efficiency of CNN in learning useful positional information in structured data by improving its channel-wise interdependencies.
3. An efficient, robust and effective deep learning-based fingerprint localization method based on the joint use of CNN and transformer module has been proposed and developed to estimate the locations of target nodes in LoRaWAN networks with acceptable levels of localization accuracies. The novel contribution of this method is the development of a hybrid CNN-transformer fingerprint-based localization model by leveraging the strengths of both CNNs and transformers. CNNs capture features from the input data at the local level, while the attention mechanism of the transformer captures features from the input data at the global level.
4. Adopting a 0.7/0.15/0.15 data split scheme for the training, validation, and test set, respectively, and using the entire LoRaWAN dataset, the CNN-SE method achieved localization accuracies of 291.51 m and 147.55 m mean and median localization errors, respectively, on the test set, using the powed data representation scheme. With the CNN-transformer method, the localization accuracy of 288.1 m and 143.7 m mean and median localization errors, respectively, were achieved using the same experimental settings. The localization accuracies achieved by these two methods have outperformed the localization accuracies of the currently available state-of-the-art fingerprint-based localization methods in the literature, evaluated using the same publicly available LoRaWAN dataset. An R^2 score of 0.93 obtained by both methods further indicates the high degree to which the proposed methods have been able to fit data in their respective regressors, enabling them to localize target nodes with satisfactory localization accuracies.
5. The optimal preprocessing techniques for the LoRaWAN dataset for improved localization performance are presented. The experimental results revealed superior performances of the proposed methods when trained on the publicly available LoRaWAN dataset transformed using a powed data representation scheme compared to positive, normalized, and exponential data representation schemes.

1.6 RESEARCH OUTPUTS

The following research articles and conference papers, which have been published or are currently under review in peer-reviewed journals and conferences, are the results of the findings of this research

study:

1.6.1 Journal articles

1. A. Lutakamale, H. Myburgh, and A. De Freitas, “RSSI-based fingerprint localization in LoRaWAN networks using CNNs with squeeze and excitation blocks,” Ad Hoc Networks, vol. 159, Jun. 2024, Art. no. 103486.
2. A. Lutakamale, H. Myburgh, and A. De Freitas, “A hybrid convolutional neural network-transformer method for received signal strength indicator fingerprinting localization in long range wide area network,” Engineering Applications of Artificial Intelligence, vol. 133, Jul. 2024, Art. no. 108349.
3. A. Lutakamale, H. Myburgh, and A. De Freitas, "A Review of Animal Localization Solutions: Enabling Technologies, Challenges and Future Direction." *Pervasive and Mobile Computing*, (**under review**).

1.6.2 Conference articles

1. A. Lutakamale, H. Myburgh, and A. De Freitas, "A Review of Indoor/Outdoor Localization Solutions in Sigfox and LoRaWAN Networks" in *Proceedings of the Southern Africa Telecommunication Networks and Applications Conference (SATNAC)*, 2023, pp. 262-267.
2. A. Lutakamale, H. Myburgh, and A. De Freitas, "Fingerprinting Localization in Sigfox Networks: Analysis of the Impact of Base Stations with a Small Number of Received Messages" (**Submitted to the Southern Africa Telecommunication Networks and Applications Conference (SATNAC-2024)**).

1.7 THESIS OVERVIEW

The remaining chapters in this thesis are organized as follows: In Chapter 2, a comprehensive review of relevant literature on wireless communication technologies and localization approaches in wireless networks is presented. The emphasis is on the LPWAN technologies being used as enabling technologies for implementing large-scale IoT applications and machine learning techniques being adopted to implement fingerprint-based localization approaches. The objective of this chapter is to identify the research gap by analyzing the strengths and weaknesses of the currently available localization approaches in LPWAN networks, focusing on LoRaWAN networks. A detailed understanding of the main building blocks of this thesis’s proposed fingerprint-based localization methods, namely the CNN model, SE block, and transformer module, is provided in Chapter 3. The focus is on discussing their main features, key network parameters associated with their implementation, and general working

mechanisms. In Chapter 4, the two fingerprint-based localization methods that have been proposed in this thesis to estimate the target node's locations in LoRaWAN networks are presented. The first method is based on CNNs enhanced with SE blocks, whereby the SE blocks have been adopted to complement the efficiency of the CNN regressor in learning useful position features in structured data by improving channel-wise interdependencies. The second fingerprint-based localization method being presented in this chapter is a hybrid CNN-transformer method to estimate the target nodes's locations in LoRaWAN networks. CNNs are adopted to complement the strengths of the transformer by adding the ability to capture local features from input data and consequently allowing the transformer, through the attention mechanism, to effectively learn global dependencies from the input data. In Chapter 5, a series of experiments carried out to evaluate the effectiveness of the proposed fingerprint-based localization methods, i.e., the branched CNN localization method enhanced with SE blocks and the hybrid CNN-transformer fingerprinting method to localize target nodes in LoRaWAN networks using a publicly available LoRaWAN dataset, are presented. Different training data extraction strategies and performance metrics have been adopted to evaluate the performance of the two proposed fingerprint-based localization methods in different scenarios to prove their robustness and effectiveness. In Chapter 6, conclusive remarks on the thesis are presented along with future research opportunities in the context of node localization research, particularly in unlicensed LPWAN networks, which are increasingly being adopted in implementing large-scale IoT applications.

CHAPTER 2 LITERATURE STUDY

2.1 CHAPTER OVERVIEW

In this chapter, a comprehensive review of short-range and long-range wireless communication technologies, which are crucial in implementing localization methods that are needed to equip IoT applications with location estimation capabilities, is presented. The goal is to help researchers understand how these technologies function as well as their strengths and limitations. A clear understanding of different aspects of these technologies will assist researchers to effectively deploy them in implementing different IoT applications to solve different societal problems. Moreover, in addition to the short-range and long-range wireless technologies, a comprehensive review of different types of range-based and fingerprint-based localization methods/techniques that have been proposed in the literature, is also presented to give researchers a clear understanding of their working mechanisms, how they are implemented, to which environment they are most effective, as well as their effectiveness in estimating the location of the target objects. Understanding these localization methods is crucial to allowing researchers to effectively deploy them in different localization tasks and to complement them with other technologies and techniques to enhance their localization performance.

The rest of this chapter is organized as follows: In Section 2.2, an overview of LPWAN technologies is presented, whereby the key technological features of licensed and unlicensed LPWAN technologies are discussed. In Section 2.3, short-range wireless communication technologies are presented, including a discussion of their key technological features, implementation, and usage. The localization process and the localization parameters being adopted in pinpointing the target object's locations are presented and discussed in Section 2.4. The range-based and fingerprint-based localization approaches in unlicensed LPWAN networks are contextualized in Section 2.5. In Section 2.6, algorithmic aspects of range-based localization methods in LPWAN networks are elaborated on, while in Section 2.7, currently available range-based localization solutions in the literature are discussed. In Section 2.8, algorithmic aspects

of fingerprint-based localization approaches in LPWAN networks are introduced. In Section 2.9, the currently available fingerprint-based localization solutions in LoRaWAN networks being proposed in the literature are discussed. Conclusive remarks of this chapter are provided in Section 2.10.

2.2 LOW-POWER WIDE AREA NETWORKS

LPWAN is a wireless network technology that offers low-cost, low-power, and long-range communication capabilities to IoT applications. Communication ranges of 10-40 km and 1-5 km in rural and urban areas, respectively, are possible in a typical LPWAN network [4, 33]. There are two types of LPWAN technologies: those which operate in licensed frequency bands, e.g., long-term evolution machine type communication (LTE-M), narrow-band internet of things (NB-IoT) and extended coverage-global systems for mobile communications-internet of things (EC-GSM-IoT), and those which operate in unlicensed frequency bands, e.g., Mioty, Weightless, Wize, Sigfox, and LoRaWAN [1–3]. LoRaWAN (which is the focus of this thesis) and Sigfox technologies, along with NB-IoT technology, are the three most adopted LPWAN technologies in IoT applications [34–36]. Table 2.1 and Table 2.2 summarize key technological features of unlicensed and licensed LPWAN networks.

2.2.1 Long-term evolution machine type communication

LTE-M is an LPWAN technology which is standardized by 3GPP and operates in the licensed LTE spectrum. It has been designed to transmit bits at low latency of a few seconds at data rates ranging from 10 kbps to 1 Mbps [37]. LTE-M can coexist in the same bandwidth with other LTE services (in-band deployment) and can support real-time and non-real-time applications. The features that make this technology suitable for IoT applications are its longer coverage, power efficiency, and low complexity. The communication architecture of LTE-M comprises user and control planes, with the former dealing in sending and receiving user data while the latter controls all the processes concerned in initiating communication and in authenticating the end devices [38]. A single base station in LTE-M can support 100000+ end devices, as well as allow the transmission of messages at a wide range of sizes [39]. This technology can be used to implement smart solutions in cities, asset tracking and health facilities, to name a few.

2.2.2 Mioty network

Mioty, which is standardized by the European Telecommunication Standard Institute (ETSI) and developed by Fraunhofer IIS, is an LPWAN technology designed to offer energy-efficient long-range wireless communications to IoT applications that can reach up to 5 km in urban areas and up to 15 km in rural areas [40]. It operates at ISM frequency bands of 868 MHz (Europe) and 915 MHz (North America) and can offer data rates of up to 0.514 Kbps at a bandwidth of 200 kHz [1]. Mioty, which is

based on ultra narrow-band-bandwidth, uses the telegram splitting multiple access (TSMA) technique, which operates by allowing transmission of radio bursts of a single message over different frequencies and times [41, 42]. The splitting of a message into shorter packages (fragments), transmitting them at different frequencies and times, and then re-assembling them at the receiver, make this technology robust to interference, subsequently ensuring reliable and stable wireless communication. Mioty has been designed to offer smart solutions in buildings, agriculture, logistics, and cities.

2.2.3 Wize network

Wize is an LPWAN technology based on the EN-13747-Wireless M-bus standard [43]. Its protocol stack consists of four layers: the physical layer (PHY), the data link layer, the presentation layer, and the application layer. At an operation frequency of 169 MHz with a bandwidth of 75 kHz comprising six 12.5 kHz channels, Wize can offer long-range wireless communication of up to 20 km at extremely low power consumption [43]. Gaussian frequency shift keying (GFSK) and 4GFSK are the two modulation techniques deployed in Wize networks. With the GFSK modulation technique, Wize can offer maximum data rates of up to 2400 bits/s or 4800 bits/s, while with the 4GFSK modulation technique, the data rates can reach 6400 bits/s [1]. Wize supports a maximum payload length of 256 octets and 104 octets for the uplink and downlink transmission, respectively [43]. Wize has been designed for use in smart city projects and has notable applications in smart metering for gas, water, and electricity.

2.2.4 Extended coverage-global system for mobile internet of things

EC-GSM-IoT is an LPWAN technology introduced by 3GPP to operate on the existing 2G cellular infrastructures to offer scalable and low complex communication capabilities to IoT applications, using the general packet radio service (GPRS) communication standard. It operates in the global system for mobile communication (GSM) frequency band at a bandwidth of 200 kHz [2]. At a maximum data rate of 70 Kbps (using the GMSK modulation technique) or 240 Kbps (using the 8PSK modulation technique) [39], EC-GSM-IoT can offer low power and long-range communication capabilities with a communication coverage of up to 15 km [2]. Its projected battery lifetime is ten-plus years, and it can support a large number of end devices, compared to the GSM technology [37].

2.2.5 Random phase multiple access

Random phase multiple access (RPMA) is an LPWAN technology proposed and developed by Ingenu to offer higher throughput, and scalable and long-range wireless communication to IoT applications. It operates in an unlicensed 2.4 GHz ISM frequency band with a maximum bandwidth of 1 MHz [2]. Direct sequence spread spectrum (DSSS) and frequency shift keying (FSK) are the two modulation

techniques that it employs for the uplink and downlink transmission, respectively, with the former allowing sharing of a single time slot among multiple transmitters [2]. At a maximum data rate of 80 Kbps [39], RPMA offers low-cost, low-power, and bi-directional long-range communication reaching up to 5 km and 15 km for the urban and rural areas, respectively, at a maximum payload of 10 KB [2]. To add reliability and security to transmission, RPMA adopts an AES-256b encryption technique [44]. RPMA can be adopted in a wide range of IoT applications, e.g., agriculture, smart grids, smart city, and asset tracking, to name a few [45].

2.2.6 Weightless network

Weightless, proposed and developed by a weightless special interest group (Weightless SIG), is an LPWAN technology designed to offer long-range communication coverage for IoT applications. Weightless exists in three standards: weightless-W, weightless-N, and weightless-P. Weightless-W adopts DBPSK and quadrature amplitude modulation (QAM) modulation techniques and operates in licensed frequency bands of 470 MHz to 790 MHz (TV white space spectrum), offering data rates of between 1 Kbps and 10 Mbps [39]. The two remaining standards, weightless-N and weightless-P, operate in unlicensed frequency bands, with the former adopting a differential binary phase shift keying (DBPSK) modulation technique and the later, Gaussian filtered minimum shift keying (GMSK) and offset quadrature phase shift keying (QPSK) modulation techniques [2]. Weightless-N offers unidirectional communication at a communication range of up to 3 km at a maximum data rate of 100 Kbps, while weightless-P offers bi-directional communication with communication coverage of up to 2 km at a data rate speed of up to 100 Kbps [39].

2.2.7 Narrow band-internet of things

NB-IoT, which is a licensed LPWAN technology standardized by the 3rd Generation Partnership Project (3GPP), is a wireless communication technology designed to offer low-power and long-range communication capabilities to IoT applications based on massive machine-to-machine (M2M) communications [2]. The narrow-band communication protocol of the NB-IoT technology is implemented at the evolved packet core (EPC) framework, which is a core network of existing long-term evolution (LTE) network infrastructure [46]. NB-IoT offers two different data transmission speeds, for the uplink (UL) and downlink transmission (DL), i.e., 20 Kbps and 250 Kbps data rates, respectively [35, 47]. Orthogonal frequency division multiple access (OFDMA) and single carrier frequency division multiple access (SC-FDMA) are two modulation techniques adopted for the downlink and uplink transmission, respectively, in an NB-IoT network [47]. NB-IoT technology can offer a communication range of up to 120 km (3GPP release 15) [1]. NB-IoT can offer three different modes of operation: standalone mode,

which utilizes a global system for mobile communications (GSM) carrier with a 200 kHz bandwidth; in-band, which deploys a narrow band of 180 kHz bandwidth within an LTE carrier; and guard-band, which uses unused 180 kHz LTE's guard band [1, 2, 48]. To equip NB-IoT technology with the ability to estimate locations of target nodes, in release 14, the 3GPP has added the ability to extract the observed time difference of arrival (OTDoA) [34]. OTDoA localization parameter can be adopted by any suitable geometrical technique to estimate the location of the target node.

2.2.8 Long-range wide area network

LoRaWAN refers to an LPWAN technology developed by Semtech and standardized by the LoRa Alliance to realize low-power and long-range wireless communication, covering a communication distance of 2-5 km and up to 15 km in urban and rural areas, respectively [2, 49]. It is a medium access control (MAC) sublayer operating on top of the LoRa physical (PHY) layer. Chirp spread spectrum (CSS) deployed in the LoRa PHY layer is the modulation technique based on varying sinusoidal pulses, which provide a trade-off between the data rate and communication quality [49, 50]. The operational bandwidth of LoRaWAN is regional-specific, i.e., in Europe, it is allowed to operate at 433 MHz and 868 MHz ISM bands; in North America, at 915 MHz ISM band, and in Asia, at 430 MHz ISM band [2, 47, 51]. In LoRaWAN, the spreading factor (SF), which ranges from 7-12, enables the data rate to be varied from 0.3 Kbps to 50 Kbps [1, 46], whereby the higher the SF, the lower the data rate. Additionally, using high SF values leads to high energy consumption, but the network becomes more resilient to interference and noise [50]. The bandwidths offered by LoRaWAN networks are 125 kHz, 250 kHz, and 500 kHz [1, 49, 50]. A typical LoRaWAN network deploys a star-of-stars network topology comprising end devices (EDs), LoRaWAN gateways (GWs), a LoRaWAN network server and LoRaWAN application servers, as illustrated in Figure 2.1. It has three types of end devices: class A, class B, and class C [2, 52]. Class A devices are provided with only two short downlink windows to receive a message after an uplink transmission, while class B devices are provided with the ability to add extra windows to receive messages in downlink transmission at scheduled times. The receiving windows for the downlink transmission in class C devices are kept open continuously. In LoRaWAN networks, all the available localization techniques (geometrical, statistical, or pattern matching techniques) that use time of arrival (ToA), time difference of arrival (TDoA) and received signal strength indicator (RSSI) as localization parameters can be utilized to find a target node's location estimates [53]. Using these localization parameters does not require modification of the hardware or firmware on the end device (transmitting tag) side. Localization of a transmitter (target node) is done by analyzing RSSI or ToA/TDoA at the receiving gateways of the LoRaWAN network

by deploying approaches such as multilateration, pattern matching, maximum likelihood estimation, and multilateral dissections [54].

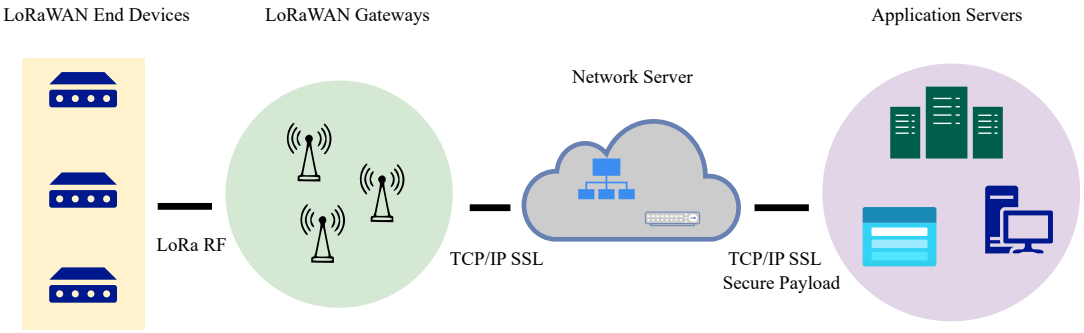

Figure 2.1. LoRaWAN network architecture. Adapted from [55], ©2019 IEEE.

2.2.9 Sigfox network

Sigfox is an LPWAN technology developed to offer end-to-end low-power and long-range communication capabilities, particularly for IoT applications. The frequency bands within which Sigfox is allowed to operate differ regionally. In Europe it operates at 868 MHz; in North America (US) it operates at 915 MHz; and in Asia it operates at 433 MHz ISM bands [2, 47]. Sigfox technology adopts two modulation techniques, both at 100 Hz ultra-narrowband (UNB), i.e., Gaussian frequency shift keying (GFSK) and differential binary phase shift keying (DBPSK), for the downlink (DL) and uplink (UL) transmissions, respectively [47, 56–58]. Data transmission speed in Sigfox networks can either be 100 bps or 600 bps, depending on the deployment region [1, 46]. On account of UNB, communication in Sigfox technology can cover a distance of 3-10 km in rural areas and 30-50 km in urban areas [2]. In a single day, Sigfox allows transmission of 140 (maximum) uplink messages (each with a maximum 12 bytes payload) and maximum four downlink messages (each with a maximum of eight bytes payload) [35, 36, 46]. The introduction of restrictions on the size and number of messages to be transmitted per single day is to minimize the energy consumption of Sigfox devices and to prolong battery life. Sigfox networks deploy a star network topology consisting of end devices (EDs), gateways (GWs), and network servers. GWs receive messages from EDs and forward the packets to Sigfox cloud for processing, followed by the transmission of the application payload to the application server [2]. Localization in the Sigfox networks can be achieved through probabilistic calculation of the most probable location of the target node by utilizing RSSI values extracted from its transmitted messages received at the Sigfox gateways [59]. With the extracted RSSI values at the gateways, any geometrical technique, such as trilateration or fingerprinting-based techniques, can be utilized to estimate the location of the target node.

Table 2.1. Key technological features of unlicensed LPWAN technologies.

Features	**Sigfox [4, 60]**	**LoRaWAN [4, 60]**	**RPMA [2, 39, 44, 61]**	**Mioty [1, 40, 42, 61]**	**Weightless-P/N [2, 39]**	**Wize [1, 43]**
Operation frequency	868 MHz (Europe), 915 MHz (North America), and 433 MHz (Asia)	868 MHz (Europe), 915 MHz (North America), and 433 MHz (Asia)	2.4 GHz	868 MHz (Europe), 915 MHz (North America)	868 MHz (Europe), 915 MHz (North America)	169 MHz
Bandwidth	100 Hz	250 kHz and 125 kHz	1 MHz	200 kHz	100 kHz	75 kHz
Modulation technique	BPSK	CSS	DSSS (UL), FSK (DL)	TSMA	DBPSK, GMSK, or QPSK	GFSK & 4GFSK
Maximum data rate	100 bps	50 Kbps	80 Kbps	0.514 Kbps	100 Kbps	2.4 Kbps/4.8 Kbps
Permitted messages per day	140 (UL), 4 (DL)	No restriction	No restriction	No restriction	No restriction	No restriction
Communication coverage	10 km (urban), 40 km (rural)	5 km (urban), 10 km (rural)	5 km (urban), 15 km (rural)	5 km (urban), 15 km (rural)	up to 2 km/3 km	Up to 20 km
Standardization	Sigfox company and ETSI	LoRa-Alliance	Ingenu	ETSI	Weightless SIG	Wireless M-Bus
Authentication & encryption	Not defined	AES 128b	AES 256b	Not defined	AES 128b	Not defined

Table 2.2. Key technological features of licensed LPWAN technologies.

Features	**NB-IoT [37,61]**	**LTE-M [37,62]**	**EC-GSM-IoT [2,39,61]**	**Weightless-W [39]**
Operation frequency	868 MHz	700 MHz	900 MHz/1800 MHz	470 MHz to 790 MHz
Bandwidth	180 kHz/ 200 kHz	1080 kHz	200 kHz	100 kHz
Modulation technique	OFDMA (DL), SC-FDMA (UL)	OFDMA	GMSK/8PSK	DBPSK/QAM
Data rate	20 Kbps (UL), 250 Kbps (DL)	10 Kbps to 1 Mbps	70 Kbps/240 Kbps	Between 1 Kbps & 10 Mbps
Permitted messages per day	No restriction	No restriction	No restriction	No restriction
Communication coverage	1 km (Urban), 10 km (rural)	Up to 5 km	Up to 15 km	up to 5 km
Standardization	3GPPP	3GPPP	3GPPP	Weightless SIG
Authentication & encryption	LTE encryption	AES-256b	Mutual authentication	AES-128b

2.3 SHORT-RANGE WIRELESS COMMUNICATION TECHNOLOGIES

Unlike low-power long-range wireless communication technologies which are ideal for large-scale IoT applications requiring broad communication coverage, short-range wireless communication technologies mostly find usage in critical IoT applications where the main requirement is wireless communication characterized by low latency and high availability [63]. As the name implies, short-range communication technologies are ideal for applications requiring short-range connectivity, typically below 100 m [64]. Because of their short-range nature, and being relatively easy to be configured and maintained, they are more preferable to be deployed in local area networks. In the context of localization in wireless communication, radio frequency identification (RFID), Wireless Fidelity

(WiFi), ZigBee, and Bluetooth technologies are the most predominant short-range wireless communication technologies. Their key technological features and characteristics are further elaborated on in subsequent sections.

2.3.1 Radio frequency identification

RFID is a short-range wireless communication technology used to identify, position, and track objects by using radio waves. An RFID system consists of three main components: a tag, a reader and a data-processing computer [65–68]. Depending on the mechanism by which an RFID tag is powered, these tags can be classified in active, passive, and semi-active tags. Active tags are battery-powered; passive tags draw their energy from the reader's power; while semi-active tags, on the other hand, though they can power their circuits, still require assistance from the readers to transmit data [65]. The reading distance of active tags is up to 100 m, while that of passive tags is up to 10 m [51]. The operation frequencies of an RFID system can be categorized as low-frequency (LF), high-frequency (HF), ultra-high frequency (UHF), and super-high frequency (SHF) if it operates at 125 kHz frequency, 13.56 MHz frequency, 433 MHz, 860-960 MHz and 2.4 GHz frequencies, and 5.8 GHz frequency, respectively [51]. LF, HF, and UHF RFID systems can offer communication ranges of up to 10 cm, between 10 cm and 1 m, and above 1 m, respectively [69]. An RFID tag location can be estimated through distance estimation, feature matching, or proximity-based techniques by leveraging RSSI, phase, and Doppler frequency shift features being extracted along with other data by an RFID reader [65, 70].

2.3.2 Wireless Fidelity

WiFi technology belongs to the IEEE 802.11 standard for WLAN. Depending on its maximum data rate, operation frequency, bandwidth, and modulation technique, WiFi is available in several standards, namely IEEE 802.11-base version, IEEE 802.11b, IEEE 802.11a, IEEE 802.11g, IEEE 802.11n, IEEE 802.11ac, and IEEE 802.11ax [46, 71], as presented in Table 2.3. WiFi has been designed to offer high data rate communication with minimum latency, covering a communication range of up to 100 m in a local area network [49]. WiFi operates in 2.4 and 5 GHz frequencies dedicated to ISM research. The latest version of the IEEE 802.11 WLAN standard (IEEE 802.11ax) can offer wireless communication at a maximum data rate of more than 10 Gbps [72]. Because of its ability to offer wireless communication at a high data rate, WiFi is very efficient in transmitting and streaming multimedia data.

2.3.3 ZigBee

ZigBee, initially standardized by ZigBee Alliance and now by Connectivity Standard Alliance (CSA) [73], is an IEEE 802.15.4 software and hardware standard for data communication designed to offer

Table 2.3. Evolution of WLAN standards. Taken from [72], ©2019 IEEE.

Version	**Release time**	**Maximum data rate**	**Frequency band (GHz)**	**Bandwidth (MHz)**	**Modulation**
IEEE 802.11-base version	Jun. 1997	2 Mbps	2.4	20	BPSK, QPSK, DSSS, FHSS
IEEE 802.11b	Sept. 1999	11 Mbps	2.4	20	BPSK, QPSK, DSSS (CCK)
IEEE 802.11a	Sept. 1999	54 Mbps	5	20	BPSK, QPSK, 16-QAM, 64-QAM, OFDM
IEEE 802.11g	Jun. 2003	54 Mbps	2.4	20	BPSK, QPSK, 16-QAM, 64-QAM, OFDM, DSSS
IEEE 802.11n	Oct. 2009	600 Mbps	2.4, 5	20, 40	BPSK, QPSK, 16-QAM, 64-QAM, OFDM
IEEE 802.11ac	Dec. 2013	6.933 Gbps	2.4, 5	20, 40, 80, 160	BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM, OFDM
IEEE 802.11ax	Approx. 2019	>10 Gbps	< 6	-	BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM, 1024-QAM, OFDM, OFDMA

scalable, low-cost, and low-power wireless connectivity for various applications. Its communication protocol comprises physical, MAC, transmission, network, and application layers [74]. ZigBee technology can be operated in three different frequencies, i.e., 868 MHz, 915 MHz, and 2.4 GHz, offering a data rate speed reaching as high as 250 Kbps [51]. Signal transmission in ZigBee networks can cover a communication range of 10 m to 100 m, depending on the propagation condition of an environment where the ZigBee nodes are deployed [75, 76]. The nodes in a ZigBee network are categorized into coordinator, router, and end devices. ZigBee supports multiple network topologies such as tree, star, and mesh networks [51, 74, 77]. Since some nodes in ZigBee networks support path attenuation, a trilateration technique can be adopted to estimate the target node's location [78]. The fingerprinting-based localization approach through feature matching is another approach that can be

adopted in estimating the target node's location in ZigBee networks.

2.3.4 Bluetooth

Bluetooth technology, with an operation frequency of 2.4 GHz, is a communication technology belonging to the IEEE 802.15.1 standard. In the quest to improve data rate and coverage range, the Bluetooth Special Interest Group (SIG) has released several Bluetooth specifications, namely classic Bluetooth (versions 1.0, 1.2, 2.0, and 3.0), Bluetooth 4.0 (Bluetooth low energy (BLE)), and Bluetooth 5.0 [79]. A key feature of BLE is the low transmission power (0.01 mW-10 mW) [46], significantly lower than the previous specifications. Another feature of BLE is its ability to offer data rates of up to 24 Mbps, covering communication ranges of between 70 m to 100 m [49], using the frequency hopping spread spectrum (FHSS) modulation technique [46]. Bluetooth protocol based on Bluetooth 5.0 specification is characterized by an increase in data rate and coverage range (communication range of up to 300 m) compared to its predecessor [79]. Bluetooth 5.1 specification (next to Bluetooth 5.0) has been provided with the ability to extract the AoA and angle of departure (AoD) of a wireless signal. In earlier releases of Bluetooth technology (before version 5.1), a target node was localized by using distance estimations from RSSI, ToA, and TDoA values released by Bluetooth devices and scene analysis approaches, using fingerprint-based localization techniques [80]. However, the ability to provide angle-related information in Bluetooth 5.1 has made it possible to use AoA and AoD localization parameters to compute the location of the target node [79, 81]. Table 2.4 presents key technological features of popular short-range communication technologies.

2.4 LOCALIZATION PROCESS

In wireless networks, localization refers to estimating the positions/locations of target nodes (objects) in indoor or outdoor settings by making use of absolute physical locations of a few nodes (reference nodes) and inter-node measurements such as distance, angle, and time of arrival of a signal [89]. In outdoor settings, the estimated locations are commonly in the form of latitude and longitude coordinates. In indoor settings, the form with which the estimated locations are presented, varies according to the localization approach being adopted, ranging from coordinates within a custom coordinate system to a grid of tiles, each representing a geographical area in space. GPS-based localization systems are only viable in outdoor settings, and in this case, the estimated locations are directly available in the form of latitude and longitude coordinates. In contrast, GPS-free localization systems may operate in outdoor or indoor settings, and the estimated locations could be in the form of a custom coordinate system or a grid of tiles, depending on the adopted localization method. Key parameters that are usually adopted in a GPS-free localization process are the angle of arrival (AoA), ToA, TDoA, RSSI, and channel state

Table 2.4. Key technological features of short-range communication technologies.

Specifications	**RFID [68, 82–84]**	**WiFi [68, 83, 85, 86]**	**ZigBee [68, 82, 83, 85, 86]**	**Bluetooth [79, 87, 88]**
Frequency band	Passive: 125-134 MHz; Active: 433 MHz	5-60 GHz	2.4 GHz	2.4 GHz
Bandwidth	200/400 kHz	2 MHz	2 MHz	1 MHz
Standards	ISO 18000-6C	IEEE 802.11a/c/b/d/g/n/ac/ax	IEEE 802.15.4	IEEE 802.15.1
Range	10 cm, 1 m, 12 m, 100 m	20-100 m	0.03-1.6 km	Up to 300 m
Data rate	40-160 Kbps	11-54 Mbps	250 Kbps	24 Mbps
Power consumption	Low	High	Low	Low
Cost	Low	High	Low	Low
Modulation	DSSS, FHSS	DSSS/CCK	DSSS	FHSS
Security	Present, hummingbird, photon, DES	AES-128 b	AES-128 b	AES-128 b
Battery life	Passive: N/A, Active: 3-5 years	Days to months	Months to years	2 days to 14 years

information (CSI) [90].

2.4.1 Angle of arrival

In wireless networks, AoA is an incidence angle at which the transmitted signal impinges on a receiver [49, 91]. This propagation direction of the transmitted signal is measured by using antenna arrays deployed at the receiver side of the communications channel. Localizing a target node by using AoA as a localization parameter, involves the use of angle measurements from at least two different transmitters to create a triangle, through which the location of the target node is estimated as the

intersection of direction lines, using the triangulation technique [69, 92, 93].

2.4.2 Time of arrival

ToA, in the context of wireless communication, refers to the duration that a signal takes to travel from a transmitter to a receiver [69, 91, 92]. Localization approaches that use ToA as a localization parameter first compute the physical distance between the transmitter (target node) and receiver (reference node) by multiplying the speed of light with the estimated ToA. For a successful location estimation using ToA, the transmitted signal from the target node must be received by at least three reference nodes [49]. The trilateration technique can then be adopted to estimate the target node's location by using the computed distances from the target node to the reference nodes [69, 93]. ToA-based localization methods require strict clock synchronization between transmitters and receivers to estimate the location of the target node accurately [69, 91].

2.4.3 Time difference of arrival

TDoA refers to the difference in the signal's arrival times at the receiver (target node) being transmitted by multiple transmitters (reference nodes) [69, 92]. The use of a TDoA localization parameter addresses the main challenge facing ToA-based methods of requiring strict clock synchronization between transmitters and receivers. Similar to ToA-based localization approaches, localization approaches that use TDoA as a localization parameter first compute the physical distances of a signal from multiple transmitters (reference nodes) to a receiver (target node) by multiplying the differences in arrival times of a signal with the speed of light. The target node's location, which is the intersection of three or more hyperboloids resulting from the physical distances being computed from at least three reference nodes, can then be estimated by using multilateration techniques [49].

2.4.4 Received signal strength indicator

In wireless communication, RSSI refers to the strength (signal power) of emitted signals measured/observed at the communication channel's receivers [47, 94]. The closer the transmitter is positioned closer to the receiver, the higher the RSSI values, while the further the transmitter is positioned from the receiver, the smaller the RSSI values [91, 95]. Using RSSI as the localization parameter requires to first establish the path loss model that defines signal propagation characteristics of an environment where the localization process is carried out and use it in the next step to compute estimated distances from the transmitter (target node) to several reference nodes (receiving nodes). Some geometrical techniques, such as trilateration, can then be adopted to estimate the target node's location. RSSI values of the transmitted signal at different known locations can also be used to build a

database of signal features (fingerprints) and train fingerprinting-based localization models for location estimation in wireless networks through feature matching [49].

2.4.5 Channel state information

CSI, which is analyzed at the physical layer of a communication channel, describes the key characteristics of a transmitter-receiver communication link through which the signal propagates, which determines how the signal is attenuated, scattered, and refracted as it travels along that communication channel [96]. For instance, in the orthogonal frequency division multiplexing (OFDM) modulation technique, a signal that travels along multiple orthogonal sub-channels of different frequencies will exhibit different signal amplitude and phase at the receiving end of the communication channel due to differences in sub-carrier frequencies [49]. The information regarding the amplitude and phase responses captured by the CSI can then be utilized to compute the target node's location. The localization approach that can be adopted in locating a target node by using CSI is the fingerprint-based localization approach [49]. This localization approach involves collecting CSI information of a transmitted signal at some designated locations, and building a fingerprint database that can be used to train machine learning-based localization methods to estimate the target node's location through feature matching.

2.4.6 Global positioning system coordinates

GPS is a wireless communication technology based on the system of satellites deployed in the earth's orbit to provide location-based information to users. User, space (satellite segment), and control segments are three discrete parts that form a GPS [97–99]. The user segment refers to anyone who receives positional information from the satellites for civil, commercial, or military purposes. The space segment comprises 24 satellites that have been deployed strategically in orbits above the earth to ensure that at any point in time, anywhere in the world, the user can receive satellite signals from at least four satellites. The control segment is the ground control station responsible for overseeing, managing, and maintaining the overall operation of the GPS. The GPS technology essentially uses a trilateration technique to estimate the location of a target node (GPS receiver) [98, 99]. The GPS receiver has to be able to receive line of sight (LOS) GPS signals from at least three satellites for its two-dimensional (2D) position (latitude and longitude coordinates) to be estimated, while it requires GPS signals from at least four satellites for its three-dimensional (3D) position (latitude and longitude coordinates and the altitude) to be estimated [100]. Because a target node (GPS receiver) has to intercept GPS signals in LOS propagation settings for its location to be accurately estimated, GPS positioning only suits the outdoor (open space) localization scenario. The accuracy of location estimations is subject to the

number of obstructions in an outdoor localization environment. In LOS signal propagation conditions, the GPS positioning system can achieve meter-level positioning accuracy or even centimeter-level accuracy when operating in real-time kinetic (RTK) mode [101]. The type of GPS receiver can also play a significant role in the localization accuracy of the GPS positioning system. Localization accuracies of a GPS localization system that are achievable when using differential GPS (D-GPS), kinematic GPS, and consumer-level GPS devices are 50 cm to 10 m, 2 cm to 3 cm, and 15 m to 20 m, respectively [102]. Figure 2.2 illustrates the key components of a GPS.

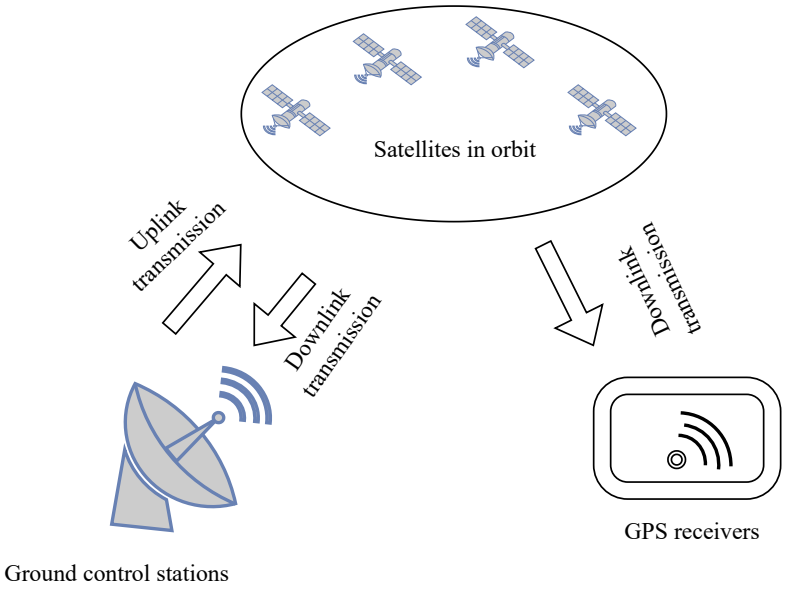

Figure 2.2. Global positioning system.

2.5 LOCALIZATION APPROACHES IN UNLICENSED LPWAN NETWORKS

In unlicensed LPWAN networks, particularly in LoRaWAN and Sigfox networks, the location of the transmitter (target node) can be estimated through the analysis of localization parameters such as RSSI, AoA, ToA or TDoA at the receiving LPWAN base stations (gateways). Range-based methods along with fingerprinting-based methods are the two localization approaches that can be adopted to estimate the locations of target nodes in unlicensed LPWANs [18]. In range-based localization approaches, a path loss model is usually applied to convert an RSS value from a specific gateway into an estimated distance to that gateway. Any available geometrical (e.g., trilateration and min-max) or statistical (e.g., Bayesian and maximum likelihood) techniques can then be utilized to estimate the final location of a node [18].

The fingerprint (matching) localization technique [103–105], which is the focus of this thesis, infers the current position of a node in a wireless network through feature matching by relying on a radio map that

has been built using signal features at pre-determined locations. The implementation of this localization method takes place in two steps. The first step involves the construction of a database (radio map) that relates scene features (such as RSSI and ToA) from videos, virtual images, or electromagnetic signals with positions of nodes in pre-determined geographic locations. In the second step of a fingerprint localization technique, the target node's real-time position is estimated by associating the measured data from the target node at a particular location with the nearest fingerprint location information being stored in the radio map being created. The localization accuracy of this localization method depends heavily on the calibration of the fingerprint database and the quality of the location-dependent measurements that have been collected in creating the database. The implementation of fingerprint-based localization methods is generally done by using machine learning techniques.

2.6 RANGE-BASED LOCALIZATION METHODS IN UNLICENSED LPWAN NETWORKS

In LPWAN networks, estimating the locations of target nodes by using range-based approaches can be done through the adoption of geometrical or statistical techniques [106]. Geometrical techniques involve first computing distances from a transmitter (target node) to several anchor nodes (reference nodes) through the use of path loss models [18]. Distances from a transmitter to anchor nodes can also be computed from multiple angle of arrivals of transmitted signals being measured at the anchor nodes. Then, depending on the adopted technique, the locations of target nodes can be estimated as the center of the intersecting areas/lines among the squares, triangles, or circles being formed, using the estimated distances. Examples of geometrical techniques/ algorithms that can be adopted in the estimation of locations of target nodes are triangulation, trilateration, multilateration, or min-max techniques [107]. In addition to geometrical techniques, statistical techniques can sometimes be adopted in estimating locations of target nodes in LPWAN networks. Examples of the adopted statistical techniques are Bayesian filtering techniques (Kalman filtering and particle filtering) and maximum likelihood techniques [108–110].

2.6.1 Trilateration

In the context of localization in wireless communication networks, trilateration is a localization technique that is employed to estimate the location of a target node, utilizing measured/computed distances from it to three non-linear anchor nodes (nodes whose physical locations are known) [111]. The initial step in the location estimation of a target node using this technique involves the use of geometric relations to plot circles (2D localization) or spheres (3D localization) by using distances from the target node to the reference nodes computed with the aid of localization parameters such as

ToA, TDoA, and RSSI [112]. The target node's location is then estimated as the point of intersection of circles/spheres, employing algebraic or geometric means [113]. A typical example of a trilateration localization process is illustrated in Figure 2.3.

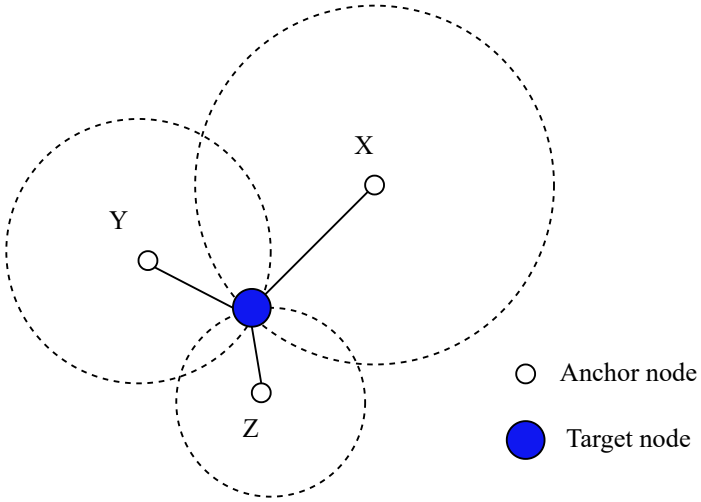

Figure 2.3. Trilateration localization technique. Adapted from [114], ©2016 IEEE.

2.6.2 Multilateration

In wireless communication, the multilateration technique is a geometrical approach to estimate the locations of target nodes. It uses at least three reference nodes (anchor nodes) and their measured/estimated distances from the target node [115]. The estimated distances from the target node to the reference nodes are computed by utilizing ToA, TDoA, or RSSI localization parameters. The location of target nodes is estimated as the intersection point of spheres and circles (derived from the computed distances from the target node to the reference nodes) for three-dimensional and two-dimensional localization scenarios, respectively. The use of more reference nodes helps to filter out outlier estimations of computed distances and consequently improves the accuracy and robustness of the multilateration localization technique [116].

2.6.3 Triangulation

Localizing a target node in LPWAN networks by using the triangulation technique requires the installation of directional antennas of the anchor nodes to measure the source signal's arrival angle. After having measured multiple angles of arrival, the triangulation technique is then used to determine the estimated location of the target node by simply using trigonometrical relationships to create a triangle, through which the location of the target node is estimated as the intersection of direction lines [69, 92, 93]. Figure 2.4 illustrates a typical triangulation process.

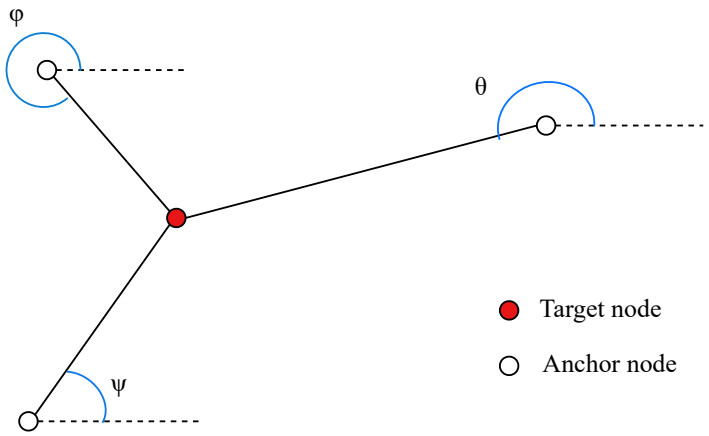

Figure 2.4. Triangulation localization technique. Adapted from [117], ©2018 IEEE.

2.6.4 Min-max and extended min-max

A min-max algorithm estimates the position of a target node from the center of the intersecting area among the squares, formed by using estimated distances between the target node and the anchor nodes [118]. On the other hand, the extended version of a min-max algorithm called e-min-max [18] estimates the location of a node based on the distances between the target node and the receiving gateways obtained through path loss models. This algorithm uses estimated distances to every receiving gateway to compose a rectangle. Then the weight of each rectangle vertex is calculated based on a specific weight function. In the final step of this algorithm, the location of the transmitter is determined by calculating the weighted average of the locations of the vertices [119].

2.6.5 Statistical techniques

Bayesian filtering techniques such as Kalman filtering and particle filtering can also be leveraged in estimating the location of a node in unlicensed LPWAN networks, as reported in [108, 120]. In the Kalman filtering algorithm [121], an unknown variable is estimated according to a joint probability distribution over the variables for each timeframe being estimated, based on a series of measurements that are being observed over time. In particle filtering techniques, Monte Carlo approximations represent a target distribution via a set of particles that have been distributed according to the target density. The weight in each particle is updated with respect to the probability which corresponds to a measurement at a receiving LPWAN gateway, and through sequential importance resampling (SIR), particles with large weights are duplicated. In contrast, those with small weights are removed [122]. Maximum likelihood (ML) is another type of statistical technique that can also be leveraged in

estimating the location of a node in LPWAN networks, as reported in [106]. This technique is based on the minimization of the mean square error (MSE) of the estimated distances between the target node and anchor nodes [123]. The localization accuracy using this technique, increases with the increase in the number of measurements by minimizing the error variance.

2.7 RANGE-BASED LOCALIZATION METHODS IN LORAWAN NETWORKS

Several research works exist in the literature that adopt a range-based localization approach to estimate the locations of target nodes in LoRaWAN networks, using either RSSI, AoA, ToA, or TDoA measurements or a combination of more than one of these parameters that have been extracted from LoRaWAN devices. The range-based localization approaches are implemented by adopting geometrical techniques (such as triangulation and multilateration) or statistical techniques (such as maximum likelihood and Bayesian filtering).

In [124], the authors proposed a localization scheme in LoRaWAN networks by using RSSI measurements. In this localization scheme, the path loss model is first established from the communication links of the target node and the anchor nodes, followed by the location estimation of a target node, by using a trilateration algorithm. In [125], a localization method is proposed whereby the location of a target node in the LoRaWAN network is computed by using TDoA values that have been measured from the signal being transmitted by the target node and received by several gateways. In this scheme, the issue of asynchronization between different gateways, which may affect the overall localization accuracy, is addressed by using an additional stationary node. The authors in [126] proposed a localization approach that utilizes TDoA measurements that were estimated from differential phase sampling being applied in a LoRaWAN uplink signal for node localization in LoRaWAN networks. In this scheme, a least square algorithm is applied to compute the location of the target node in the back-end server by integrating the anchor's reference positions and the TDoA values.

In [108], an approach to use TDoA measurements to localize a node in LoRaWAN networks is proposed and implemented in NS-3. In order to increase the accuracy of the proposed localization scheme, a Kalman filter is used to remove clock synchronization errors before adopting the Chan algorithm to infer the location of the target node in two scenarios: one involving three gateways and another involving more than three gateways. In [109], a super-resolution localization scheme based on AoA is proposed to localize a node in LoRaWAN networks. In order to improve the localization performance of the proposed scheme, bandwidth is increased through the synchronization of multiple

communication channels first, before adopting an ESPRIT algorithm to compute the location of a LoRaWAN transmitter. The authors in [127] proposed a localization scheme in LoRaWAN networks based on the combination of TDoA and AoA parameters. In this scheme, two probability density maps, one for TDoA measurements and another for AoA measurements, are built first and then combined into a new map whereby the final location of the target node is computed from the intersection of the merged AoA and hyperbola resulting from TDoA measurements from two gateways.

Two key advantages of deploying range-based localization approaches exist. Firstly, they can easily be deployed and used as soon as the path loss model has been established. Secondly, if the requirement of the presence of at least three gateways to cover an entire localization area has been met, these approaches can be deployed anywhere. However, the adoption of range-based localization approaches in LoRaWAN networks is less attractive to researchers for several reasons. The first reason is the requirements of dedicated hardware in their implementation. For instance, AoA-based approaches require the installation of an array of antennas for angle measurements, which can be very expensive, while ToA and TDoA-based localization approaches require accurate clock synchronization among anchor nodes. The second reason is the poor performance of range-based localization approaches due to fluctuations in localization parameters such as RSSI being caused by shadowing and fading phenomena due to multipath propagation [106]. For these localization approaches to achieve acceptable levels of localization performances, researchers must find ways to limit the effect of a significant number of outliers on the localization parameters that may occur due to the heterogeneous nature of the outdoor or indoor environment, as well as wireless propagation effects, such as fading and shadowing due to multipath propagation [18].

2.8 FINGERPRINT-BASED LOCALIZATION APPROACHES

As pointed out in [51, 103–105], the fingerprint-based technique is a localization technique that consists of two phases, namely offline and online phases. In the offline phase, a scene analysis is conducted to obtain scene features from videos, virtual images, or electromagnetic signals, and to build a database that relates these features with the positions of nodes in pre-determined physical locations. The constructed database of fingerprints constitutes the radio map for the area where the localization process is conducted. For example, the analyzed scene features could be RSSI, ToA, TDoA, or CSI. The online phase estimates the target node's real-time position by comparing the measured data from the target node at a particular location with the nearest fingerprinting location information from the constructed radio map.

Fingerprint-based localization approaches are implemented by using machine learning techniques. Machine learning refers to an evolving branch of computational algorithms being designed to emulate human intelligence by drawing ideas from several disciplines, such as artificial intelligence, computer science, probability and statistics, information theory, psychology, control theory, and philosophy [128]. Machine learning algorithms are categorized into supervised, unsupervised, semi-supervised, and reinforcement learning algorithms [129, 130]. In supervised learning, a training dataset containing examples for the input and labeled answers or target values for the output is required, while in the case of unsupervised learning, an algorithm learns by detecting patterns without the need for any pre-existing labels or specifications [131]. Semi-supervised learning requires only partially labeled training data (a lot of unlabeled data and a small quantity of labeled data) [132]. Semi-supervised learning can be viewed as the combination of both supervised and unsupervised learning, where part of the dataset is partially labeled to infer the unlabeled portion of the dataset. Reinforcement learning algorithms do not require to be presented with input and output pairs; instead, they require the specification of a goal and provide a list of allowable actions and their environmental constraints for their outcomes to enable the model to experience the process of achieving the goal by itself by using the principle of trial and error to maximise a reward [133].

Common supervised machine learning algorithms reported in the literature are decision trees (DT), support vector machines (SVM), Bayesian algorithms, k-nearest neighbor (KNN), random forest (RF), association rule (AR) algorithms and neural networks, while k-means clustering, principal component analysis (PCA), autoencoders and restricted Boltzmann machines are typical examples of unsupervised learning [129, 134–137]. Semi-supervised learning algorithms often combine unsupervised and supervised algorithms, such as deep belief networks.

In the context of fingerprint-based localization methods in unlicensed LPWAN networks, the predominant machine learning techniques adopted are KNN, SVM, DT, RF, k-means, multi-layer perceptron, linear regression algorithms, and deep learning. The sections below elaborate on these techniques.

2.8.1 K-nearest neighbors

The KNN algorithm, which falls under the category of non-parametric algorithms, is a machine learning algorithm being used predominantly for tackling classification tasks. The first step in implementing this algorithm is to categorize sample points of a dataset into different groups, depending on the overall

features of the dataset. The next step is to determine to which group the new data points (from the test set) belong by computing their corresponding Euclidean distances to the training set data points [138]. The decision to classify new data points as to which group they belong is determined by the number of 'neighboring' data points (k-data points) closest to the new data points, which are considered (from all the dataset categories) in the implementation of the algorithm. For instance, if k is set to 1, then one neighboring data point to the new data points, each from all the dataset categories, are selected, and their Euclidean distances to the new data points are then computed with the closest neighboring data point determining the group to which the new data points belong. The same applies when k is set to 2, 3, 4, etc., where the neighboring data points that will be considered will correspondingly be 2, 3, 4, etc. Figure 2.5 illustrates the working principle of the KNN algorithm. Suppose a blue point is to be classified as to whether it belongs to the red dots category or the green dots category. If k is set to 1, then the KNN algorithm will classfy the blue dot as belonging to the green category. If k is set to 2, the blue dot will be classified as belonging to the green dots category. However, when k is set to 3 or 4, the blue dot will be classified as belonging to the red dots category. The most challenging part in implementing the KNN algorithm is the choice of the number of neighboring data points ('k') which will give optimal classification results. It is always recommended to conduct cross-validation by picking several values of 'k' and running the KNN algorithm, using each one of them to find the one value that gives optimal results [139], a process that makes this algorithm computationally demanding and time-consuming [140].

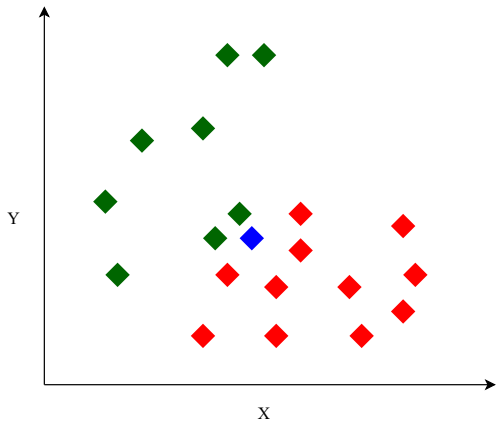

Figure 2.5. KNN algorithm. Adapted from [134], ©2020 IEEE.

2.8.2 Support vector machines

SVM is a machine-learning algorithm mostly used to tackle classification problems, though with some adjustments, it can also solve regression-based problems. In this algorithm, samples of a dataset of 'n'

features are represented in an n-dimensional space, with the values of each feature corresponding to the values of specific coordinates [141]. This higher-dimensional representation of data points is done through the use of kernel functions such as linear, polynomial, or radial basis function (RBF). The choice of the ideal kernel function to tackle a particular task is crucial to getting the best out of an SVM algorithm; otherwise, it might give unreliable results. A rule of thumb in selecting the best possible kernel function to solve classification/regression tasks by using the SVM algorithm is to use the linear function for the training dataset whose data points are linearly related and to use polynomial or RBF functions for non-linear data points [129]. This algorithm works by maximizing the distance between different classes by identifying large-margin hyperplanes among different classes/values [142]. Its working principle is to extract subsets of data points from a training set that are closest to the decision surface and to find a separating hyperplane in the feature space with the largest margin possible that maximizes the distance between different classes. In the context of an SVM algorithm, these data points closest to the decision surface are referred to as support vectors, and their corresponding distance to the hyperplane is referred to as the margin [129]. Being the closest data points to the decision surface increases the difficulty in classifying the support vectors, so finding the largest margin between them is key to getting the best possible decision region for each class (for a classification task) or value (for a regression task) [142]. According to [135, 139], the SVM algorithm better suits a classification/regression task for datasets with many features but with a relatively small number of data samples. Figure 2.6 illustrates an SVM classifier built by using linear mapping.

2.8.3 Decision tree

DT is a machine-learning algorithm for solving regression and classification tasks. Its implementation involves dividing a dataset into several subsets of data points according to a certain set of rules to form a tree-like structure that is used to make decisions by evaluating the feature vector [143]. The vertex nodes of the created tree represent the features in a dataset, the branches contain corresponding values each vertex can have in a test sample, while the leaf nodes represent the final outcome (class) for each instance [139]. A decision node of a DT algorithm associates test attributes to desired choices in the branch nodes, as well as to the final outcomes (classes) on the leaf nodes. The DT algorithm is implemented by starting with a tree with unoccupied nodes, and then a root node (vertex) of the tree is decided by the feature which best splits the training sample. Based on the information gain and Gini index [139], branches and leaves of the tree will consequently be created. The tree will keep on 'growing' as long as its maximum depth is yet to be attained or when there is a possibility of creating new leaf nodes that can decrease the entropy (evaluation criterion of a DT algorithm); otherwise, it

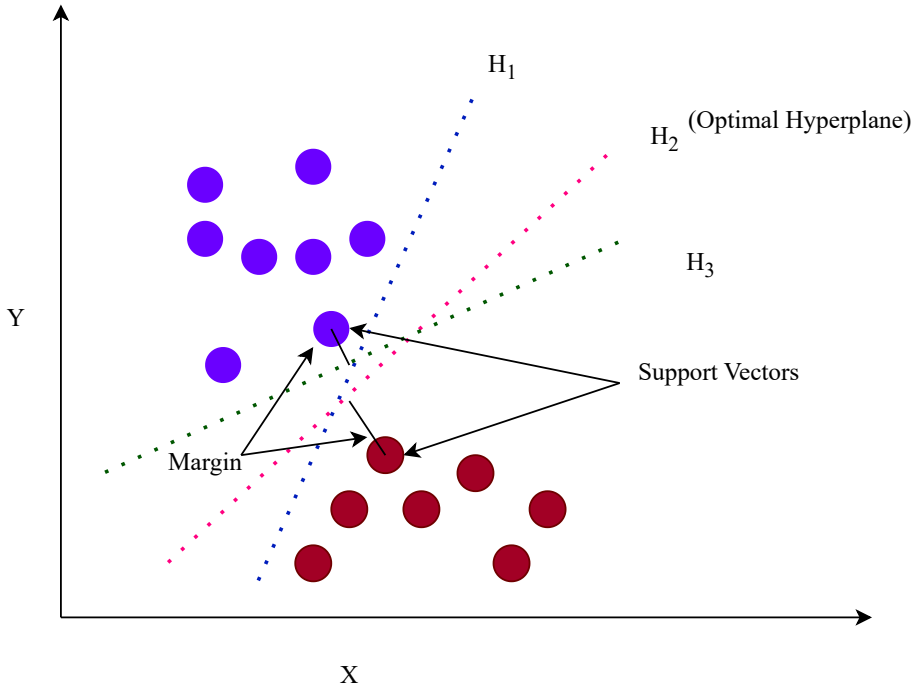

Figure 2.6. SVM algorithm. Adapted from [129], ©2019 IEEE.

ceases to grow any further. After the creation of the tree, the next step is to get rid of all branches with outliers through tree-pruning. At the inference stage of the DT algorithm, the best class (classification), or value (regression) of a test sample is decided through the implementation of a series of 'if-then' rules at the decision and leaf nodes of the created tree, which represent 'choices' and 'final outcomes', respectively [144]. Then, a leaf node with the smallest entropy (cost function) is selected as the final outcome of a DT regressor or classifier. The cost function in a DT algorithm is used to find a group of branches with the most similar features while its depth helps to control the model's training to prevent it from over-fitting [135]. Figure 2.7 illustrates an example of a tree in a DT algorithm.

2.8.4 Random forest

RF is an ensemble machine-learning algorithm that trains multiple decision trees in parallel through bootstrapping to improve the generalization ability of a machine-learning regressor/classifier [144]. The first step in implementing an RF algorithm is to create multiple decision trees based on the features of the training dataset by observing a set of if-then rules similar to what is done in DT algorithms. Some trees will produce accurate results, while others will produce wrong results. The second step is to aggregate results from the multiple uncorrelated trees and pick up the most dominant value

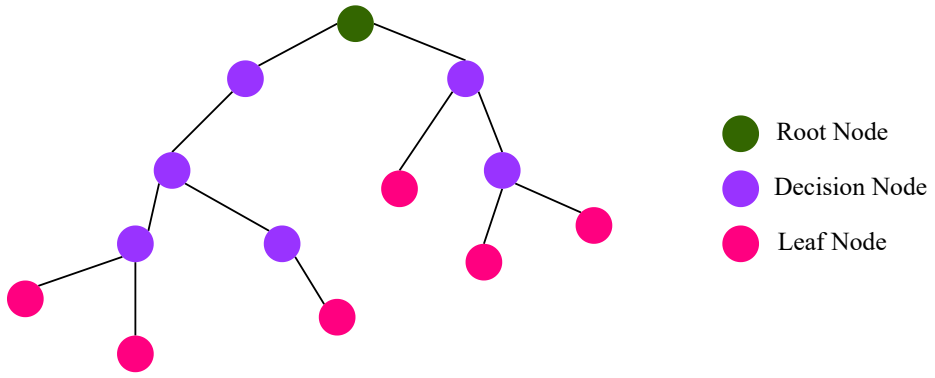

Figure 2.7. DT algorithm. Adapted from [140], ©2020 IEEE.

(regression problem) or class (classification problem) from the multiple trees as the final outcome of the RF regressor/classifier [134, 135]. The fact that the final decision being made by the RF algorithm is a result of aggregated decisions from multiple unique decision trees makes this algorithm less prone to over-fitting the training set. Consequently, it leads to better results than using only a single tree, which is the case for a DT algorithm. Figure 2.8 illustrates an example of a random forest.

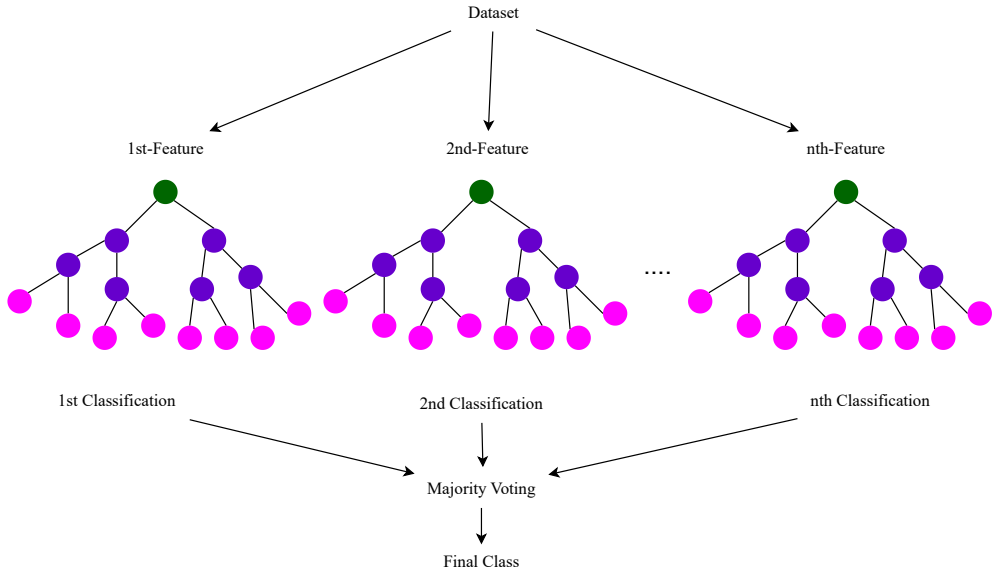

Figure 2.8. RF algorithm. Adapted from [135], ©2021 IEEE.

2.8.5 K-means

K-means clustering (with ‘k’ and ‘mean’ referring to the number of clusters and mean of attributes, respectively) is a machine-learning algorithm whose implementation involves partitioning a dataset into k non-overlapping clusters, each comprising samples with similar characteristics. [135, 139]. It

assigns each new data point to one of the k-clusters based on the related features [139]. Upon creation of k-clusters, each cluster will have data points of similar characteristics. The implementation of this algorithm requires two input data sets: the number of clusters ('k') and the training dataset, which comprises individual samples with a set of features to be learned. As put forward in [129, 134, 139], assigning data samples to their appropriate clusters follows several steps. The first step is to randomly initialize the centroids of all k-clusters, followed by the computation of squared Euclidean distances between a data sample and each of the cluster centroids. Then, based on the squared Euclidean distance, the data sample will be placed in the cluster whose centroid is closest to the data sample. Once all the samples have been placed in their relevant clusters, the algorithm updates each cluster's centroid iteratively through computation of the mean of all samples that it contains until the convergence condition is valid. Figure 2.9 illustrates a working principle of a k-means algorithm.

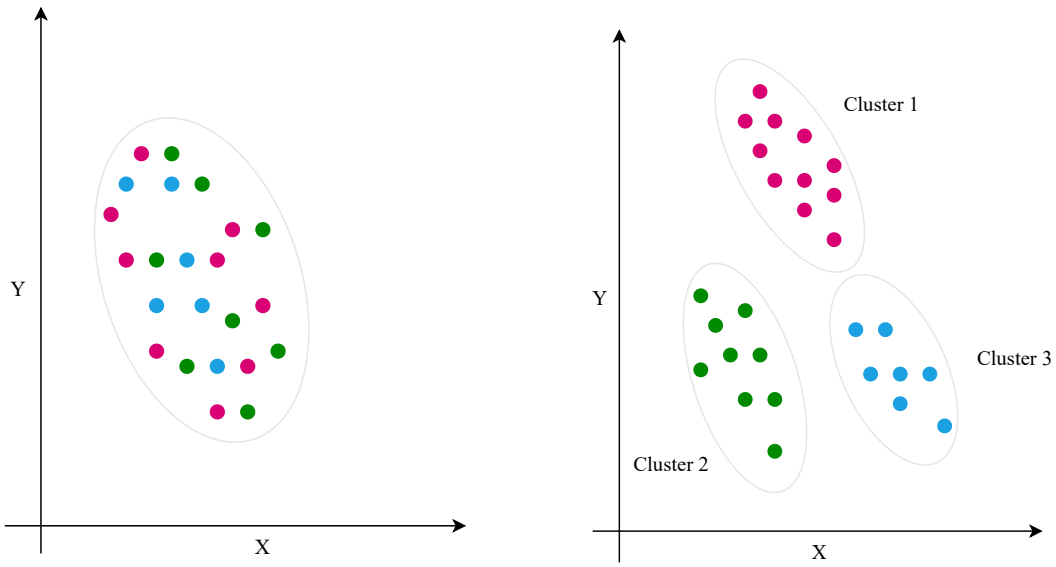

Figure 2.9. K-means algorithm. Adapted from [139], ©2021 IEEE.
(a) Before clustering. (b) After clustering.

2.8.6 Artificial neural network/Multi-layer perceptron

ANN is a machine learning algorithm developed by imitating how the biological neurons of the human brain work. An ANN model consists of an input layer, at least one hidden layer (comprising several neurons), and an output layer. Neurons, which are non-linear activation functions [145], connect to

one another to share relevant information to enable the mapping of the input values to an output value. The information from multiple neurons in hidden layers is aggregated in a layer and transformed according to some mathematical computations to form new values that are used as input data for the next layer [143]. During the training process of an ANN model, the weights and biases (attributes that control the learning process of a model) of subsequent layers are updated based on the input data from the preceding layers. At the same time, the values in each neuron are computed as well. The model's training follows a sequence of forward pass and backpropagation circles to minimize the cost function and enhance the model's performance [140, 143]. Figure 2.10 illustrates an example of a feedforward ANN. In a feedforward ANN, as the name implies, the information flow is in one direction only, i.e., from the input layer through to the hidden layers and subsequently to the output layers. A typical example of a feedforward ANN is a multi-layer perceptron (MLP).

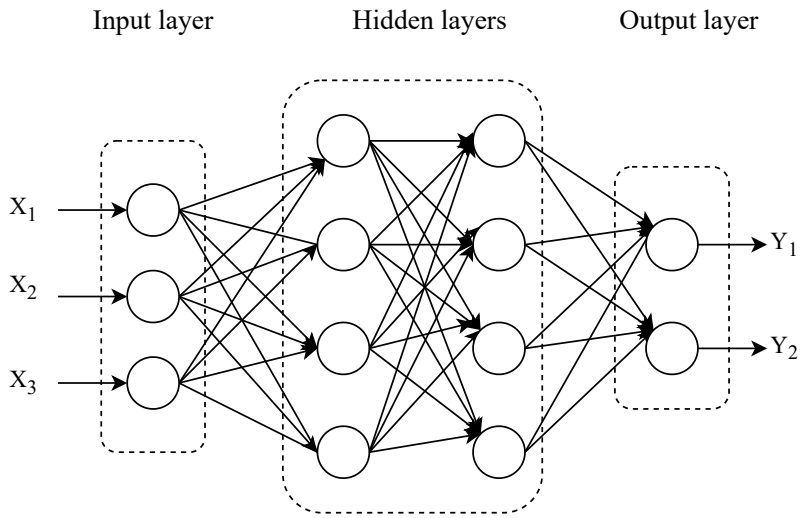

Figure 2.10. ANN algorithm. Adapted from [146], ©2022 IEEE.

2.8.7 Linear regression algorithms

A linear regression algorithm is a machine learning algorithm that models a target variable (dependent variable) based on one or more independent variables. It is a statistical-based technique deployed to predict a value based on certain inputs by plotting a straight line that best describes the relationship existing between the target variable and the independent variables [141]. The best-fitting line (regression line) is usually established by minimizing errors between the actual and the predicted/estimated data points [135]. As reported in [18], a linear regression model is expressed as

$$y = w[0] * x[0] + w[1] * x[1] + ... + w[n] * x[n] + b, \tag{2.1}$$

where x is a feature vector of size n with w and b being the learnable parameters. The ordinary least square (OLS), ridge, least absolute shrinkage and selection operator (LASSO), elastic net, stochastic gradient descent (SGD), and polynomial regression are typical examples of variations of linear regression algorithms adopted for fingerprinting localization in LPWAN networks [18]. The OLS algorithm works by minimizing the mean squared error (MSE) loss/cost function (residual sums of squared error) between the estimated/predicted data points and the actual data points [147, 148].

Ridge is a variant of linear regression algorithms, which, unlike the OLS algorithm, has a bias/penalty term to control over-fitting. It is based on the $L2$ regularization technique since the added bias is the squared sum of the weights. As reported in [149], mathematically, the ridge algorithm minimizes a loss/cost function expressed as

$$Loss_{Ridge} = MSE(y_{true}, y_{pred}) + \alpha \sum_{i=1}^{n} \psi_i^2, \tag{2.2}$$

where $MSE(y_{true}, y_{pred})$ is the mean squared error between predicted and actual values, $\sum_{i=1}^{n} \psi_i^2$ is a bias term which is a squared sum of weights, α which is >=0, is a coefficient value that regulates shrinkage. Increasing this value to higher values leads to more shrinkage (under-fitting) of the model while decreasing this value leads to less shrinkage of the model [149, 150]. The ridge algorithm is known to be effective in solving regression tasks involving datasets with relatively small sample data but with a relatively high number of features [151].

LASSO is another form of linear regression algorithm that also works by adding a bias/penalty term to the sum of squared residuals between the predicted and the actual values to control over-fitting of the model [147, 150]. Unlike in the ridge algorithm, where the added bias term is α multiplied by the sum of squares of the weights, in the LASSO algorithm, the added bias term is the sum of the absolute value of the weights multiplied by a factor α ($L1$ regularization). As pointed out in [149], mathematically, the LASSO algorithm attempts to minimize a loss function expressed as

$$Loss_{LASSO} = MSE(y_{true}, y_{pred}) + \alpha \sum_{i=1}^{n} |\psi_i|, \tag{2.3}$$

where $MSE(y_{true}, y_{pred})$ is the mean squared error between predicted and actual values, $\sum_{i=1}^{n} |\psi_i|$ is a bias term which is the sum of absolute values of the weights, α which is >=0, is a coefficient value that regulates shrinkage. The learning mechanism adopted in this algorithm is to shrink data values to the central value (mean), with α being the parameter that controls the degree of shrinkage.

The elastic net is a variant of linear regression algorithms that combines and adds both the LASSO's

$L1$ and ridge's $L2$ norms (penalty terms) to the sum of the squared residuals between the predicted and actual values [149, 152]. According to [151], its working principle is to minimize a loss function expressed as

$$Loss_{ElasticNet} = MSE(y_{true}, y_{pred}) + \alpha_1 \sum_{i=1}^{n} |\psi_i| + \alpha_2 \sum_{i=1}^{n} \psi_i^2. \tag{2.4}$$

Gradient descent is one of the popular optimization techniques used to train artificial neural networks and linear regression models by finding optimal values that minimize a cost function through iteratively updating the initial learnable parameters [149]. In the SGD algorithm, the choice of initial learnable parameters is done randomly to reduce the computational complexity of machine learning models that have been trained by using relatively large datasets [18].

The polynomial regression algorithm is a form of linear regression algorithm being implemented by fitting a large variety of nonlinear data into a higher-dimensional space [18, 149]. According to [153], in this algorithm, dependent and independent variables are modeled into a high-order polynomial model expressed as

$$y_i = \alpha_0 + \alpha_1 x_i + \alpha_2 x_i^2 + \alpha_3 x_i^3 + ... + \alpha_k x_i^k + e_i, \tag{2.5}$$

where α_j is a jth coefficient of the regression model, k is the degree of the polynomial, and e_i is a random error.

2.8.8 Deep learning

Deep learning is a machine learning subdomain that uses mathematical computations to map the input to the output through multi-layer transformations [135]. The superior performance of deep learning algorithms over traditional machine learning algorithms in solving different types of tasks is attributed to their ability to perform automatic feature extraction from the training datasets. Deep learning algorithms can learn hidden patterns within data and consequently make relevant predictions as a result of accelerated computational power and massive data [135]. The emphasis in deep learning is on learning successive layers that have increasingly meaningful representations. The term 'deep' in deep learning refers to the multiple successive non-linear layers of interconnected artificial neurons that are used to find useful patterns from the training dataset [154, 155].

In literature, different types of deep neural networks exist, which act as the baseline for the pre-trained models in deep learning; the common ones are ANN, CNN, and recurrent neural networks (RNN) [146]. Because inputs are processed in the forward direction, ANN is termed a deep feed-forward neural network that is capable of learning non-linear functions. The RNNs were designed to address the

looping constraint of ANN in hidden layers. Because of its capability to capture sequential information available in the input data, deep recurrent neural networks can solve problems related to audio, text, and time-series data. The filters or kernels are the key building blocks of convolutional neural networks. With the help of conventional operations, the kernels in CNN models can extract relevant and correct features from the input data [143]. Other deep learning networks that are increasingly gaining popularity are long-short-term memory networks (LSTMs), deep belief networks (DBNs), generative adversarial networks (GANs), restricted Boltzmann machines (RBMs), radio basis function networks (RBFNs), and autoencoders [139].

Deep learning finds applications in almost every field, such as finance, image recognition, health care, and localization/positioning, to name a few. In the context of node localization in unlicensed LPWAN networks, LoRaWAN and Sigfox in particular, the popular deep learning models being adopted to estimate locations of target nodes through fingerprint-based approaches, are the LSTMs and CNNs. In Chapter 3, CNNs, along with transformer modules, which are the main building blocks of this thesis's proposed fingerprinting-based localization methods, will be elaborated upon.

2.9 FINGERPRINT-BASED LOCALIZATION METHODS IN LORAWAN NETWORKS

Researchers are increasingly being attracted to adopting fingerprint-based localization approaches compared to the range-based counterpart, due to their robustness in challenging environments with multi-path and NLOS phenomena [156] and being relatively more accurate. This is attributed to their ability to learn useful positional information even from noisy data [11].

The authors in [9] took part in a large-scale outdoor measurement campaign to create fingerprint databases for Sigfox and LoRaWAN networks to equip researchers with a tool to verify the performance of their localization algorithms. In addition to these datasets, the authors implemented a KNN fingerprint-based localization method and evaluated it, using their LoRaWAN dataset (version 1.1 of their urban LoRaWAN dataset), achieving a localization accuracy of 398.4 m mean error. The follow-up research, which adopted the same dataset version of the LoRaWAN dataset to evaluate their fingerprint-based localization methods, is presented in [11, 157]. In [157], the authors implemented KNN, extra trees and MLP fingerprinting methods, reporting localization accuracy of 357 m mean errors with the best-performing MLP fingerprinting model. The researchers in [11] implemented three fingerprint localization methods, reporting a localization performance of 191.53 m with the LSTM

method, which outperformed the other two methods based on ANN and CNN. The difference between the methods proposed in [11] compared to the two methods proposed in this thesis is that in [11], the authors proposed three fingerprint-based interpolation-aided localization methods whereby the first method used only LSTM, the second method used only CNN and the third method used only ANN. The two methods proposed in this thesis, in addition to the use of CNN, SE blocks and transformer modules are used to improve the efficiency and localization accuracies of the proposed methods.

Fingerprint-based localization methods reported in [18, 158–160] were evaluated using version 1.2 of the urban LoRaWAN dataset presented in [9]. Authors in [18] implemented ten different types of regression algorithms along with the extended min-max algorithm and reported the best localization performance of 340 m mean errors achieved with the RF algorithm. In [158], a KNN-RF method was implemented by utilizing hybrid data and it achieved a localization accuracy of 332.63 m mean errors. The researchers in [159] implemented and trained RF and MLP fingerprint-based localization methods by using a RSSI-TDoA differential database, achieving better performance with the MLP method, reporting a mean error of 310 m on the test set. In [160], a hierarchical clustering-based technique was proposed for fingerprinting localization in the LoRaWAN network. With the weighted kernel regressor, the proposed localization approach was able to achieve 346.03 m mean error.

The authors in [161] presented a case study, using data augmentation techniques to improve the localization performance of SVR, KNN, extra trees, and MLP fingerprinting algorithms that were trained on small datasets. The best localization accuracy of 12 m mean localization error was recorded, using the KNN algorithm at an outdoor urban area covering 8 km^2. The researchers in [162] performed a comprehensive evaluation of different strategies that could be used to improve the spatial resolution of small radio maps by using large radio maps through the adoption of inter-technology knowledge transfer. The evaluated methods were interpolation methods using radius-inverse distance weighting (IDW), Gaussian radial basis function (RBF), and regression methods using RF and neural networks.

Research work in [163] implemented and evaluated SVR and Gaussian process regression fingerprinting localization methods to localize a node in a sandstorm environment. From the experimental results, the SVR method was reported to have better localization performance than the Gaussian process regression method. In [10], through the implementation of the KNN fingerprinting localization algorithm, the authors analyzed factors that could influence the accuracy of fingerprint-based localization approaches in an outdoor setting. This work analyzed accuracy dependencies based on the number of deployed

gateways, coverage area and the distance from one measurement point to another. The authors in [5] collected outdoor RSSI fingerprints and deployed path-loss and different machine-learning models to improve RSSI-to-distance representation. The optimal model was able to achieve localization performance of between 6 and 15 m mean errors in the deployment area. Table 2.5 presents a summary of key features of range and fingerprint-based localization methods currently available in the literature, proposed to estimate locations of target nodes in LoRaWAN networks.

The robustness and effectiveness of fingerprint-based localization methods are attributed by the use of machine learning/deep learning algorithms which have proven very effective in tackling different tasks because of the ever-increasing processing and computational powers of computers. As a result, fingerprint-based localization approaches can lead to relatively high localization accuracy compared to range-based localization approaches due to their ability to effectively learn useful positional information even from RSSI (or any other localization parameter) data being collected in NLOS environment [11]. However, several disadvantages are associated with fingerprint-based localization approaches, which may limit their implementation in one way or another. Firstly, fingerprint-based localization approaches require an outdoor/indoor training database, which takes significant time, effort, and cost to build. Secondly, in the context of unlicensed LPWAN networks, the messages whose data is used to build the fingerprint databases for localization purposes, need to be collected with respect to the uplink duty cycle regulations, which may necessitate more extended data collection periods, even for months. In unlicensed LPWAN networks, Sigfox and LoRaWAN in particular, crowd-sourced fingerprinting initiatives such as ‘the things network’ can assist in reducing the cost and time of building large outdoor fingerprint databases [18].

2.10 CHAPTER SUMMARY

In this chapter, a comprehensive review of short and long-range wireless communication technologies adopted to implement IoT applications was presented. The focus was on their technological aspects and adoption in IoT applications. Key localization parameters used to implement the localization methods/techniques being adopted to estimate the target object’s location, were also discussed. Additionally, two types of localization approaches, namely range-based and fingerprint-based localization approaches, which can be adopted in estimating locations of target nodes in unlicensed LPWAN networks, particularly in LoRaWAN networks, were presented and discussed. For each of the localization approaches, types of localization methods/techniques associated with it and currently available in the literature, were presented along with a discussion on their implementation, their strengths and

Table 2.5. A Summary of the key features of the reported localization methods in the related works.

Research work	Localization approach	Localization parameter	Nature of the dataset	Localization environment	Adopted algorithm(s)
[124]	Range-based	RSSI	Real	Outdoor	Trilateration
[125]	Range-based	TDoA	Simulated	Outdoor	Multilateration
[126]	Range-based	TDoA	Real	Indoor	Least Square
[108]	Range-based	TDoA	Simulated	Indoor/outdoor	Chan
[109]	Range-based	AoA	Real	Indoor/outdoor	ESPRIT
[127]	Range-based	TDoA and AoA	Real	Outdoor	Triangulation/Trilateration
[9]	Fingerprinting	RSSI	Real	Outdoor	kNN
[157]	Fingerprinting	RSSI	Real	Outdoor	kNN, MLP, Extra Trees
[11]	Fingerprinting	RSSI	Real	Indoor/outdoor	LSTM, ANN, CNN
[18]	Fingerprinting/ Range-based	RSSI	Real	Outdoor	Ten different regression algorithms plus extended min-max algorithm
[158]	Fingerprinting	Fused RSSI-TDoA	Real	Outdoor	kNN-RF
[159]	Fingerprinting	Fusion of differential RSSI-TDoA	Real	Outdoor	MLP, RF
[160]	Fingerprinting	RSSI	Real	Outdoor	k-means + weighted kernel regression
[161]	Fingerprinting	RSSI	Real	Outdoor	Support vector regression, extra trees, KNN, MLP
[162]	Fingerprinting	RSSI	Real	Outdoor	Its implementation is based on inter-technology knowledge transfer, using classical machine learning algorithms
[163]	Fingerprinting	RSSI	Real	Outdoor	Support vector regression and Gaussian process regression
[10]	Fingerprinting	RSSI	Real	Outdoor	kNN
[5]	Fingerprinting/ Range-based	RSSI	Real	Outdoor	Trilateration, DTs, KNN, SVM

weaknesses, and their overall accuracy in locating a target node in unlicensed LPWAN networks. Currently available range and fingerprint-based solutions proposed in the literature were also presented and discussed to contextualize the status of node localization research in LoRaWAN networks.

CHAPTER 3 BACKGROUND METHODS

3.1 CHAPTER OVERVIEW

In this chapter, a background knowledge of the main building blocks of this thesis's proposed fingerprinting-based localization methods is presented. Firstly, a typical CNN model is presented, then a description of its main features, followed by a discussion on how to improve the generalization abilities of CNN models by fighting the over-fitting phenomenon (a prevailing challenge facing CNN models) through the use of regularization techniques. Key network parameters needed to be defined in the compilation step of CNN models are discussed as well. A brief description of how a CNN model extracts important features from input data is also provided to get a summarized insight into its working principle. Additionally, the role of the squeeze and excitation block (SE) in enhancing the performances of CNN models by improving channel-wise interdependencies is introduced and discussed. Lastly, the transformer module is introduced and discussed, focusing on its main features and general working mechanism.

The remaining sections of this chapter are organized as follows: In Section 3.2, CNNs are introduced, whereby their key features, regularization techniques, and compilation steps are discussed. In this section, the SE block, which has proven very effective in improving channel-wise interdependencies in CNN models, is also presented and discussed. In Section 3.3, the transformer module is presented, whereby its architectural composition, computational mechanism, and the different ways that it can be adopted in solving different tasks are some of the insights featured in this section. Conclusive remarks on this chapter are provided in Section 3.4.

3.2 CONVOLUTIONAL NEURAL NETWORK

CNN is a special case of feedforward neural networks that has proven to be very effective in tackling regression/classification problems/tasks, using grid-structured datasets [141]. The assumption that

the input data is in a grid-like structure makes it easier to encode certain properties to the CNN architectures, allowing effective extraction of important features from the input data [164]. This eventually improves their learning capabilities, on account of having neurons with learnable weights and biases. A traditional CNN model comprises at least one convolutional layer, at least one pooling layer, and at least one fully connected layer. These layers are stacked together sequentially in such a way that each layer can transform the activations/outputs of the previous layer, using a back propagation mechanism [164]. In addition to these layers, it comprises an activation unit to implement non-linearity to each element in a feature set. The state-of-the-art performance achieved by CNN when applied in application areas (ranging from image/video processing/recognition, image segmentation, object tracking/localization, object detection, speech and natural language processing, to name a few) is attributed to their robustness to positional shifts of features in the training data through sharing of weights among different convolutional kernels/filters, subsequently creating feature maps of equivalent translation [141]. Figure 3.1 illustrates an example of a CNN architecture.

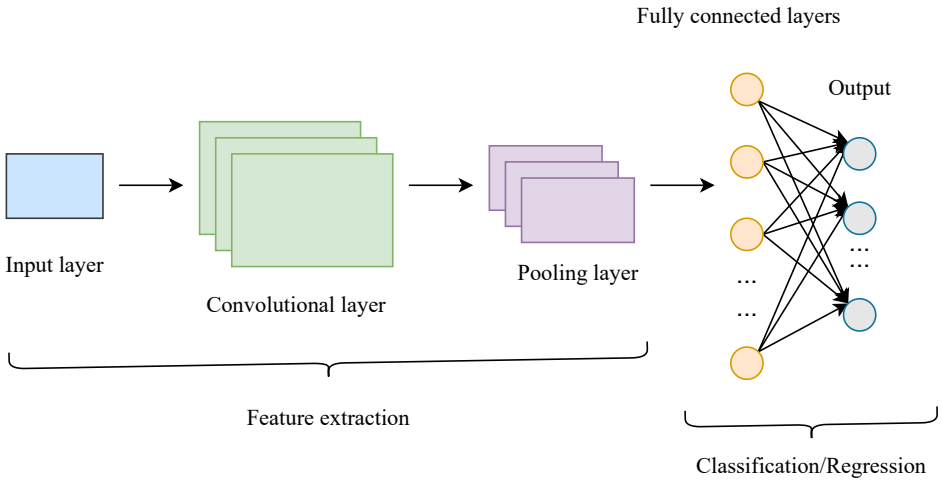

Figure 3.1. A typical CNN architecture. Adapted from [165], ©2024 IEEE.

3.2.1 Features of a convolutional neural network

The features of CNN models are specific to the problems to be solved; however, most CNN models consist of convolutional layers, convolution filters/kernels, pooling layers, fully connected layers, and activation functions.

Convolutional layer: The convolutional layer, the main building block of CNNs, is used as a feature extractor to learn and extract key representations that best describe the distinctive spatial features of the input data by using learnable tiny matrices, referred to as convolution kernels/filters. This layer

produces different feature maps by sliding the kernels/filters across the input feature map (output feature map of the previous layer) through computation of element-wise products of corresponding values of the kernel/filter and receptive field, and adding up the resulting products to get the activation value at that particular point [164, 166]. During the training of a CNN model, by using the learnable filters/kernels, the lower (initial) convolution layers capture high-level spatial features of the input feature map, while the higher (deeper) convolution layers capture low-level spatial features of the input feature map [166].

Pooling layer: In CNN, pooling layers are used for one main purpose: spatially down-sampling the convolutional layers' output feature maps to relatively small sizes so that the subsequent convolutional layers can effectively extract useful/relevant features from them [154]. Since the pooling layers shrink larger feature maps to smaller ones, they reduce the number of trainable parameters as a result, as well as lowering the network computation burden [167], [168], making the network translation invariant [166, 169]. Down-sampling of large feature maps can be implemented by adopting any of the following popular pooling techniques, i.e., through finding the maximum value (max pooling), the minimum value (min pooling), the average value (average pooling), or gated value (gated pooling), corresponding to each rectangular neighborhood of a specified region of the feature map to be pooled [164, 166]. The pooling operation can also be implemented by extracting hierarchical feature information through partitioning a particular feature map in a tree-like structure (tree pooling).

Fully connected layer: The fully connected (FC) layers (sometimes referred to as dense layers) take as input the flattened outputs of either the final pooling layer, the final convolutional layer, or the output of a fellow fully connected layer, if the CNN has multiple fully connected layers. FC layers form the last part of a CNN and, depending on the problem to be solved, could form a CNN classifier or regressor that generates the final model output through successive interconnection of each neuron of an FC layer to each neuron of its preceding layer [166]. Unlike the convolutional layer, which acts as a feature extractor layer, and pooling layers, which act as dimensionality reduction layers [164], the FC layers are deployed to learn complex and non-linear spatial relationships that exist between different features in an input feature map [170].

Activation function: In CNN, adding activation functions after learnable layers (convolution or FC layers) introduces non-linearity to the network [170]. This move prevents the network from only learning linear transformations of the input data, restricting CNN from learning more complex relationships.

Instead, adding non-linearity allows the network to learn much more complex representations, using a stack of multiple learnable layers [154]. The activation functions should be differentiable to allow backpropagation to occur during the training of CNN models [166]. Currently, different types of activation functions exist that can be used to introduce non-linearity into CNN, with the most popular ones being the Sigmoid (the input is a real number, the output is a value in the range of [0,1]), Tanh (the input is a real number, the output is a value in the range of [-1,1]), Rectified Linear Unit (ReLU) (the input is a real number, the output is a positive real number), and Leaky ReLU (the input is a real number, the output is a positive real number or a downscaled negative real number), which are defined as

Sigmoid:

$$f(x) = \frac{1}{1+e^{-x}}, \tag{3.1}$$

Tanh:

$$f(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}, \tag{3.2}$$

ReLU:

$$f(x) = max(0,x), \tag{3.3}$$

Leaky ReLU:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ mx & \text{if } x \leqslant 0, \end{cases} \tag{3.4}$$

where x is an input variable and m is a constant value (leaky factor).

3.2.2 Regularization techniques of convolutional neural network models

One of the main challenges facing deep learning models, particularly CNN models, is their inability to generalize well to previously unseen data, causing them to either under-fit or over-fit the unseen data. Over-fitting refers to a phenomenon where the CNN model performs well on training data, but performs poorly when subjected to previously unseen data (test data) [154]. On the other hand, under-fitting refers to the failure of the CNN model to learn enough representations from the training data, leading to the model's underperformance when subjected to test data [166]. A well-designed CNN model generalizes well when subjected to unseen data, i.e., performs better on training and test data. Figure 3.2 illustrates these phenomena for a binary classification problem.

Over-fitting is the most predominant challenge facing CNN models, and to overcome this problem, several techniques can be deployed, with popular ones being weight regularization using $L1$ and $L2$ regularization techniques, regularization using dropout technique, data augmentation, early-stopping and batch normalization [166].

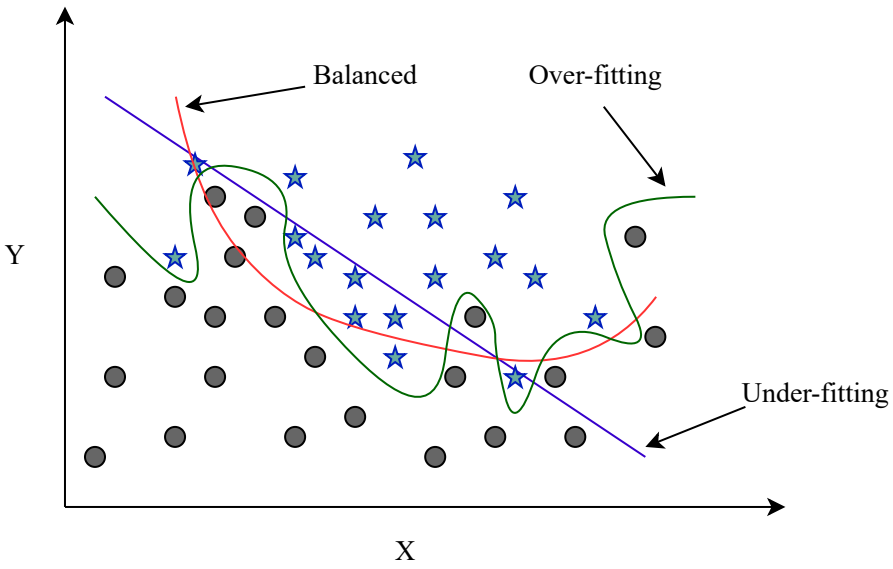

Figure 3.2. Illustration of a balanced, under-fitted, and over-fitted model for a binary classification problem.

*L*1 ***and*** *L*2 ***regularization***: *L*1 and *L*2 regularization techniques fight the over-fitting phenomenon of deep learning models by adding a 'penalty term' to the cost (loss) function to force large weights to take small values. The difference between *L*1 and *L*2 regularization techniques is in the type of the added penalty term, i.e., for a network of parameter X, cost function $C(X)$ and hyperparameters β_1 and β_2, implementing *L*1 is equivalent to adding to $C(X)$ the absolute value of the weight coefficients, while implementing *L*2 is equivalent to adding to $C(X)$ the square of the value of the weight coefficients [154, 164, 166]. The loss function for an *L*1 regularized model is expressed as

$$CostFunction_{L1} = C(X) + \beta_1.\|X\|_1, \tag{3.5}$$

while for the *L*2 (sometimes referred to as weight decay) regularized model, the cost function is expressed as

$$CostFunction_{L2} = C(X) + \frac{\beta_2}{2}.\|X\|_2^2, \tag{3.6}$$

where $\|X\|$ is the matrix norm of the network weights [166], $\|X\|_1 = \max_{1\leq j\leq n}(\Sigma_{i=1}^{n}|x|_{ij})$, $\|X\|_2 = \sqrt{\Sigma_{i=1}^{n}\Sigma_{j=1}^{n}(x_{ij})^2}$.

Dropout: Dropout is another regularization technique that has proven very effective in mitigating over-fitting problems that affect deep learning models. This regularization technique works by randomly

zeroing (dropping out) some output features of the layer (neurons) during training and keeping the neurons intact during the inference stage [164,166]. To implement the dropout regularization technique, a scale called 'dropout rate' has to be defined, which will determine how many neurons are to be dropped out. In essence, randomly dropping out some neurons during the training of deep learning models prevents the network from memorizing less significant patterns from the data, which might negatively affect the network's ability to generalize well [154].

Data augmentation: Having limited data to train deep learning models like CNNs can also cause a network to over-fit as there are very few samples that it can learn from to draw inferences, rendering it ineffective to generalize well. Data augmentation can be adopted to address this issue, which is simply a process of increasing the size of training data by adding some modifications to the current data. It can be implemented in several ways, i.e., by rotating, re-scaling, re-adjusting the brightness, cropping, and flipping the samples, to name a few [166, 171]. Data augmentation has proven effective in fighting over-fitting, particularly when image datasets are used [154].

Early-stopping: Early-stopping is another approach that can be adopted to mitigate over-fitting problems when training deep learning models. It is an act of stopping the model's training process once a target evaluation metric has ceased to improve for a certain number of consecutive training rounds (epochs) [154]. It is implemented by extracting a portion of data from training data to use as a validation set, whereby a model is trained on the training data and validated on the validation data. While validating the model, the performance of the model on the validation set is monitored, interrupting the training process once there has been no improvement in the model's performance in the validation set for a defined number of consecutive training epochs [166]. Figure 3.3 illustrates when to execute early-stopping during model training.

Batch normalization: Batch normalization can also play a big role in reducing the chances of over-fitting of a CNN due to its effectiveness in improving gradient flow and regularization of network weights in deeper networks [166, 171]. When implemented, it normalizes data by maintaining an exponential moving average of means and variances being computed in each min-batch [154].

3.2.3 Compilation of convolutional neural network models

During CNN model training, the complete learning process of the network is configured at the compilation step of the model, where key network parameters are defined, i.e., a loss function,

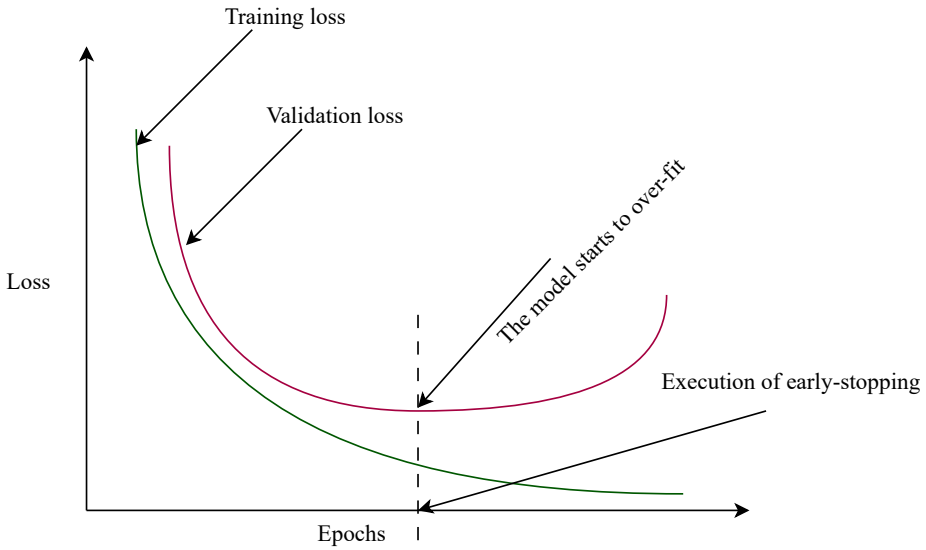

Figure 3.3. Execution of early-stopping during model training.

optimizer, and evaluation metrics.

Loss Function: The loss function, sometimes referred to as the objective function, assists the network to measure its performance during the training process. It measures how well the network performs by computing the 'loss' value, which indicates how much the model's predicted values deviate from their corresponding real/actual values [172]. Different types of loss functions exist, that can be adopted in training the CNN model, with popular ones being the mean squared error (MSE), mean absolute error (MAE), binary cross-entropy, categorical cross-entropy, and hinge loss, to name a few. Adopting any of these loss functions is determined by the nature of the training dataset and the type of task intended to be solved.

Optimizer: An optimizer is central to neural network success (CNN in particularly), as it controls the learning process of a network by establishing a mechanism through which the network will update its weights, based on the observed network performances being measured by the loss function. Its role is to strive to minimize the 'loss' value, which is an error depicting the difference between predicted and actual values. An optimizer computes the gradient of the loss by using the backpropagation algorithm; then, based on the computed gradients, it adjusts/updates the learnable weights of a network in a direction that is likely to lower the loss value [154, 166]. Examples of optimizers that can be adopted

during the compilation of CNN are SGD, root mean square propagation (RMSProp), adaptive moment estimation (Adam), AdaGrad, and AdaDelta.

Evaluation Metrics: These are the performance metrics being monitored during model training. They are specific to the problem to be solved. Regression-based problems use different evaluation metrics compared to classification-based problems. A popular evaluation metric for a regression-based problem is mean absolute error (MAE). For classification problems, commonly used evaluation metrics are accuracy, the area under the receiver operating characteristics curve, precision, and recall, to name a few [154].

3.2.4 Working principle of a convolutional neural network model

As explained in detail in [154], the adoption of CNNs to solve different types of machine learning tasks involves the execution of four key steps. The first step is to transform the training dataset into an input feature map formatted as 3D tensors, where the height and width of the feature map represent the spatial axes, and the depth of the feature map represents the channel axis. The second step involves the extraction of small patches from the input feature map through the use of a convolution operation. The convolution operation refers to the extraction of small 3D patches (of shape (WindowHeight, WindowWidth, InputDepth)) of surrounding features by sliding 2D windows (kernels) of a specific size over the 3D input feature map, covering all possible locations of features. In the third step, a dot product operation (involving each of the extracted small 3D patches and the same learned weight matrix) is executed to transform each of the small 3D patches into a 1D vector of shape (OutputDepth,). In the final step, all the resulting 1D vectors are combined spatially to create the new 3D output feature map (of shape (Height, Width, OutputDepth)), where OutputDepth is the number of filters computed by the convolution operation. Depending on the type of task to be solved, this 3D output feature map has to be flattened and connected to several FC layers. Figure 3.4 summarizes these key steps.

3.2.5 Performance enhancement of convolutional neural network models using squeeze and excitation blocks

The superior performances of the CNN-based models in solving different machine-learning tasks are attributed to their effectiveness in learning useful spatial and channel-wise information encoded within receptive fields of feature maps by a convolution operation. Many attempts to enhance the performances of CNN-based models by improving their representational powers focus on two main aspects: spatial representations or channel-wise representations. The SE block, which was initially proposed in [173], focuses on the latter, i.e., it was proposed with a focus on improving the channel-wise ability of CNN

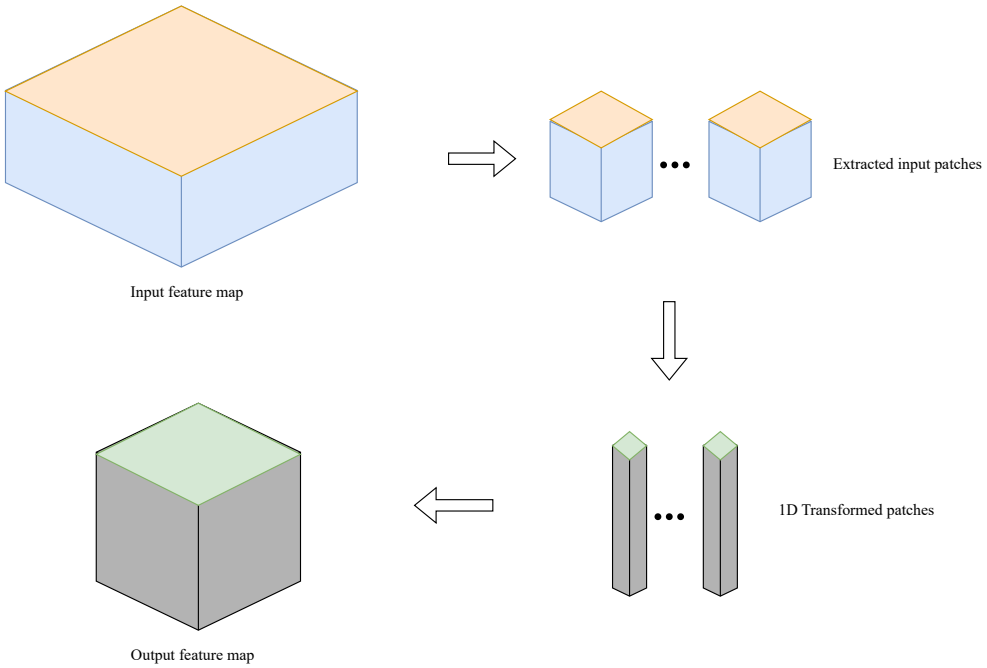

Figure 3.4. A convolution process.

models to extract and learn useful information by adaptively recalibrating the convolutional feature responses to model their channel-wise interdependencies explicitly. This channel-wise recalibration of convolutional features helps the CNN models to concentrate on more impactful features than on the less useful ones. Consequently, it improves the overall learning capabilities of the models.

As elaborated in [174], the implementation of the SE block is based on executing two operations: squeeze and excitation operations. In the former operation, the contextual information not being included in the local receptive fields is used to establish channel-wise statistics by using a global average pooling approach. The role of the latter operation is to develop channel-wise interdependencies through the aggregation of channel-wise statistics being created from the former operation.

According to [173] and as illustrated in Figure 3.5, implementing an SE block involves three key steps. In the first step, the features in a transformed input feature map U are passed through a squeeze operation with the goal of spatially (across HxW dimensions) aggregating the feature maps to create a channel descriptor that encodes a global distribution of channel-wise feature responses. The lower layers can access the information from the global receptive fields through this descriptor. In the second step, the excitation operation executes a self-gating mechanism (established by using FC layers

being activated by ReLU and Sigmoid activation functions) to activate channel-wise sample-specific information. The final step involves reweighting feature maps U to create the output feature map of the SE block.

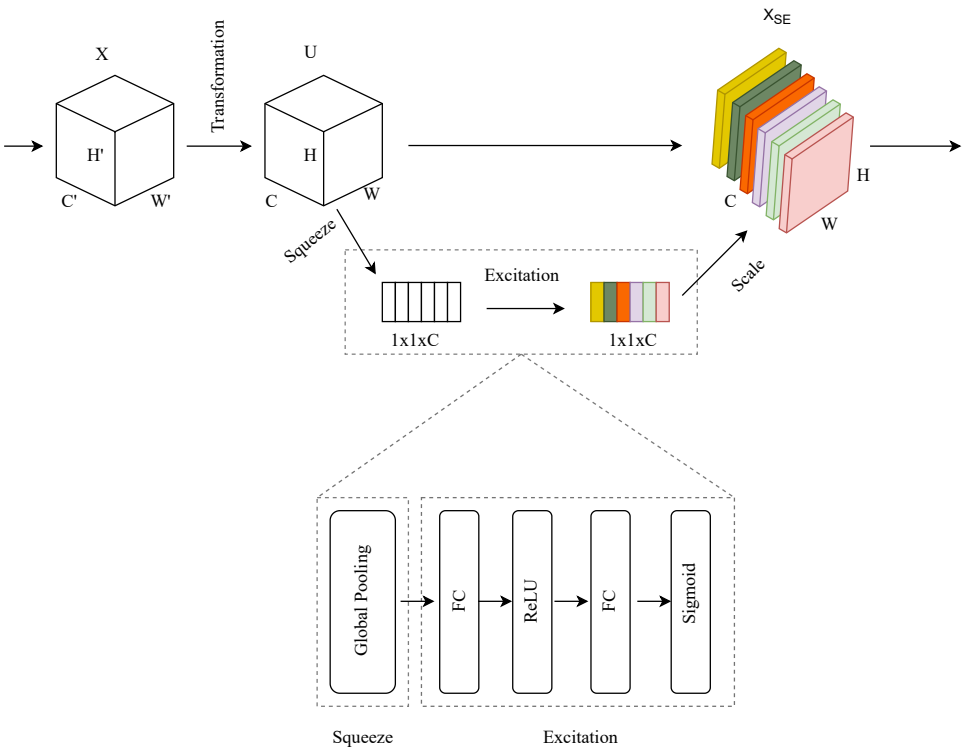

Figure 3.5. SE block. Adapted from [173], ©2018 IEEE.

3.3 TRANSFORMERS

The transformer is a deep learning architecture that has been developed primarily for sequence-to-sequence modeling. Unlike traditional recurrence networks, which are sequential in nature and limited to learning short-term dependency of input sequences, transformers, through attention mechanism, can learn long-term dependencies between input sequences in parallel [175]. The successful adoption of transformer models in solving different natural language processing (NLP) tasks has drawn the interest of researchers in applying them to a wide range of tasks across different fields, such as computer vision, remote sensing, time series, and speech processing, to name a few [176].

The original transformer architecture, famously known as the vanilla transformer [177] as illustrated in Figure 3.6, comprises an encoder and decoder modules; however, depending on the problem to be solved, a transformer model may be developed using just an encoder module or just a decoder module.

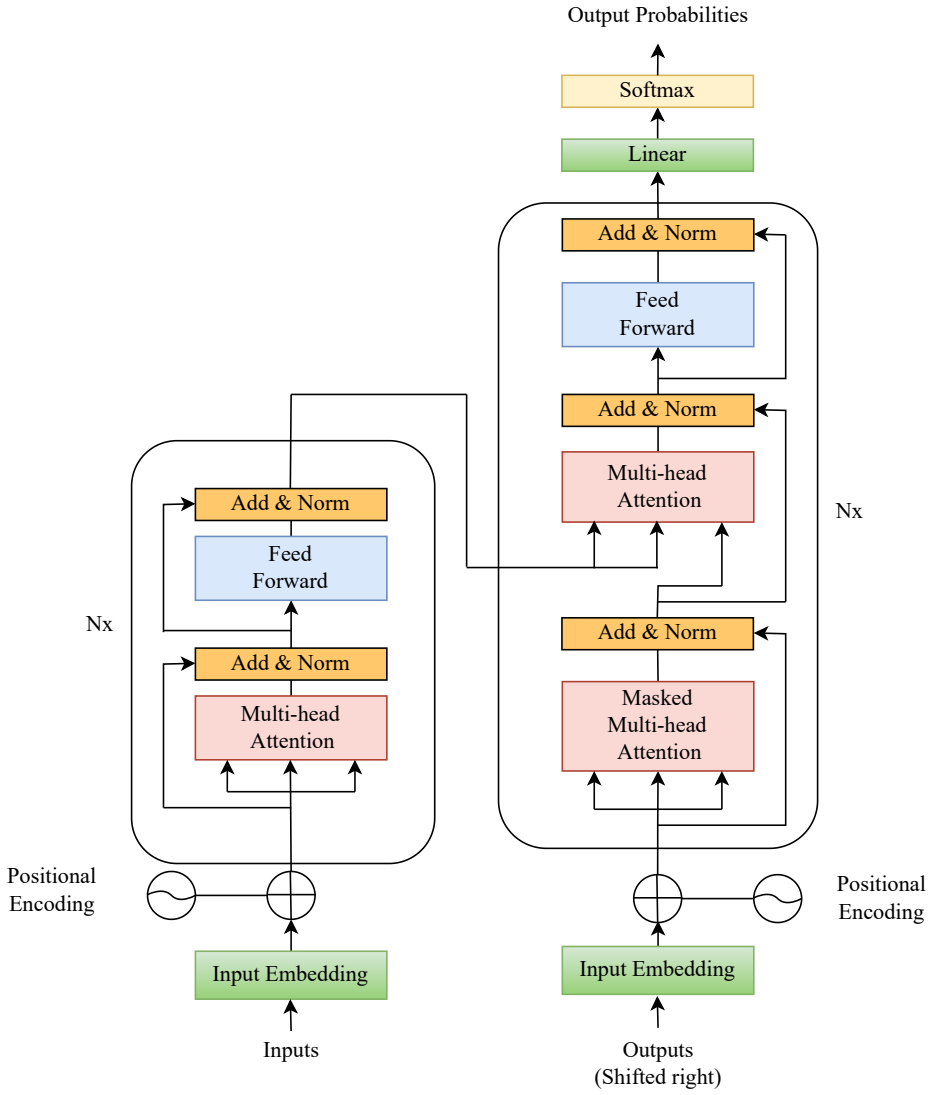

Figure 3.6. Vanilla transformer. Adapted from [178], ©2024 IEEE.

The building blocks of the transformer encoder/decoder are the embedding layer, positional embedding layer, self-attention block, multi-head self-attention block (MHSA), and point-wise feedforward networks (FFN). The embedding layer creates a dense vector of representation sequences from tokens generated from input sequences being fed to the attention block. The positional embedding layer allows the transformers to know the position and order of appearance of each word in the input sequences [179]. The self-attention block is used to implement a self-attention mechanism which gives the transformer model the ability to learn word dependencies in parallel across each of the input sequences, offering a

pairwise correlation and long-range dependencies between word tokens in the input sequences [179]. Through the computation of attention weights of each word in the input sequences, which forms the basis of the importance/relevance of a particular word token relative to others in an input sequence, the attention mechanism allows the transformer model to concentrate on specific aspects of an input sequence, enabling it to effectively learn the long-term dependencies and correlation between word tokens in the input sequences [180]. As pointed out in [178], mathematically, the attention mechanism is expressed as

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \tag{3.7}$$

where Q, K, and V are the query (the current focus of the attention mechanism to be compared to each element in an input sequence), key (an element of the input sequence compared to the query), and value (learned score to each element of the input sequence) tensors respectively, d_k is the output dimension of the key matrix, and T is the transpose of a vector. According to [176, 181], for an input sequence X ($X \in {}^{n\times d}$), where n is the total number of word tokens, and d is the dimension of the embedding layer, Q, K, and V are the linear transformations of the input sequence X expressed as

$$Q = X \cdot W_q, \tag{3.8}$$

$$K = X \cdot W_k, \tag{3.9}$$

$$V = X \cdot W_v, \tag{3.10}$$

where $W_q \in {}^{c\times d}$, $W_k \in {}^{c\times d}$ and $W_v \in {}^{c\times d}$ are the weight matrices to generate the Q, K, and V vector matrices.

As elaborated in [182], the multi-head self-attention (MHSA) block, which is expressed as

$$MHSA(Q,K,V) = Concatenation(Head_0, Head_1, ..., Head_{h-1})W^0, \tag{3.11}$$

where,

$$Head_i = softmax\left(\frac{QW^{Qi}(KW^{K_i})^T}{\sqrt{D_q/h}}\right)VW^{V_i}, \tag{3.12}$$

$$W^0 \in {}^{h\cdot D_v \times N}, \tag{3.13}$$

comprises multiple concatenated self-attention blocks referred to as 'Heads,' each having its own learnable weight matrices W^{Q_i}, W^{K_i}, and W^{V_i}, for the Q, K, and V tensors, respectively. h denotes the number of heads in a multi-head self-attention block, $i = 0, 1, 2, \ldots, (h-1)$, Dq and D_v denotes the dimension of the Q and V, respectively, and N denotes the number of embeddings.

The point-wise FFN, which is a series of two FC layers activated by either the ReLU or GELU activation function, allows the transformer encoder/decoder to learn position-specific information for

each input sequence [179].

In the original transformer architecture (vanilla transformer), the encoder module was deployed to extract useful features from the input sequences and generate representations to be fed to the decoder to solve an NLP task. The encoder structure of the vanilla transformer is a stack of multiple identical layers, each consisting of two main blocks, the MHSA and the FFN blocks. The output of the positional embedding layer is fed to the MHSA whose output is passed to the normalization layer and then to the FFN block. The output of the MHSA is then added to the output of the FFN block through a residual connection with a resulting output vector being normalized to generate the overall output of the encoder. The decoder, on the other hand, is a stack of multiple identical layers, each comprising three main blocks, i.e., masked MHSA, multi-head cross attention (MHCA), and the FFN blocks. The input to the masked MHSA is the output of the positional embedding layer; the output of the masked MHSA is combined with the output of the positional embedding layer through a residual connection, resulting in an output vector that is combined with the output of the encoder to generate the input vector to the MHCA block. The normalized output of the masked MHSA block is added to the output of the MHCA block through a residual connection, resulting in an output vector that is normalized and fed to the FFN block. The normalized output of the MHCA block is added to the output of the FFN block through a residual connection, resulting in an output vector that is normalized and fed to the fully connected layer to form the output of the decoder.

3.4 CHAPTER SUMMARY

In this chapter, the main building blocks of CNN, namely, the convolutional, pooling, and FC layers, were discussed, along with popular approaches used to address the main challenge facing deep learning models (CNN in particular), which is an over-fitting phenomenon. Additionally, key network parameters, namely the loss function, optimizer, and evaluation metric that need to be defined at the compilation step of the network, were discussed, and some popular choices of these parameters were pointed out. A general working principle of a CNN model was also provided, focusing on data transformation, extraction of input patches from a transformed input feature map, a dot product transformation of extracted input patches using a learned weight matrix, and generation of the output feature map ready to be processed further by a classifier or regressor. The SE block was also introduced, and its role in enhancing the performances of CNN models by improving channel-wise interdependencies was discussed. Lastly, the transformer module was presented and discussed, focusing on its main features and working principles.

CHAPTER 4 FINGERPRINTING LOCALIZATION USING DEEP LEARNING METHODS

4.1 CHAPTER OVERVIEW

In this chapter, two fingerprint-based localization methods are proposed to estimate the locations of target nodes in LoRaWAN networks. The first method is a fingerprint-based branched CNN localization method enhanced with SE blocks, where the joint use of CNNs and SE blocks was to improve channel-wise interdependencies. The second method is a hybrid CNN-transformer fingerprinting method to localize a node in LoRaWAN networks, whereby the CNNs are adopted to complement the strengths of the transformer by adding the ability to capture local features from input data and consequently allowing the transformer, through the attention mechanism, to effectively learn global dependencies from the input data. The model architectures for the two proposed methods are presented along with their implementation procedures, using a publicly available LoRaWAN dataset. This chapter is based on two published journal articles, one titled: "RSSI-based fingerprint localization in LoRaWAN networks using CNNs with squeeze and excitation blocks" [89] and another one titled: "A hybrid convolutional neural network-transformer method for received signal strength indicator fingerprinting localization in Long-Range Wide Area Network" [183].

The remaining sections of the chapter are as follows: In Section 4.2, the background information on the status of the currently available fingerprint-based localization solutions in LoRaWAN networks and the justification for suggesting the proposed localization solutions is presented. In Sections 4.3 and 4.4, the two proposed fingerprint-based localization methods are explored. Implementation procedures of the proposed methods by using a publicly available LoRaWAN dataset are discussed in Section 4.5, and the conclusive remarks on this chapter are provided in Section 4.6.

4.2 BACKGROUND

As pointed out in [32], in estimating locations of target nodes in unlicensed LPWAN networks, particularly LoRaWAN, the fingerprint-based localization approach can achieve relatively high localization accuracies compared to the range-based localization approach. This is attributed to the use of machine learning algorithms, which are central in implementing fingerprint-based localization methods. Thanks to the advancement made in machine learning algorithms, capped by the increased processing and computational power of computers, fingerprint-based localization methods, unlike range-based localization methods, can extract and learn useful positional features even from relatively noisy data collected in NLOS environment settings, consequently enabling them to efficiently and effectively infer locations of the target nodes with satisfactory localization accuracies [11].

Most of the currently available fingerprint-based localization methods being proposed and developed to estimate locations of target nodes in LPWAN networks, particularly LoRaWAN, use classical 'shallow' machine learning techniques such as the KNN algorithm (and its variants, such as weighted KNN), SVM, RF, and DT. The efficiency and performance of these models rely heavily on the size of the training dataset. They tend to give satisfactory localization accuracies under specific conditions, particularly when the training dataset is relatively small; however, their complexity tends to increase as the size of training datasets is increasing (which is necessary if a large outdoor environment is to be covered for localization purposes), ultimately resulting in a decline in localization performance [11].

The success brought about by deep learning models based on CNN in solving different types of computer vision tasks has attracted the interest of researchers to adopt them in developing fingerprint-based localization methods to localize target nodes in various kinds of wireless communication networks, as in [184–194].

Leveraging the efficiency and effectiveness of the CNNs in tackling regression and classification tasks using relatively large datasets, two deep learning-based localization methods based on CNNs to address the drawbacks of the currently available fingerprinting localization methods in LoRaWAN networks are proposed in this chapter. The first proposed method is a CNN-based regressor enhanced with SE blocks to estimate the locations of target nodes in LoRaWAN networks. Two reasons drive the choice of CNN regressor. The first reason is the complexity of training classical machine learning-based localization models in unlicensed LPWAN networks that is experienced when relatively large datasets

are used, which can sometimes lead to degradation in localization performance. The second reason is CNN models' efficiency in learning useful position information in structured data, as reported in [195]. The joint use of CNN and SE blocks is to improve channel-wise interdependencies. The second proposed method is a hybrid CNN-transformer fingerprint-based method to localize target nodes within LoRaWAN networks. Transformers are becoming popular in computer vision and natural language processing tasks largely on account of their ability to capture global dependencies from the input data; however, they are less equipped to capture local dependencies from the input data [196–198]. CNNs, on the other hand, are relatively poor in establishing global dependencies of features in the training data because of the locality of the convolution operation [197]. The joint use of a CNN and a transformer in this proposed method complement each other, resulting in a robust, efficient and effective localization model.

4.3 FINGERPRINTING LOCALIZATION USING CONVOLUTIONAL NEURAL NETWORK WITH SQUEEZE AND EXCITATION BLOCKS

Inspired by an inception module [199], the first fingerprint-based localization method proposed in this chapter is a three-branched two-dimensional CNN regressor with six convolutional layers. Deep learning networks based on inception modules comprise subgraphs of layers with several convolutional branches which process the same input data in parallel followed by the merging of the outputs of each branch into a single output which depending on the type of machine learning task to be solved, can then be processed further using a classifier or regressor. The useful aspect of this branched CNN architecture is that it allows processing information at various scales to extract features simultaneously at the next stage after aggregation. This model structure can also prevent a blow-up in computational complexity even if the number of processing units at each stage increases significantly [199]. Unlike in [199] where the authors present an inception model developed to process image data to solve classification and detection tasks using a network of layers with several parallel convolutional branches comprising a range of kernel sizes, i.e., 1x1, 3x3, 5x5 convolutions activated by rectified linear units and adopting max-pooling layers for resolution reduction, the proposed fingerprint-based localization model (refer to Figure 4.1) is developed to process RSSI signal data to tackle the regression task. The first branch has a single convolutional layer, whereas the second and third branches have two and three convolutional layers, respectively. Each convolutional layer computes eight filters over its input with a 1x1 kernel size. The leaky ReLU activation function is used to activate the convolutional layers with a 0.3 constant gradient, allowing for a small, non-zero gradient when the unit is saturated and inactive. In order to improve channel-wise interdependencies, after each convolutional layer, an SE block with eight filters

is connected. An SE block [173] is a channel-wise attention mechanism widely used to improve the overall performance of CNNs. It has been used to improved success in computer vision tasks. The three branches of the convolutional layers are then concatenated and connected to six FC layers. The first five FC layers are activated by the ReLU activation function. The first FC layer has 512 units, the second has 256 units, the third has 128 units, the fourth has 64 units, and the fifth has 32 units. The last FC layer with two units is activated by a linear activation function for regression purposes. The decision to use five FC layers before the last two-unit FC layer (by halving the number of units in each subsequent layer) prevents information loss and thus improves the localization performance. In order to prevent over-fitting and subsequently improve the model performance, several hyperparameters necessary in the model's training were tuned. The optimal hyperparameters adopted in this work are a 0.15 drop-out ratio and a 0.01 $L1$ kernel regularizer. At the compiling stage of the model, the Adam optimizer with a dynamic learning rate is used. The learning rate is first set at 0.001 and keeps decreasing by a factor of 0.1 if the validation loss does not improve for ten consecutive epochs. The adopted loss function is the mean absolute error (MAE). The location estimation error and R^2 score metrics are the key performance evaluation criteria for this proposed method. Figure 4.1 is the illustration of the proposed localization approach.

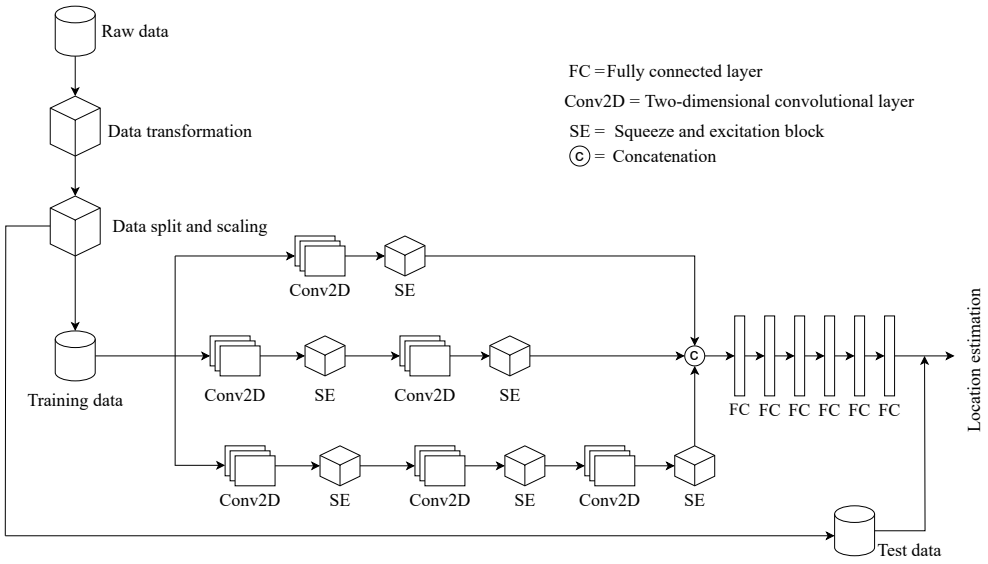

Figure 4.1. Proposed model architecture for the fingerprinting localization method using CNN with SE blocks.

4.4 FINGERPRINTING LOCALIZATION USING A HYBRID CONVOLUTIONAL NEURAL NETWORK-TRANSFORMER METHOD

A hybrid fingerprint-based localization model using CNNs and transformer modules is the second method being proposed in this chapter to estimate the locations of target nodes in LoRaWAN networks. This method has been inspired by research reported in [200] and [201]. The authors in [200] proposed a transformer-inspired sequential neural network model to tackle a classification problem that involves signal modulation of radar pulses. The top layers in this model comprise three convolutional layers, each with a kernel size of 11, stride of 2 and 64 output channels activated by ReLU activation function, whereas its bottom layers comprise 2 transformer encoder layers each utilizing 4 attention heads followed by a classifier to generate the final output. In [201], the authors present a deep learning model based on CNN and transformer modules to solve classification problems of three medical datasets, that is, cerebral emboli, electrocardiogram and heartbeat categorization datasets using raw and time-frequency representation signals. The presented model comprises two structures, one responsible for the processing of raw signals which is a 1D CNN-transformer model and the second structure which is responsible for extraction of frequency representation signals which is a 2D CNN model. The final part of this model is a classifier that combines the outputs of the two models into the final output. The second fingerprint-based localization method being proposed in this chapter is developed to tackle a regression task involving RSSI signals of a publicly available LoRaWAN dataset. In this method, CNNs are adopted to complement the strengths of the transformer by adding the ability to capture local features from input data and consequently allowing the transformer, through the attention mechanism, to effectively learn global dependencies from the input data. The original transformer architecture, commonly referred to as the vanilla transformer [177], features encoder and decoder structures to process sequences from two different types of data (two language types, to be precise) to achieve a sequence-to-sequence machine translation [202]. Instead of using the complete structure of the vanilla transformer, this method adopts only the encoder part, since only one type of data is processed. The features of the transformer encoder are enough to learn global dependencies from the input sequences and output representations, which can be processed further by a classifier or regressor for performing classification and regression tasks, respectively. A slight modification is made to the transformer encoder by opting to process the input data using a stack of three one-dimensional convolutional (1D-CNN) layers instead of positional encoding. A stack of three 1D-CNN layers replaces the transformer encoder's positional-wise feed-forward neural network block to capture local context within the encoder module [203]. For all the 1D-CNN layers, the number of filters used is eight, with a kernel size of one. The transformer encoder's embedding and dense dimensions are set to eight.

Figure 4.2 illustrates the proposed model architecture. The input data is first processed by the 1D-CNN layers to learn local dependencies before being fed into the transformer encoder to learn the global dependencies from the input data. Information learned at the local level is then concatenated with the information learned at the global level to form the output of the first part of the proposed method. The second part of the proposed method comprises a stack of four FC layers, with 512, 256, 128 and 2 hidden units, all activated by the ReLU activation function except for the last two-units FC layer, which is activated by a linear activation function for regression purposes. The output of the first part of the localization model is then flattened and fed into the second part for the final location estimation. To improve the learning capabilities of the method, a small dropout ratio of 0.1 is introduced for the 512 and 256-units FC layers. At the compilation stage of the localization model, the Adam optimizer was used with a learning rate initially set at 0.001 and reduced by a factor of 0.1 after ten successive epochs of unimproved validation loss. The mean absolute error (MAE) is adopted as the loss function to train the model. Similar to the first proposed method, the location estimation error and R^2 score metrics are the adopted metrics to evaluate the performance of this method. Model checkpoints and early-stopping callbacks are also introduced to better optimize the training duration of the localization model.

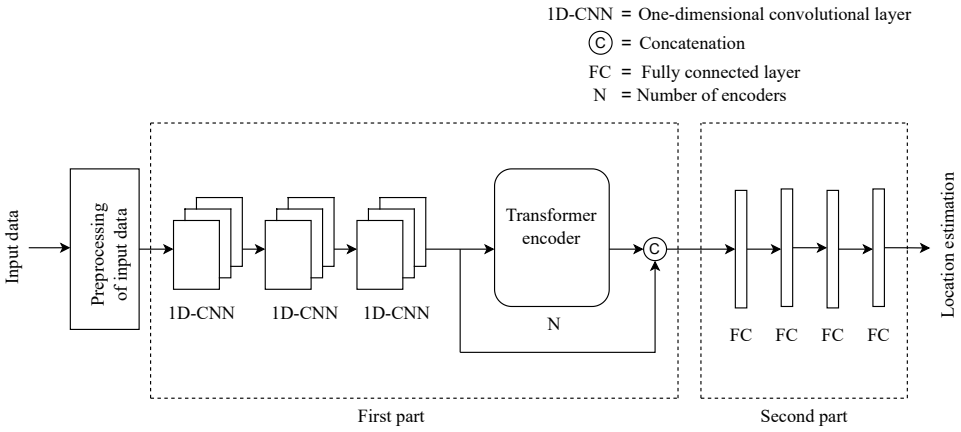

Figure 4.2. The proposed model architecture for the fingerprinting-based localization method using a hybrid CNN-transformer method.

4.5 IMPLEMENTATION PROCEDURES

The LoRaWAN dataset used to validate the performance of the two proposed fingerprint-based localization methods is version 1.2 of the publicly available urban LoRaWAN dataset reported in [9], which was collected in Antwerp, Belgium, in 2019. This dataset contains a total of 130,430 messages (number of samples), each with six unique attributes (RSSI values in dBm from 72 gateways, SF, receiving time, horizontal dilution of precision (HDOP), latitude and longitude). Initially, four attributes were

extracted from each message, including the RSSI values from 72 gateways, SF values, latitudes and longitudes. The other two attributes were not included because they require more preprocessing procedures to effectively be processed by the proposed methods, which could significantly increase model complexity. During the construction of this database, an out-of-reach RSSI value of -200 dBm was given to a gateway that failed to receive the transmitted message. Upon scanning the LoRaWAN database to find out which gateways had failed to receive at least a single transmitted message, 28 gateways were found to have never received a single transmitted message, so they were removed from the dataset. So, together with the SF column, the remaining dataset used to validate the performance of the proposed methods in this thesis has 47 features. The first 45 features (gateway columns and SF column) were used as training data, while the last two features (latitudes and longitudes columns) were used as target labels. Therefore, the resulting dataset has 130,430 samples (total number of transmitted messages), each with 45 features. For this dataset to be successfully fed into CNNs, it is re-shaped to tensors of shape (1,45,1) corresponding to processing a single message (sample) at a time. The ground truth references, on the other hand, consist of 130,430 samples with two features (latitudes and longitudes).

Before feeding a machine learning model with RSSI-based training data, a preprocessing procedure has to be performed on the training data to ease the learning process of the model. The first step is to search for the smallest RSSI value received from the dataset ($RSSI_{min}$), followed by replacing the out-of-reach RSSI values with '$\tau = RSSI_{min} - 1$' [157]. The last preprocessing step involves transforming the resulting training dataset into optimal representations by using any of the four commonly adopted data representation techniques, namely positive, normalized, powed and exponential data representation schemes [204–206], expressed as

$$Positive_i(x) = RSSI_i - \tau, \tag{4.1}$$

$$Normalized_i(x) = \frac{Positive_i(x)}{-\tau}, \tag{4.2}$$

$$Exponential_i(x) = \frac{e^{\frac{Positive_i(x)}{\alpha}}}{e^{\frac{-\tau}{\alpha}}}, \tag{4.3}$$

$$Powed_i(x) = \left(\frac{Positive_i(x)}{-\tau}\right)^{\beta}, \tag{4.4}$$

where i and $RSSI_i$ represent the gateway identifier and RSSI value at gateway i, respectively. The α and β in the exponential and powed data representation schemes are the parameters defined according to how RSSI values are distributed in a dataset. In [206], they were originally set at 24 and e, respectively, with e being a mathematical constant for RSSI values collected indoors by using WiFi signals. The

α and β parameters adopted in this thesis are 60 and 1.1, respectively, re-adjusted in [157] for the outdoor RSSI values in the LoRaWAN network.

Since the training labels are in latitudes and longitudes coordinates, the Haversine formula is used for the computation of the equivalent distance between two points on the earth's surface at the location estimation stage of the model. The Haversine formula [207] is defined as

$$Hav\left(\frac{D}{R}\right) = Hav(\gamma_2 - \gamma_1) + cos(\gamma_1)cos(\gamma_2)Hav(\vartheta_2 - \vartheta_1), \tag{4.5}$$

where 'Hav' represents the Haversine function, expressed as

$$Hav(C) = sin^2\left(\frac{C}{2}\right) = \frac{1 - cos(C)}{2}. \tag{4.6}$$

In Equation (4.5), D and R represent the distance between two coordinates and the sphere's radius, respectively, γ_1 and γ_2 represent the latitudes of coordinates 1 and 2, respectively, and ϑ_1 and ϑ_2 represent the longitudes of coordinates 1 and 2, respectively, all in radians.

In this thesis, the proposed fingerprint-based localization models are implemented using the Keras and Scikit-Learn Python libraries alongside TensorFlow as the backend. Additionally, Google collaboratory Jupyter notebooks were used to run the experiments on a 32 GB RAM Core i7 LG computer workstation.

Since the features in training data are in different scales, Sklearn's StandardScaler is used to re-scale them to values with zero mean and unit standard deviation. The labels are re-scaled to the range of [0, 1] using Sklearn's MinMaxScaler. Re-scaling the training data is recommended to prevent biases towards the features with large values during model training [154].

4.6 CHAPTER SUMMARY

In this chapter, the two deep learning-based fingerprinting localization methods to estimate the locations of target nodes in LoRaWAN networks were proposed and developed. The first method was based on CNNs enhanced with SE blocks, and the second one was a hybrid CNN-transformer localization model. Detailed structural settings of the two proposed methods and four data representation schemes commonly used to improve the learning efficiency of machine learning models trained on the RSSI values were discussed. Additionally, key procedures for implementing the proposed methods to estimate locations of target nodes in LoRaWAN networks using a publicly available dataset were presented. With those detailed implementation procedures, which included setting up experiments

and preprocessing the training data, the robustness and effectiveness of the proposed fingerprint-based localization methods in estimating the locations of target nodes in LoRaWAN networks could be validated.

CHAPTER 5 RESULTS AND DISCUSSION

5.1 CHAPTER OVERVIEW

In this chapter, a series of experiments being carried out to evaluate the effectiveness of the fingerprint-based branched CNN localization method, enhanced with SE blocks and the hybrid CNN-transformer fingerprinting method for localizing a node in LoRaWAN networks by using a publicly available LoRaWAN dataset, are presented. Different training data extraction strategies and performance metrics have been adopted in the performance evaluation of the two proposed fingerprint-based localization methods in different scenarios to prove their robustness and effectiveness. The experimental results, analysis and discussion presented in this chapter have been published as part of two published journal articles: one titled: "RSSI-based fingerprint localization in LoRaWAN networks using CNNs with squeeze and excitation blocks" and another one titled: "A hybrid convolutional neural network-transformer method for received signal strength indicator fingerprinting localization in Long Range Wide Area Network" [89, 183].

The remaining sections of this chapter are organized as follows: In Section 5.2, the performance evaluation of a fingerprinting localization method based on CNN with SE blocks is presented. In Section 5.3, the performance evaluation of a hybrid CNN-transformer fingerprinting localization method in different experimental settings is presented. Comparative remarks regarding the localization performances of the two fingerprint-based methods being proposed in this thesis are provided in Section 5.4. A performance comparison of the two proposed methods with other fingerprint-based localization methods being proposed in the literature that were evaluated by using the same dataset being used in this thesis, is provided in Section 5.5. In Section 5.6, conclusive remarks on the chapter are provided.

5.2 PERFORMANCE EVALUATION OF THE FINGERPRINTING LOCALIZATION BASED ON CONVOLUTIONAL NEURAL NETWORK WITH SQUEEZE AND EXCITATION BLOCKS

The following experiments were conducted to validate the localization performance of the proposed fingerprinting method based on CNN with SE blocks, which, for simplicity, will be referred to as the CNN-SE method. In each experiment, a proposed fingerprint localization model is trained for 150 epochs in mini-batches of 512 samples. Before splitting the dataset into training, validation and test sets, a random shuffling of the data with a fixed seed of 42 has first been conducted. Shuffling the training data helps to prevent any bias during the training, prevents the model from learning the training order, and consequently improves the generalization abilities of machine learning models.

5.2.1 Performance on different data representation schemes

To evaluate the performance of the proposed method on different data representations, the dataset is first transformed into the powed data representation scheme and is then split into training, validation and test sets according to a 0.7/0.15/0.15 data split ratio. After the dataset split, the localization model is trained by using the training and validation sets, followed by the location estimation step by utilizing the test set. This experiment is repeated for 0.8/0.1/0.1 and 0.85/0.1/0.05 data split ratios. The reason for trying different data split ratios was to determine if the use of different data split ratios would have any impact on the overall localization performance of the method.

Furthermore, the dataset is also transformed into exponential, positive and normalized data representation schemes and all the experiments conducted using the powed data representation scheme, are repeated. Table 5.1 presents the performance of the proposed method for each data representation scheme. From the results, it is clear that the powed data representation scheme slightly outperforms all three other data representation schemes over the three data split ratios, achieving mean localization errors (m) of 292.04, 290.2, and 284.57 on the 0.7/0.15/0.15, 0.8/0.1/0.1 and 0.85/0.1/0.05 data split ratios, respectively.

These results show that a reduction in the size of the test set leads to improved localization results. This outcome is because the model is being subjected to more and more training data. However, the ratio of the testing data relative to training and validation data used to validate the performance of machine learning models should not be too small for it to generalize well on the unseen data. The experimental results also reveal slightly better localization performance of the proposed method trained

Table 5.1. Performance of the CNN-SE method on different data representation schemes (optimal results are in bold).

Data split ratio	Powed	Exponential	Positive	Normalized
0.7/0.15/0.15	**292.04**	294.46	298.13	294.43
0.8/0.1/0.1	**290.2**	290.86	291.78	293.72
0.85/0.1/0.05	**284.57**	286.12	289	284.53

using the dataset transformed according to a powed data representation scheme, compared to the other data representation schemes. Overall, the positive data representation scheme achieved poor localization results compared to the rest of the data representation schemes. The better localization results obtained by using powed, as well as exponential data representation schemes over the three different data split ratios, could be attributed to the non-linear nature of these two data representation schemes. Adopting any of these two data representation schemes to similar machine learning-based localization tasks, using RSSI data, will likely yield better results than the other data representation schemes, as reported in [18, 157, 159], where authors adopted similar datasets to the one that has been used in this thesis.

Since the powed data representation scheme has yielded the best results for the proposed localization method, it is the adopted data representation scheme for the remaining experiments. Additionally, unless otherwise stated, a 0.7/0.15/0.15 ratio is adopted to split the dataset into training, validation, and test sets (extracting first the training set, then the validation set and lastly, the test set) for the remaining experiments. This data split ratio is the most adopted data split ratio in the literature, so it is adopted in the rest of the experiments to have a fair comparison of the localization performance with other related works.

5.2.2 Performance for each retained percentage of PCA variances

To analyze the performance of the method on different percentages of the retained principal component analysis (PCA) variances, PCA was performed on the training dataset and a series of experiments were conducted. The experiments on 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 percentages of retained PCA variances of the training dataset were conducted one by one to yield localization performance in terms of mean localization error, as shown in Table 5.2 and Figure 5.1. Table 5.2 presents the performance of the proposed method on each percentage of retained PCA variances, as well as the

trainable parameters as a result of the reduction in variances. Figure 5.1 shows full and zoomed-in versions of figures presenting cumulative distribution function (CDF) curves of test errors for each of the retained percentages of the variances. A training dataset with only 10% of the retained PCA variance produced the worst localization result of 549.05 m mean localization error; which is expected, given the large amount of information being removed from the training dataset. As the percentage of the retained PCA variances was increased, the localization results also improved, reaching the highest localization accuracy of 292.04 m mean localization error with 100% retained PCA variance. The same trend is also seen regarding the number of trainable parameters. The use of reduced percentages of PCA variance of the training data reduces the computational burden of the method because of the reduced number of training parameters but at the expense of reduced localization accuracies. These results show that the proposed localization model can achieve acceptable levels of localization accuracies even with a 70% retained PCA variance of the training dataset, consequently reducing the number of training parameters and, hence, the computational burden of the method. Performing PCA is necessary because, in the LoRaWAN fingerprint databases, the number of messages received by different gateways that are distributed in an area for localization purposes, differs. Depending on the transmitter's location, some gateways will receive many messages, some will receive fewer, and some will even fail to receive a single message. This situation necessitates performing a PCA analysis on the dataset (before using it to train the model to remove any components that have very little influence on the data) to increase the model's training efficiency without compromising the localization performance. The proposed method's satisfactory localization performance, even with a 30% reduction in retained variances of the training data, proves its effectiveness and suitability to localize a node in LoRaWAN networks.

5.2.3 Performance on different subsets of training, validation and test data

In this section, experiments conducted to determine the impact that different sample sizes of training data have on the performance of the CNN-SE method are presented. The procedure followed is that the portion of the training sample size in different percentages (20%, 40%, 60%, 80%, and 100%) is extracted first and then, by using a ratio of 0.7/0.15/0.15, the training, validation and test sets are extracted to train and validate the method. This training procedure is carried out to determine to what extent the reduced sizes of the training samples will impact the overall localization performance of the proposed method. Experiments for each of the sample sizes being extracted are then performed, yielding localization results as presented in Table 5.3 and Figure 5.2, whereby in Table 5.3, localization performance in terms of mean localization error (m) on the test set for each extracted dataset size are presented. Figure 5.2, on the other hand, presents full and zoomed-in CDF curves of test errors for

Table 5.2. Performance of the CNN-SE method on different percentages of retained PCA variances.

Retained variances (%)	**Mean localization error (m)**	**Number of trainable parameters**
10	549.05	200,230
20	378.42	224,806
30	338.36	261,670
40	317.01	310,822
50	311.27	372,262
60	309.66	421,414
70	304.58	482,854
80	303.34	544,294
90	299.29	630,310
100	292.04	728,614

each extracted dataset size. As indicated in Table 5.3 and Figure 5.2, it is clear that these results are in accordance with the general trend of machine learning models of exhibiting an improved performance when the size of the training set is increased. The use of only 20% of the original dataset to train and test the proposed method produced the worst localization accuracy of 325.91 m mean localization error; however, the performance improved with the increase in sample sizes of the training data to 292.04 m mean localization error when the full size of the dataset was used. From these results, it is clear that the proposed method is able to yield acceptable levels of localization performance when trained using at least 70% of the size of the original dataset. The fact that the proposed method can give satisfactory localization performance even with a 30% reduction in the original sample size of the data being used for the training and testing of the method, validates its potential to be applied in applications with even smaller datasets.

5.2.4 Performance for different sample sizes of training data with fixed test set

In this section, unlike in Section 5.2.3, 15% of the original sample size of the training dataset is extracted and fixed to be used to test the proposed localization method's performance. For the remaining 85% of the sample size, a portion of the data in different percentages of 40, 60, 80 and 100 are extracted to train and validate the proposed localization method. For each extracted sample, a ratio of 0.8/0.2 is adopted

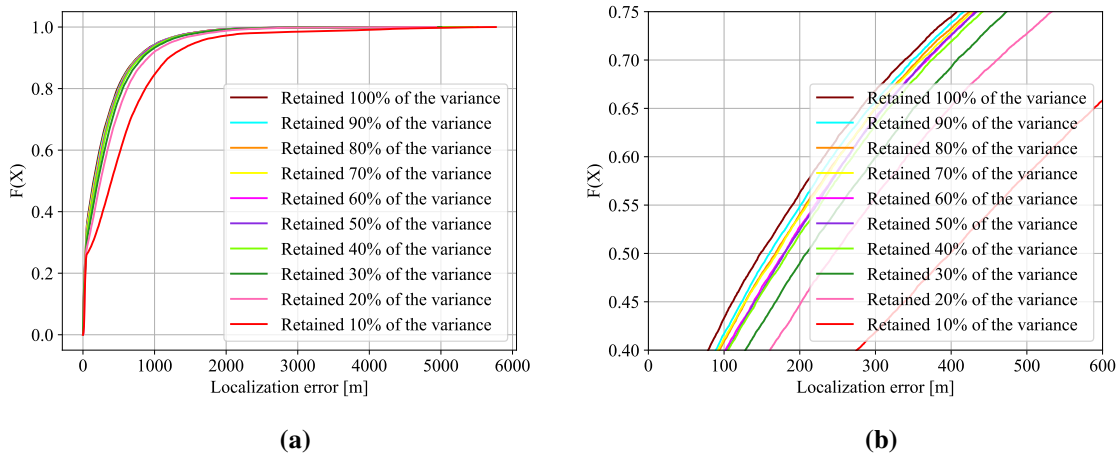

Figure 5.1. CDFs of localization errors (m) for the CNN-SE method when trained on datasets with different retained PCA variances.
(a) Full CDF curves. (b) Zoomed-in CDF curves.

Table 5.3. Performance of the CNN-SE method on different training, validation and test sets.

Used fraction of dataset (%)	**Mean localization error (m)**
20	325.91
40	310.69
60	305.83
80	300.34
100	292.04

to split it into training and validation data. The localization model is then trained on each sample size, using the training and validation data. This data splitting and localization model training strategy is adopted to determine how well the proposed method will generalize when trained on reduced sizes of training and validation sets but tested on the same test set. As a general trend, the generalization ability of a machine learning algorithm is expected to improve with an increase in the training dataset; however, a well-designed algorithm can generalize well even with reduced sizes of the training samples. Table 5.4 shows the localization performance in mean errors (m) for each percentage of the remaining sample size. Figure 5.3 shows the full and zoomed-in CDF curves of test errors for each percentage

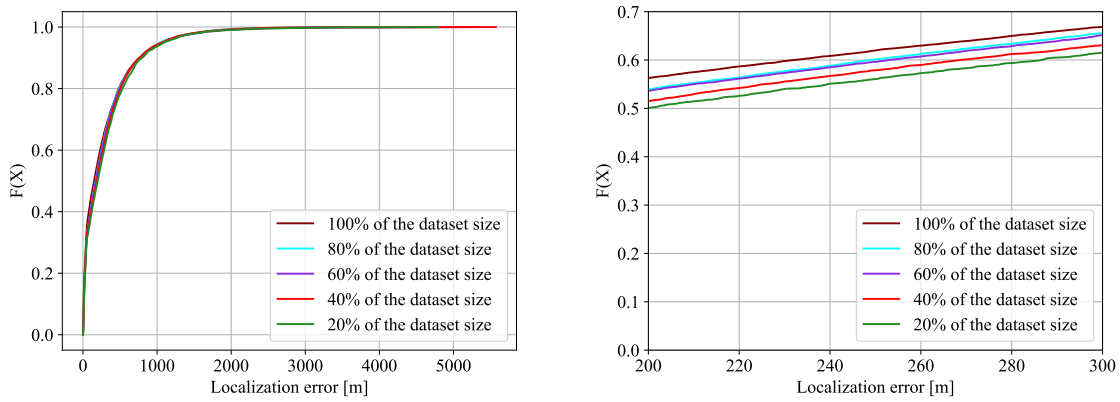

Figure 5.2. CDFs of localization errors (m) for the CNN-SE method when trained on different dataset sizes.
(a) Full CDF curves. (b) Zoomed-in CDF curves.

of the remaining sample size. It is clear from the results that the reduced sizes of the training and validation sets impact the generalization ability of the proposed method. However, the fact that the proposed method was able to obtain an accuracy of 299.79 m mean localization error, when trained using a training sample size of as little as 60% of the remaining dataset and tested on the unchanged test set, is a clear proof of the potential of the method in localizing a node in LoRaWAN networks with acceptable levels of localization accuracies.

Table 5.4. Performance of the CNN-SE method on fixed test set using different subsets of training and validation sets of the remaining dataset.

Used fraction of the remaining dataset (%)	**Mean localization error (m)**
40	310.59
60	299.79
80	298.56
100	296.11

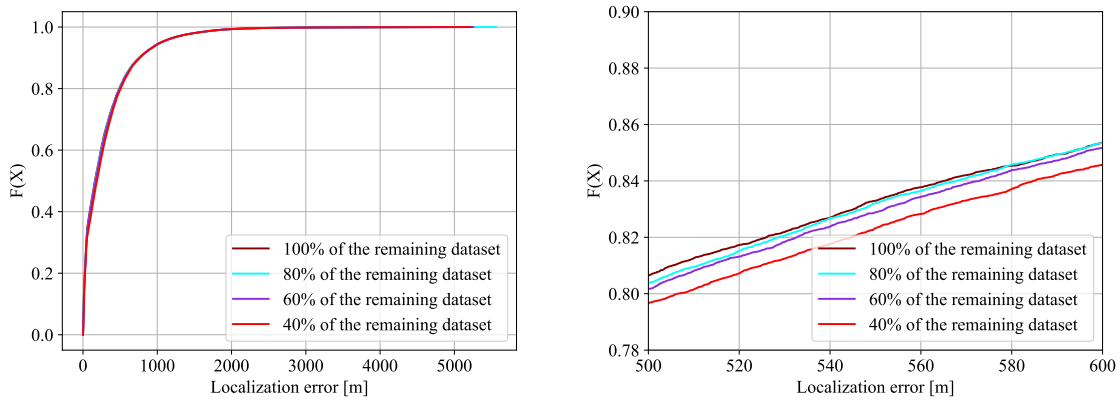

Figure 5.3. CDFs of localization errors (m) for the CNN-SE method trained on different sample sizes and tested on a fixed training set.
(a) Full CDF curves. (b) Zoomed-in CDF curves.

5.2.5 Performance for different data shuffling seeds

In this section, different training data shuffling seeds were used to shuffle the training data. Then, experiments were conducted to evaluate how they would impact the localization performance of the proposed method. In machine learning, shuffling is performed to change the order of how the individual samples appear in the dataset. For a machine learning algorithm to be robust, its performance is not expected to change significantly when it is subjected to different ordering of the sample elements in the same dataset. In order to reproduce a particular order of data elements, the same seed value is used as a base number when generating random numbers. The training data was shuffled with 1, 2, 3, 5, 10, 14, 16, 22, 42, 64, 128, and 512 random seeds before splitting into training, validation, and test sets. Because the dataset was randomly shuffled by using different seeds, the contents of the training, validation, and test sets for each of the settings will differ, thereby impacting the final localization performance, which is evidenced in Table 5.5, which shows how the proposed method performs in terms of mean errors (m) for each shuffling seed. It is noted that shuffling seeds 42 and 64 give slightly better results compared to the rest of the shuffling seeds. However, despite the different shuffling seeds, the proposed method still produced consistent results, proving its robustness.

Table 5.5. Localization performance of the CNN-SE method in terms of mean errors (m) on different training data shuffling seeds.

Data shuffling seeds	**Localization performance**
1	293.79
2	293.35
3	294.35
5	294.38
10	294.37
14	293.33
16	293.81
22	293.79
42	292.04
64	292.41
128	294.36
512	294.04

5.3 PERFORMANCE EVALUATION OF THE HYBRID CONVOLUTIONAL NEURAL NETWORK-TRANSFORMER FINGERPRINTING LOCALIZATION

The following experiments were carried out to evaluate the performance of the proposed hybrid CNN-transformer fingerprinting-based localization method, herein referred to as the CNN-transformer method. All experiments that were conducted adopted the powed data representation scheme to transform the training data into an optimal form. This data representation scheme is adopted because of its non-linear nature that has proved to be effective in improving the performance of fingerprinting localization methods, trained on the datasets similar to the one used in this thesis [18, 157, 159]. Before carrying out the experiments, the dataset was shuffled with a random seed of 42 and then split into training, validation and test sets containing 70, 15, and 15 percentages of the dataset, respectively. The training set was extracted first, followed by the validation set, and lastly, the test set. The dataset was randomly shuffled to make the proposed localization model robust by preventing it from learning the order in which the individual samples appear in the training dataset, which might lead to performance biases. Each experiment is run for 120 epochs, using 512-sized mini-batches of training samples.

5.3.1 Performance with multiple attention heads/transformer encoders

In this section, results from the experiments being carried out to determine the optimal structure of the transformer encoder, which yields the best localization performance when trained on the LoRaWAN dataset (introduced in Section 4.5), are presented. Two structural changes to the localization model were made concerning the transformer encoder: the number of attention heads and stacked encoders. The number of attention heads was varied between one and four heads. In each instance, the localization model was trained by using the LoRaWAN dataset, yielding localization results as shown in Table 5.6. The proposed method achieved performance of 290.71 m, 292.66 m, 292.93 m, and 292.21 m mean localization errors, respectively, on the test set. Based on the results, increasing the number of attention heads does not result in a significant performance improvement in the localization accuracy.

Table 5.6. Performance of the CNN-transformer method with multiple attention heads.

Number of attention heads	**Mean localization errors (m)**		
	Training set	**Validation set**	**Test set**
1	257.43	287.89	290.71
2	260.63	288.66	292.66
3	261.69	290.21	292.93
4	258.24	289.27	292.21

To determine the impact of stacking more than one encoder on the proposed CNN-transformer localization method, the number of encoders (each with a single attention head) was set to one, two, three, and four, and consequently trained on the LoRaWAN dataset, yielding localization results as indicated in Table 5.7. The mean localization errors of 290.71 m, 291.66 m, 292.99 m, and 292.13 m achieved on the test set indicate that stacking more than one encoder will not improve localization performance and only lead to additional computational complexity. The slight variations being observed in localization performance of the proposed method when the number of attention heads and encoders was varied, are due to the structural changes made to the model configuration, which introduced different representations of features for the model to learn and deduce distance estimation from them at the inference stage. With regard to the CNN structure, multiple experiments were run to determine the optimal number of layers, whereby a CNN structure with three layers was enough to yield acceptable levels of localization accuracy. Therefore, unless otherwise stated, a hybrid structure (made of a three-layered

CNN structure, a single transformer encoder with one attention head and a regressor with four fully connected layers) is adopted for the rest of the experiments.

Table 5.7. Performance of the CNN-transformer method with multiple transformer encoders.

Number of transformer encoders	Mean localization errors (m)		
	Training set	Validation set	Test set
1	257.43	287.89	290.71
2	262.85	289.04	291.66
3	256.97	288.53	292.99
4	257.45	288.10	292.13

5.3.2 Performance for each retained percentage of PCA variances

In this section, similar to what had been done for the CNN-SE method, PCA was conducted on the training dataset to extract samples of datasets with 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 percentages of retained PCA variances. The proposed method was then trained on each of the fractions of the dataset with different numbers of PCA components to determine the extent to which the reduction in the number of components affects the overall performance of the method. Table 5.8 presents the localization performance of the method in mean localization errors (m) for each retained PCA variance, as well as the number of trainable parameters corresponding to the percentage of the retained variance. On the other hand, Figure 5.4 presents full and zoomed-in CDF curves of test errors for each of the retained PCA variances. As indicated by Table 5.8, the localization accuracy of the method became increasingly poor as the percentages of the retained PCA variances became even smaller due to the information loss as a result of reducing the number of PCA components. The localization accuracy decreased from 290.71 m mean localization error when the dataset had 100% of retained PCA variances, to 504.28 m mean localization error when the dataset had just 10% of retained PCA variance. Regarding the trainable parameters of the method, the decrease in percentages of the retained PCA variances leads to a decrease in the number of trainable parameters due to the increasing number of PCA components being removed from the dataset. The reduction in the trainable parameters lowers the computational burden of the method, though it negatively affects its localization accuracy. However, the proposed method was able to learn useful positional information from the dataset with as small as 60% of retained PCA variance, consequently achieving satisfactory localization accuracy of 307.54 m mean

localization error with only 331,371 trainable parameters, proving its effectiveness and robustness. Though performing PCA on the training dataset could help in increasing the computational efficiency of the localization method by getting rid of PCA components with very little influence on the training data, the number of removed PCA components should not be excessive, not to severely affect the overall localization accuracy of the method.

Table 5.8. Performance of the CNN-transformer method on different percentages of retained PCA variances.

Retained variances (%)	**Mean localization error (m)**	**Number of trainable parameters**
10	504.28	183,915
20	348.63	200,299
30	319.39	224,875
40	311.33	257,643
50	309.35	298,603
60	307.54	331,371
70	306.45	372,331
80	305.38	413,291
90	301.34	470,635
100	290.71	536,171

5.3.3 Performance on different subsets of training, validation and test data

In order to explore how the proposed method performs when trained on a reduced sample size, smaller subsets of the data were first extracted from the dataset and split into training, validation, and test sets. This also leads to unique subsets of data in each scenario, which reduces the impact of spurious artefacts in the dataset on the resulting performance. The proposed method was then trained on the new sample sizes, yielding localization results in terms of mean localization errors (m) as indicated in Table 5.9. Figure 5.5 shows the full and zoomed-in CDF curves of localization errors for all the sample sizes. As observed from these results, the proposed method achieved the lowest localization accuracy of 633.59 m mean localization error when 20% of the original dataset was used. In this instance, the low accuracy stems from the limited number of training samples, which hindered the

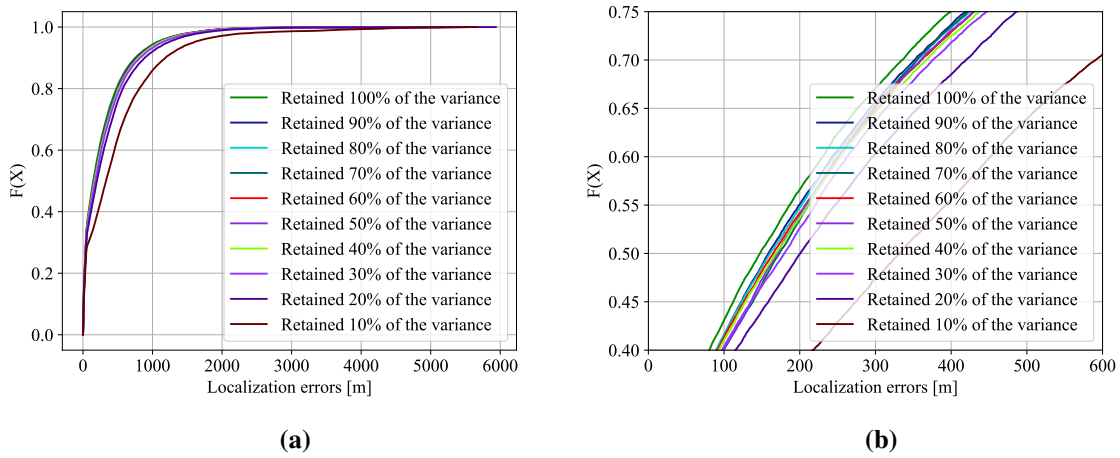

Figure 5.4. CDFs of localization errors (m) for the CNN-transformer method when trained on datasets with different retained PCA variances.
(a) Full CDF curves. (b) Zoomed-in CDF curves.

model from capturing sufficient patterns in the training data. This limitation prevented the efficient learning of meaningful representations of features in individual samples and introduced dependencies with other samples in the training set. The localization performance improved to 345.14 m and then to 304.12 m mean localization errors when the sample size was increased to 40% and 60% of the original dataset, respectively. This improvement in the localization accuracy is attributed to exposing the model to more training samples, which increased the generalization ability of the proposed method. The localization accuracy further improved to 294.14 m and 290.71 m mean localization errors when 80% and 100% of the original dataset were used, respectively, which further supports the notion that larger dataset sizes boost the performance of deep learning-based localization models in the context of LoRaWAN networks. The variations in the localization accuracies being observed in each of the adopted data split strategies, stem from variations in the size of training data, which either limits the model from learning all useful features when the size of the data is relatively small, leading to lower localization accuracies or enables the model to capture more sufficient patterns when the size of the training data increases, which eventually improves the localization accuracies. Despite using different sets of training, validation, and test data to train, validate and test the accuracy of the proposed method, satisfactory localization results were obtained even with a 40% reduction in the sample size, further proving its effectiveness and robustness in localizing a node in LoRaWAN networks.

Table 5.9. Performance of the CNN-transformer method on different subsets of training, validation and test data.

Used fraction of dataset (%)	**Mean localization error (m)**
20	633.59
40	345.14
60	304.12
80	294.14
100	290.71

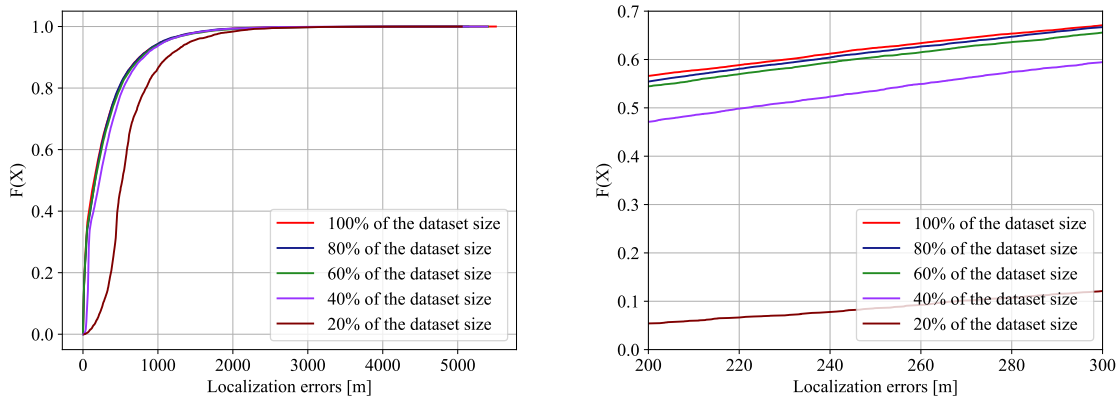

Figure 5.5. CDFs of localization errors (m) for the CNN-transformer method when trained on different sample sizes.
(a) Full CDF curves. (b) Zoomed-in CDF curves.

5.3.4 Performance for different sample sizes of training data with fixed test set

In this section, unlike in Section 5.3.3, a fixed 15% of the original LoRaWAN dataset was extracted and set aside to test the performance of the proposed localization method. Using a fixed test set, ensures consistency in the evaluation process across the different variations of training/validation sets, allowing for a fair and reliable performance comparison. For the remaining 85% of the LoRaWAN dataset, fractions of 40, 60, 80, and 100 percentages were extracted, and each was split into a training set containing 80% of the remaining samples and a validation set containing 20% of the remaining samples.

The proposed method was then trained on the new training and validation sets and tested on the fixed test set, yielding localization results in mean localization errors (m), as indicated in Table 5.10. Figure 5.6 shows the full and zoomed-in CDF curves of localization errors for all the extracted fractions of the LoRaWAN dataset. The lowest localization accuracy of 310.31 m mean localization error being observed when 40% of the remaining dataset was used, is a result of exposing the proposed method to a relatively small training sample size, which limited the ability of the method to capture more useful information from the training data to increase its generalization ability on the unseen data. The localization accuracy improved to 304.08 m and then to 294.07 m mean localization errors when 60% and 80% of the remaining dataset was used, respectively, due to the increase in training samples. When the entire remaining dataset was used, the performance of the proposed method improved slightly to 294.04 m mean localization error. This slight improvement in the localization performance is due to the fact that the model has learned nearly all of the useful information from the training dataset to infer the node's location from the fixed test set. Similar to what was observed in the previous section, the observed variations in the localization accuracies in this section are mainly due to variations in the size of the training data, which either limit or improve the learning capabilities of the proposed localization model. These results show that the proposed method can yield satisfactory localization performance when using at least 60% of the remaining LoRaWAN dataset, further indicating the importance of training deep learning models by using relatively large datasets for improved performances. Overall, these results have proved the effectiveness, robustness and hence the potential of the proposed method to localize a node in LoRaWAN networks.

Table 5.10. Performance of the CNN-transformer method on fixed test set using different subsets of training and validation sets of the remaining dataset.

Used fraction of the remaining dataset (%)	**Mean localization error (m)**
40	310.31
60	304.08
80	294.07
100	294.04

5.3.5 Performance for different data shuffling seeds

In this section, the impact of training the proposed method by using datasets that were shuffled using different shuffling seeds, is analyzed. Similar to what had been done with the first method, the training

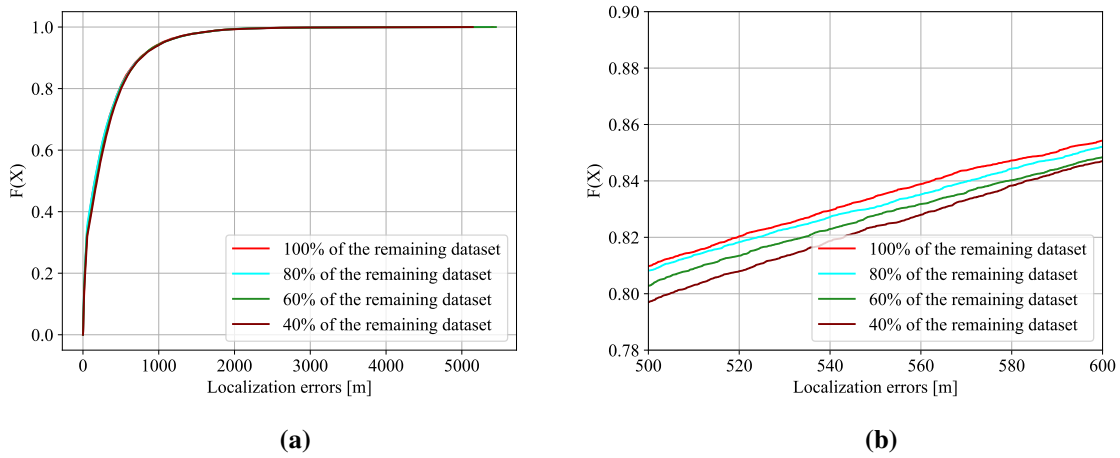

Figure 5.6. CDFs of localization errors (m) for the CNN-transformer method when trained on different sample sizes and tested on a fixed test set.
(a) Full CDF curves. (b) Zoomed-in CDF curves.

data was shuffled with 1, 2, 3, 5, 10, 14, 16, 22, 42, 64, 128, and 512 random seeds before splitting into training, validation, and test sets. The slight variations in the localization accuracies shown in Table 5.11 (with 1 and 42 seeds achieving the highest localization accuracies) are due to differences in the data values of the extracted training, validation, and test sets, as a result of using different shuffling seeds, which in essence changes the order of how the individual samples appear in the dataset. However, the effectiveness and robustness of the method are evident from the results presented in Table 5.11, as there are no significant differences in the localization accuracies.

5.4 COMPARATIVE REMARKS ON THE PERFORMANCE OF THE PROPOSED CONVOLUTIONAL NEURAL NETWORK-SQUEEZE AND EXCITATION AND CONVOLUTIONAL NEURAL NETWORK-TRANSFORMER FINGERPRINTING LOCALIZATION METHODS

In this section, comparative remarks on the performances of the two fingerprint-based localization methods being proposed in this thesis to estimate the locations of target nodes in LoRaWAN networks, are presented. The comparison is made with respect to three aspects: when the two methods were trained using datasets with different percentages of retained PCA variances; when the methods were trained using different subsets of the training, validation and test sets; and when the methods were trained using different sample sizes of training data and tested on the same fixed test set.

Table 5.11. Localization performance of the CNN-transformer method in terms of mean errors (m) when trained on datasets shuffled with different shuffling seeds.

Data shuffling seeds	**Localization performance**
1	290.84
2	292.96
3	291.07
5	292.95
10	293.93
14	293.03
16	291.31
22	292.33
42	290.71
64	292.12
128	294.68
512	291.07

5.4.1 Performance comparison when trained on different retained PCA variances

In this section, localization accuracies of the two proposed methods as presented in Sections 5.2.2 and 5.3.2, when the methods were trained on different percentages of PCA variances, are compared. The bar plot in Figure 5.7 shows the localization performances of the CNN-transformer and CNN-SE methods in terms of mean localization error (m) when trained on different retained percentages of the PCA variances. Both methods experience degradation in localization accuracy with reduced percentages of the retained PCA variances, which is understandable, since more and more PCA components have been removed. As indicated in Figure 5.7, on average, the CNN-transformer outperforms the CNN-SE method with the increase in percentages of retained PCA variances, reaching high localization accuracy of 290.71 m mean localization error when 100% of the PCA variances were retained, compared to 292.04 m mean localization error, which was achieved by the CNN-SE method with the same percentage of retained PCA variances. Additionally, with 536,171 trainable parameters, compared to 728,614 trainable parameters of the CNN-SE method (when the methods were trained on 100% of retained PCA variances), the CNN-transformer method is evidently lighter and hence less demanding

computationally than the CNN-SE method.

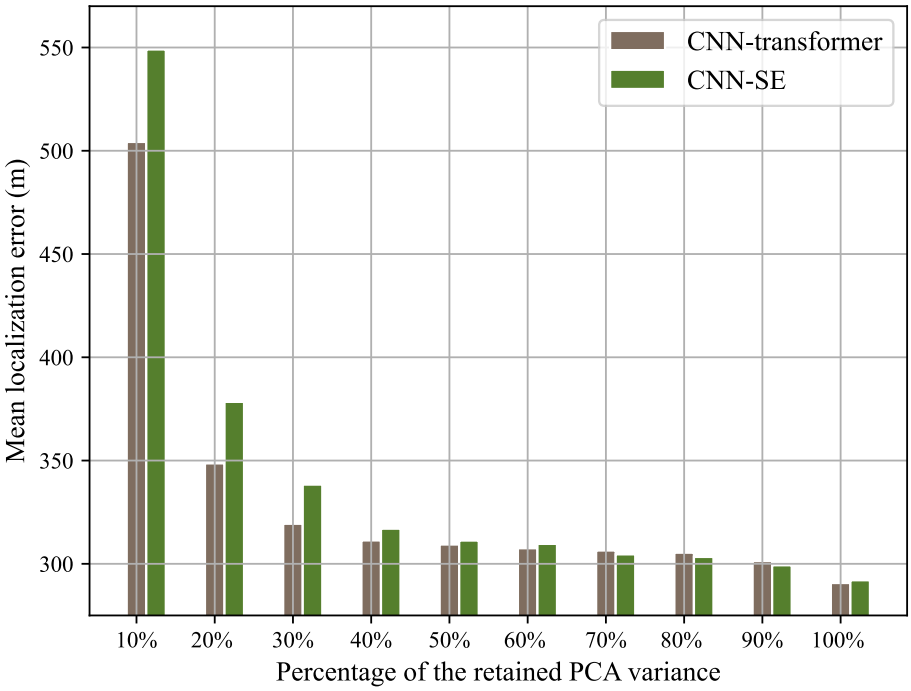

Figure 5.7. Performance of the proposed methods when trained on different retained PCA variances.

5.4.2 Performance comparison when trained on different subsets of training, validation and test sets

Figure 5.8 compares localization performances of the CNN-SE and CNN-transformer methods, presented in Sections 5.2.3 and 5.3.3, when the methods were trained on different subsets of training, validation and test sets. As indicated in Figure 5.8, the CNN-SE method outperforms the CNN-transformer with a reduction in the size of the training dataset; however, when both methods are exposed to increased sizes of the training dataset, the CNN-transformer method appears to generalize better than the CNN-SE method. This is attributed to the learning mechanism of the CNN-transformer method, which relies on the self-attention block, which excels with the increase in the training dataset.

5.4.3 Performance comparison when trained on different sample sizes of training data with fixed test set

In this section, localization performances of the two methods in terms of mean localization error (m) that were reported in Sections 5.2.4 and 5.3.4 (when the methods were trained on different sample sizes of the remaining dataset after fixing a portion of the original dataset for testing the localization performances of the methods), are compared. As indicated in Figure 5.9, similar performance trends of

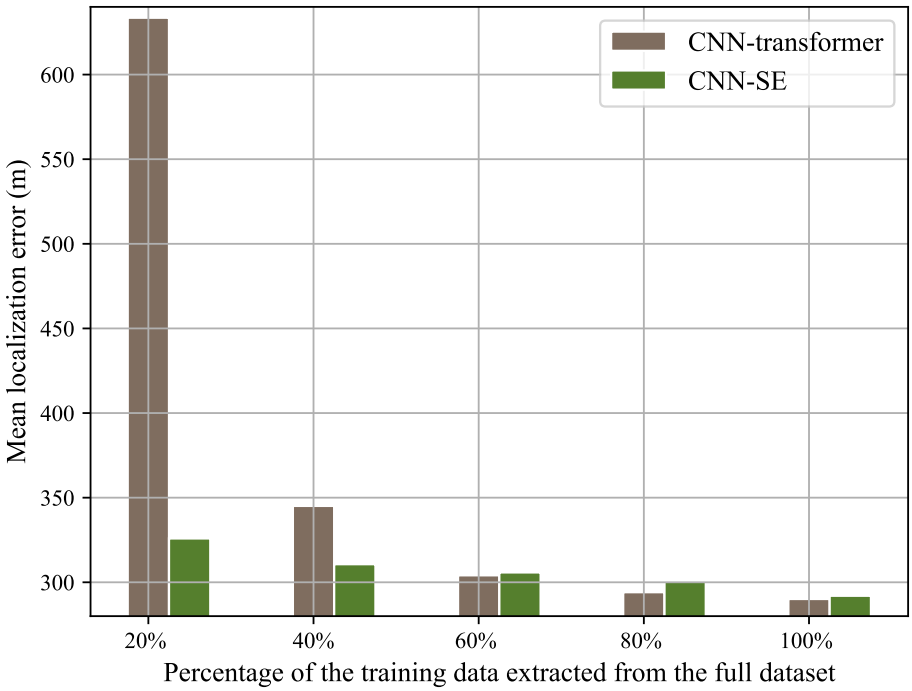

Figure 5.8. Performance of the proposed methods when trained on different subsets of training, validation and test sets.

the two methods that were observed in Section 5.4.2 are also observed here, i.e., the CNN-transformer method shows its superiority compared to the CNN-SE method, when the methods are exposed to more and more training data. A conclusive remark that can be drawn from the localization results presented in Figure 5.8 and Figure 5.9 is that, the effectiveness of the CNN-transformer method, compared to the CNN-SE in estimating locations of target nodes in LoRaWAN networks, will be more evident when a large number of features is present in the training dataset. Since the working principle of the CNN-transformer method is centered on the implementation of a self-attention mechanism that enables it to learn positional dependencies in parallel across each of the input sample sequences, by offering a pairwise correlation and long-range dependencies between tokens in the input sequences, it is likely to perform better and better with the increase in the number of features, as there will be more information to enable it to infer locations with relatively high accuracy.

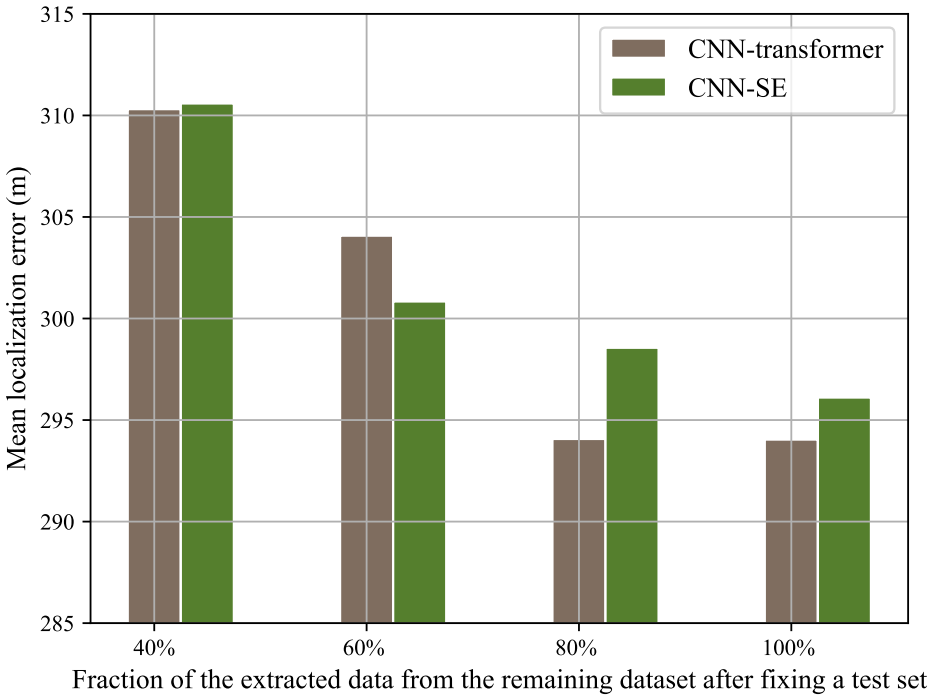

Figure 5.9. Performance of the proposed methods when trained on different sample sizes of training data with fixed test set.

5.5 PERFORMANCE COMPARISON WITH METHODS PROPOSED IN THE LITERATURE TRAINED USING THE SAME DATASET

In this section, the localization performance of the two proposed methods, i.e., CNN-SE and CNN-transformer, are compared with related works proposed in [18, 158–160], which were evaluated using the same dataset as the proposed methods. The metrics used to evaluate these related works vary from one work to another. Mean localization error, R^2 score, and validation time were adopted in [18], while mean and median errors were adopted in [159] and [160]. In [158], only mean localization error was reported.

In [18], the authors proposed several fingerprinting-based localization methods; among them, RF achieved the best results of 340 m mean localization error and 0.91 R^2 score. In [158], a KNN-RF ensemble method is proposed for fingerprinting localization, achieving localization accuracy of 332.63 m mean localization error. The authors in [159] proposed a fingerprinting localization method based on MLP, achieving localization accuracy of 57 m and 310 m median and mean localization errors,

respectively. In [160], on the other hand, a fingerprinting localization method based on k-means and weighted kernel regression is proposed, achieving localization accuracy of 158.48 m and 346 m median and mean localization errors, respectively. In all these research works, the same data split ratio of 0.7/0.15/0.15 for training, validation, and test sets was adopted. Additionally, in all these research works, GPS coordinates in latitudes and longitudes were used as ground truth references.

For a fair comparison, in addition to using the same dataset, the experimental environments and procedures should be the same for all the compared methods. Fulfilling this condition is challenging due to various reasons, including the unavailability of source codes, missing key information regarding experimental settings and procedures, and variations in the choice of metadata used for training the proposed methods in the related works. The comparison conducted in this thesis is thereby limited to the final localization performance reported in the related works, which is justifiable, since the ground truth references (latitudes and longitudes) were used in all of the proposed methods. Table 5.12 summarizes the key experimental settings and parameters associated with the compared related works.

The configuration of the fingerprinting-based methods proposed in this thesis, whose results were used for comparison, adopted a powed data representation scheme, using training data shuffled with 42 seeds with training data, validation data, and test data extracted by using a ratio of a 0.7/0.15/0.15 (same data split ratio as in the related works). In order to obtain a more general localization performance of the proposed methods, the localization results were evaluated by repeatedly randomizing the splitting of the entire dataset into three subsets. In each instance the subsets represented the training, validation and test sets. With respect to the first method i.e., CNN-SE method: In the first randomized experiment, a localization accuracy of 148.06 m and 292.04 m median and mean localization errors, respectively, were achieved on the test set. In the second randomised experiment a localization accuracy of 148.86 m and 288.82 m median and mean localization errors, respectively, were achieved. Finally, in the third randomized experiment, a localization performance of 145.72 m and 293.62 m median and mean localization errors, respectively, were achieved. In each experiment, the obtained R^2 score was 0.93. Therefore, on average, the proposed CNN-SE method achieved a localization accuracy of 147.55 m and 291.51 m median and mean localization errors, respectively, on the test set and an R^2 score of 0.93. Each randomized experiment consisted of 728,614 trainable parameters, and trained through 150 iterations. Figure 5.10 illustrates the mapping of the test set's true latitude and longitude coordinate pairs and the estimated latitude and longitude coordinate pairs for one of the randomized

Table 5.12. Summary of key experimental settings of the proposed methods and related works evaluated using the same dataset.

Research work	**Python libraries used**	**Experimental environment**	**Localization parameters**	**Ground truth references**	**Distance formula**
[18]	Only Scikit-Learn is mentioned	Virtual machine with 32 GB RAM memory and 10 CPU cores	RSSI	Latitudes and longitudes	Haversine[1]
[158]	Only Scikit-Learn is mentioned	Not mentioned	RSSI and TDoA	Latitudes and longitudes	Not mentioned
[159]	Only Scikit-Learn is mentioned	Not mentioned	Differential RSSI and TDoA	Latitudes and longitudes	Not mentioned
[160]	Not mentioned	Not mentioned	RSSI	Latitudes and longitudes	Not mentioned
CNN-SE method	Keras+Scikit-Learn+TensorFlow	Google Colab using a 32 GB RAM Core i7 LG computer	RSSI and SF	Latitudes and longitudes	Haversine
CNN-transformer method	Keras+Scikit-Learn+TensorFlow	Google Colab using a 32 GB RAM Core i7 LG computer	RSSI and SF	Latitudes and longitudes	Haversine

[1] The Haversine distance is not explicitly mentioned in [18]; however, it is described in the first author's PhD thesis [208].

experiments.

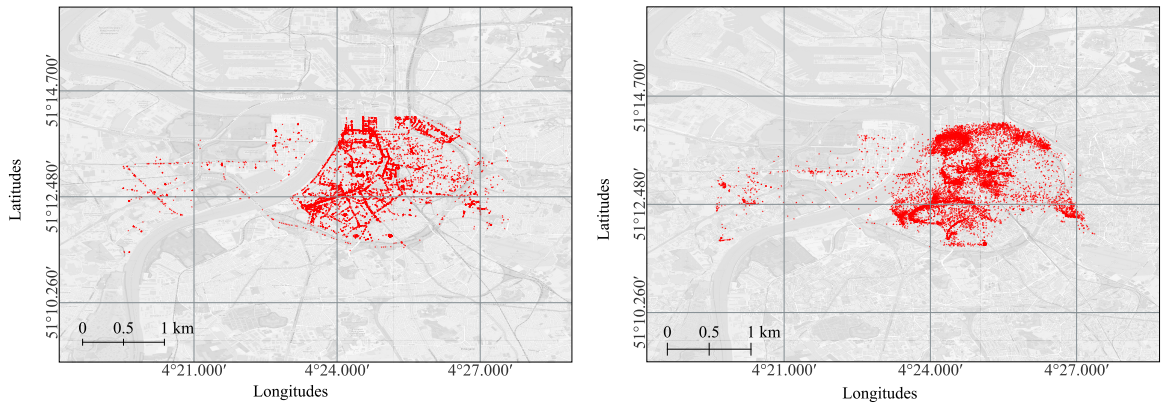

Figure 5.10. Mapping of the true latitude and longitude coordinate pairs of the test set and the estimated latitude and longitude coordinate pairs for the CNN-SE method in one of the randomized experiments. (a) True latitude and longitude coordinate pairs from the test set. (b) Estimated latitude and longitude coordinate pairs.

Regarding the second method, i.e., CNN-transformer method: The first randomized experiment achieved a localization accuracy of 283.79 m and 137.69 m mean and median localization errors, respectively; the second randomized experiment achieved a localization accuracy of 289.80 m and 146.08 m mean and median localization errors, respectively; the third randomized experiment, on the other hand, achieved localization accuracy of 290.71 m and 147.34 m mean and median localization errors, respectively. Similar to the first proposed method, an R^2 score of 0.93 was obtained in each of the three randomized experiments. Each randomized experiment consisted of 536,171 trainable parameters and was trained through 120 iterations. With these results, on average, the CNN-transformer method achieved a localization accuracy of 288.1 m and 143.7 m mean and median localization errors, respectively. Figure 5.11 shows the spatial distribution of the data points of the true latitude and longitude coordinate pairs of the test set and the estimated latitude and longitude coordinate pairs for one of the randomized experiments.

The performance comparison between the two fingerprinting localization methods being proposed in this thesis and the related works in terms of mean localization error, median localization error, and R^2 score is presented in Table 5.13. The variations in the reported localization results between the proposed method and the related works are mainly due to the structural differences in the way that

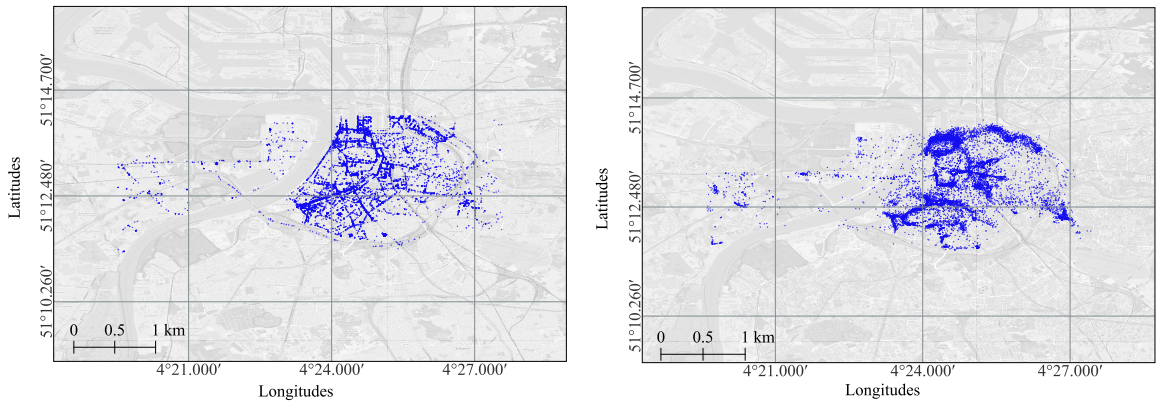

Figure 5.11. Spatial distribution of the data points of the true latitude and longitude coordinate pairs of the test set and the estimated latitude and longitude coordinate pairs for the CNN-transformer method in one of the randomized experiments.
(a) True latitude and longitude coordinate pairs from the test set. (b) Estimated latitude and longitude coordinate pairs.

the localization models were developed and trained, the machine learning technique being adopted, as well as preprocessing techniques adopted, which bring different model learning capabilities which determine the final outcome regarding the localization accuracies.

As indicated in Table 5.13, the two fingerprinting-based localization methods proposed in this thesis outperformed the best performing RF method among ten implemented methods in [18] in terms of mean error and R^2 score; outperformed the KNN-RF method proposed in [158], the MLP method proposed in [159] and a hybrid method that used k-means and weighted kernel regression proposed in [160], in terms of mean localization error. The variation in median errors, i.e., 57 m for the method proposed in [159], and 158.48 m for the method proposed in [160], compared to 147.55 m and 143.7 m obtained in this thesis for the first and second proposed methods, respectively, is due to variation in the number of outliers present in the error lists of the proposed methods. However, since the mean and median errors in each of the methods that have been proposed in this thesis are closer to each other than the mean and median errors of the other two methods, the distribution of errors computed by the proposed methods is less skewed, compared to the distribution of errors of the other proposed methods. Research works proposed in [158] and [159] require more computation time and memory space due to the use of training data containing fused RSSI and TDoA values, unlike the proposed methods in this

Table 5.13. Localization performances of the proposed methods compared to related works trained on the same dataset.

Research work	**Scheme**	**Mean localization error (m)**	**Median localization error (m)**	R^2 **Score**
[18]	Random forest	340	Not reported	0.91
[158]	KNN-RFR	332.63	Not reported	Not reported
[159]	MLP	310	57	Not reported
[160]	k-means + weighted kernel regression	346	158.48	Not reported
CNN-SE method	**CNN + SE**	**291.51**	**147.55**	**0.93**
CNN-transformer method	**CNN + transformer**	**288.1**	**143.7**	**0.93**

thesis, which used only RSSI values and the SF. The methods proposed in [18] and [160] are likely to require less computation time and memory space, compared to the proposed methods, due to a PCA that was performed on the training data in [18] to remain with 95% of the variance, and the use of only 27 gateways in [160] as a result of the removal of gateways with less than 1% visibility.

The superior localization accuracies achieved by the two fingerprinting-based localization methods proposed in this thesis, when compared to the related works, particularly in terms of mean localization error, and the robustness and effectiveness demonstrated by them when they were trained by using different sample sizes of training data as well as different data extraction strategies, is a clear proof of their potential to localize target nodes in LoRaWAN networks. An R^2 score of 0.93 obtained by both methods further indicates the high degree to which the proposed methods were able to fit data into the machine learning regressors, enabling them to localize target nodes with satisfactory localization accuracies. The resulting 728,614 (for the CNN-SE method) and 536,171 (for the CNN-transformer method) trainable parameters when the two methods were trained on the entire LoRaWAN dataset indicate that the two methods are relatively inexpensive regarding computational complexity, further affirming their suitability for deployment in real-world localization applications.

5.6 CHAPTER SUMMARY

In this chapter, a series of experiments were carried out to evaluate the localization performances of two fingerprinting-based localization methods that had been proposed in this thesis for estimating the location of target nodes in LoRaWAN networks, using a publicly available LoRaWAN dataset. Adopting a 0.7/0.15/0.15 data split scheme for the training, validation, and test set, respectively, and using the entire LoRaWAN dataset, the CNN-SE fingerprinting localization method achieved localization accuracies of 291.51 m and 147.55 m mean and median errors, respectively, on the test set using the powed data representation scheme. With the CNN-transformer method, the localization accuracy of 288.1 m and 143.7 m mean and median localization errors, respectively, were achieved by using the same experimental settings. The localization accuracies that were achieved by these two methods have outperformed the localization accuracies of the currently available state-of-the-art fingerprint-based localization methods in the literature, that were evaluated by using the same publicly available LoRaWAN dataset, further proving their effectiveness and robustness in estimating locations of target nodes in LoRaWAN networks with satisfactory localization accuracies.

CHAPTER 6 CONCLUSION

6.1 CHAPTER OVERVIEW

The main goal of this chapter is to provide researchers with a brief but comprehensive summary of the key insights into the main achievements and contributions being made in this thesis. Guided by the research questions and the specific research objectives, this thesis succeeded in proposing two novel fingerprinting-based localization methods, which highlight its main achievement. The first one is a branched CNN localization method enhanced with SE blocks, and the second one is a hybrid CNN-transformer fingerprinting method. These two methods were proposed to address the limitations of currently available fingerprinting-based localization solutions in unlicensed LPWAN networks (particularly LoRaWAN), which predominantly rely on conventional 'shallow' machine learning models. While such models may yield satisfactory results under specific conditions, their complexity tends to increase as the size of the training datasets increases, ultimately resulting in a decline in localization performance.

In this thesis, the presentation of the technical research contributions is in accordance with the fundamental principles guiding the scientific presentation of technical reports within the electronics, computer engineering, and wireless communication networks disciplines. The conclusive remarks presented in this chapter provide the research community with a summarized insight into the key concepts and ideas, and highlight the key findings and contributions made to the body of knowledge in the field of node localization in unlicensed LPWAN networks. Overall, in this chapter, researchers are provided with a consistent and coherent reporting of the main research achievements that will bring about a clear understanding of what is presented in the entire thesis.

6.2 SUMMARY OF CHAPTERS

In Chapter 1, justifications for conducting this research are provided, including the research problem that ought to be solved. Research questions that will guide the investigation of scientific methods to address the identified research problem are presented. In addition, the main research objective and several specific research objectives are presented to set out the goals that must be achieved at the end of this research. Moreover, a list of contributions to the scientific body of knowledge and a list of research articles that have been published or are still under review at the time of this thesis being written, are also provided.

In Chapter 2, key localization parameters used to implement the localization methods/techniques being adopted to estimate the target node's location are discussed. Additionally, two types of localization approaches, namely range-based and fingerprint-based approaches, which can be leveraged to estimate target nodes' locations in unlicensed LPWAN networks, particularly in LoRaWAN networks, are presented and discussed. For each of the localization approaches, types of localization methods/techniques currently available in the literature that are associated with it are presented along with a discussion on their implementation, their strengths and weaknesses, and their overall accuracy in locating a target node. Currently available localization solutions for both range and fingerprint-based approaches, that have been proposed in the literature to estimate locations of target nodes in LoRaWAN networks are also presented and discussed to give researchers an up-to-date status of localization research in LoRaWAN networks.

In Chapter 3, the network architecture of a general CNN model is discussed, describing its main features, regularization techniques that can be adopted to fight the over-fitting phenomenon (a prevailing challenge facing CNN models) and improve the generalization ability of CNN models, key network parameters needed to be defined in the compilation step of CNN models, as well as a brief description of its working principle. Additionally, two modules, namely squeeze and excitation blocks and a transformer module, which are central in the development of the proposed fingerprinting-based methods are presented and discussed, focusing on their main features and general working mechanism.

In Chapter 4, the model structural settings of the two proposed fingerprinting-based localization methods and four data representation schemes commonly used to improve the learning efficiency of machine learning models trained on RSSI values are presented, along with a description of the proposed methods's training strategy. Key procedures regarding their implementation to estimate

the locations of target nodes in LoRaWAN networks using a publicly available dataset are presented. These detailed implementation procedures, which include setting up experiments and preprocessing the training data, explain how the robustness and effectiveness of the proposed fingerprinting localization methods in estimating the locations of target nodes in LoRaWAN networks can be validated.

In Chapter 5, performance evaluation of the two proposed fingerprint-based methods, i.e., the CNN-SE method and CNN-transformer method, in localizing the target nodes in LoRaWAN networks using a publicly available LoRaWAN dataset, are presented. Regarding the former method, adopting a 0.7/0.15/0.15 data split scheme for the training, validation, and test set, respectively, and using the entire LoRaWAN dataset, the achieved localization accuracies were 291.51 m, and 147.55 m mean and median localization errors, respectively, on the test set, using the powed data representation scheme. The latter method achieved localization accuracy of 288.1 m and 143.7 m mean and median localization errors, respectively, using the same experimental settings. The localization accuracies achieved by these two methods have outperformed the localization accuracies of the currently available state-of-the-art fingerprint-based localization methods in the literature, evaluated by using the same publicly available LoRaWAN dataset, further proving their effectiveness and robustness in estimating locations of target nodes in LoRaWAN networks with acceptable levels of localization accuracies. Additionally, an R^2 score of 0.93 obtained by both methods further indicates the high degree to which the proposed methods have been able to fit data in their respective regressors, enabling them to localize target nodes with satisfactory localization accuracies.

6.3 SUMMARY OF CONTRIBUTIONS

The contribution of this research to the existing body of knowledge can be summarized as follows:

1. The comprehensive literature review on LPWAN networks, short-range wireless communication technologies, localization solutions in wireless networks, currently available machine learning-based fingerprint localization approaches in unlicensed LPWAN networks, as well as the exploration of recent advancements in machine learning algorithms (deep learning in particular) and data preprocessing techniques, have provided valuable knowledge and tools crucial to assisting researchers in developing effective and accurate localization solutions in unlicensed LPWAN networks, particularly fingerprinting-based approaches.
2. An efficient, robust and effective deep learning-based fingerprint localization method based on

convolutional neural network and squeeze and excitation blocks (CNN-SE) has been proposed and developed to estimate the locations of target nodes in LoRaWAN networks with acceptable levels of localization accuracies. The use of SE blocks is to complement the efficiency of CNN in learning useful positional information in structured data by improving its channel-wise interdependencies.

3. An efficient, robust and effective deep learning-based fingerprint localization method based on the joint use of CNN and transformer module has been proposed and developed to estimate the locations of target nodes in LoRaWAN networks with acceptable levels of localization accuracies. The novel contribution of this method is the development of a hybrid CNN-transformer fingerprint-based localization model by leveraging the strengths of both CNNs and transformers. CNNs capture features from the input data at the local level, while the attention mechanism of the transformer captures features from the input data at the global level.
4. Adopting a 0.7/0.15/0.15 data split scheme for the training, validation, and test set, respectively, and using the entire LoRaWAN dataset, the CNN-SE method achieved localization accuracies of 291.51 m and 147.55 m mean and median localization errors, respectively, on the test set, using the powed data representation scheme. With the CNN-transformer method, the localization accuracy of 288.1 m and 143.7 m mean and median localization errors, respectively, were achieved using the same experimental settings. The localization accuracies achieved by these two methods have outperformed the localization accuracies of the currently available state-of-the-art fingerprint-based localization methods in the literature, evaluated using the same publicly available LoRaWAN dataset. An R^2 score of 0.93 obtained by both methods further indicates the high degree to which the proposed methods have been able to fit data in their respective regressors, enabling them to localize target nodes with satisfactory localization accuracies.
5. The optimal preprocessing techniques for the LoRaWAN dataset for improved localization performance are presented. The experimental results revealed superior performances of the proposed methods when trained on the publicly available LoRaWAN dataset transformed using a powed data representation scheme compared to positive, normalized, and exponential data representation schemes.

6.4 FUTURE RESEARCH OPPORTUNITIES

Node localization research in LPWAN networks is still a relatively new field of research with many future research opportunities to explore. In the context of this thesis, future research directions can follow/take two major paths. The first one is building upon the two proposed novel localization

methods, i.e., a branched CNN localization method enhanced with SE blocks and a hybrid CNN-transformer fingerprinting method, by attempting to explore ways and techniques to improve the overall localization performance. The second one is to explore other aspects associated with node localization in unlicensed LPWAN networks.

Regarding the first path, the proposed methods can be extended as follows:

1. One could perform an analysis to assess how the overall performances of the proposed fingerprinting localization methods are influenced by including alternative localization parameters derived from the same publicly available LoRaWAN dataset, such as the fusion of differential RSSI-TDoA fingerprints.
2. One may adopt Bayesian optimization techniques to derive optimal hyperparameters for the proposed fingerprint-based localization methods to shorten the time spent in the 'try and error' search of optimal model hyperparameters.
3. Given the data-driven nature of the proposed methods, one could explore the use of interpolation techniques enhanced with autoencoders. This could effectively expand the size of the training dataset, thereby enhancing the localization performance of the proposed methods.
4. Concerning the practical implementation of the proposed methods in real-world IoT use cases, an opportunity exists to capitalize on the evolving capabilities of modern computing technologies like cloud, fog, and edge computing. Leveraging these technologies for processing location-related data can help to meet the quality-of-service requirements in real-time applications.

With regard to the second path, the following research opportunities can be explored:

1. Explore the use of transfer learning techniques to transfer the knowledge acquired in large fingerprint datasets of a source unlicensed LPWAN network (such as LoRaWAN) to increase the radio map's spatial resolution of relatively small fingerprint datasets of a related target unlicensed LPWAN technology (such as Sigfox) with a goal of improving the localization accuracy in the targeted domain.
2. Because fingerprint-based localization methods in unlicensed LPWAN networks are data-driven ones whose accuracy depends heavily on how the fingerprint databases have been built, one

could explore different data cleaning strategies, such as z-scores, to remove outliers from the datasets that might, in one way or another, hinder the performance of the proposed methods.

3. Filtering techniques, such as particle filters and Kalman filters, can be incorporated into fingerprinting-based localization methods to improve the spatial representation of individual data points in fingerprint databases and consequently improve the performances of proposed fingerprinting-based localization methods.

By exploring these open research issues, further advancement can be made in the field of node localization in unlicensed LPWAN networks by enhancing the performance of existing fingerprint-based localization methods or by contributing to proposing and developing other more efficient and reliable fingerprint-based localization methods.

In conclusion, this thesis has contributed to the understanding and improvement of fingerprinting-based localization methods in unlicensed LPWAN networks. The proposed branched CNN localization method enhanced with SE blocks and a hybrid CNN-transformer fingerprinting method present novel methods of estimating the locations of target objects through the use of a fingerprinting-based localization approach. By exploring the other research opportunities, more significant strides can be made in advancing this relatively new but important and very demanding research domain.

REFERENCES

[1] C. Milarokostas, D. Tsolkas, N. Passas, and L. Merakos, "A comprehensive study on LPWANs with a focus on the potential of LoRa/LoRaWAN systems," *IEEE Communications Surveys and Tutorials*, vol. 25, no. 1, pp. 825–867, Jul. 2023.

[2] Y. Chen, Y. A. Sambo, O. Onireti, and M. Imaran, "A survey on LPWAN-5G integration: Main challenges and potential solutions," *IEEE Access*, vol. 10, pp. 32 132–32 149, Mar. 2022.

[3] M. Chen, Y. Miao, X. Jian, X. Wang, and I. Humar, "Cognitive-LPWAN: Towards intelligent wireless services in hybrid low power wide area networks," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 2, pp. 409–417, Oct. 2019.

[4] K. Mekki, E. Bajic, F. Chaxel, and F. Meyer, "A comparative study of LPWAN technologies for large-scale IoT deployment," *ICT Express*, vol. 5, no. 1, pp. 1–7, Mar. 2019.

[5] M. Anjum, M. Khan, S. Hassan, H. Jung, and K. Dev, "Analysis of time-weighted LoRa-based positioning using machine learning," *Computer Communications*, vol. 193, pp. 266–278, Sep. 2022.

[6] M. Stusek *et al.*, "Accuracy assessment and cross-validation of LPWAN propagation models in urban scenarios," *IEEE Access*, vol. 8, pp. 154 625–154 636, Aug. 2020.

[7] Y. Song, J. Lin, M. Tang, and S. Dong, "An Internet of Energy Things based on wireless LPWAN," *Engineering*, vol. 3, pp. 460–466, Aug. 2017.

[8] M. Stusek *et al.*, "LPWAN coverage assessment planning without explicit knowledge of base station locations," *IEEE Internet of Things Journal*, vol. 9, no. 6, pp. 4031–4050, Aug. 2022.

[9] M. Aernouts, R. Berkvens, W. Xia, K. Vlaenderen, and M. Weyn, "Sigfox and LoRaWAN datasets for fingerprint localization in large urban and rural areas," *Data*, vol. 3, no. 2, Apr. 2018, Art. no. 13.

[10] E. Svertoka *et al.*, "Evaluation of real-life LoRaWAN localization: Accuracy dependencies analysis based on outdoor measurement datasets," in *Proceedings of the 14th International Conference on Communications (COMM)*, Jun. 2022, pp. 1–7.

[11] J. Purohit, X. Wang, S. Mao, X. Sun, and C. Yang, "Fingerprinting-based indoor and outdoor localization with LoRa and deep learning," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM-2020)*, Dec. 2020, pp. 1–6.

[12] Z. Li, X. Li, G. Mou, D. Jiang, X. Bao, and Y. Wang, "Design of localization system based on ultra-wideband and long range wireless," in *Proceedings of the IEEE 11th International Conference on Advanced Infocomm Technology (ICAIT)*, Oct. 2019, pp. 142–146.

[13] M. Anjum, M. Khan, S. Hassan, A. Mahmood, H. Qureshi, and M. Gidlund, "RSSI fingerprinting-based localization using machine learning in LoRa networks," *IEEE Internet of Things Magazine*, vol. 3, no. 4, pp. 53–59, Dec. 2020.

[14] A. Loganathan, N. Ahmad, and P. Goh, "Self-adaptive filtering approach for improved indoor localization of a mobile node with ZigBee-based RSSI and odometry," *Sensors*, vol. 19, no. 21, Nov. 2019, Art. no. 4748.

[15] B. Xu, B. Hussain, Y. Wang, H. Cheng, and C. Yue, "Smart home control system using VLC and Bluetooth-enabled AC light bulb for 3D indoor localization with centimeter-level precision," *Sensors*, vol. 22, no. 21, Oct. 2022, Art. no. 8181.

[16] C. Basri and A. Khadimi, "Survey on indoor localization system and recent advances of WiFi fingerprinting technique," in *Proceedings of the 5th International Conference on Multimedia*

Computing and Systems (ICMCS), Sep. 2016, pp. 253–259.

[17] C. Liu, C. Wang, J. Luo, and Q. He, "A cooperative indoor localization enhancement framework on edge computing platforms for safety-critical applications," in *Proceedings of the 15th International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*, Dec. 2019, pp. 372–377.

[18] T. Janssen, R. Berkvens, and M. Weyn, "Benchmarking RSS-based localization algorithms with LoRaWAN," *Internet of Things*, vol. 11, Sep. 2020, Art. no. 100235.

[19] Z. Shen, K. Xu, and X. Xia, "2D fingerprinting-based localization for mmWave cell-free massive MIMO systems," *IEEE Communications Letters*, vol. 25, no. 11, pp. 3556–3560, Sep. 2021.

[20] C. Liao, K. Xu, X. Xia, W. Xie, and M. Wang, "AOA-assisted fingerprint localization for cell-free massive MIMO system based on 3D multipath channel model," in *Proceedings of the IEEE 6th International Conference on Computer and Communications (ICCC)*, Dec. 2020, pp. 602–607.

[21] X. Sun, C. Wu, X. Gao, and G. Li, "Fingerprint-based localization for massive MIMO-OFDM system with deep convolutional neural networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 10 846–10 857, Sep. 2019.

[22] C. Wei *et al.*, "Joint AOA-RSS fingerprint based localization for cell-free massive MIMO systems," in *Proceedings of the IEEE 6th International Conference on Computer and Communications (ICCC)*, Dec. 2020, pp. 590–595.

[23] X. Sun, X. Gao, G. Li, and W. Han, "Single-site localization based on a new type of fingerprint for massive MIMO-OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6134–6145, Mar. 2018.

[24] X. Song *et al.*, "A novel convolutional neural network based indoor localization framework with WiFi fingerprinting," *IEEE access*, vol. 7, pp. 110 698–110 709, Aug. 2019.

[25] T. Koike-Akino, P. Wang, M. Pajovic, H. Sun, and P. Orlik, "Fingerprinting-based indoor localization with commercial MMWave WiFi: A deep learning approach," *IEEE access*, vol. 8, pp. 84 879–84 892, Apr. 2020.

[26] Z. Chen, H. Zou, J. Yang, H. Jiang, and L. Xie, "WiFi fingerprinting indoor localization using local feature-based deep LSTM," *IEEE Systems Journal*, vol. 14, no. 2, pp. 3001–3010, Jun. 2019.

[27] N. Kuxdorf-Alkirata, G. Maus, M. Gemci, and D. Brückmann, "A passive fingerprinting approach for device-free surveillance and localization applications using a Bluetooth low energy infrastructure," in *Proceedings of the IEEE 24th International Conference on Intelligent Engineering Systems (INES)*, Jul. 2020, pp. 31–36.

[28] S. Hu, K. He, X. Yang, and S. Peng, "Bluetooth fingerprint based indoor localization using Bi-LSTM," in *Proceedings of the IEEE 31st Wireless and Optical Communications Conference (WOCC)*, Aug. 2022, pp. 161–165.

[29] J. An and L. Choi, "Inverse fingerprinting: Server side indoor localization with Bluetooth low energy," in *Proceedings of the IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sep. 2016, pp. 1–6.

[30] M. Zhang and W. Cai, "Multivariate polynomial interpolation based indoor fingerprinting localization using Bluetooth," *IEEE sensors letters*, vol. 2, no. 4, pp. 1–4, Oct. 2018.

[31] J. Niu, B. Wang, L. Shu, T. Duong, and Y. Chen, "ZIL: An energy-efficient indoor localization system using ZigBee radio to detect WiFi fingerprints," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1431–1442, May. 2015.

[32] M. Aernouts, B. Bellekens, R. Berkvens, and M. Weyn, "A comparison of signal strength localization methods with Sigfox," in *Proceedings of the 15th Workshop on Positioning, Navigation and Communications (WPNC)*, Oct. 2018, pp. 1–6.

[33] B. Chaudhari, M. Zennaro, and S. Borkar, "LPWAN technologies : Emerging application

characteristics , requirements , and design considerations," *Future Internet*, vol. 12, no. 3, Mar. 2020, Art. no. 46.

[34] T. Janssen, A. Koppert, R. Berkvens, and M. Weyn, "A survey on IoT positioning leveraging LPWAN, GNSS, and LEO-PNT," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11 135–11 159, Feb. 2023.

[35] N. Islam, B. Ray, and F. Pasandideh, "IoT based smart farming : Are the LPWAN technologies suitable for remote communication?" in *Proceedings of the IEEE International Conference on Smart Internet of Things (SmartIoT)*, Aug. 2020, pp. 270–276.

[36] G. Stanco, A. Botta, F. Frattini, U. Giordano, and G. Ventre, "On the performance of IoT LPWAN technologies: The case of Sigfox, LoRaWAN and NB-IoT," in *Proceedings of the IEEE International Conference on Communications (ICC)*, May 2022, pp. 2096–2101.

[37] M. Hossain and J. Markendah, "Comparison of LPWAN technologies: Cost structure and scalability," *Wireless Personal Communications*, vol. 121, no. 1, pp. 887–903, Nov. 2021.

[38] M. Islam, H. Jamil, S. Pranto, R. Das, A. Amin, and A. Khan, "Future industrial applications: Exploring LPWAN-driven IoT protocols," *Sensors*, vol. 24, no. 8, Apr. 2024, Art. no. 2509.

[39] B. Chaudhari, M. Zennaro, and S. Borkar, "LPWAN technologies: Emerging application characteristics, requirements, and design considerations," *Future Internet*, vol. 12, no. 3, Mar. 2020, Art. no. 46.

[40] M. Theissen, L. Kern, T. Hartmann, and E. Clausen, "Use-case-oriented evaluation of wireless communication technologies for advanced underground mining operations," *Sensors*, vol. 23, no. 7, Mar. 2023, Art. no. 3537.

[41] U. Dang, H. Zessin, C. Rohde, and F. Leschka, "Energy-efficient user terminals for Internet of Things applications over satellite," in *Proceedings of the 37th International Communications Satellite Systems Conference (ICSSC)*, Oct. 2019, pp. 1–13.

[42] V. Orfanos, S. Kaminaris, P. Papageorgas, D. Piromalis, and D. Kandris, "A comprehensive review of IoT networking technologies for smart home automation applications," *Journal of Sensor and Actuator Networks*, vol. 12, no. 2, Apr. 2023, Art. no. 30.

[43] Wize Alliance, "Why Wize?" Wize-Alliance.com, Accessed: May 21, 2024. [Online]. Available: https://www.wize-alliance.com/Why-Wize/Technology

[44] N. Naik, "LPWAN technologies for IoT systems: Choice between ultra narrow band and spread spectrum," in *Proceedings of the IEEE International Systems Engineering Symposium (ISSE)*, Oct. 2018, pp. 1–8.

[45] E. Ogbodo, A. Abu-Mahfouz, and A. Kurien, "A survey on 5G and LPWAN-IoT for improved smart cities and remote area applications: From the aspect of architecture and security," *Sensors*, vol. 22, no. 16, Aug. 2022, Art. no. 366313.

[46] Q. Qadir, T. Rashid, N. Al-Salihi, B. Ismael, A. Kist, and Z. Zhang, "Low power wide area networks: A survey of enabling technologies, applications and interoperability needs," *IEEE Access*, vol. 6, pp. 77 454–77 473, Nov. 2018.

[47] A. Moradbeikie, A. Keshavarz, H. Rostami, S. Paiva, and S. Lopes, "GNSS-free outdoor localization techniques for resource-constrained IoT architectures: A literature review," *Applied Sciences*, vol. 11, Nov. 2021, Art. no. 10793.

[48] J. Sanchez-Gomez *et al.*, "Integrating LPWAN technologies in the 5G ecosystem : A survey on security challenges and solutions," *IEEE Access*, vol. 8, pp. 216 437–216 460, Nov. 2020.

[49] F. Zafari, A. Gkelias, and K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 3, pp. 2568–2598, Apr. 2019.

[50] Q. He, Y. Lan, W. Liu, and Y. Liu, "A comparison study of LPWAN technologies based on onsite measurements in a converter station hall," in *Proceedings of the 4th International Conference on Energy, Electrical and Power Engineering (CEEPE)*, Apr. 2021, pp. 842–846.

[51] S. Asaad and H. Maghdid, "A comprehensive review of indoor/outdoor localization solutions in IoT era : Research challenges and future perspectives," *Computer Networks*, vol. 212, Jul. 2022, Art. no. 109041.

[52] R. Chall, S. Lahoud, and M. Helou, "LoRaWAN network: Radio propagation models and performance evaluation in various environments in Lebanon," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2366–2378, Mar. 2019.

[53] J. Pospisil, R. Fujdiak, and K. Mikhaylov, "Investigation of the performance of TDoA-based localization over LoRaWAN in theory and practice," *Sensors*, vol. 20, no. 19, Sep. 2020, Art. no. 5464.

[54] I. Daramouskas, D. Mitroulias, I. Perikos, M. Paraskevas, and V. Kapoulas, "Localization in LoRa networks based on time difference of arrival," in *Studies in Computational Intelligence*, J. Kacprzyk, Ed. Cham, Switzerland: Springer, 2022, pp. 130–143.

[55] W. Ayoub, A. Samhat, F. Nouvel, M. Mroue, and J. Prevotet, "Internet of Mobile Things: Overview of LoRaWAN, DASH7, and NB-IoT in LPWANs standards and supported mobility," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 2, pp. 1561–1581, Oct. 2019.

[56] K. Nolan, W. Guibene, and M. Kelly, "An evaluation of low power wide area network technologies for the Internet of Things," in *Proceedings of the International Wireless Communications and Mobile Computing Conference (IWCMC)*, Sep. 2016, pp. 439–444.

[57] N. Akhmedov, H. Khujamatov, A. Lazarev, and M. Seidullarew, "Application of LPWAN technologies for the implementation of IoT projects in the Republic of Uzbekistan," in *Proceedings of the International Conference on Information Science and Communications Technologies (ICISCT)*, Nov. 2021, pp. 1–4.

[58] Y. Lykov, A. Paniotova, V. Shatalova, and A. Lykova, "Energy efficiency comparison LPWANs: LoRaWAN vs Sigfox," in *Proceedings of the IEEE International Conference on Problems of Infocommunications, Science and Technology*, Oct. 2020, pp. 485–490.

[59] Sigfox Build, "Sigfox network location," Build.Sigfox.com, Accessed: Mar. 14, 2024. [Online]. Available: https://build.sigfox.com/geolocation-technologies

[60] M. Iqbal, A. Abdullah, and F. Shabnam, "An application based comparative study of LPWAN technologies for IoT environment," in *Proceedings of the 2020 IEEE Region 10 Symposium (TENSYMP)*, Jun. 2020, pp. 1857–1860.

[61] M. Ivezic, "IoT wireless protocols," Society5.com, Accessed: May 23, 2024. [Online]. Available: https://society5.com/miot-5g/iot-wireless-protocols-spreadsheet

[62] Onomondo, "A complete guide to LTE-M for IoT," Onomondo.com, Accessed: May 23, 2024. [Online]. Available: https://onomondo.com/blog/lte-m-iot-guide

[63] B. Shilpa, R. Radha, and P. Movva, "Comparative analysis of wireless communication technologies for IoT applications," in *Artificial Intelligence and Technologies: Lecture Notes in Electrical Engineering*, R. Raje, F. Hussain, and R. Kannan, Eds. Gateway East, Singapore: Springer, 2022, pp. 383–394.

[64] B. Baker, J. Woods, M. Reed, and M. Afford, "A survey of short-range wireless communication for ultra-low-power embedded systems," *Journal of Low Power Electronics and Applications*, vol. 14, no. 2, May 2024, Art. no. 27.

[65] R. Balaji, R. Malathi, M. Priya, and K. Kannammal, "A comprehensive nomenclature of RFID localization," in *Proceedings of the International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 2020, pp. 1–9.

[66] K. Duan and S. Cao, "Emerging RFID technology in structural engineering: A review," *Structures*, vol. 28, pp. 2404–2414, Dec. 2020.

[67] M. Boros, "Reliability testing of RFID reading devices as part of localization systems," in *Proceedings of the 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, Jul. 2023, pp. 1–6.

[68] W. Tao, L. Zhao, G. Wang, and R. Liang, "Review of the Internet of Things communication technologies in smart agriculture and challenges," *Computers and Electronics in Agriculture*, vol. 189, Oct. 2021, Art. no. 106352.

[69] S. Hayward, K. Lopik, C. Hinde, and A. West, "A survey of indoor location technologies, techniques and applications in industry," *Internet of Things*, vol. 20, Nov. 2022, Art. no. 100608.

[70] S. Fu, S. Zhang, D. Jiang, and X. Liu, "Real-time and accurate RFID tag localization based on multiple feature fusion," in *Proceedings of the 16th International Conference on Mobility, Sensing and Networking (MSN)*, Dec. 2020, pp. 694–699.

[71] L. Hamami and B. Nassereddine, "Application of wireless sensor networks in the field of irrigation: A review," *Computers and Electronics in Agriculture*, vol. 179, Dec. 2020, Art. no. 105782.

[72] J. Yao *et al.*, "Comprehensive study on MIMO-related interference management in WLANs," *IEEE Communication Surveys and Tutorials*, vol. 21, pp. 2087–2110, Jan. 2019.

[73] A. Zohourian *et al.*, "IoT ZigBee device security: A comprehensive review," *Internet of Things*, vol. 22, Jul. 2023, Art. no. 100791.

[74] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on Internet of Things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1125–1142, Mar. 2017.

[75] C. Prakash, L. Singh, A. Gupta, and S. Lohan, "Advancements in smart farming: A comprehensive review of IoT, wireless communication, sensors, and hardware for agricultural automation," *Sensors and Actuators A: Physical*, vol. 362, Aug. 2023, Art. no. 114605.

[76] A. Rao *et al.*, "Real-time monitoring of construction sites: Sensors, methods, and applications," *Automation in Construction*, vol. 136, Apr. 2022, Art. no. 104099.

[77] M. Raj *et al.*, "A survey on the role of Internet of Things for adopting and promoting agriculture 4.0," *Journal of Network and Computer Applications*, vol. 187, Aug. 2021, Art. no. 103107.

[78] C. Li, J. Cheng, and K. Chen, "Top 10 technologies for indoor positioning on construction sites," *Automation in Construction*, vol. 118, Oct. 2020, Art. no. 103309.

[79] Y. Zhuang, C. Zhang, Y. Li, L. Chen, and R. Chen, "Bluetooth localization technology: Principles, applications, and future trends," *IEEE Internet of Things*, vol. 9, no. 23, pp. 23 506–23 524, Sep. 2022.

[80] I. Sergi, M. Gammariello, T. Montanaro, E. Imperiale, A. Shumba, and L. Parono, "A literature review on outdoor localization systems based on the Bluetooth technology," in *Proceedings of the 7th International Conference on Smart and Sustainable Technologies (SpliTech)*, Jul. 2022, pp. 1–5.

[81] Z. Hajiakhondi-Meybodi, M. Salimibeni, A. Mohammadi, and K. Plataniotis, "Bluetooth low energy and CNN-based angle of arrival localization in presence of Rayleigh fading," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Jun. 2021, pp. 7913–7917.

[82] W. Kassab and K. Darabkh, "A–Z survey of Internet of Things: Architectures, protocols, applications, recent advances, future directions and recommendations," *Journal of Network and Computer Applications*, vol. 163, Aug. 2020, Art. no. 102663.

[83] L. Danys *et al.*, "Visible light communication and localization: A study on tracking solutions for industry 4.0 and the operator 4.0," *Journal of Manufacturing Systems*, vol. 64, pp. 535–545, Jul. 2022.

[84] ECO, "Annex 11: Radio frequency identification applications," EFIS.CEPT.org, Accessed: May 24, 2024. [Online]. Available: https://efis.cept.org/adhoc-grabber.jsp?annex=14

[85] A. Bhujel *et al.*, "Sensor systems for greenhouse microclimate monitoring and control: A review," *Journal of Biosystems Engineering*, vol. 45, pp. 341–361, Dec. 2020.

[86] Z. Alavikia and M. Shabro, "A comprehensive layered approach for implementing Internet of Things-enabled smart grid: A survey," *Digital Communications and Networks*, vol. 8, pp. 388–410, Jun. 2022.

[87] J. Liu, C. Chen, Y. Ma, and Y. Xu, "Energy analysis of device discovery for Bluetooth low energy," in *Proceedings of the IEEE 78th Vehicular Technology Conference (VTC-Fall)*, Sep. 2013, pp. 1–5.

[88] Z. Li and Y. Chen, "Bluefi: Physical-layer cross-technology communication from Bluetooth to WiFi," in *Proceedings of the IEEE 40th International Conference on Distributed Computing Systems (ICDCS)*, Nov. 2020, pp. 399–409.

[89] A. Lutakamale, H. Myburgh, and A. De Freitas, "RSSI-based fingerprint localization in LoRaWAN networks using CNNs with squeeze and excitation blocks," *Ad Hoc Networks*, vol. 159, Jun. 2024, Art. no. 103486.

[90] L. Cheng, Y. Li, M. Xue, and Y. Wang, "An indoor localization algorithm based on modified joint probabilistic data association for wireless sensor network," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 63–72, Mar. 2021.

[91] Y. Zhuang *et al.*, "A survey of positioning systems using visible LED lights," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 3, pp. 1963–1988, Feb. 2018.

[92] A. Abba, J. Sanusi, O. Oshiga, and S. Mikail, "A review of localization techniques in wireless sensor networks," in *Proceedings of the 2nd International Conference on Multidisciplinary Engineering and Applied Science (ICMEAS)*, Nov. 2023, pp. 1–5.

[93] M. Maheepala, A. Kouzani, and M. Joordens, "Light-based indoor positioning systems: A review," *IEEE Sensors Journal*, vol. 20, no. 8, pp. 3971–3995, Jan. 2020.

[94] S. Sivasakthiselvan and V. Nagarajan, "Localization techniques of wireless sensor networks: A review," in *Proceedings of the International Conference on Communication and Signal Processing (ICCSP)*, Jul. 2020, pp. 1643–1648.

[95] N. Alrajeh, M. Bashir, and B. Shams, "Localization techniques in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2013, pp. 1–9, Jun. 2013.

[96] W. Liu *et al.*, "Survey on CSI-based indoor positioning systems and recent advances," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sep. 2019, pp. 1–8.

[97] G. Mintsis, S. Basbas, P. Papaioannou, C. Taxiltaris, and I. Tziavos, "Applications of GPS technology in the land transportation system," *European Journal of Operational Research*, vol. 152, pp. 399–409, Jan. 2004.

[98] S. Kumar and K. Moore, "The evolution of global positioning system (GPS) technology," *Journal of Science Education and Technology*, vol. 11, no. 1, pp. 59–80, Mar. 2002.

[99] C. Rizos, "Trends in GPS technology and applications," in *Proceedings of the 2nd International Workshop on Location Based Systems*, Jul. 2003, pp. 1–41.

[100] M. Abinaya and R. Devi, "Intelligent vehicle control using wireless embedded system in transportation system based on GSM and GPS technology," *International Journal of Computer Science and Mobile Computing*, vol. 3, no. 9, pp. 244–258, Sep. 2014.

[101] J. Luo, Z. Yin, L. Gui, and X. Yang, "Accurate localization for indoor and outdoor scenario by GPS and UWB fusion," in *Proceedings of the 9th International Conference on Control, Automation and Robotics (ICCAR)*, Apr. 2023, pp. 411–416.

[102] E. Zhang and N. Masoud, "Increasing GPS localization accuracy with reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2615–2625, Feb. 2021.

[103] X. Gong, X. Yu, X. Liu, and X. Gao, "Machine learning-based fingerprint positioning for massive MIMO systems," *IEEE Access*, vol. 10, pp. 89 320–89 330, Aug. 2022.

[104] A. Khalajmehrabadi, N. Gatsis, and D. Akopian, "Modern WLAN fingerprinting indoor positioning methods and deployment challenges," *IEEE Communications Surveys and Tutorials*, vol. 19, no. 3, pp. 1974–2002, Mar. 2017.

[105] A. Nessa, B. Adhikari, F. Hussain, and X. Ferdando, "A survey of machine learning for indoor positioning," *IEEE Access*, vol. 8, pp. 214 945–214 965, Nov. 2020.

[106] E. Goldoni, L. Prando, A. Vizziello, P. Savazzi, and P. Gamba, "Experimental data set analysis of RSSI-based indoor and outdoor localization in LoRa networks," *Internet Technology Letters*, vol. 2, no. 1, Jan. 2019.

[107] C. Bouras, A. Gkamas, V. Kokkinos, and N. Papachristos, "Time difference of arrival localization study for SAR systems over LoRaWAN," in *Proceedings of the 15th International Conference on Future Networks and Communications (FNC)*, Jan. 2020, pp. 292–299.

[108] M. Chen, H. Zhao, C. Shi, X. Chen, and D. Niu, "Multi-scene LoRa positioning algorithm based on Kalman filter and its implementation on NS3," *Ad Hoc Networks*, vol. 141, Mar. 2023, Art. no. 103097.

[109] J. Liu, J. Gao, S. Jha, and W. Hu, "Seirios: Leveraging multiple channels for LoRaWAN indoor and outdoor localization," in *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, Oct. 2022, pp. 656–669.

[110] W. Bakkali, M. Kieffer, M. Lalam, and T. Lestable, "Kalman filter-based localization for Internet of Things LoRaWAN end points," in *Proceedings of the IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–6.

[111] J. Luomala and I. Hakala, "Adaptive range-based localization algorithm based on trilateration and reference node selection for outdoor wireless sensor networks," *Computer Networks*, vol. 210, Jun. 2022, Art. no. 108865.

[112] M. Rusli, M. Ali, N. Jamil, and M. Din, "An improved indoor positioning algorithm based on

RSSI-trilateration technique for Internet of Things (IoT)," in *Proceedings of the International Conference on Computer and Communication Engineering (ICCCE)*, Jul. 2016, pp. 72–77.

[113] S. Pradhan, Y. Bae, J. Pyun, N. Ko, and S. Hwang, "Hybrid ToA trilateration algorithm based on line intersection and comparison approach of intersection distances," *Energies*, vol. 12, no. 9, May 2019, Art. no. 1668.

[114] G. Han, J. Jiang, C. Zhang, T. Duong, M. Guizani, and G. Karagiannidis, "A survey on mobile anchor node assisted localization in wireless sensor networks," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 3, pp. 2220–2243, Mar. 2016.

[115] Y. Li *et al.*, "Toward location-enabled IoT (LE-IoT): IoT positioning techniques, error sources, and error mitigation interaction feature," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4035–4062, Sep. 2020.

[116] B. Ando, S. Baglio, R. Crispino, and V. Marletta, "An introduction to indoor localization techniques. Case of study: A multi-trilateration-based localization system with user–environment interaction feature," *Applied Sciences*, vol. 11, no. 16, Aug. 2021, Art. no. 7392.

[117] C. Laoudias, A. Moreira, S. Kim, S. Lee, L. Wirola, and C. Fischione, "A survey of enabling technologies for network localization, tracking, and navigation," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 4, pp. 3607–3644, Jul. 2018.

[118] G. Zanca, F. Zorzi, A. Zanella, and M. Zorzi, "Experimental comparison of RSSI-based localization algorithms for indoor wireless sensor networks," in *Proceedings of the Workshop on Real-World Wireless Sensor Networks (REALWSN'08)*, Apr. 2008, pp. 1–5.

[119] C. Karney, "Algorithms for geodesics," *Journal of Geodesy*, vol. 87, no. 1, pp. 43–55, Jul. 2013.

[120] W. Bakkali, M. Kieffer, M. Lalam, and T. Lestable, "Kalman filter-based localization for Internet of Things LoRaWAN end points," in *Proceedings of the IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–6.

[121] Y. Kim and H. Bang, "Introduction to Kalman filter and its applications," in *Introduction and Implementations of the Kalman Filter*, F. Govaers, Ed. London, United Kingdom: IntechOpen, 2019, pp. 1–16.

[122] M. Aernouts *et al.*, "Combining TDoA and AoA with a particle filter in an outdoor LoRaWAN network," in *Proceedings of the IEEE/ION Position, Location and Navigation Symposium (PLANS)*, Apr. 2020, pp. 1060–1069.

[123] J. Desay and U. Tureli, "Evaluating performance of various localization algorithms in wireless and sensor networks," in *Proceedings of the IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2007, pp. 1–5.

[124] A. Vazquez-Rodas, F. Astudillo-Sainas, C. Sanchez, B. Arpi, and L. Minchala, "Experimental evaluation of RSSI-based positioning system with low-cost LoRa devices," *Ad Hoc Networks*, vol. 105, Aug. 2020, Art. no. 102168.

[125] R. Muppala, A. Navnit, D. Devendra, E. Materay, N. Accetturaz, and A. Hussain, "Feasibility of standalone TDoA-based localization using LoRaWAN," in *Proceedings of the International Conference on Localization and GNSS (ICL-GNSS)*, Jun. 2021, pp. 1–7.

[126] D. Guo, C. Gu, L. Jiang, W. Luo, and R. Tan, "ILLOC: In-hall localization with standard LoRaWAN uplink frames," *ACM Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 1, pp. 1–26, Mar. 2022.

[127] M. Aernouts, N. Bnilam, R. Berkvens, and M. Weyn, "TDAoA: A combination of TDoA and AoA localization with LoRaWAN," *Internet of Things*, vol. 11, Sep. 2020, Art. no. 100236.

[128] I. ElNaqa, R. Li, and M. Murphy, *Machine Learning in Radiation Oncology*. Cham, Switzerland: Springer International Publishing, 2015.

[129] J. Xie *et al.*, "A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 1, pp. 393–430, Aug. 2019.

[130] M. Hossan, R. Noor, K. Yau, S. Azzuhri, M. Z'Aba, and I. Ahmedy, "Comprehensive survey of machine learning approaches in cognitive radio-based vehicular Ad Hoc networks," *IEEE Access*, vol. 8, pp. 78 054–78 108, Apr. 2020.

[131] V. Mirama, L. Diez, A. Bahillo, and V. Quintero, "A survey of machine learning in pedestrian localization systems: Applications, open issues and challenges," *IEEE Access*, vol. 9, pp. 120 138–120 157, Aug. 2021.

[132] A. Qayyum, J. Qadir, M. Bilal, and A. Al-Fuqaha, "Secure and robust machine learning for healthcare: A survey," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 156–180, Jul. 2021.

[133] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electron Mark*, vol. 31, pp. 685–695, Sep. 2021.

[134] M. Al-Garadi, A. Mohamed, A. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Communications Surveys and Tutorials*, vol. 22, no. 3, pp. 1646–1685, Apr. 2020.

[135] M. Ridwan, N. Radzi, F. Abdullah, and Y. Jalil, "Applications of machine learning in networking: A survey of current issues and future challenges," *IEEE Access*, vol. 9, pp. 52 523–52 556, Mar. 2021.

[136] S. Aslam, A. Jikani, J. Sultana, and L. Almutairi, "Feature evaluation of emerging e-learning systems using machine learning: An extensive survey," *IEEE Access*, vol. 9, pp. 69 573–69 587, May 2021.

[137] E. Bouti, V. Loscri, and A. Gallais, "How machine learning changes the nature of cyberattacks on IoT networks: A survey," *IEEE Communications Surveys and Tutorials*, vol. 24, no. 1, pp. 248–279, Nov. 2022.

[138] M. Mahmood, M. Matin, P. Sarigiannidis, and S. Goudos, "A comprehensive review on artificial intelligence/machine learning algorithms for empowering the future IoT towards 6G era," *IEEE*

Access, vol. 10, pp. 87 535–87 562, Aug. 2022.

[139] A. Jamalipour and S. Murali, "A taxonomy of machine-learning-based intrusion detection systems for the Internet of Things: A survey," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9444–9466, Nov. 2021.

[140] K. Shaukat, S. Luo, V. Varadharajan, I. Hameed, and M. Xu, "A survey on machine learning techniques for cyber security in the last decade," *IEEE access*, vol. 8, pp. 222 310–222 354, Dec. 2020.

[141] N. Choudhry, J. Abawajy, S. Huda, and I. Rao, "A comprehensive survey of machine learning methods for surveillance videos anomaly detection," *IEEE Access*, vol. 11, pp. 114 680–114 713, Oct. 2023.

[142] P. Klaine, M. Imran, O. Onireti, and R. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE communications surveys and tutorials*, vol. 19, no. 4, pp. 2392–2431, Jul. 2017.

[143] M. Ozkan-Ozay *et al.*, "A comprehensive survey: Evaluating the efficiency of artificial intelligence and machine learning techniques on cyber security solutions," *IEEE Access*, vol. 12, pp. 12 229–12 256, Jan. 2024.

[144] T. Kim, L. Vecchietti, K. Choi, S. Lee, and D. Har, "Machine learning for advanced wireless sensor networks: A review," *IEEE Sensors Journal*, vol. 21, no. 11, pp. 12 379–12 397, Nov. 2020.

[145] L. Ruan, M. Dias, and E. Wong, "Enhancing latency performance through intelligent bandwidth allocation decisions: A survey and comparative study of machine learning techniques," *Journal of Optical Communications and Networking*, vol. 12, no. 4, pp. B20–B32, Apr. 2020.

[146] S. Band *et al.*, "When smart cities get smarter via machine learning: An in-depth literature review," *IEEE Access*, vol. 10, pp. 60 985–61 015, Jun. 2022.

[147] M. Khanesar, M. Yan, W. Syam, S. Piano, R. Leach, and D. Branson, "A neural network separation approach for the inclusion of static friction in nonlinear static models of industrial robots," *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 6, pp. 3294–3304, Apr. 2023.

[148] E. Laftchiev and D. Nikovski, "An IoT system to estimate personal thermal comfort," in *Proceedings of the IEEE 3rd World Forum on Internet of Things (WF-IoT)*, Dec. 2016, pp. 672–677.

[149] R. Almejrb, O. Sallabi, and A. Mohamed, "Applying C Atboost regression model for prediction of house prices," in *Proceedings of the International Conference on Engineering and MIS (ICEMIS)*, Jul. 2022, pp. 1–7.

[150] A. Modak, T. Chatterjee, S. Nag, R. Roy, B. Tudu, and R. Bandyopadhyay, "Linear regression modelling on epigallocatechin-3-gallate sensor data for green tea," in *Proceedings of the 4th International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, Nov. 2018, pp. 112–117.

[151] S. Das, A. Paramane, U. Rao, and P. Rozga, "A hybrid regression model to estimate remaining useful life of transformer liquid," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 31, no. 2, pp. 1062–1069, Oct. 2023.

[152] I. Ghosh, M. Sanyal, R. Jana, and P. Dan, "Machine learning for predictive modeling in management of operations of EDM equipment product," in *Proceedings of the 2nd International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, Sep. 2016, pp. 169–174.

[153] X. Liu, W. Zhao, P. Li, S. Jiao, and Q. Liu, "Scalable poisoning against regression-type edge computing applications via an approximate optimization strategy," in *Proceedings of the IEEE Conference on Computer Communications Workshops*, Apr. 2019, pp. 1–6.

[154] F. Chollet, *Deep Learning with Python*. Shelter Island, New York, NY, USA: Manning Publications Co, 2018.

[155] A. Syed, D. Sierra-Sosa, A. Kumar, and A. Elmaghraby, "IoT in smart cities: A survey of technologies, practices and challenges," *Smart Cities*, vol. 4, no. 2, pp. 429–475, Mar. 2021.

[156] H. Zhang, Z. Wang, W. Xia, Y. Ni, and H. Zhao, "Weighted adaptive KNN algorithm with historical information fusion for fingerprint positioning," *IEEE Wireless Communications Letters*, vol. 11, no. 5, pp. 1002–1006, Feb. 2022.

[157] G. Anagnostopoulos and A. Kalousis, "A reproducible comparison of RSSI fingerprinting localization methods using LoRaWAN," in *Proceedings of the 16th Workshop on Positioning, Navigation and Communications (WPNC)*, Oct. 2019, pp. 1–6.

[158] Z. Pandangan and M. Talampas, "Hybrid LoRaWAN localization using ensemble learning," in *Proceedings of the Global Internet of Things Summit (GIoTS)*, Jun. 2020, pp. 1–6.

[159] G. Ferreras and M. Talampas, "LoRa-based differential fingerprint localization in outdoor environments," in *Proceedings of the 7th IEEE World Forum on Internet of Things (WF-IoT)*, Jun. 2021, pp. 710–715.

[160] Y. Li, J. Barthelemy, S. Sun, P. Perez, and B. Moran, "Urban vehicle localization in public LoRaWAN network," *IEEE Internet Things Journal*, vol. 9, no. 12, pp. 10 283–10 294, Oct. 2022.

[161] L. Marquez and M. Calle, "Machine learning for localization in LoRaWAN: A case study with data augmentation," in *Proceedings of the IEEE Future Networks World Forum (FNWF)*, Oct. 2022, pp. 422–427.

[162] A. Pimpinella, A. Redondi, M. Nikoli, and M. Cesana, "Machine learning-based localization of LoRaWAN devices via inter-technology knowledge transfer," in *Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops)*, Jun. 2020, pp. 1–6.

[163] I. Aqeel, M. Iorkyase, H. Zangoti, C. Tachatatzis, R. Atkinson, and I. Aondonovic, "LoRaWAN-implemented node localisation based on received signal strength indicator," *IET Wireless Sensor Systems*, pp. 1–16, Aug. 2022.

[164] J. Teuwen and N. Moriakov, "Convolutional neural networks," in *Handbook of Medical Image Computing and Computer Assisted Intervention*, S. Zhou, D. Rueckert, and G. Fichtinger, Eds. San Diego, California, USA: Academic Press (Elsevier), 2020, pp. 481–501.

[165] Y. Ghadi *et al.*, "Machine learning solutions for the security of wireless sensor networks: A review," *IEEE Access*, vol. 12, pp. 12 699–12 719, Jan. 2024.

[166] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti, and D. De, "Fundamental concepts of convolutional neural network," in *Recent Trends and Advances in Artificial Intelligence and Internet of Things*, V. Balas, R. Kumar, and R. Srivastava, Eds. Cham, Switzerland: Springer, 2020, pp. 519–567.

[167] P. Shruti and R. Rekha, "A review of convolutional neural networks, its variants and applications," in *Proceedings of the International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS)*, Feb. 2023, pp. 31–36.

[168] N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," in *Proceedings of the International Conference on Communication and Signal Processing (ICCSP)*, Apr. 2017, pp. 588–592.

[169] S. Indolia, A. Goswami, S. Mishra, and P. Asopa, "Conceptual understanding of convolutional neural network: A deep learning approach," in *Proceedings of the Procedia Computer Science*, Jan. 2018, pp. 679–688.

[170] V. Jadeja, A. Rao, A. Srivastava, S. Singh, P. Chaturvedi, and G. Bhardwaj, "Convolutional neural networks: A comprehensive review of architectures and application," in *Proceedings of the 6th International Conference on Contemporary Computing and Informatics*, Sep. 2023, pp. 460–467.

[171] S. Joshi, M. Manu, and A. Mittal, "A review of the evolution and applications of convolutional neural network (CNN)," in *Proceedings of the 2nd International Conference on Edge Computing and Applications (ICECAA)*, Jul. 2023, pp. 1109–1114.

[172] D. Dai, "An introduction of CNN: Models and training on neural network models," in *Proceedings of the International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR)*, Nov. 2021, pp. 135–138.

[173] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2018, pp. 7132–7141.

[174] S. Mekruksavanich, P. Jantawong, and A. Jitpattanakul, "A lightweight deep convolutional neural network with squeeze-and-excitation modules for efficient human activity recognition using smartphone sensors," in *Proceedings of the 2nd International Conference on Big Data Analytics and Practices (IBDAP)*, Aug. 2021, pp. 23–27.

[175] S. Islam *et al.*, "A comprehensive survey on applications of transformers for deep learning tasks," *Expert Systems with Applications*, vol. 241, Nov. 2024, Art. no. 122666.

[176] S. Nerella *et al.*, "Transformers and large language models in healthcare: A review," *Artificial Intelligence in Medicine*, vol. 154, Sep. 2024, Art. no. 102900.

[177] A. Vaswani *et al.*, "Attention is all you need," in *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS)*, Dec. 2017, pp. 1–11.

[178] S. Bengesi, H. El-Sayed, M. Sarker, Y. Houkpati, J. Irungu, and T. Oladunni, "Advancements in generative AI: A comprehensive review of GANs, GPT, autoencoders, diffusion model, and transformers," *IEEE Access*, vol. 12, pp. 69 812–69 837, May 2024.

[179] K. Chitty-Venkata, S. Mittal, M. Emani, V. Vishwanath, and A. Somani, "A survey of techniques for optimizing transformer inference," *Journal of Systems Architecture*, vol. 144, Sep. 2023, Art. no. 102990.

[180] S. Casola, I. Lauriola, and A. Lavelli, "Pre-trained transformers: An empirical comparison," *Machine Learning with Applications*, vol. 9, Sep. 2022, Art. no. 100334.

[181] J. Li, J. Chen, Y. Tang, C. Wang, B. Landman, and S. Zhou, "Transforming medical imaging with transformers? A comparative review of key properties, current progresses, and future perspectives," *Medical Image Analysis*, vol. 85, Apr. 2023, Art. no. 102762.

[182] F. Shamshad *et al.*, "Transformers in medical imaging: A survey," *Medical Image Analysis*, vol. 88, Aug. 2023, Art. no. 102802.

[183] A. Lutakamale, H. Myburgh, and A. De Freitas, "A hybrid convolutional neural network-transformer method for received signal strength indicator fingerprinting localization in long range wide area network," *Engineering Applications of Artificial Intelligence*, vol. 133, Jul. 2024, Art. no. 108349.

[184] B. Soro and C. Lee, "Joint time-frequency RSSI features for convolutional neural network-based indoor fingerprinting localization," *IEEE Access*, vol. 7, pp. 104 892–104 899, Aug. 2019.

[185] J. Jang and S. Hong, "Indoor localization with WiFi fingerprinting using convolutional neural network," in *Proceedings of the 10th International Conference on Ubiquitous and Future Networks (ICUFN)*, Jul. 2018, pp. 753–758.

[186] G. Zhang, P. Wang, H. Chen, and L. Zhang, "Wireless indoor localization using convolutional neural network and Gaussian process regression," *Sensors*, vol. 19, no. 11, May 2019, Art. no. 2508.

[187] M. Stahlke, T. Feigl, M. Garcia, R. Stirling-Gallachery, J. Seitz, and C. Mutschler, "Transfer learning to adapt 5G AI-based fingerprint localization across environments," in *Proceedings of the IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, Jun. 2022, pp. 1–5.

[188] S. Yang, C. Sun, and Y. Kim, "Indoor 3D localization scheme based on BLE signal fingerprinting and 1D convolutional neural network," *Electronics*, vol. 10, no. 15, Jul. 2021, Art. no. 1758.

[189] X. Song *et al.*, "A novel convolutional neural network based indoor localization framework with WiFi fingerprinting," *IEEE Access*, vol. 7, pp. 110 698–110 709, Aug. 2019.

[190] D. Sun, E. Wei, and S. Xu, "Improving fingerprint indoor localization using convolutional neural networks," *IEEE Access*, vol. 8, pp. 193 396–193 411, Oct. 2020.

[191] Z. Liu, B. Dai, X. Wan, and X. Li, "Hybrid wireless fingerprint indoor localization method based on a convolutional neural network," *Sensors*, vol. 19, no. 20, Oct. 2019, Art. no. 4597.

[192] W. Njima, I. Ahriz, R. Zayani, M. Terre, and R. Bouallegue, "Deep CNN for indoor localization in IoT-sensor systems," *Sensors*, vol. 19, no. 14, Jul. 2019, Art. no. 3127.

[193] R. Sinha and S. Hwang, "Comparison of CNN applications for RSSI-based fingerprint indoor localization," *Electronics*, vol. 8, no. 9, Sep. 2019, Art. no. 989.

[194] F. Qin, T. Zuo, and X. Wang, "CCpos: WiFi fingerprint indoor positioning system based on CDAE-CNN," *Sensors*, vol. 21, no. 4, Feb. 2021, Art. no. 1114.

[195] S. De-Bast, A. Guevara, and S. Pollin, "CSI-based positioning in massive MIMO systems using convolutional neural networks," in *Proceedings of the IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, May 2020, pp. 1–5.

[196] L. Yang *et al.*, "FusionNet: A convolution–transformer fusion network for hyperspectral image classification," *Remote Sensing*, vol. 14, no. 16, Aug. 2022, Art. no. 4066.

[197] X. Li, B. Qiu, G. Cao, C. Wu, and L. Zhang, "A novel method for ground-based cloud image classification using transformer," *Remote Sensing*, vol. 14, no. 16, Aug. 2022, Art. no. 3978.

[198] R. Shao, X. Bi, and Z. Chen, "A novel hybrid transformer-CNN architecture for environmental microorganism classification," *PLoS ONE*, vol. 17, no. 11, pp. 1–22, Nov. 2022.

[199] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 1–9.

[200] M. Ziemann and C. Metzler, "Transformers for robust radar waveform classification," in *Proceedings of the 56th Asilomar Conference on Signals, Systems, and Computers*, Oct. 2022, pp.

759–763.

[201] Y. Vindas, B. Guepie, M. Almar, E. Roux, and P. Delachartre, "An hybrid CNN-transformer model based on multi-feature extraction and attention fusion mechanism for cerebral emboli classification," in *Proceedings of the 7th Machine Learning for Healthcare Conference*, Aug. 2022, pp. 270–296.

[202] T. Lin, X. Liu, and X. Qiu, "A survey of transformers," *AI Open*, vol. 3, pp. 111–132, Jan. 2022.

[203] Z. Wu, Z. Liu, J. Lin, Y. Lin, and S. Han, "Lite transformer with long-short range attention," 2020, *arXiv:2004.11886v1*.

[204] T. Janssen, M. Aernouts, R. Berkvens, and M. Weyn, "Outdoor fingerprinting localization using Sigfox," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sep. 2018, pp. 1–6.

[205] G. Anagnostopoulos and A. Kalousis, "A reproducible analysis of RSSI fingerprinting for outdoor localization using Sigfox: Preprocessing and hyperparameter tuning," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sep. 2019, pp. 1–8.

[206] J. Torres-Sospedra, R. Montoliu, S. Trilles, O. Belmonte, and J. Huerta, "Comprehensive analysis of distance and similarity measures for WiFi fingerprinting indoor positioning systems," *Expert Systems with Applications*, vol. 42, no. 23, pp. 9263–9278, Dec. 2015.

[207] T. Monawar, S. Mahmud, and A. Hira, "Anti-theft vehicle tracking and regaining system with automatic police notifying using Haversine formula," in *Proceedings of the 4th International Conference on Advances in Electrical Engineering (ICAEE)*, Sep. 2017, pp. 775–779.

[208] T. Janssen, "Energy-efficient positioning for the Internet of Things," Ph.D. dissertation, Faculty of Applied Engineering, University of Antwerp, Belgium, 2023.