Engineering Applications of Artificial Intelligence 133 (2024) 108349

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Research paper

A hybrid convolutional neural network-transformer method for received signal strength indicator fingerprinting localization in Long Range Wide Area Network

Albert Selebea Lutakamale *,[1], Herman C. Myburgh[2], Allan de Freitas[3]

University of Pretoria, Department of Electrical, Electronic and Computer Engineering, Private Bag x20, Hatfield, 0028, Pretoria, South Africa

ARTICLE INFO

Keywords:
Fingerprint localization
Deep learning
Convolutional neural networks
Internet of Things
Wireless communications

ABSTRACT

In recent years, low-power wide area networks (LPWANs), particularly Long-Range Wide Area Network (LoRaWAN) technology, are increasingly being adopted into large-scale Internet of Things (IoT) applications thanks to having the ability to offer cost-effective long-range wireless communication at low-power. The need to provide location-stamped communications to IoT applications for meaningful interpretation of physical measurements from IoT devices has increased demand to incorporate location estimation capabilities into LoRaWAN networks. Fingerprint-based localization methods are increasingly becoming popular in LoRaWAN networks because of their relatively high accuracy compared to range-based localization methods. This work proposes hybrid convolutional neural networks (CNNs)-transformer fingerprinting method to localize a node in a LoRaWAN network. CNNs are adopted to complement the strengths of the Transformer by adding the ability to capture local features from input data and consequently allow the Transformer, through the attention mechanism, to effectively learn global dependencies from the input data. Specifically, the proposed method works by first learning the local location features from the input data using the CNNs and passing the resulting information to the transformer encoder to learn global features from the input data. The output of the transformer encoder is then concatenated with information learned at the local level and then passed through the regressor for the final location estimation. With a localization performance of 290.71 m mean error achieved, the proposed method outperformed similar state-of-the-art works in the literature evaluated on the same publicly available LoRaWAN dataset.

1. Introduction

Localization in wireless networks refers to the process undertaken to estimate the location of a target node/object deployed in indoor or outdoor environments. This process usually involves the exchange of positional signal parameters between the anchor node (nodes whose physical location is known) and other nodes to establish the location of a desired node (Kumari et al., 2019; Bhatti, 2018; Alomari et al., 2018).

In the context of outdoor environments, Global Positioning System (GPS) has been the most widely adopted technology for location estimation purposes for many years (Obeidat et al., 2021). For this technology to provide highly accurate localization performance, an object whose location is to be estimated must establish a clear line-of-sight (LOS) with GPS satellites. However, the high implementation cost and power-hungry nature of GPS-based localization systems make them less attractive in large-scale IoT applications where the emphasis is on energy efficiency and cost-effectiveness (C.E. et al., 2018; Singh and Sharma, 2018). Another key reason for the increased interest in adopting emerging wireless technologies other than GPS in developing localization systems is the poor performance of GPS-based localization systems in urban canyons or environments with many obstructions, such as tall trees.

Short-range wireless communications technologies such as Bluetooth, WiFi and ZigBee are common in IoT applications; however, applying them in large-scale IoT applications is not economically feasible due to their short-range nature. To meet the requirements for low-cost, low-power and long-range communications in large-scale IoT applications, researchers have shifted their focus to long-range LPWAN

* Corresponding author.
E-mail addresses: albert.lutakamale@tuks.co.za (A.S. Lutakamale), herman.myburgh@up.ac.za (H.C. Myburgh), allan.defreitas@up.ac.za (A. de Freitas).
[1] Albert Selebea Lutakamale is with University of Dodoma currently working towards a Ph.D. degree in Electronic Engineering at the University of Pretoria.
[2] Herman C. Myburgh is with University of Pretoria, South Africa.
[3] Allan de Freitas is with University of Pretoria, South Africa.

https://doi.org/10.1016/j.engappai.2024.108349
Received 1 November 2023; Received in revised form 25 February 2024; Accepted 26 March 2024
Available online 8 April 2024

technologies (Anjum et al., 2020; Miles et al., 2020; Lalle et al., 2019). LoRaWAN, along with Sigfox, are the two most adopted LPWAN technologies in large-scale IoT applications thanks to their ability to offer cost-effective long-range communications using batteries that can last for many years (Queralta et al., 2019; Mkkhaylov et al., 2020; Ikpehai et al., 2019).

In recent years, IoT-related technologies have proven to be key enabling technologies to provide smart solutions to nearly all spheres of human life, such as in industries, healthcare facilities, agricultural-related projects, and logistic applications, to name a few. Remote control and telemetry, asset tracking, machine control, safety monitoring, and implementation of smart projects in homes and cities are notable examples of use cases of IoT-related technologies (Lalle et al., 2019; Cho et al., 2019; Akhmedov et al., 2021; Perez et al., 2022; Durand et al., 2019). Incorporating location estimation capabilities into these large-scale IoT applications will enable the provision of location-stamped communication, which is crucial in extracting useful information from physical measurements obtained from IoT devices.

In LPWAN networks, particularly LoRaWAN networks, localization of a target node is achieved through either fingerprinting or range-based approaches (Janssen et al., 2020a; Marquez and Calle, 2023; Janssen et al., 2020b; Islam et al., 2023; Anagnostopoulos and Kalousis, 2021). Range-based approaches involve the application of a path loss model to provide distance estimation of a device to a nearby gateway from received signal strength indicator measurements (RSSI), which are then used to pinpoint its probable location (Janssen et al., 2020b). Fingerprinting-based methods, on the other hand, estimate the location of a target node by using a database of earlier collected signal features (fingerprints) through feature matching (Janssen et al., 2020b; Islam et al., 2023). Compared to range-based localization approaches, fingerprinting-based approaches are increasingly being applied to IoT applications because of their better localization performance relative to their range-based counterparts (Janssen et al., 2020a; Aernouts et al., 2018a). The effectiveness of fingerprinting-based methods is attributed to the utilization of machine learning models, which are able to learn useful positional information even from noisy data collected in non-LOS (NLOS) environmental settings (Purohit et al., 2020).

In the literature, the majority of fingerprint-based localization methods in LoRaWAN networks are implemented using standalone or a hybrid of classical 'shallow' machine learning algorithms, notably k-nearest neighbours (kNN), Random Forests (RFs), Decision Trees (DTs) and Support Vector Machines (SVMs). When the training data is relatively large, these classical machine learning algorithms become computationally expensive to train, and their performances tend to degrade. Therefore, these algorithms are not ideal for fingerprinting-based localization using large outdoor fingerprint databases, which are necessary if a large outdoor environment is to be covered (Janssen et al., 2020a).

In the recent past, convolutional neural networks (CNNs) have performed extremely well in computer vision tasks due to their effectiveness in learning local dependencies from the input data (Chollet, 2018). However, CNNs are poor in establishing global dependencies between the input data because of the locality of the convolution operation (Li et al., 2022b). On the other hand, transformers are becoming popular in computer vision and natural language processing tasks thanks largely to the ability to capture global dependencies from the input data; however, they are less equipped to capture local dependencies from the input data (Yang et al., 2022; Li et al., 2022b; Shao et al., 2022).

The novel contribution of this work is the development of a hybrid CNN-transformer fingerprinting-based localization method in LoRaWAN networks 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. The optimal preprocessing techniques for the LoRaWAN dataset for improved localization performance are presented. With a localization performance of 290.71 m mean error obtained when evaluated on a publicly available LoRaWAN dataset (Aernouts et al., 2018b), the proposed method outperformed similar state-of-the-art methods in the literature evaluated on the same dataset.

The rest of the paper is structured as follows: Sections 2 and 3 provide an overview of related works and preliminary concepts, respectively. The proposed method is detailed in Section 4, while Section 5 is devoted to the description of experimental settings and procedures. In Section 6, the experimental results are presented and discussed. Finally, conclusions are drawn and future works discussed in Section 7.

2. Related works

In LoRaWAN networks, by analysing RSSI, AoA, ToA or TDoA or a combination of these parameters received by LoRaWAN gateways, the location of a target node can be estimated by using fingerprinting-based or range-based localization approaches (Janssen et al., 2020a). 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 Vazquez-Rodas et al. (2020), the authors proposed a localization scheme in LoRaWAN networks 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 using a trilateration algorithm. In Muppala et al. (2021), a localization method is proposed whereby the location of a target node in the LoRaWAN network is computed using TDoA values measured from the signal 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 using an additional stationary node. The authors in Guo et al. (2022) proposed a localization approach that utilizes TDoA measurements estimated from differential phase sampling 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 Chen et al. (2023), an approach to using 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 Liu et al. (2022), 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 first increased through the synchronization of multiple communication channels before adopting an ESPRIT algorithm to compute the location of a LoRaWAN transmitter. The authors in Aernouts et al. (2020) 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 one for AoA measurements, are first built 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.

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 caused by shadowing and fading phenomena due to multipath propagation (Goldoni et al., 2019). Researchers are increasingly being attracted to adopting fingerprinting-based localization approaches due to their robustness in challenging environments with multipath and NLOS phenomena (Zhang et al., 2022), and being relatively more accurate. This is attributed to their ability to learn useful positional information even from noisy data (Purohit et al., 2020). The authors in Aernouts et al. (2018b) 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 fingerprinting-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 fingerprinting-based localization methods, is presented in Anagnostopoulos and Kalousis (2019b) and Purohit et al. (2020). In Anagnostopoulos and Kalousis (2019b), the authors implemented kNN, Extra Trees and Multilayer Perceptron (MLP) fingerprinting methods, reporting localization accuracy of 357 m mean errors with the best-performing MLP fingerprinting model. The researchers in Purohit et al. (2020) implemented three fingerprinting localization methods, reporting localization performance of 191.53 m with the long short-term memory method, which outperformed the other two methods based on the artificial neural network and CNN. In this work, the architectural structure of the CNN-based localization method comprises two CNN layers and two dense layers. In contrast, the proposed method is a hybrid CNN-Transformer-based fingerprinting approach, incorporating CNNs to enhance the transformer's capacity to effectively capture local features from the input data.

Fingerprinting-based localization methods reported in (Janssen et al., 2020a; Pandangan and Talampas, 2020; Ferreras and Talampas, 2021; Li et al., 2022a) were evaluated using version 1.2 of the urban LoRaWAN dataset presented in Aernouts et al. (2018b). Authors in Janssen et al. (2020a) 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 Pandangan and Talampas (2020), a kNN-RF method was implemented utilizing hybrid data and achieved a localization accuracy of 332.63 m mean errors. The researchers in Ferreras and Talampas (2021) implemented and trained RF and MLP fingerprinting-based localization methods 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 Li et al. (2022a), a hierarchical clustering-based technique is 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 Marquez and Calle (2022) present a case study using data augmentation techniques to improve the localization performance of Support Vector Regression (SVR), kNN, Extra Trees, and MLP fingerprinting algorithms trained on small datasets. The best localization accuracy of 12 m mean error was recorded using the kNN algorithm at an outdoor urban area covering 8 km^2. The researchers in Pimpinella et al. (2020) performed a comprehensive evaluation of different strategies that can 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 Aqeel et al. (2023) 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 Svertokat et al. (2022), through the implementation of the kNN fingerprinting localization algorithm, the authors analysed factors that could influence the accuracy of fingerprinting-based localization approaches in an outdoor setting. This work analysed accuracy dependencies based on the number of deployed gateways, coverage area and the distance from one measurement point to another. The authors in Anjum et al. (2022) 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 1 presents a summary of key features of the reported localization approaches in the related works.

The motivation to undertake this study is driven by the strong demand to incorporate location estimation capabilities into LoRaWAN networks, which are increasingly being adopted into large-scale IoT applications thanks to having the ability to offer cost-effective long-range wireless communication at low power. Factors such as high power consumption, high implementation costs, and poor localization performance in urban canyons or environments with many obstructions make 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. The limitation of another category of localization solutions based on range-based approaches, such as the need to install dedicated hardware like expensive antenna arrays and the requirement for accurate clock synchronization among anchor nodes, is another reason to opt for a fingerprinting-based localization method. In contrast to the prevailing fingerprinting-based localization methods for LoRaWAN networks, which predominantly rely on conventional 'shallow' machine learning models, this study proposes a hybrid fingerprinting-based localization method designed to accurately localize nodes within LoRaWAN networks with a CNN-transformer architecture. This deep-learning approach is introduced to overcome the limitations of shallow-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 performance (Purohit et al., 2020).

3. Preliminaries

3.1. LoRaWAN technology

LoRaWAN technology, which operates on top of the LoRa physical layer (PHY), is a medium access control protocol (MAC) proposed by Semtech and maintained by the LoRa Alliance. A Chirp Spread Spectrum (CSS) modulation scheme is used by LoRa PHY on which LoRaWAN resides. This proprietary modulation technique enables long-range communications between 2 to 5 km and up to 15 km in urban and rural areas, respectively (Chen et al., 2022). The CSS technique modulates signals through frequency-varying chirp pulses with the ability to counter the effects of interference, multipath, and Doppler shifts (Perez et al., 2022; Zafari et al., 2019). In LoRaWAN, channel bandwidth along with the spreading factor (SF) parameter, with values ranging from 7 to 12, are used to adjust the modulated data rate of the transmitted signal. The data rate varies from 300 bps to 50 kbps (Sassi and Fourati, 2022). In North America, LoRaWAN is specified to operate at 915 MHz; in Europe, it operates at 868 MHz, while in Asia, it operates at 433 MHz sub-GHz unlicensed industrial, scientific, and medical (ISM) bands (Perez et al., 2022; Chen et al., 2022; Stusek et al., 2020).

The network topology adopted in LoRaWAN is a star topology where end devices establish single-hop connections with gateways (Chen et al., 2022). A fully LoRaWAN architecture consists of one or more gateways, LoRaWAN servers, application servers and end devices, as illustrated in Fig. 1. The long-range, low-energy and low-cost communications features of LoRaWAN make it an ideal technology for large-scale IoT applications where the emphasis is on energy efficiency and cost-effectiveness.

Table 1
A Summary of the key features of the reported localization approaches in the related works.

Research work	Localization approach	Localization parameter	Nature of the dataset	Localization environment	Adopted algorithm(s)
Vazquez-Rodas et al. (2020)	Range-based	RSSI	Real	Outdoor	Trilateration
Muppala et al. (2021)	Range-based	TDoA	Simulated	Outdoor	Multilateration
Guo et al. (2022)	Range-based	TDoA	Real	Indoor	Least Square
Chen et al. (2023)	Range-based	TDoA	Simulated	Indoor/outdoor	Chan
Liu et al. (2022)	Range-based	AoA	Real	Indoor/outdoor	ESPRIT
Aernouts et al. (2020)	Range-based	TDoA and AoA	Real	Outdoor	Triangulation/Trilateration
Aernouts et al. (2018b)	Fingerprinting	RSSI	Real	Outdoor	kNN
Anagnostopoulos and Kalousis (2019b)	Fingerprinting	RSSI	Real	Outdoor	kNN, MLP, Extra Trees
Purohit et al. (2020)	Fingerprinting	RSSI	Real	Indoor/outdoor	LSTM, ANN, CNN
Janssen et al. (2020a)	Fingerprinting/ Range-based	RSSI	Real	Outdoor	Ten different regression algorithms plus extended Min-Max algorithm
Pandangan and Talampas (2020)	Fingerprinting	Fused RSSI-TDoA	Real	Outdoor	kNN-RF
Ferreras and Talampas (2021)	Fingerprinting	Fusion of differential RSSI-TDoA	Real	Outdoor	MLP, RF
Li et al. (2022a)	Fingerprinting	RSSI	Real	Outdoor	k-means + weighted kernel regression
Marquez and Calle (2022)	Fingerprinting	RSSI	Real	Outdoor	Support Vector Regression, Extra Trees, kNN, MLP
Pimpinella et al. (2020)	Fingerprinting	RSSI	Real	Outdoor	Its implementation is based on inter-technology knowledge transfer using classical machine learning algorithms
Aqeel et al. (2023)	Fingerprinting	RSSI	Real	Outdoor	Support Vector Regression and Gaussian Process Regression
Svertokat et al. (2022)	Fingerprinting	RSSI	Real	Outdoor	kNN
Anjum et al. (2022)	Fingerprinting/ Range-based	RSSI	Real	Outdoor	Trilateration, DTs, kNN, SVM

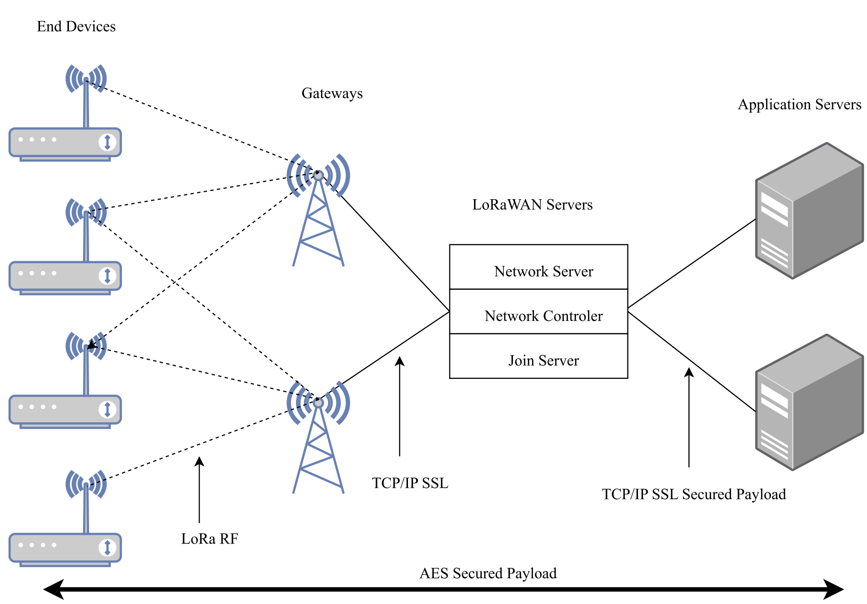


Fig. 1. LoRaWAN network architecture.
Source: Redrawn from Chen et al. (2022).

3.2. The transformer model

Transformer is a prominent natural language processing (NLP) model first proposed in Vaswani et al. (2017) to perform sequence-to-sequence modelling for machine translation tasks. Since their inception, along with other NLP tasks such as classification and language modelling, transformers have also achieved great success in computer vision and audio processing tasks (Lin et al., 2022). Due to their versatility to fit into different machine learning tasks as long as the input data is formatted accordingly, researchers are increasingly adopting transformers to build high-performing machine learning models.

The original transformer architecture, famously known as the Vanilla transformer, as illustrated in Fig. 2, is made up of the encoder and decoder blocks. Each encoder block is built using multi-head self-attention and feed-forward network (FFN) modules. Implementing a residual connection just before layer normalization at the output of

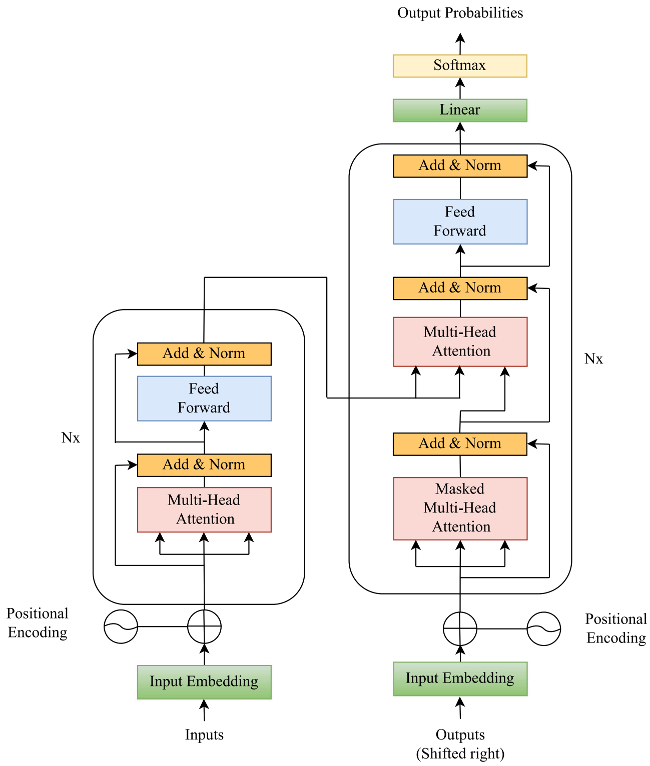


Fig. 2. Model architecture of the Vanilla Transformer.
Source: Redrawn from Vaswani et al. (2017).

the attention and FFN modules helps prevent vanishing gradient phenomena when building very deep model architectures. The transformer decoder, on the other hand, in addition to all the modules contained in the transformer encoder, has a cross-attention module to allow encoder and decoder features to influence each other.

Key to the performance of transformers is the attention mechanism, which is a graph-like inductive bias which employs a pooling operation to relevantly connect each word in a sequence (Tay et al., 2022). The self-attention mechanism allows word tokens in the same sequence to modify each other's representations. On the other hand, the cross-attention mechanism allows word tokens in the encoder and decoder to influence each other's representations.

As reported in Vaswani et al. (2017) and Lin et al. (2022), the scaled dot-product attention mechanism, which is given by

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{D_k}}\right)V, \tag{1}$$

where matrices $Q \in \mathbb{R}^{N \times D_k}$, $K \in \mathbb{R}^{M \times D_k}$, and $V \in \mathbb{R}^{M \times D_v}$, is implemented using a function which maps three vector matrices to the output, namely query (Q), key (K), and value (V) vector matrices. N is the length of queries, M is the length of keys and values, D_v is the dimension of values and D_k is the dimension of queries and keys. $\sqrt{D_k}$ plays the role of tackling the gradient vanishing problem.

Depending on the machine learning task to be attended, transformer models usually apply a multi-head attention mechanism with m-dimensional Q, K and V, given by

$$MultiHeadAttention(Q, K, V) = Concatenation(Head_1, \dots, Head_H)W^0, \tag{2}$$

where, $Head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)$. $W_i^Q \in \mathbb{R}^{m \times D_k}$, $W_i^K \in \mathbb{R}^{m \times D_k}$, $W_i^V \in \mathbb{R}^{m \times D_v}$, and $W^0 \in \mathbb{R}^{HD_v \times m}$ with H being the number of projections. Eq. (1) is used to compute the output of each of the projected Q, K, and V, followed by concatenation of all of the outputs into the original m-dimensional representation.

4. Proposed method

Inspired by Ziemann and Metzler (2022) and Vindas et al. (2022), this work proposes a hybrid model using CNNs and transformer modules for fingerprinting-based localization 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. The original transformer architecture, commonly referred to as the Vanilla transformer (Vaswani et al., 2017), features encoder and decoder structures in order to process sequences from two different types of data (two language types, to be precise) to achieve a sequence-to-sequence machine translation (Lin et al., 2022). Instead of using the complete structure of the Vanilla transformer, this work 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 further processed 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. The transformer encoder's positional-wise feed-forward neural network block is replaced by a stack of three 1D-CNN layers to capture local context within the encoder module (Wu et al., 2020). For all the 1D-CNN layers, the number of filters used is eight, with a kernel size of one. The embedding and dense dimensions of the transformer encoder are both set to eight. Fig. 3 illustrates the proposed model architecture. The input data is processed first by the 1D-CNN layers to learn local dependencies before being fed into the transformer encoder for the purpose of learning 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 fully connected (FC) layers, with 512, 256, 128 and 2 hidden units, all activated by the ReLu activation function except for the last 2-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. The mean location estimation error is adopted as the metric to evaluate the performance of the proposed method. Model checkpoints and early stopping callbacks are also introduced to better optimize the training duration of the localization model.

5. Experimental settings and procedures

The LoRaWAN dataset used to validate the performance of the proposed fingerprinting-based localization method is version 1.2 of the publicly available urban LoRaWAN dataset reported in Aernouts et al. (2018b), 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, spreading factor (SF), receiving time, horizontal dilution of precision (HDOP), latitude and longitude). Initially in this work, 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 for them to effectively be processed by the proposed method, which could significantly increase model complexity. During the construction

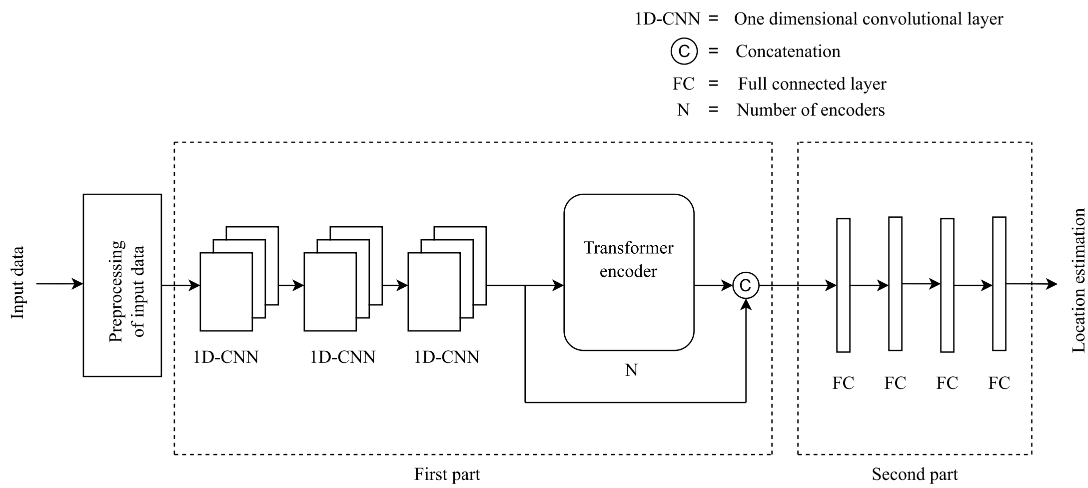


Fig. 3. The proposed model architecture.

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 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 in this work 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 received RSSI value from the dataset ($RSSI_{min}$), followed by replacing the out-of-reach RSSI values with '$\tau = RSSI_{min} - 1$' (Anagnostopoulos and Kalousis, 2019b). The last preprocessing step involves transforming the resulting training dataset into optimal representations using any of the four commonly adopted data representation techniques, namely Positive, Normalized, Powed and Exponential data representations (Janssen et al., 2018; Anagnostopoulos and Kalousis, 2019a; Torres-Sospedra et al., 2015), given by

$$Positive_i(x) = RSSI_i - \tau, \tag{3}$$

$$Normalized_i(x) = \frac{Positive_i(x)}{-\tau}, \tag{4}$$

$$Exponential_i(x) = \frac{e^{\frac{Positive_i(x)}{\alpha}}}{e^{\frac{-\tau}{\alpha}}}, \tag{5}$$

$$Powed_i(x) = \left(\frac{Positive_i(x)}{-\tau}\right)^{\beta}, \tag{6}$$

where i and $RSSI_i$ stand for 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 Torres-Sospedra et al. (2015), they were originally set at 24 and e, respectively, with e being a mathematical constant for RSSI values collected indoors using WiFi signals. The α and β parameters adopted in this work are 60 and 1.1, respectively, re-adjusted in Anagnostopoulos and Kalousis (2019b) for the outdoor RSSI values in the LoRaWAN network.

This work adopted the Powed data representation scheme to transform the training data to an optimal form. This data representation scheme is adopted because of its non-linearity 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 work (Janssen et al., 2020a; Anagnostopoulos and Kalousis, 2019b; Ferreras and Talampas, 2021). 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 prediction stage of the model. The Haversine formula (Monawar et al., 2017) is defined as

$$Hav\left(\frac{y}{x}\right) = Hav(\gamma_2 - \gamma_1) + cos(\gamma_1)cos(\gamma_2)Hav(\vartheta_2 - \vartheta_1), \tag{7}$$

where 'Hav' represents the Haversine function, given by

$$Hav(C) = sin^2\left(\frac{C}{2}\right) = \frac{1 - cos(C)}{2}, \tag{8}$$

y and x stand for the distance between two coordinates and the sphere's radius, respectively, γ_1 and γ_2 stand for the latitudes of coordinates 1 and 2, respectively, and ϑ_1 and ϑ_2 stand for the longitudes of coordinates 1 and 2, respectively, all in radians.

This work used Keras and Scikit-Learn Python libraries and TensorFlow backend to implement the machine learning models. Additionally, Google Colaboratory 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. Rescaling the training data is recommended to prevent biases towards the features with large values during model training (Chollet, 2018).

6. Experimental results and discussion

The following experiments were carried out to evaluate the performance of the proposed fingerprinting-based localization method. 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 percent of the training samples, respectively. 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 may lead to performance biases. Each experiment is run for 120 epochs using 512-sized mini-batches of training samples.

Table 2
Performance of the proposed 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

Table 3
Performance of the proposed 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

6.1. Performance of the proposed method with multiple attention heads/ transformer encoders

In this section, experiments were 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 5). Two structural changes were made to the transformer encoder, namely the number of attention heads and the number of stacked encoders in the localization model. The number of attention heads was varied between one and four heads. In each instance, the localization model was trained using the LoRaWAN dataset, yielding localization results as shown in Table 2. 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.

To determine the impact of stacking more than one encoder on the proposed 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 3. 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 observed slight variations 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 give satisfactory 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.

6.2. Performance of the proposed method on different subsets of training, validation and test data

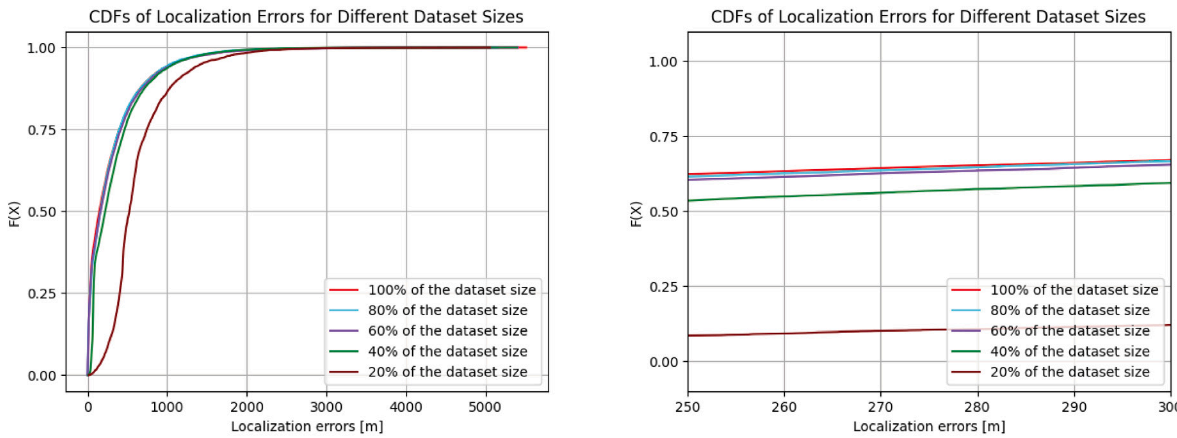
In order to explore how the proposed method performs when trained on a reduced sample size, smaller subsets of the data were first extracted from the dataset and split into training, validation, and test sets. This also leads to unique subsets of data in each scenario, which reduces the impact of spurious 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 4. Fig. 4 shows the full and enlarged cumulative distribution function (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 percent of the original dataset was used. The low accuracy in this instance stems from the limited number of training samples, which hindered the 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 percent and 60 percent 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 improved further to 294.14 m and 290.71 m mean localization errors when 80 percent and 100 percent of the original dataset were used, respectively, which further supports the findings 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 observed in each of the data split strategies adopted stems 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 low localization accuracies or enables the model to capture more sufficient patterns when the size of the training data increases which eventually improves the localization performance. 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 percent reduction in the sample size, further proving its effectiveness and robustness in localizing a node in LoRaWAN networks.

Table 4
Performance of the proposed 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

6.3. Performance of the proposed method on fixed test set

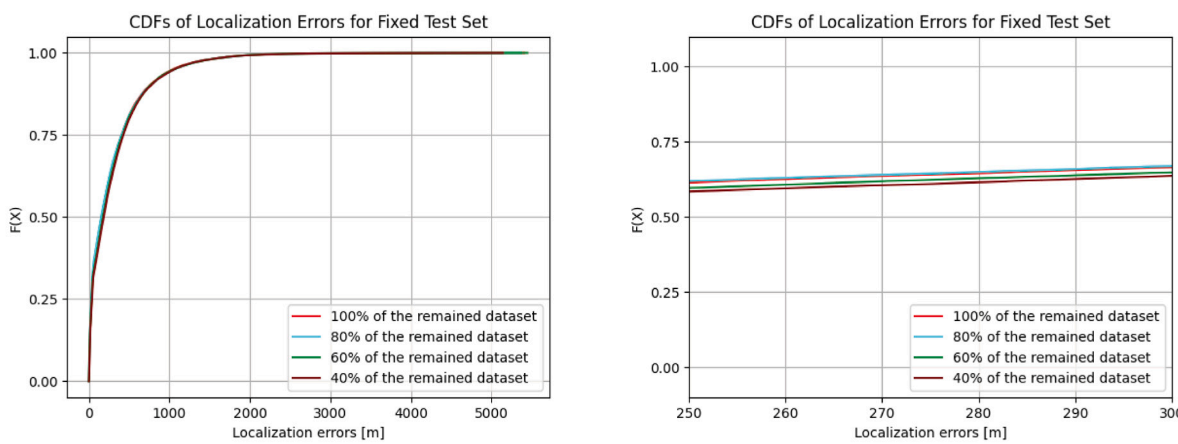
In this section, unlike in Section 6.2, a fixed 15 percent of the original LoRaWAN dataset was extracted and set aside to test the performance of the proposed localization method. Using a fixed test set ensures consistency in the evaluation process across the different variations of training/validation sets allowing for a fair and reliable performance comparison. For the remaining 85 percent of the LoRaWAN dataset, fractions of 40, 60, 80, and 100 percent were extracted, and each was split into a training set containing 80 percent of the remaining samples and a validation set containing 20 percent 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. Fig. 5 shows the full and enlarged cumulative distribution function curves of localization errors for all the extracted fractions of the LoRaWAN dataset. The lowest localization accuracy of 310.31 m mean localization error observed when 40 percent 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



(a) Full CDF curves

(b) Enlarged CDF curves

Fig. 4. CDFs of localization errors (m) for different sample sizes.



(a) Full CDF curves

(b) Enlarged CDF curves

Fig. 5. CDFs of localization errors (m) for different sample sizes with fixed test set.

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 percent of the remaining dataset was used, respectively, due to the increase in training samples. When the whole 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 percent of the remaining LoRaWAN dataset, further indicating the importance of training deep learning models 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
Performance of the proposed method on fixed test set using different subsets of training and validation sets of the remained dataset.

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

6.4. Performance comparison of the proposed method with methods proposed in the literature trained using the same dataset

In this section, the localization performance of the proposed method is compared with the localization performances of the fingerprinting localization methods reported in Janssen et al. (2020a), Pandangan and Talampas (2020), Ferreras and Talampas (2021), and Li et al. (2022a), which adopted the same LoRaWAN dataset. In Janssen et al. (2020a), 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 Pandangan and Talampas (2020), a kNN-RF ensemble method is proposed for fingerprinting localization, achieving localization accuracy of 332.63 m mean localization error. The authors in Ferreras and Talampas (2021) proposed a fingerprinting localization method based on MLP, achieving localization accuracy of 57 m and 310 m median and mean localization errors, respectively. In Li et al. (2022a), on the other hand, a fingerprinting localization method based on K-means and Weighted Kernel Regression is proposed,

Table 6
A Summary of the key experimental settings and parameters associated with the related works whose performances are compared with the proposed method.

Research work	Python libraries used	Experimental environment	Localization parameters	Ground truth references	Performance metrics
Janssen et al. (2020a)	Only Scikit-Learn is mentioned	Virtual machine with 32 GB RAM memory and 10 CPU cores	RSSI	Latitudes and Longitudes	Mean localization error, R^2 score and execution time
Pandangan and Talampas (2020)	Only Scikit-Learn is mentioned	Not mentioned	RSSI and TDoA	Latitudes and Longitudes	Mean localization error
Ferreras and Talampas (2021)	Only Scikit-Learn is mentioned	Not mentioned	Differential RSSI and TDoA	Latitudes and Longitudes	Median and Mean localization errors
Li et al. (2022a)	Not mentioned	Not mentioned	RSSI	Latitudes and Longitudes	Median and Mean localization errors

achieving localization accuracy of 158.48 m and 346 m median and mean localization errors, respectively. In all these 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 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 general unavailability of source code, missing key information about experimental settings and procedures, and variations in the choice of metadata used for training in related works. The comparison conducted in this section is limited to the final localization performance reported in the related works, which is justifiable given that the ground truth references (latitudes and longitudes) were used in all of the proposed methods. Table 6 summarizes the key experimental settings and parameters associated with the compared related works.

The localization results of the proposed method, which were compared with the localization results reported in the related works, were obtained by training the proposed method for 120 iterations using the full LoRaWAN dataset split into training, validation, and test sets according to a 0.7/0.15/0.15 ratio. With this setup, the proposed method resulted in 536,171 trainable parameters, taking 374.76 s to train, yielding mean and median localization errors of 290.71 m and 147.34 m, respectively. Fig. 6 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. Table 7 presents the performance comparison between the proposed method and the related works in terms of mean and median localization errors. The variations in the localization results reported between the proposed method and the related works are mainly due to the structural differences in the way the localization models were built and trained, the machine learning technique adopted as well as pre-processing techniques adopted, which brings different model learning capabilities which determine the final localization accuracies.

As indicated in Table 7, the proposed method outperforms all the related works in terms of mean localization error. The 147.34 m median localization error obtained using the proposed method is better than 158.48 m reported in Li et al. (2022a); however, it is inferior to 57 m reported in Ferreras and Talampas (2021). This difference in median errors obtained by the proposed method and the method proposed in Ferreras and Talampas (2021) could be due to variations in the number of outliers present in the localization errors computed by both methods. However, the closeness of the mean and median localization errors obtained by the proposed method shows that the degree of skewness in the distribution of localization errors in the proposed method is small in comparison to Ferreras and Talampas (2021).

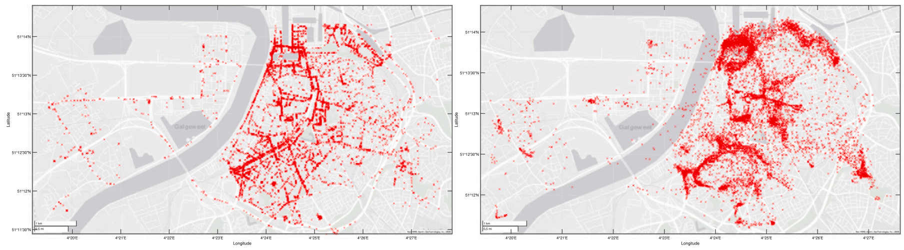
Based on this performance comparison, the proposed method has achieved a 6.22% increase in localization accuracy in terms of mean localization error compared to the currently available best-performing method in the literature evaluated using the same LoRaWAN dataset. Additionally, the proposed method's computational efficiency is underscored by a relatively small number of trainable parameters, 536,171, and a training duration of 374.76 s. Thereby affirming its suitability for deployment in real world localization applications.

7. Conclusion

This work proposed a hybrid CNN-transformer model to localize a node in LoRaWAN networks. Upon analysing the optimal structure of the proposed fingerprinting-based localization model, increasing the number of transformer encoders and attention heads did not significantly improve the localization accuracy of the proposed method for the adopted training dataset due to having a relatively small number of features. The performance of the proposed method was also analysed by subjecting it to different sample sizes of the training dataset with fixed and different test sets. The results showed much improvement in the localization accuracy as the dataset size increased. The proposed method with a single transformer encoder having one attention head and trained on the fully LoRaWAN dataset, achieved a performance of 290.71 m mean localization error on the test set, which is a 6.22% increase compared to the currently available state-of-the-art fingerprinting-based localization method in the literature. The performance of the proposed method proves its effectiveness in localizing a node in LoRaWAN networks with acceptable levels of localization accuracy. In the future, this study can be extended in several ways. Firstly, an analysis could be conducted to assess how the overall performance of the proposed fingerprinting localization method is influenced by including alternative localization parameters derived from the same publicly available LoRaWAN dataset, such as the fusion of differential RSSI-TDoA fingerprints. Secondly, given the data-driven nature of the proposed method, 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 method. Lastly, concerning the practical implementation of the proposed method in real-world IoT use cases, there is an opportunity to capitalize on the evolving capabilities of modern computing technologies like cloud, fog, and edge computing as reported in Gill et al. (2024). Leveraging these technologies for processing location-related data can help meet the quality-of-service requirements in real-time applications.

Funding

The authors received no direct funding for this research.



(a) True latitude and longitude coordinate pairs from the test set

(b) Estimated latitude and longitude coordinate pairs

Fig. 6. 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.

Table 7
Performance comparison of the proposed method with methods proposed in the literature trained on the same LoRaWAN dataset.

Research work	Scheme	Mean localization errors (m)	Median localization error (m)
Janssen et al. (2020a)	RF	340	Not reported
Pandangan and Talampas (2020)	kNN-RF	332.63	Not reported
Ferreras and Talampas (2021)	MLP	310	57
Li et al. (2022a)	K-means and Weighted Kernel Regression	346	158.48
Proposed Method	**CNN + Transformer**	**290.71**	**147.34**

CRediT authorship contribution statement

Albert Selebea Lutakamale: Writing – original draft, Software, Methodology, Formal analysis, Conceptualization, Investigation, Validation. **Herman C. Myburgh:** Writing – review & editing, Validation, Supervision, Resources, Investigation, Funding acquisition. **Allan de Freitas:** Writing – review & editing, Validation, Supervision, Resources, Investigation.

Declaration of competing interest

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

Data availability

The dataset used in this work is a publicly available dataset.

Acknowledgements

Albert Selebea Lutakamale received financial support from the Ministry of Education, Science and Technology of the United Republic of Tanzania under the MoEST-HEET Project fund.

References

Aernouts, M., Bellekens, B., Berkvens, R., Weyn, M., 2018a. A comparison of signal strength localization methods with Sigfox. http://dx.doi.org/10.1109/WPNC.2018.8555743, Paper Presented at the 2018 15th Workshop on Positioning, Navigation and Communications (WPNC), 25-26 October 2018, Bremen, Germany.

Aernouts, M., Berkvens, R., Xia, W., Vlaenderen, K., Weyn, M., 2018b. Sigfox and LoRaWAN datasets for fingerprint localization in large urban and rural areas. Data 3 (13), 1–15. http://dx.doi.org/10.3390/data3020013.

Aernouts, M., Bnilam, N., Berkvens, R., Weyn, M., 2020. TDAoA: A combination of TDoA and AoA localization with LoRaWAN. Internet Things 11, 100236. http://dx.doi.org/10.1016/j.iot.2020.100236.

Akhmedov, N., Khujamatov, H., Lazarev, A., Seidullayev, M., 2021. Application of LPWAN technologies for the implementation of IoT projects in the Republic of Uzbekistan. http://dx.doi.org/10.1109/ICISCT52966.2021.9670339, Paper Presented at the 2021 International Conference on Information Science and Communications Technologies (ICISCT), 03-05 November 2021, Tashkent, Uzbekistan.

Alomari, A., Comeau, F., Phillips, W., Aslam, N., 2018. New path planning model for mobile anchor-assisted localization in wireless sensor networks. Wirel. Netw. 24, 2589–2607. http://dx.doi.org/10.1007/s11276-017-1493-2.

Anagnostopoulos, G., Kalousis, A., 2019a. A reproducible analysis of RSSI fingerprinting for outdoor localization using Sigfox: Preprocessing and hyperparameter tuning. http://dx.doi.org/10.1109/IPIN.2019.8911792, Paper Presented at the 2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 30 September 2019 - 03 October 2019, Pisa, Italy.

Anagnostopoulos, G., Kalousis, A., 2019b. A reproducible comparison of RSSI fingerprinting localization methods using LoRaWAN. http://dx.doi.org/10.1109/WPNC47567.2019.8970177, Paper Presented at the 2019 16th Workshop on Positioning, Navigation and Communications (WPNC), 23-24 October 2019, Bremen, Germany.

Anagnostopoulos, G., Kalousis, A., 2021. Analysing the data-driven approach of dynamically estimating positioning accuracy. http://dx.doi.org/10.1109/ICC42927.2021.9500369, Paper Presented at the ICC 2021 - IEEE International Conference on Communications, 14-23 June 2021, Montreal, QC, Canada.

Anjum, M., Khan, M., Hassan, S., Jung, H., Dev, K., 2022. Analysis of time-weighted LoRa-based positioning using machine learning. Comput. Commun. 193, 266–278. http://dx.doi.org/10.1016/j.comcom.2022.07.010.

Anjum, M., Khan, M., Hassan, S., Mahmood, A., Qureshi, H., Gidlund, M., 2020. RSSI fingerprinting-based localization using machine learning in LoRa networks. IEEE Internet Things Mag. 3 (4), 53–59. http://dx.doi.org/10.1109/IOTM.0001.2000019.

Aqeel, I., Iorkyase, E., Zangoti, H., Tachtatzis, C., Atkinson, R., Andonovic, I., 2023. LoRaWAN-implemented node localisation based on received signal strength indicator. IET Wirel. Sens. Syst. 13 (4), 117–132. http://dx.doi.org/10.1049/wss2.12039.

Bhatti, G., 2018. Machine learning based localization in large-scale wireless sensor networks. Sensors 18, 4179. http://dx.doi.org/10.3390/s18124179.

C.E., O., Dhillon, H., Buehrer, R., 2018. A statistical characterization of localization performance in wireless networks. IEEE Trans. Wireless Commun. 17 (9), 5841–5856. http://dx.doi.org/10.1109/TWC.2018.2850310.

Chen, Y., Sambo, A., Onireti, O., Imran, M., 2022. A survey on LPWAN-5G integration: Main challenges and potential solutions. IEEE Access 10, 32132–32149. http://dx.doi.org/10.1109/ACCESS.2022.3160193.

Chen, M., Zhao, H., Shi, C., Chen, X., Niu, D., 2023. Multi-scene LoRa positioning algorithm based on Kalman filter and its implementation on NS3. Ad Hoc Netw. 141, 103097. http://dx.doi.org/10.1016/j.adhoc.2023.103097.

Cho, J., Hwang, D., Kim, K., 2019. Improving TDoA based positioning accuracy using machine learning in a LoRaWAN environment. http://dx.doi.org/10.1109/ICOIN.2019.8718160, Paper Presented at the 2019 International Conference on Information Networking (ICOIN), 09-11 January 2019, Kuala Lumpur, Malaysia.

Chollet, F., 2018. Deep Learning with Python. Manning Publications Co, Shelter Island, NY 11964.

Durand, T., Visagie, L., Booysen, M., 2019. Evaluation of next-generation low-power communication technology to replace GSM in IoT-applications. IET Commun. 13 (16), 2533–2540. http://dx.doi.org/10.1049/iet-com.2019.0168.

Ferreras, G., Talampas, M., 2021. LoRa-based differential fingerprint localization in outdoor environments. http://dx.doi.org/10.1109/WF-IoT51360.2021.9595880, Paper Presented at the 7th IEEE World Forum on Internet of Things, WF-IoT 2021, 14 June 2021 - 31 July, New Orleans, LA, USA.

Gill, S., Wu, H., Patros, P., Ottaviani, C., Arora, P., Pujol, V., Haunschild, D., Parlikad, A., Cetinkaya, O., Lutfiyya, H., Stankovski, V., Li, R., Ding, Y., Qadir, J., Abraham, A., Ghosh, S., Song, H., Sakellarious, R., Rana, O., Rodrigues, J., Kanhere, S., Dustdar, S., Uhling, S., Ramamohanarao, K., Buyya, R., 2024. Modern computing: Vision and challenges. Telemat. Inform. Rep. 13, 100116. http://dx.doi.org/10.1016/j.teler.2024.100116.

Goldoni, E., Prando, L., Vizziello, A., Savazzi, P., Gamba, P., 2019. Experimental data set analysis of RSSI-based indoor and outdoor localization in LoRa networks. Internet Technol. Lett. 2 (1), http://dx.doi.org/10.1002/itl2.75.

Guo, D., Gu, C., Jiang, L., Luo, W., Tan, R., 2022. ILLOC: In-hall localization with standard LoRaWAN uplink frames. ACM Interact. Mob. Wearable Ubiquitous Technol. 6 (1), 1–26. http://dx.doi.org/10.1145/3517245.

Ikpehai, A., Adebisi, B., Rabie, K., Anoh, K., Ande, R., Gacanin, H., Mbanaso, U., 2019. Low-Power Wide Area network technologies for Internet-of-Things: A comparative review. IEEE Internet Things J. 6 (2), 2225–2240. http://dx.doi.org/10.1109/JIOT.2018.2883728.

Islam, K., Murray, D., Diepeveen, D., 2023. Machine learning-based LoRa localisation using multiple received signal features. IET Wirel. Sens. Syst. 13, 133–150. http://dx.doi.org/10.1049/wss2.12063.

Janssen, T., Aernouts, M., Berkvens, R., Weyn, M., 2018. Outdoor fingerprinting localization using Sigfox. http://dx.doi.org/10.1109/IPIN.2018.8533826, Paper Presented at the 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 24-27 September 2018, Nantes, France, 1–6.

Janssen, T., Berkvens, R., Weyn, M., 2020a. Benchmarking RSS-based localization algorithms with LoRaWAN. Internet Things (Netherlands) 11, 100235. http://dx.doi.org/10.1016/j.iot.2020.100235.

Janssen, T., Berkvens, R., Weyn, M., 2020b. RSS-based localization and mobility evaluation using a single NB-IoT cell. Sensors 20, 6172. http://dx.doi.org/10.3390/s20216172.

Kumari, J., Kumar, P., Singh, S., 2019. Localization in three-dimensional wireless sensor networks: A survey. J. Supercomput. 75, 5040–5083. http://dx.doi.org/10.1007/s11227-019-02781-1.

Lalle, Y., Fourati, L., Fourati, M., Barraca, J., 2019. A comparative study of LoRaWAN, SigFox, and NB-IoT for smart water grid. http://dx.doi.org/10.1109/GIIS48668.2019.9044961, Paper Presented at the 2019 Global Information Infrastructure and Networking Symposium (GIIS), 18-20 December 2019, Paris, France.

Li, Y., Barthelemy, J., Sun, S., Perez, P., Moran, B., 2022a. Urban vehicle localization in public LoRaWan network. IEEE Internet Things J. 9 (12), 10283–10294. http://dx.doi.org/10.1109/JIOT.2021.3121778.

Li, X., Qiu, B., Cao, G., Wu, C., Zhang, L., 2022b. A novel method for ground-based cloud image classification using transformer. Remote Sens. 14, 3978. http://dx.doi.org/10.3390/rs14163978.

Lin, T., Liu, X., Qiu, X., 2022. A survey of transformers. AI Open 3, 111–132. http://dx.doi.org/10.1016/j.aiopen.2022.10.001.

Liu, J., Gao, J., Jha, S., Hu, W., 2022. Seirios: Leveraging multiple channels for LoRaWAN indoor and outdoor localization. http://dx.doi.org/10.1145/3447993.3483256, Paper Presented at the 27th Annual International Conference on Mobile Computing and Networking, January 31-February 4, 2022, New Orleans, LA, USA.

Marquez, L., Calle, M., 2022. Machine learning for localization in LoRaWAN: A case study with data augmentation. http://dx.doi.org/10.1109/FNWF55208.2022.00081, Paper Presented at the 2022 IEEE Future Networks World Forum (FNWF), 10-14 October 2022, Montreal, QC, Canada.

Marquez, L., Calle, M., 2023. Understanding LoRa-based localization: Foundations and challenges. IEEE Internet Things J. 10 (13), 11185–11198. http://dx.doi.org/10.1109/JIOT.2023.3248860.

Miles, B., Bourennane, E., Boucherkha, S., Chikhi, S., 2020. A study of LoRaWAN protocol performance for IoT applications in smart agriculture. Comput. Commun. 164, 148–157. http://dx.doi.org/10.1016/j.comcom.2020.10.009.

Mkkhaylov, K., Stusek, M., Masek, P., Fujdiak, R., Mozny, R., 2020. Communication performance of a real-life wide-area low-power network based on Sigfox technology. http://dx.doi.org/10.1109/ICC40277.2020.9148645, Paper Presented at the ICC 2020 - 2020 IEEE International Conference on Communications (ICC), 07-11 June 2020, Dublin, Ireland.

Monawar, T., Mahmud, S., Hira, A., 2017. Anti-theft vehicle tracking and regaining system with automatic police notifying using Haversine formula. http://dx.doi.org/10.1109/ICAEE.2017.8255459, Paper Presented at the 2017 4th International Conference on Advances in Electrical Engineering (ICAEE), 28-30 September 2017, Dhaka, Bangladesh.

Muppala, R., Navnit, A., Devendra, D., Materay, E., Accetturaz, N., Hussain, A., 2021. Feasibility of standalone TDoA-based localization using LoRaWAN. http://dx.doi.org/10.1109/ICL-GNSS51451.2021.9452311, Paper Presented at the 2021 International Conference on Localization and GNSS (ICL-GNSS), 01-03 June 2021, Tampere, Finland.

Obeidat, H., Shuaieb, W., Obeidat, O., Abd-Alhameed, R., 2021. A review of indoor localization techniques and wireless technologies. Wirel. Pers. Commun. 119, 289–327. http://dx.doi.org/10.1007/s11277-021-08209-5.

Pandangan, Z., Talampas, M., 2020. Hybrid LoRaWAN localization using ensemble learning. http://dx.doi.org/10.1109/GIOTS49054.2020.9119520, Paper Presented at the 2020 Global Internet of Things Summit (GIoTS), 03 June 2020, Dublin, Ireland.

Perez, M., Sierra-Sanchez, F., Chaparro, F., Chaves, D., Paez-Rueda, C., Galindo, G., Fajardo, A., 2022. Coverage and energy-efficiency experimental test performance for a comparative evaluation of unlicensed LPWAN: LoRaWAN and SigFox. IEEE Access 10, 97183–97196. http://dx.doi.org/10.1109/ACCESS.2022.3206030.

Pimpinella, A., Redondi, A., Nikoli, M., Cesana, M., 2020. Machine learning based localization of LoRaWAN devices via inter-technology knowledge transfer. http://dx.doi.org/10.1109/ICCWorkshops49005.2020.9145033, Paper Presented at the 2020 IEEE International Conference on Communications Workshops (ICC Workshops), 07-11 June 2020, Dublin, Ireland.

Purohit, J., Wang, X., Mao, S., Sun, X., Yang, C., 2020. Fingerprinting-based indoor and outdoor localization with LoRa and deep learning. http://dx.doi.org/10.1109/GLOBECOM42002.2020.9322261, Paper Presented at the GLOBECOM 2020 - 2020 IEEE Global Communications Conference, 07-11 December 2020, Taipei, Taiwan.

Queralta, J., Gia, T., Zou, Z., Tenhunen, H., Westerlund, T., 2019. Comparative study of LPWAN technologies on unlicensed bands for M2M communication in the IoT: beyond LoRa and LoRaWAN. http://dx.doi.org/10.1016/j.procs.2019.08.049, Paper Presented at the 14th International Conference on Future Networks and Communications (FNC), 19-21 August 2019.

Sassi, M., Fourati, L., 2022. Comprehensive survey on air quality monitoring systems based on emerging computing and communication technologies. Comput. Netw. 209, 108904. http://dx.doi.org/10.1016/j.comnet.2022.108904.

Shao, R., Bi, X., Chen, Z., 2022. A novel hybrid transformer-CNN architecture for environmental microorganism classification. PLoS ONE 17 (11), 1–22. http://dx.doi.org/10.1371/journal.pone.0277557.

Singh, S., Sharma, S., 2018. A PSO based improved localization algorithm for wireless sensor network. Wirel. Pers. Commun. 98, 487–503. http://dx.doi.org/10.1007/s11277-017-4880-1.

Stusek, M., Moltchanov, D., Masek, P., Mikhaylov, K., Zeman, O., Roubicek, M., Koucheryavy, Y., Hosek, J., 2020. Accuracy assessment and cross-validation of LPWAN propagation models in urban scenarios. IEEE Access 8, 154625–154636. http://dx.doi.org/10.1109/ACCESS.2020.3016042.

Svertokat, E., Marghescut, I., Rusu-Casandrat, A., Burget, R., Hosek, J., Masek, P., Ometov, A., 2022. Evaluation of real-life LoRa WAN localization: Accuracy dependencies analysis based on outdoor measurement datasets. http://dx.doi.org/10.1109/COMM54429.2022.9817302, Paper Presented at the 2022 14th International Conference on Communications (COMM), 16-18 June 2022, Bucharest, Romania.

Tay, Y., Dehghani, M., Metzler, D., 2022. Efficient transformers: A survey. arXiv:2009.06732v3.

Torres-Sospedra, J., Montoliu, R., Trilles, S., Belmonte, O., Huerta, J., 2015. Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems. Expert Syst. Appl. 42, 9263–9278. http://dx.doi.org/10.1016/j.eswa.2015.08.013.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., Polosukhin, I., 2017. Attention is all you need. Paper Presented at the 31st Conference on Neural Information Processing Systems (NIPS 2017), December 4 - 9, Long Beach, CA, USA.

Vazquez-Rodas, A., Astudillo-Sainas, F., Sanchez, C., Arpi, B., Minchala, L., 2020. Experimental evaluation of RSSI-based positioning system with low-cost LoRa devices. Ad Hoc Netw. 105, 102168. http://dx.doi.org/10.1016/j.adhoc.2020.102168.

Vindas, Y., Guepie, B., Almar, M., Roux, E., Delachartre, P., 2022. 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, vol. 182, pp. 270–296.

Wu, Z., Liu, Z., Lin, J., Lin, Y., Han, S., 2020. Lite transformer with long-short range attention. arXiv:2004.11886v1.

Yang, L., Yang, Y., Yang, J., Zhao, N., Wu, L., Wang, L., Wang, T., 2022. FusionNet: A convolution–transformer fusion network for hyperspectral image classification. Remote Sens. 14, 4066. http://dx.doi.org/10.3390/rs14164066.

Zafari, F., Gkelias, A., Leung, K., 2019. A survey of indoor localization systems and technologies. IEEE Commun. Surv. Tutor. 21 (3), 2568–2599. http://dx.doi.org/10.1109/COMST.2019.2911558.

Zhang, H., Wang, Z., Xia, W., Ni, Y., Zhao, H., 2022. Weighted adaptive KNN algorithm with historical information fusion for fingerprint positioning. IEEE Wirel. Commun. Lett. 11 (5), 1002–1006. http://dx.doi.org/10.1109/LWC.2022.3152610.

Ziemann, M., Metzler, C., 2022. Transformers for robust radar waveform classification. http://dx.doi.org/10.1109/IEEECONF56349.2022.10052000, Paper Presented at the 2022 56th Asilomar Conference on Signals, Systems, and Computers, 31 October 2022 - 02 November 2022, Pacific Grove, CA, USA.