Smartphone Sensor Fusion For Indoor Localization: A Deep LSTM Approach

by

Zhitao Yu

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
December 15, 2018

Keywords: Bimodal data; deep learning; deep long short-term memory (LSTM); indoor localization; light intensity; magnetic field.

Approved by

Shiwen Mao, Samuel Ginn Endowed Professor of Electrical and Computer Engineering
Thaddeus Roppel, Associate Professor of Electrical and Computer Engineering
Xiaowen Gong, Assistant Professor of Electrical and Computer Engineering

Abstract

With the fast increasing demands of location-based service and the proliferation of smartphones and other mobile devices, accurate indoor localization has attracted great interest. In this thesis, we present DeepML, a deep long short-term memory (LSTM) based system for indoor localization using smartphone magnetic and light sensors. We verify the feasibility of using bimodal magnetic and light data for indoor localization through experiments. We then design the DeepML system, which first builds bimodal images by data preprocessing, and then trains a deep LSTM network to extract the location features. Newly received magnetic field and light intensity data are then exploited for estimating the location of the mobile device using an improved probabilistic method. Our extensive experiments verify the effectiveness of the proposed DeepML system.

Acknowledgments

I would first like to thank my advisor Prof. Shiwen Mao for his patient guidance and enthusiastic help, without which this thesis would not have been possible. I would also like to thank my committee members, Prof. Xiaowen Gong and Prof. Thaddeus Roppel, for serving as my committee members. I also want to thank you for letting my defense be an enjoyable moment, and for your brilliant comments and suggestions.

I would especially like to thank my team member, Xuyu Wang, as well as other friends, who gave lots of assistance throughout my thesis work. Xuyu, who led me into this field, taught me a lot about deep learning. Thank Chao Yang, Runze Huang for their cooperative support and continual assistance through this research.

Finally, I would also like to thank my colleagues in the RFID Lab, Mr. Justin Patton, Dr. Jian Zhang, Xiangyu Wang and Yibo Lyu. They were always open to discussions about my ongoing research. I learned a lot from them about how to conduct research tasks.

This work is supported in part by the US NSF under Grant CNS-1702957, and by the Wireless Engineering Research and Education Center (WEREC) at Auburn University.

Table of Contents

Abstract ii

Acknowledgments iii

1 Introduction 1

1.1 Approach 2

1.2 Organization 4

2 Background 5

2.1 Positioning System With Visible Light 5

2.1.1 Visible Light Communication 5

2.1.2 Visible Light Positioning 6

2.2 Positioning System with Magnetic Field 10

2.3 Deep Learning 10

2.3.1 Recurrent Neural Networks (RNNs) 11

2.3.2 Long Short-Term Memory (LSTM) 11

3 Preliminaries and Motivation 14

3.1 Magnetic Field Preliminaries 14

3.2 Light Intensity Preliminaries 15

3.3 Fusion of Magnetic Field and Light Intensity Data 16

4 The Proposed DeepML System 18

4.1 DeepML System Architecture 18

4.2 Data Preprocessing 20
4.3 Offline Training 20
4.3.1 Feature Extraction 20
4.3.2 Deep LSTM Network 21
4.3.3 Softmax Classifier 22
4.4 Online Location Estimation 22

5 Experimental Study 24
5.1 Experiment Setup 24
5.1.1 Lab Scenario 24
5.1.2 Corridor Scenario 25
5.2 Localization Performance 25
5.3 Impact of Different Parameters 28
5.3.1 Impact of Test Data Size 28
5.3.2 Impact of the Number of LSTM Layers 29
5.3.3 Impact of Hidden Units 30
5.3.4 Impact of Window Size 32
5.3.5 Impact of Batch Size 33
5.3.6 Impact of Batch Epoch 35

6 Conclusions and Future Work 37
6.1 Conclusions 37
6.2 Future Work 37

References 39

List of Figures

2.1 System architecture of the ByteLight. 6

2.2 System architecture of the LiTell. 8

2.3 System architecture of the NaviLight. 9

2.4 Recurrent Neural Network (RNN) Structure. 11

2.5 Long Short-term Memory (LSTM) Structure. 12

3.1 Characteristics of magnetic field data. 15

3.2 Characteristics of light intensity data. 16

3.3 Confusion matrix of the bimodal data with magnetic field vector and light intensity. 17

4.1 The DeepML system architecture. 19

4.2 Structure of Deep LSTM network. 21

5.1 Layout of the computer lab scenario: training locations are marked as red squares. 25

5.2 Layout of the corridor scenario: training locations are marked as red squares. . 25

5.3 Training errors for the laboratory and corridor experiments. 26

5.4 CDF of localization error of the lab experiment. 26

5.5 CDF of localization error of the corridor experiment. 27

5.6 The average distance error for different size of test data. 28

5.7 The average distance error for different number of LSTM layers. 29

5.8 The average training time for different number of LSTM layers. 29

5.9 The average distance error for different hidden units. 31

5.10 The average training time for different number of hidden units. 31

5.11 The average distance error for different window size. 32

5.12 The average training time for different window size. 33

5.13 The average distance error for different batch size. 34

5.14 The average training time for different batch size. 34

5.15 The average distance error for different epoch. 35

5.16 The average training time for different epoch. 35

Chapter 1

Introduction

Indoor localization has been a research hotspot for decades [1–3]. However, unlike outdoor GPS navigation systems, there are still no robust indoor localization systems widely adopted by now. In fact, people still cannot use the popular Google Maps to navigate to a meeting room in an unfamiliar office building. Recently, there is considerable new interest in indoor localization techniques, driven by the proliferation of smartphones and other mobile devices, which, on one hand, makes it possible to enable many location based services, and on the other hand, provides an array of embedded sensors that can be exploited for indoor localization. Specifically, many researchers focus on WiFi [4] based fingerprinting indoor localization using received signal strength (RSS) [5–7] or Channel State information (CSI) [8–18]. These methods can achieve robust meter-level accuracy but cannot work effectively when the WiFi signal is weak or not available in some scenarios, such as underground parking areas.

In contrast, the geomagnetic field is *omnipresent* and thus can be considered as a ubiquitous signature for indoor localization. In the past, geomagnetism basically needs to be used with special equipments for robot tracking [19] and navigation [20]. In [20], researchers employ the leader-follower model in a navigation system, where customized magnetic sensing devices are used for blind people. On the other hand, for magnetic field based localization with smartphones, the authors in [21] use mobile phones to measure magnetic field intensity and use them as magnetic signatures for identifying locations and rooms. However, according to its strategy, this system depends heavily on pillars in the building and only achieves room-level accuracy. Recently, the Magicol system combines magnetism and WiFi RSSI to build a fingerprint map, which is designed with a particle-filtering based inertial measurement unit (IMU) engine for localization and tracking [22]. Other systems based on magnetic sequences

matching are proposed for improving tracking accuracy [23]. The above magnetic field based smartphone localization systems require the user to walk around for data collection and online localization.

In addition, visible light is also omnipresent and has been exploited for localization, due to the density and stability of lighting infrastructures. For example, visible light intensity in an underground parking area usually does not changed over time, and is not influenced by the outdoor sunlight, which can be thus leveraged for indoor localization. Existing visible light localization systems, such as polarized LEDs [24] and collocated LEDs [25], require customized LED drivers to emit identity beacons, which increases the system cost. To eliminate the need for customized LEDs, LiTell system [26] extracts high-frequency features from fluorescent light for localization. Other visible light localization systems for smartphones are based on particle-filtering and light intensity data sequence, for which there is still room for improvement by exploiting movement sensors [27, 28].

1.1 Approach

In this thesis, we exploit bimodal magnetic field and ambient light data for indoor localization with a deep learning approach. The proposed scheme is motivated by the following observations. First, the magnetic field and light intensity at each location are highly stable and robust over time. Second, magnetic field and light intensity are complementary to each other at many locations. For example, magnetic field does not perform well at some locations, while these locations may have different light intensities, which can be used to distinguish them. Using the bimodal data can enhance magnetic field based indoor localization schemes. Third, using bimodal data with magnetic field and light intensity can increase the size of input data, thus improving location diversity and recognition performance. Moreover, we incorporate a deep long-short term memory (LSTM) network to train the bimodal data, which is a popular recurrent neural network (RNN) to deal with long-range dependencies [29, 30]. The deep LSTM network has been successfully employed for speech recognition [31] and human activity recognition [32]. Compared to conventional fingerprinting based methods, the deep LSTM network only requires one group of weights trained for all training locations, instead of creating

a database for each training location. This feature can accelerate location prediction and reduce the data storage requirement.

In particular, we design DeepML, a **Deep** LSTM network based indoor localization system using smartphone **M**agnetic and **L**ight sensors. The proposed DeepML system includes a data preprocessing module for collecting magnetic field and light intensity data, and to create bimodal image data with a sliding window method. DeepML also has an offline training phase that includes feature extraction, the deep LSTM network, and a softmax classifier. A fully connected layer is implemented for extracting features from bimodal image data. The deep LSTM network consists of two layers of LSTM networks to achieve a stonger learning and representation ability. The softmax classifier employs the cross-entropy to measure the difference between true labeled data and the normalized output data, as well as the L2 regularization hyperparameter to avoid over-fitting. The back propagation through time (BPTT) algorithm, which is a gradient-based technique for training certain types of RNNs, is used for training the deep LSTM network. For online location prediction, an improved probabilistic method is leveraged for estimating the location of the target smartphone using newly received magnetic field and light intensity data.

The main contributions of this paper include:

- We experimentally validate the feasibility of using magnetic field and light intensity data for indoor localization. We show that both data are stable over time, and the fusion of magnetic field and light intensity data can improve location diversity and accuracy. To the best of our knowledge, this is the first work to employ bimodal magnetic field and light intensity data for indoor localization with a deep LSTM network approach.

- We present the DeepML system design, which first builds bimodal images to train the deep LSTM network, and then employs newly received magnetic field and light data for estimating the location of the target mobile device.

- We implement the proposed DeepML system with Android smartphones, and validate its performance in two typical indoor environments with extensive experiments. DeepML

outperforms the baseline scheme that uses magnetic field data only with considerable gains in all the experiments.

1.2 Organization

In the remainder of this paper, we present the research background in Chapter 2. Preliminaries and motivation are showed in Chapter 3. We describe the DeepML design in Chapter 4 and demonstrate its performance in Chapter 5. Chapter 6 presents the conclusion and discuss our future work.

Chapter 2

Background

Although indoor positioning technology is called extension of outdoor positioning technology, because of its application scenarios and accuracy requirements, it has not achieved much commercial value for a long time. In recent years, especially due to the development of consumer-grade market positioning technology led by Ibeacon and the marketization of centimeter-level ultra-precision positioning technology led by UWB technology, the indoor positioning technology has attracted more and more researchers' attention. At present, researchers have proposed radio frequency identification (RFID), Bluetooth, WLAN (Wireless Local Area Networks), ultra-wideband (Ultra-Wideband, UWB),visible light, magnetic and other technologies based systems. In this section, we will emphatically introduce visible light and magnetic field based indoor localization technologies.

2.1 Positioning System With Visible Light

2.1.1 Visible Light Communication

With the development and popularization of smart devices, wireless network technology has been a frontier hot research topic [33–52], such as 5G technology [53–59], Radio Frequency technology and visible light communication (VLC). The visible light communication technology uses a high-speed light and dark flashing signal which are invisible to the naked eye to transmit information, such as a fluorescent lamp with 10 kbit/s or a light-emitting diode (LED) for up to 500 Mbit/s over short distances. Compared with traditional radio frequency technology in which the spectrum is controlled strictly, VLC leverages the rich visible light spectrum to

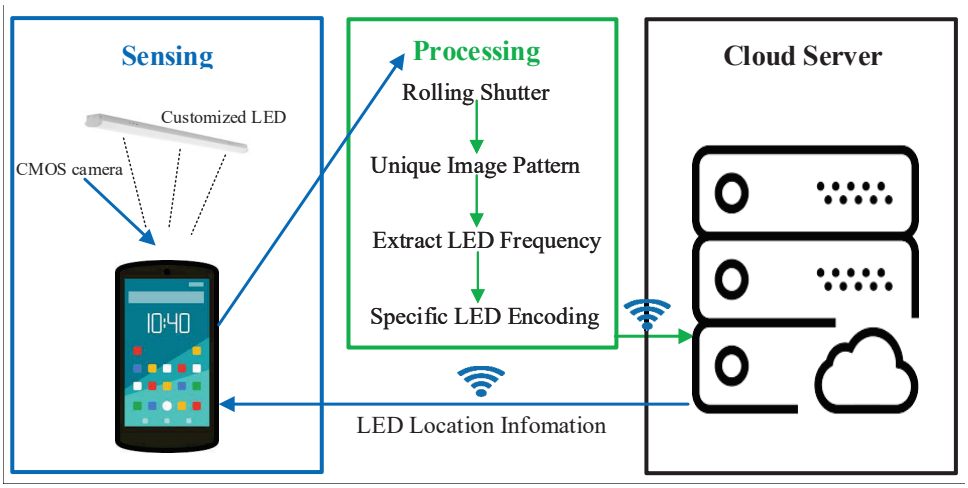

Figure 2.1: System architecture of the ByteLight.

communicate. This spectrum corresponds to the vicinity of 430790 THz in terms of frequency which is almost ten thousands times of radio spectrum. Moreover, VLC is able to transfer data at extremely high speed. In [60], Wang et al. proposed a hybrid post equalizer in a high-order s carrier-less amplitude and phase modulation based VLC system, which successfully achieved data rate of 8 Gbit/s over 1-meter indoor free space transmission. Visible light communication can be applied to the fields of lighting Internet access, visible light radio broadcasting, precise positioning of visible light, etc.

2.1.2 Visible Light Positioning

As a result of lacking GPS signal in indoor environment, the visible light positioning (VLP) has been an appealing research topic in recent years. Many researchers believe VLP technology could be a key solution to 5 billion ”indoor localization market”. VLP technologies applied on smartphones can be roughly divided into two categories according to its beaconing devices: light-emitting diodes (LEDs) based and fluorescent lights (FLs) based.

LEDs based VLP System:

VLP utilizes modulated light-emitting diode (LED) transmitters at known locations to estimate the position of VLC receivers. Besides the primary fuction of providing illumination, LEDs are also used for data transmission and localization because of its long life span, low cost and power consumption. Additionally, LEDs can be switched to different light intensity level very

fast and modulated to various frequencies, which is suitable for building fingerprinting maps. ByteLight [61] is one of most successful application of using LEDs for indoor localization. In fact, this system is based on an imaging principle of the complementary metaloxidesemiconductor (CMOS) camera which is called rolling shutter. CMOS camera does not capture the whole picture at one time but scans across the picture rapidly, either horizontally or vertically.Therefore, the rolling shutter effect causes predictable pictures' distortion when capturing fast moving objects or high frequency flashing lights. Then based on this rolling shutter, it can be found that the images obtained by CMOS cameras also have distortions in captured images under the LEDs with certain frequency information. Images obtained from different frequencies of LEDs have their unique patterns, such as the spacing between shadows on the image. As we all know, almost all of commercial-off-the-shelf(COTS) smartphones are implemented with CMOS cameras. Then the frequency information can be obtained by doing a refined image processing with performing fast fourier transform (FFT). ByteLight enables smartphones to demodulate the frequency information and communicates with the server to finally get the specific location. As shown in Fig.2.1, ByteLight is basically consists of three modules. This type of VLP system has relatively sub-meter level high positioning accuracy. However, due to customized LEDs are needed to beam identity frequency beacons, implementation and replacement of system's hardware usually cost a lot.

Fluorescent Lights based VLP System:

Compared with LEDs, the biggest advantage of fluorescent lights is the universality. Although LEDs technology has achieved great improvement, regardless of indoors or outdoors, the first choice of illumination is still fluorescent light because of its better floodlight and low price. In [26], the authors proposed a simple and robust indoor localization system which is called LiTell. The system architecture is shown in the Fig.2.2. They leverages unmodulated fluorescent lights(FLs) as location landmarks and commercial off-the-shelf(COTS) smartphones' cameras as light sensors. The key point of this LiTell system is the hypothesis that each fluorescent light has its unique and stable characteristic frequency(CF). As their observation, the FLs waveforms display like quasi-periodic patterns at both ms and μs scale. The periodicity

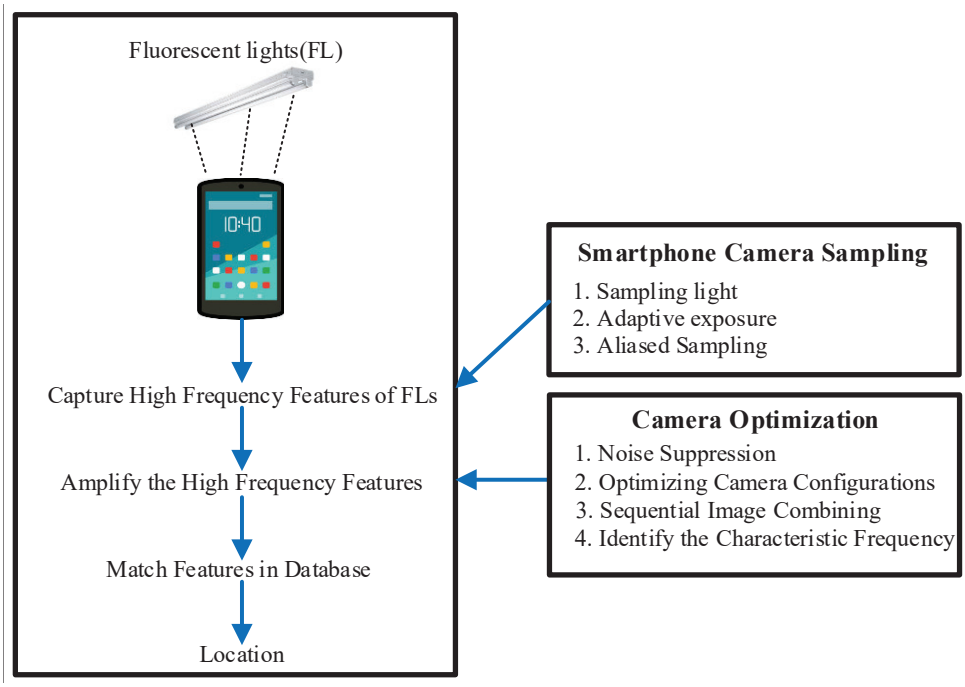

Figure 2.2: System architecture of the LiTell.

becomes noticeable in frequency domain. The researchers discovered the dominate frequency always occurs at 2 times of the fundamental frequency and denoted that as characteristic frequency. The characteristic frequency can keep stable and only have small amount of collisions with an acceptable range. Due to low sampling rate and sensitivity to high frequency light signals of COTS smartphones, another crucial contribution of this paper is the designed sampling and signal amplification mechanism which is used to enable smartphone's camera to extract high-frequency features from FLs. Finally, the researchers utilized corresponding FLs' locations and characteristic frequencies to build simple fingerprinting map on server for indoor localization. In comparison with the ByteLight, LTtell does not need any extra hardware implemented in both transmitter and receiver sides which means it is zero cost because everyone have smartphones.

Universal Light Sources based VLP System:

Unlike the two VLP systems aforementioned, this basically has no light source limitation and mostly uses light intensity to make fingerprints. This kind of VLP system does not need any modified devices in both transmiter and receiver sides. Zhao et al. proposed a generic light intensity based indoor localization and navigation system called NaviLight, which is the most

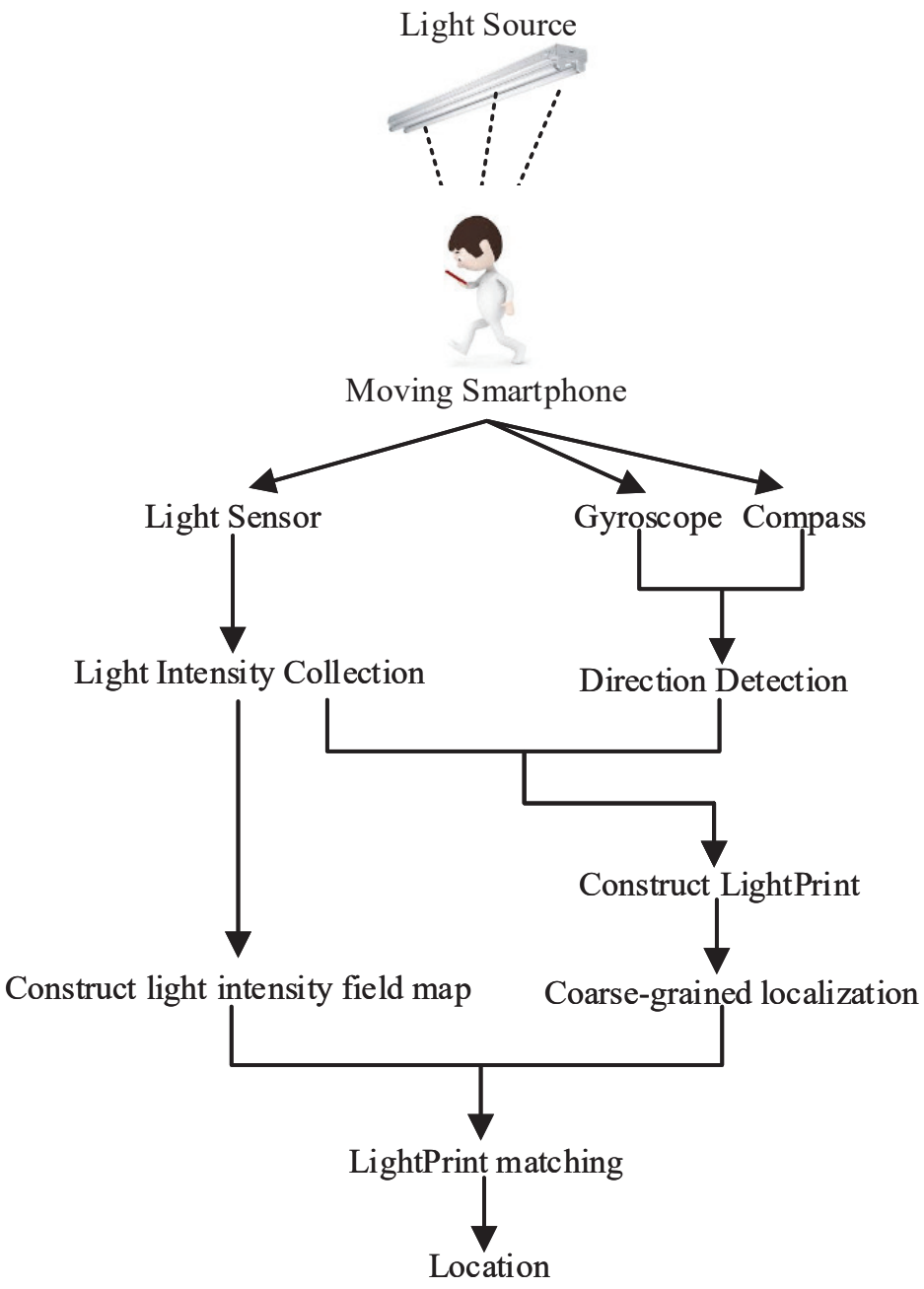

Figure 2.3: System architecture of the NaviLight.

representative one of universal light sources based VLP systems. Figure. 2.3 depicts how it works vividly. The researchers leveraged a vector of multiple light intensity values collected by smartphone during users walks to create fingerprint map which denoted as LightPrint. Meanwhile, they use collected light intensity value to construct a light intensity field map(LIF). Moreover, the localization process was divided into two parts. One is the coarse-grained localization via which is aimed to determine which subarea of the floor the LightPrint belongs to. Then they matched the LightPrint with the pre-populated LIFs. Due to uncontrollable walk

speed of human when collecting the LightPrint, it is impossible to be the same as that when LIFs were built. In order to eliminate this uncontrollable variable, they calculated the Dynamic Time Warping(DTW) distance to align and measure the similarity between two time sequences with different speed. Finally, the other part of localization process is based on the computed DTW distance to get the fine-grained localization results.

2.2 Positioning System with Magnetic Field

The Earth itself is a giant magnet that forms a basic magnetic field between the north and south poles of the geography. The earth's geomagnetic field can be disturbed by metal objects, especially when passing through a reinforced concrete structure. Buildings' metal structure disturb and distort the original geomagnetic field, which forms a unique specific indoor magnetic field. Moreover, if the building does not undergo structural changes in the steel body, the characteristics of the indoor magnetic field are robust. The indoor geomagnetic positioning is accomplished by extracting the characteristic features of the "indoor magnetic field" and matching with specific location and orientation information. However, the localization systems based on single geomagnetic field input can not achieve sub-meter level positioning. Therefore, more and more bimodal strategies have been proposed. The most practical and appealing scheme is combining Wifi signals and geomagnetic field as location landmarks [22]. The accuracy of Wifi in positioning and the sensitivity of the magnetic field in the orientation forms a great complementary relationship in positioning system.

2.3 Deep Learning

The concept of deep learning stems from the study of artificial neural networks. It is an algorithm that attempts to do high-level abstraction of data by using multiple processing layers consisting of complex structures or multiple nonlinear transformations. "Deep Learning" has been a term which is rebranding of neural network to some extent. The advantage of deep learning is to replace the manual acquisition of features with unsupervised or semi-supervised feature learning and hierarchical feature extraction efficient algorithms. In recent years, Deep

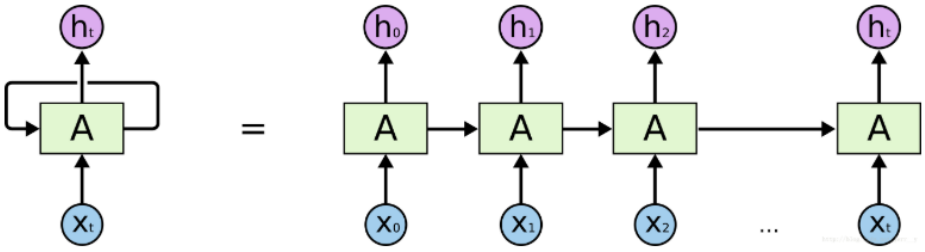

Figure 2.4: Recurrent Neural Network (RNN) Structure.

learning has been applied in computer vision, speech recognition, indoor localization [62], wireless network [63] and has achieved excellent results.

2.3.1 Recurrent Neural Networks (RNNs)

The purpose of RNNs is to process sequence data. In a traditional neural network we assume that all inputs (and outputs) are independent of each other. Moreover, the layers are fully connected from the input layer to the hidden layer to the output layer, and the nodes between each layer are disconnected. However, it is a bad idea and powerless for many tasks. For example, if you want to predict what the next word of a sentence is, you usually need to use the previous word due to the words in a sentence are not independent. RNNs are called recurrent neural networks because they perform the same task for every element of a sequence, where the current output depends on the previous computations. The specific form of expression is that the network memorizes the previous information and applies it to the calculation of the current output, that is, the nodes between the hidden layers are no longer unconnected. Furthermore, the input of the hidden layer includes not only the output of the input layer but also the output of the hidden layer at the previous moment. Another way to think about RNNs is that they have a memory which captures information about what has been calculated so far. In theory RNNs can make use of information in arbitrarily long sequences. However, they are limited to looking back only a few steps practically.

2.3.2 Long Short-Term Memory (LSTM)

As mentioned at Section 2.3.1, RNN can not handle long-term dependencies when processing variable-length sequence inputs. The dependencies cause the gradient of the loss functions to

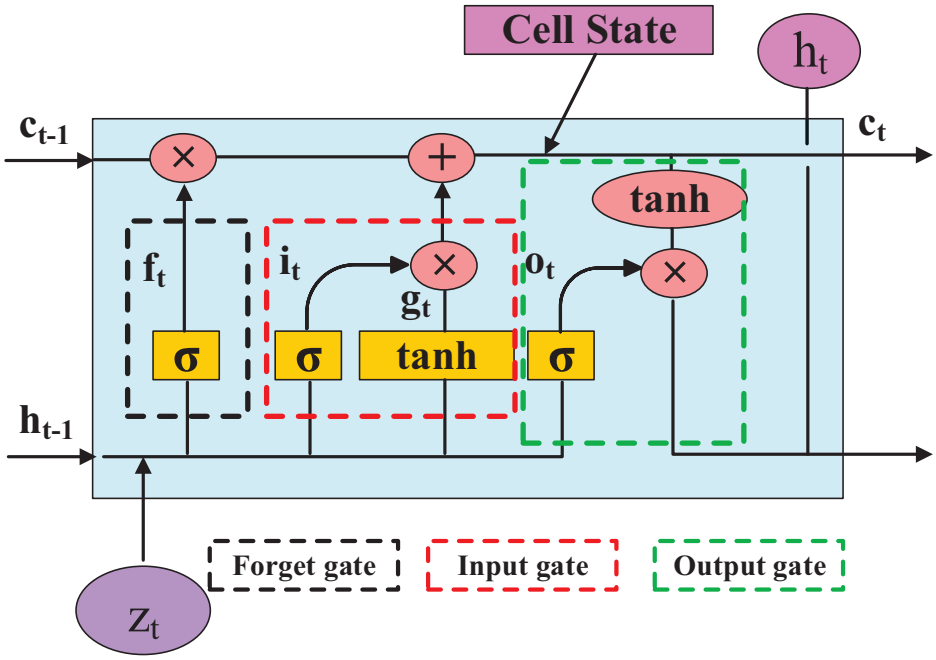

Figure 2.5: Long Short-term Memory (LSTM) Structure.

be either diffuse or explode which makes the network could not be trained successfully. In order to counter such drawbacks, researchers proposed long short-term memories architecture [64]. The biggest difference between RNN and LSTM is that there is a message conveyor called cell state at the top of LSTM, which is actually the place of information memory. Certainly, the conveyor is unable to determine what information to be added or deleted by itself, but through the structure called control gates. As shown in Fig. 2.5 LSTM network has three control gates: forget gate, input gate and output gate respectively. The forget gate determines which information from last cell state to continue to pass through the current cell. The input gate controls if a new data could flow into the memory and updates the cell state. The output gate decides which part of the cell state to be exported as output. These three control gates confirm the feasibility of using gradient-based optimization methods and avoid the diffusion and explosion of gradient. The LSTM network has been widely used in speech recognition, language modeling, sentiment analysis, human activity identification and text prediction. In recent years, many researchers introduced many LSTM-like deep learning networks, such as Gated Recurrent Unit (GRU) proposed by J.Chung et al [65]. at 2014. It combines the forget gate and the input gate into a separate ”update gate”. It also incorporates the cell state and the

hidden state and makes some other changes. The resulting model is simpler than the standard LSTM model and is becoming more popular.

Chapter 3

Preliminaries and Motivation

3.1 Magnetic Field Preliminaries

The magnetic field of the earth, i.e., the geomagnetic field, is ubiquitous on the surface of the earth, with magnitude ranging from 0.25 to 0.65 Gauss. The magnetometer in most smartphones can measure the magnetic field, in the form of a vector with three elements (m_x, m_y, m_z), which describes the magnetic field component for north, east, and vertical directions, respectively. For studying the stability and location diversity of magnetic field data, we measure the magnetic field data (m_x, m_y, m_z) at 10 different locations selected in a corridor of 20 meters long in the Broun Hall on Auburn University Campus. We obtain five different datasets collected at five different times. Fig. 3.1 shows the magnetic field data components (m_x, m_y, m_z) for different locations and times. We find that for any fixed location, all the three elements (m_x, m_y, m_z) exhibit small variations over time, as indicated by the negligible error bars. This validates the stability of magnetic field data with respect to location, which can guarantee the reliability of fingerprinting based indoor localization using magnetic field data.

In addition, we also find good diversity of magnetic field data for different locations. In Fig. 3.1, the magnetic field data exhibits sufficient variations for different locations. For example, each of the three elements (m_x, m_y, m_z) has different values for locations 1 and 2. Specifically, we can see that at least one element of the magnetic field data changes for a different location. The indoor magnetic field has local anomalies (or, local disturbances), because modern buildings generally have many ferromagnetic structures. The ambient magnetic field leads to geomagnetic anomalies, which can be leveraged for accurate indoor localization.

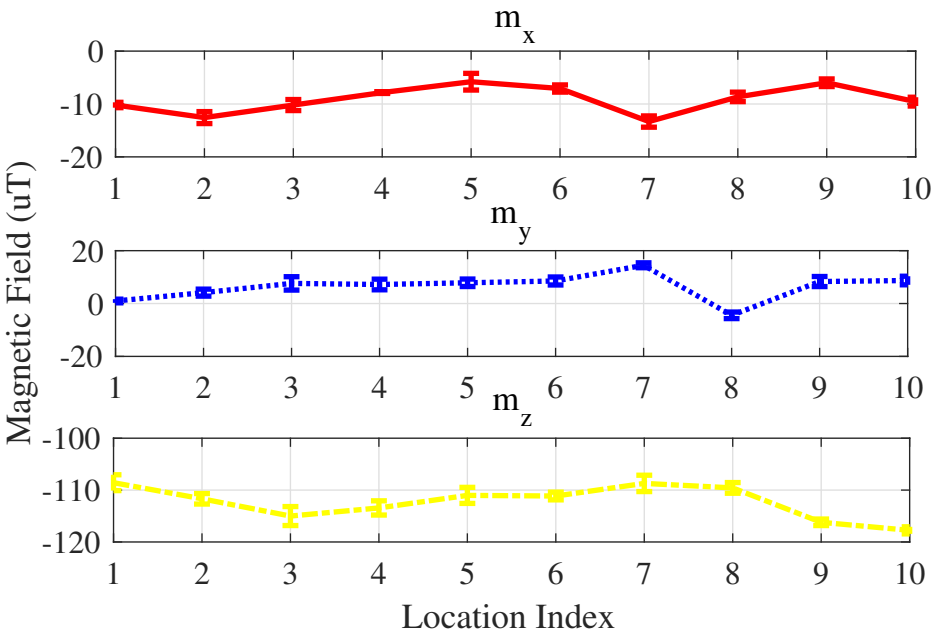

Figure 3.1: Characteristics of magnetic field data.

3.2 Light Intensity Preliminaries

Modern buildings usually use several types of light bulbs, such as the compact fluorescent lamp (CFL) and light-emitting diode (LED) [66]. Most smartphones can capture light intensity from such bulbs. In fact, light propagates in the air from the light bulbs to the smartphone light receiver, with different radiant intensity measurements for different locations, which are susceptible to the indoor propagation environment, such as shadowing, scattering, and reflection for different surfaces. This motivates the work on light intensity based fingerprinting localization technique [67].

To study the stability and location diversity of light intensity at different times and locations, we measure the light intensities at 10 different locations in the same corridor, and collected five datasets at different times. Fig. 3.2 presents the characteristics of light intensity data at different locations and times. Similarly, we find light intensity data is quite stable for any given location, as indicated by the negligible error bars. Furthermore, light intensity measurements take different values for some different locations, e.g., see neighboring locations 1 and 2, 4 and 5, and 9 and 10. But for some other neighboring locations, e.g., 3 and 4, and 7 and 8, the light intensity values are very close. Thus, unlike magnetic field data, it is difficult to use light intensities only as fingerprints for indoor localization.

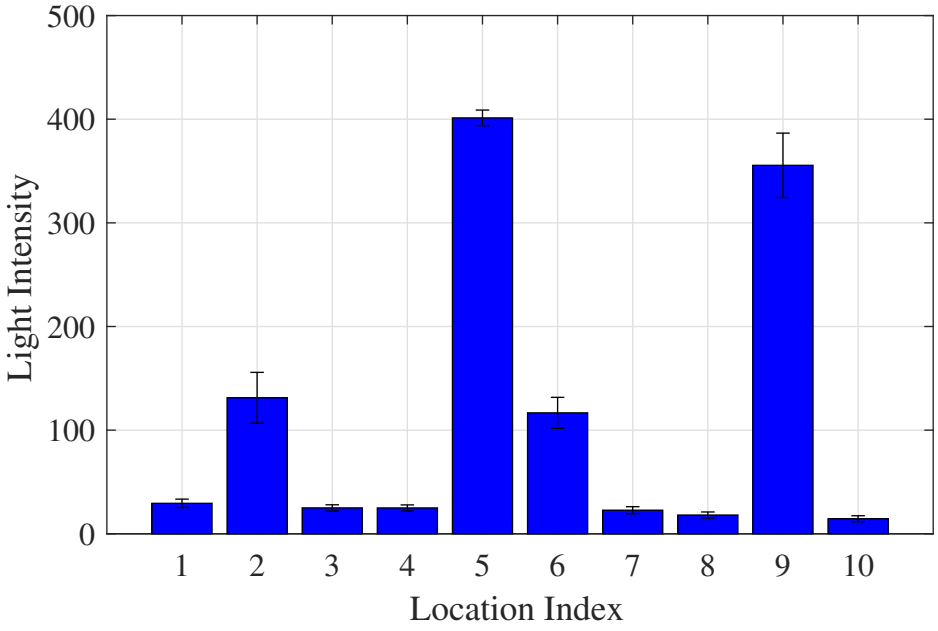

Figure 3.2: Characteristics of light intensity data.

3.3 Fusion of Magnetic Field and Light Intensity Data

Our measurement study of magnetic field and light intensity motivates us to use them as bimodal data for indoor localization. This is because we can use the different light intensities at different locations to improve the accuracy of magnetic field based indoor localization. By fusing the magnetic field and light intensity data, the dimension of input data is increased, making it suitable for the proposed deep LSTM based scheme, to strengthen the uniqueness of location features.

To measure the location diversity of the bimodal data with magnetic field and light intensity, we define the confusion matrix for N different locations as

$$\boldsymbol{D} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & \dots & d_{1N} \\ d_{21} & d_{22} & d_{23} & \dots & d_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{N1} & d_{N2} & d_{N3} & \dots & d_{NN} \end{bmatrix}, \tag{3.1}$$

where d_{ij} denotes the Euclidean distance between the two signal vectors of locations i and j, which can be computed by $d_{ij} = ||S_i - S_j||_2$, where S_i is the signal vector of location i,

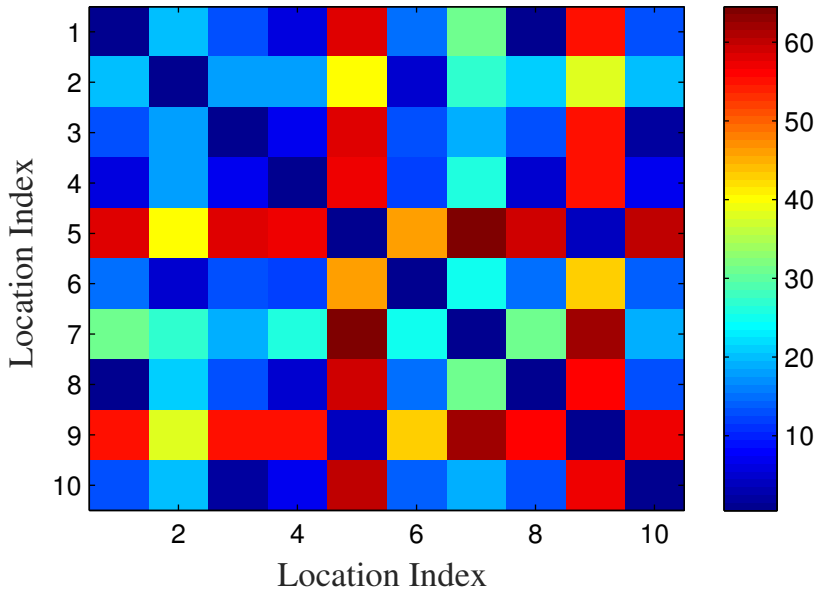

Figure 3.3: Confusion matrix of the bimodal data with magnetic field vector and light intensity.

including both the magnetic field vector and light intensity. To measure the performance of different datasets, we need to normalize the confusion matrix with the same metric.

Fig. 3.3 presents the confusion matrix of the bimodal data with magnetic field vector and light intensity for an experiment with 10 locations in the corridor. We can see that the fusion of magnetic field and light intensity achieves great location diversity with large distances for most location pairs, which is different from using magnetic field vector only or light intensity only in Sections 3.1 and 3.2. Such enhanced diversity is highly desirable for the training and location estimation of the proposed deep LSTM network for indoor localization.

Chapter 4

The Proposed DeepML System

4.1 DeepML System Architecture

We design the DeepML system and prototype it with a Samsung Galaxy S7 Edge smartphone with an Android 7.0 platform. The Android application is developed with Android Studio 2.3.3 for data collection and preprocessing. The proposed DeepML system employs both magnetic field data and ambient light for two main reasons. First, as discussed in Chapter 3, the variance in magnetic field and light intensity at each location is generally very small; they are both highly stable over time for each given position. Second, magnetic field measurements may not show sufficient location diversity in some areas. Incorporating the bimodal data could exploit the different light intensities for enhanced location diversity for such areas. Magnetic field and light intensity are complementary to each other for many locations. Using the bimodal data can improve the localization performance.

The design of DeepML is presented in Fig. 4.1. The most salient features include the use of bimodal magnetic and light data, and the deep LSTM network used for extracting location features from the bimodal data. DeepML first performs *data preprocessing* of collected magnetic field and light data, to build bimodal images using a sliding window method. During the *offline training* phase, we implement feature extraction for the bimodal images for effectively training the deep LSTM network. Compared to conventional fingerprinting based methods, DeepML does not need to establish a database for each training location, where either raw data or extracted features are stored as fingerprints. Rather, our DeepML system only requires one

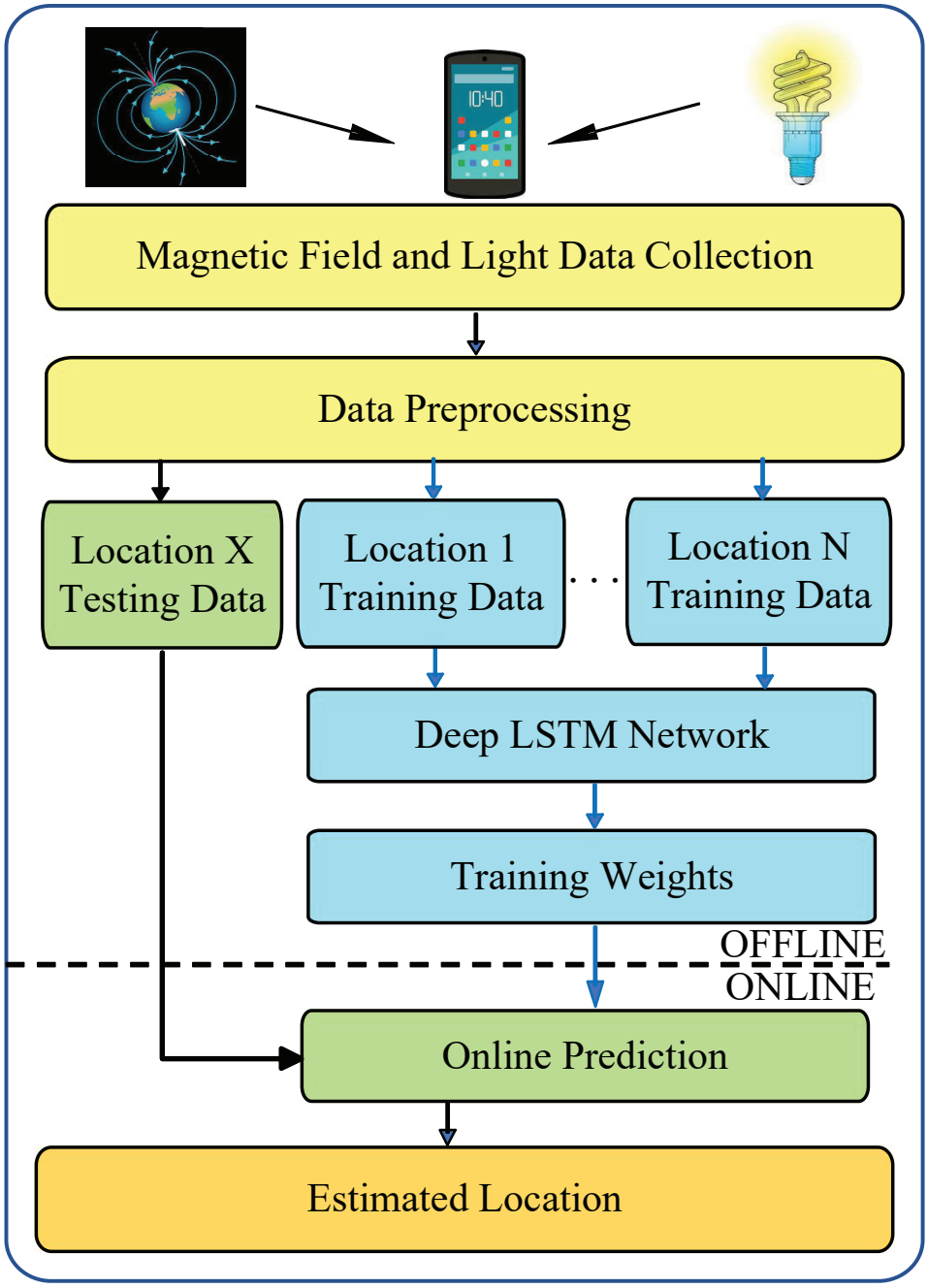

Figure 4.1: The DeepML system architecture.

group of weights to be trained for all training locations. In the *online testing* phase, we incorporate an improved probabilistic approach for location estimation, based on newly received magnetic and light bimodal data from the target mobile device.

4.2 Data Preprocessing

We first collect and record real-time readings from the smartphone magnetic field ambient light sensors. Due to the sequence size requirement of the deep LSTM network, we need to reduce the sampling rate of the magnetic field sensor, to make its readings in-sync with that from the ambient light sensor. Specifically, we obtain 1500 rows of magnetic field and light intensity combined patterns for each training location. For example, the size of training data is about 15000 rows for the 10 training locations in the corridor, and 18000 rows for the 12 train locations in the laboratory scenario. For online location estimation, the data size is about 400 rows for each test location.

We next employ a sliding window to build bimodal image data. For both training and testing phases, we set the size of the sliding window to 20. Thus, we can obtain the bimodal image data with size 20×4, with 20 measured data points in the column dimension and 4 feature values in the row dimension, including m_x, m_y, m_z for magnetic field and l for light.

4.3 Offline Training

For offline training, we propose a deep LSTM approach to extract location features from the bimodal image data with magnetic field and light intensity. The offline training module includes feature extraction, the deep LSTM network, and the Softmax classifier.

4.3.1 Feature Extraction

For better feature extraction, we implement one fully connected layer for extracting features from raw magnetic fields and light intensity data, which is formulated as

$$z_t = \text{ReLU}(Wx_t + b), \tag{4.1}$$

where x_t and z_t are the input and output of the fully connected layer, respectively, W and b are the weights and biases of the fully connected layer, respectively. $\text{ReLU}(\cdot)$ is the *rectified linear unit*, which is considered as the activation function with $\text{ReLU}(x) = \max(x, 0)$. The $\text{ReLU}(\cdot)$

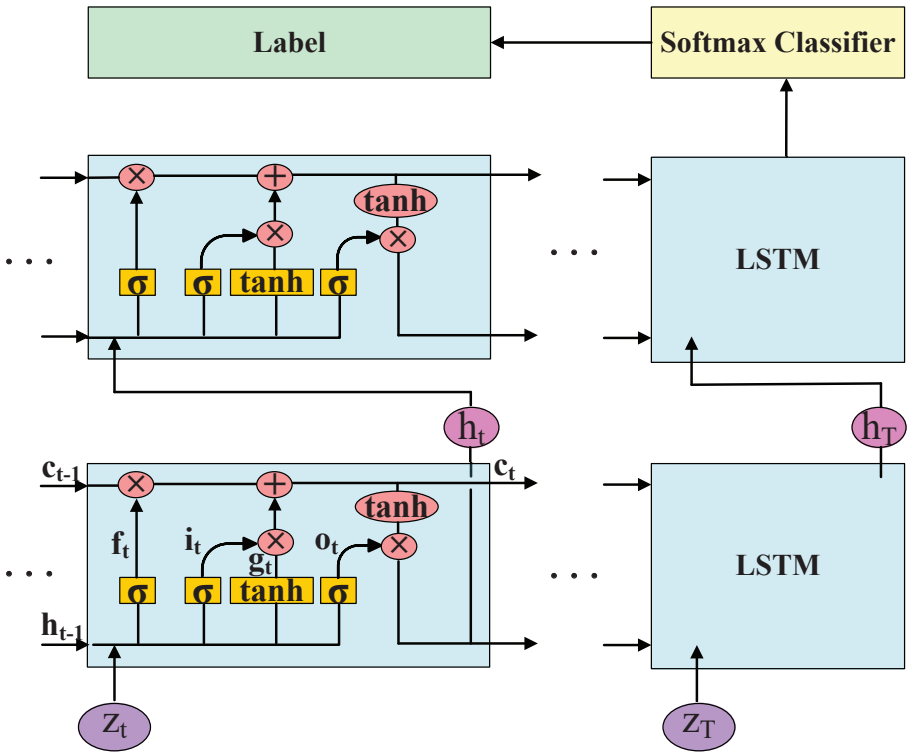

Figure 4.2: Structure of Deep LSTM network.

function has several advantages such as sparse representations, efficient gradient propagation and computation.

4.3.2 Deep LSTM Network

After feature extraction, we next use the deep LSTM algorithm for training optimal weights, where the LSTM network is a popular recurrent neural network (RNN) that can effectively deal with long-range dependency [29, 30]. It can solve the problems of exploding or vanishing gradients found in RNNs. Moreover, LSTM can exploit temporal information of magnetic field and light intensity data through recursively mapping the input sequence to output label by using the hidden LSTM units. Each LSTM unit has a built-in memory cell to store information over time using non-linear gate units, which can control the change of values and memory contents. For the proposed DeepML system, we stack two layers of the LSTM network to obtain a stronger learning and representation ability for magnetic and light sensor data, thus improving the localization performance. Figure. 4.2 clearly exhibits what the structure of Deep LSTM network looks like.

4.3.3 Softmax Classifier

The output of the final cell's hidden state in the second LSTM network is the input to a fully connected layer, which uses a basic neural network with one hidden layer to train the output data using the *Softmax classifier*. Moreover, the input data to the Softmax function is in the form of a N dimensional vector $q = [q_1, q_2, ..., q_N]$, where N is the number of training locations. The ith input data can be obtained as $q_i = h_f^T w_i$, where h_f is the output vector of the final cell's hidden state in the second LSTM network, and w_i is the weight vector of the fully connected layer. The Softmax function then maps the N dimensional vector to normalized data $p = [p_1, p_2, ..., p_N]$, that is

$$p_i = \frac{e^{q_i}}{\sum_{n=1}^{N} e^{q_n}} = \frac{e^{h_f^T w_i}}{\sum_{n=1}^{N} e^{h_f^T w_n}}, \quad \text{for } i = 1, 2, ..., N. \tag{4.2}$$

Let $J(\theta)$ be the *loss function* with the weight parameter θ. We adopt the *cross-entropy* to measure the difference between the true labeled data and the normalized output data, and use the L2 regularization hyperparameter to avoid over-fitting. To obtain the optimal weights, the training loss is minimized as

$$\min_{\theta} J(\theta) = -\sum_{i=1}^{N} y_i \log(p_i) + \frac{\lambda}{2} \|\theta\|_2^2, \tag{4.3}$$

where y_i denotes the true labeled data for the ith location, and λ is the L2 regularization hyperparameter. We then train the parameters in the deep LSTM using Backpropagation Through Time (BPTT) of LSTM. We also use the Adam Optimizer for improving the efficiency of optimization [68].

4.4 Online Location Estimation

For online location test, we first build M bimodal images with magnetic and light sensor data (as shown in the data preprocessing section), each of which has the same size as training images. Then, we leverage a probabilistic method for estimating the location of the target mobile device by feeding the M bimodal images to the trained deep LSTM network.

Let $\boldsymbol{\omega}$ denote the output results of the Softmax classifier using the deep LSTM network for N training locations with M newly measured bimodal images. We have

$$\boldsymbol{\omega} = \begin{bmatrix} \omega_{11} & \omega_{12} & \omega_{13} & \dots & \omega_{1M} \\ \omega_{21} & \omega_{22} & \omega_{23} & \dots & \omega_{2M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega_{N1} & \omega_{N2} & \omega_{N3} & \dots & \omega_{NM} \end{bmatrix}. \tag{4.4}$$

We then compute the average result for M output data at every location, thus reducing the variance of the output results. Let $\bar{\omega}_n$ be the mean of the output data vector $[\omega_{n1}, \omega_{n2}, ..., \omega_{nM}]$ in the nth row. The mean vector can be obtained as $\bar{\omega} = [\bar{\omega}_1, \bar{\omega}_2, ..., \bar{\omega}_N]$.

Finally, the position of the target mobile device is estimated as a weighted average of all the N training locations, as

$$\hat{L} = \sum_{i=1}^{N} l_n \times \bar{\omega}_n, \tag{4.5}$$

where l_n is the nth training location.

Chapter 5

Experimental Study

5.1 Experiment Setup

We prototype the DeepML system with a Samsung Galaxy S7 Edge smartphone on the Android 7.0 platform. Moreover, we implement an Android application with Android Studio 2.3.3 for data collection and preprocessing. We compare DeepML with a benchmark that uses magnetic field data only. To guarantee a fair comparison, we use for these two approaches the same magnetic field dataset and the same deep LSTM parameters to estimate the location of the mobile device. We experiment with the two methods in two different indoor scenarios.

5.1.1 Lab Scenario

Laboratory: This is a 6×12 m^2 computer laboratory in Broun Hall on the Auburn University campus. The lab is a cluttered environment with tables, chairs, and computers. The floor plan is shown in Fig. 5.1. We choose 12 training locations, which are marked as red squares. The distance between two neighboring training locations is 1.6 m. We collect 1800 rows of light intensity and magnetic field combined pattern for each training location, and 400 rows of data for each test location. Note that each test location is different from the known training locations.

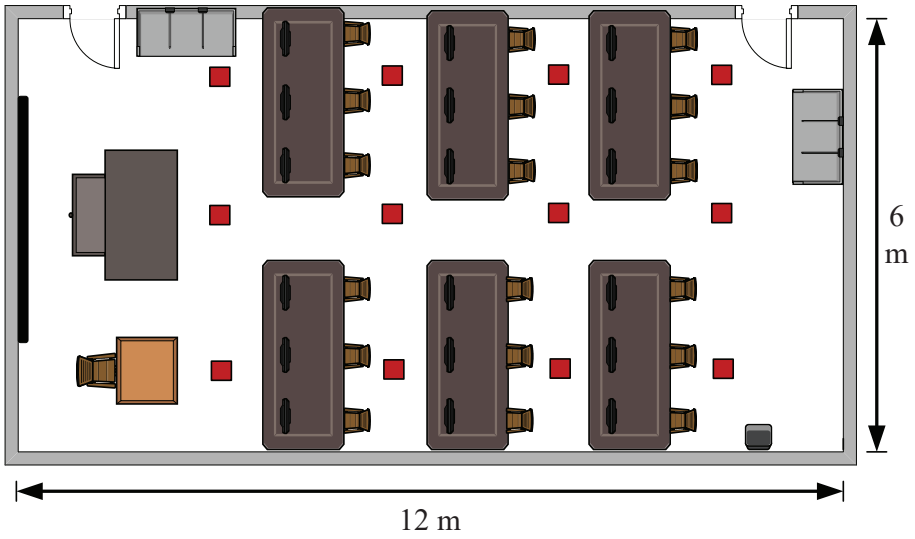

Figure 5.1: Layout of the computer lab scenario: training locations are marked as red squares.

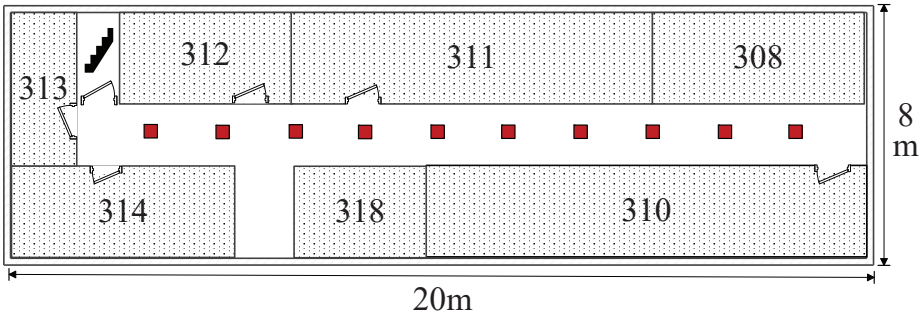

Figure 5.2: Layout of the corridor scenario: training locations are marked as red squares.

5.1.2 Corridor Scenario

Corridor: This is a 2.4 $\times$ 20 m^2 corridor in Broun Hall. As shown in Fig. 5.2, we employ 10 training locations along a straight line with 1.6 m separation. The training data size is 1500 rows and the testing data size is 400 rows.

For online test, we leverage an LSTM deep leaning model using Tensorflow on a computer with CPU 4720HQ and then integrate it with the data collection Android application, which achieves localization estimation in real-time.

5.2 Localization Performance

Figure. 5.3 illustrates the training loss over epoches of the DeepML for the laboratory and corridor scenarios. We set the epoch to 200 for preventing overfitting of training network with

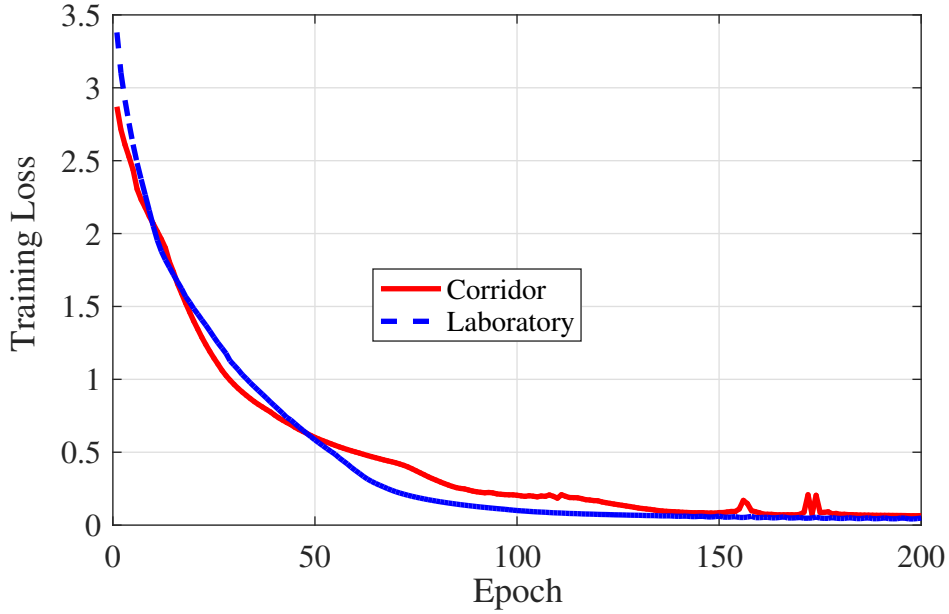

Figure 5.3: Training errors for the laboratory and corridor experiments.

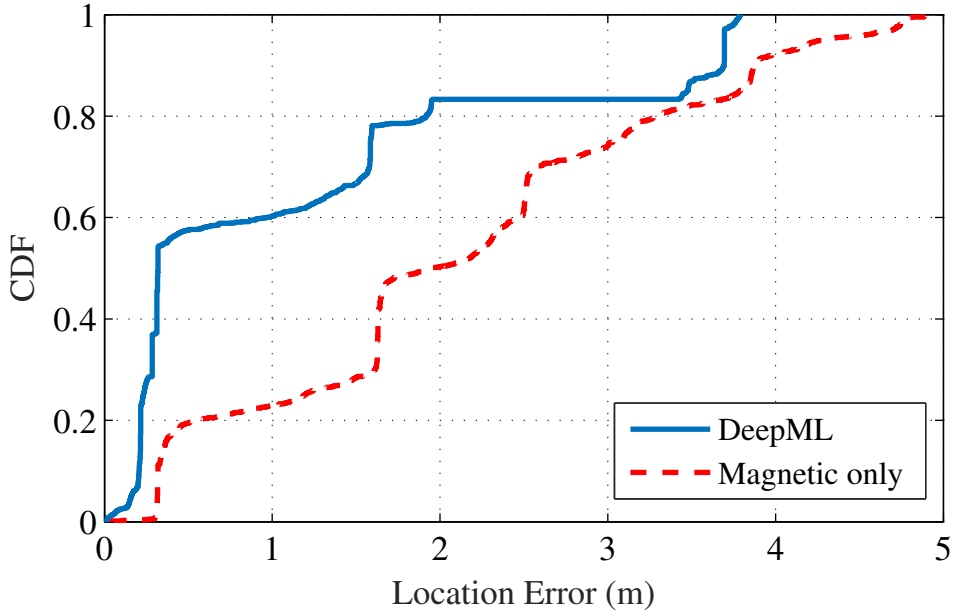

Figure 5.4: CDF of localization error of the lab experiment.

light intensity and magnetic field datasets. As depicted in Fig. 5.3, the train loss curve almost reaches about 0.05 for both the corridor and the lab scenarios.

Figure 5.4 plots the cumulative distribution function (CDF) of localization errors of the two schemes in the lab experiment. For this environment with complex light intensity and magnetic field distribution, DeepML is able to leverage bimodal magnetic-light features to predict location accurately. Figure. 5.4 shows that about 58% of the location errors with the

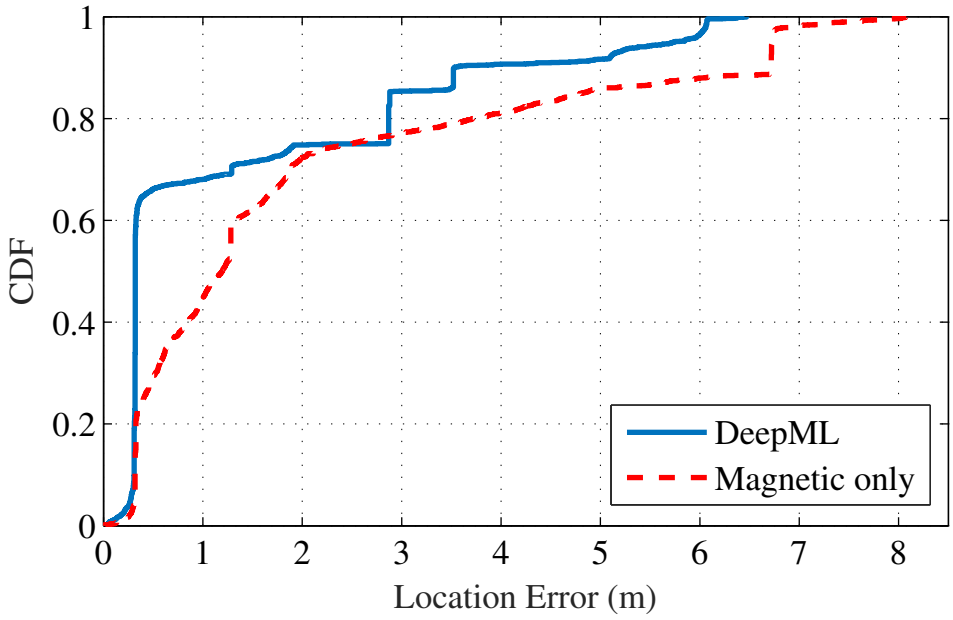

Figure 5.5: CDF of localization error of the corridor experiment.

proposed DeepML system are under 0.5 m, while 20% of the location errors with the magnetic field only scheme are under 0.5 m. Moreover, DeepML has 82% of the test locations with location errors less than or equal to 2 m, while it is 50% for the magnetic field only scheme. DeepML achieves a maximum error of 3.7 m, which is much better than the 5 m maximum error of the benchmark scheme. Apparently, the proposed DeepML system is more accurate for the cluttered lab environment.

Figure. 5.5 presents the CDF of localization errors of both schemes in the corridor scenario. There are about 65% of the test locations that have an estimation error less than or equal to 0.4 m for DeepML, while it is 25% for the magnetic only scheme. Additionally, we find DeepML has 87% of the test locations achieving an error under 3 m, comparing to 78% of the magnetic only scheme. Moreover, for the corridor scenario, the maximum location errors for DeepML and the magnetic only scheme are 6.5 m and 8.2 m, respectively. The proposed DeepML system is more robust than the baseline scheme.

The experiments validate that DeepML outperforms the benchmark scheme in both experiment scenarios. The main reason is that dual-module fingerprint has stronger location diversity, which carries more location features. In many cases magnetic field data and light intensity are complementary to each other. Furthermore, the proposed deep LSTM network can effectively

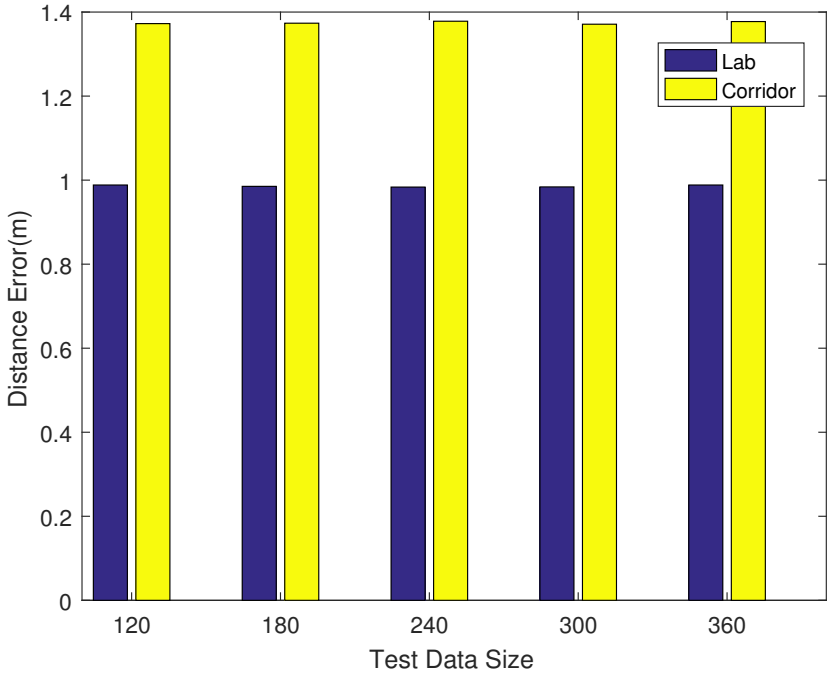

Figure 5.6: The average distance error for different size of test data.

extract the rich location features from the bimodal data, to achieve enhanced localization performance.

5.3 Impact of Different Parameters

5.3.1 Impact of Test Data Size

We test DeepML with test data sized to 120, 180, 240, 300 and 360 to determine how test data size impacts the accuracy of the indoor localization. Considering variable control, all other parameters are set same for every training position. Epoch, window size, hidden units and batch size are set to 200,10,40 and 1500, respectively. As is shown in Fig 5.6, distance errors of both scenarios are almost same for different test data size. The errors in the lobby and the lab are about 0.98m meters and 1.36 meters, respectively. This result indicates that the localization performance of DeepML is robust enough to the test data size. DeepML could achieve localization with high precision. Due to the testing time of each position is less than 0.1 second, we do not discuss the effect of test data size on time consumption.

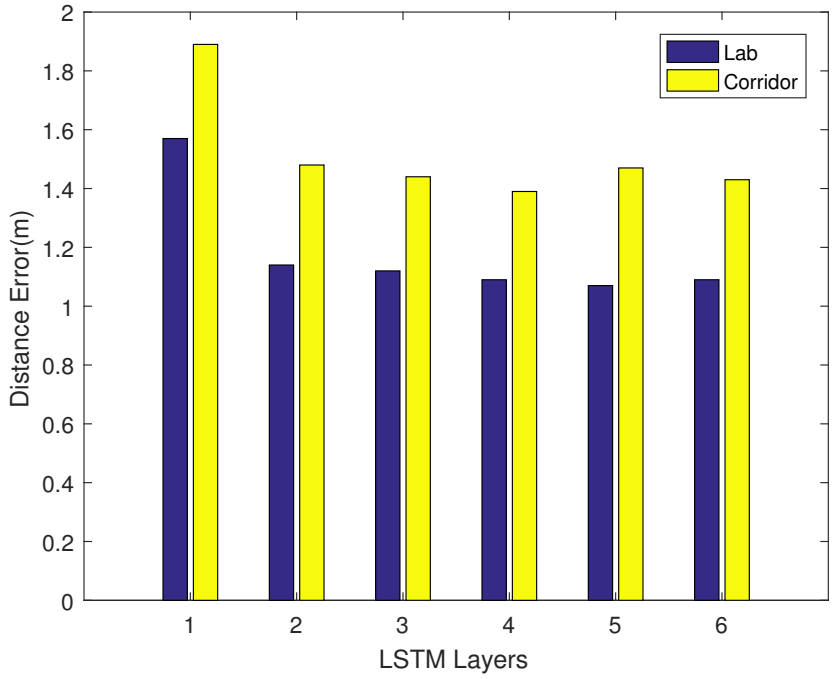

Figure 5.7: The average distance error for different number of LSTM layers.

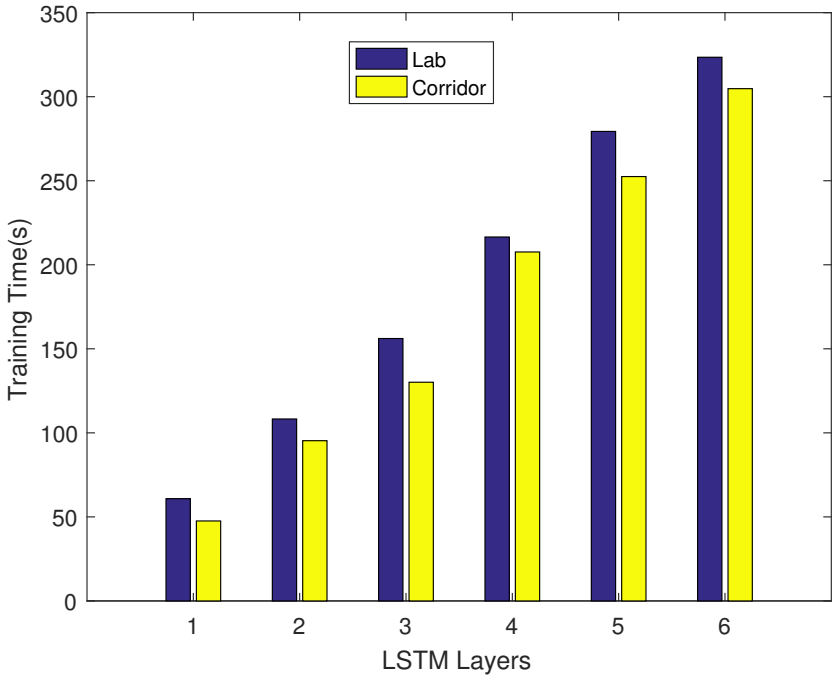

Figure 5.8: The average training time for different number of LSTM layers.

5.3.2 Impact of the Number of LSTM Layers

To evaluate how the number of LSTM layers affects the distance error, we build 5 datasets with different number of LSTM layers in every position. As shown in Fig 5.7, the distance error

has a downward trend. The lowest distance errors, 1.07 meters for the lab and 1.39 meters for the corridor, are obtained when the number of LSTM layers is 5 and 4 respectively. This result indicates that the number of LSTM layers has relation with the localization accuracy. Furthermore, we notice that the distance error in the corridor is more sensitive to the number of LSTM layers. When the number of LSTM layer exceeds 1, We can see that all distance errors for the lab are smaller than 1.18 meter and distance errors for the lab are lower than 1.48 meter. The result demonstrates the robustness of different numbers of LSTM layers. In addition, it is necessary to use deep LSTM networks with multiple layers because of the bad performance of one layer LSTM network.

Figure 5.8 presents the training time across all datasets with different number of LSTM layers. It directly shows that the number of LSTM layers has great impact on network training time. We can tell that the training time for corridor is much longer than the training time for the lab with same number of LSTM layers. Considering the localization performance which are discussed above, we still select the dataset with 4 LSTM layers pictures in every training point as the input of DeepML even though it has long training time. Another reason is that we also take into account the possibility of overfitting. It is well known that higher number of neural network layers usually leads to more over-fitting possiblllity. We can notice that the slight increasing trend of mean distance error which may caused by overfitting. In order to maintain the robustness of our system, setting the number of LSTM layers to 4 is a better and secure choice.

5.3.3 Impact of Hidden Units

The number of hidden units represents the dimensionality of the 'hidden state' of the lstm. In order to explore the effect on localization accuracy, we also implement our DeepML model with five different number of hidden units. The environments where we did comparison experiments are same with other sections, a long corridor and a teaching laboratory. We can find that the mean distance error decreases with the increase with the number of hidden units from Fig 5.9. Obviously, The best localization performance occurred at the 40 hidden units for both lab and corridor scenarios, 1.02 meters for lab and 1.33 meters for the corridor. Correspondingly,

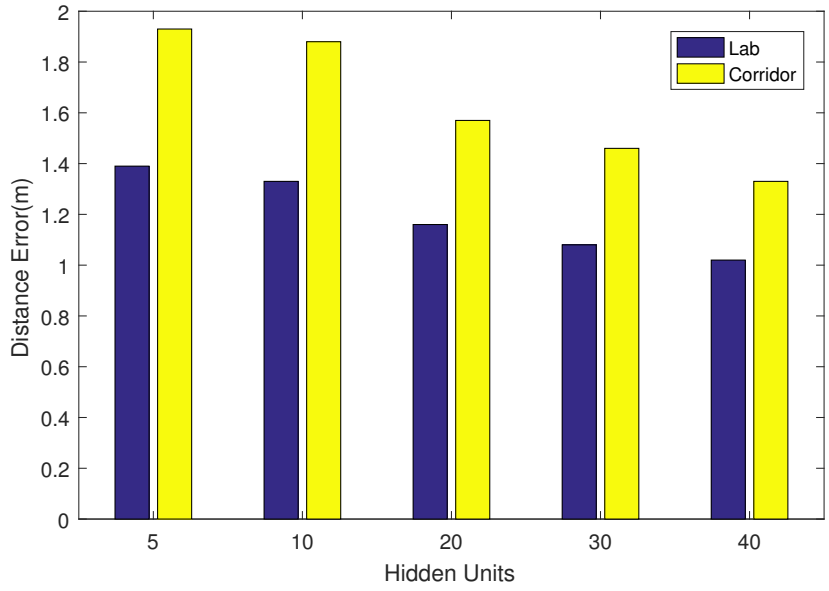

Figure 5.9: The average distance error for different hidden units.

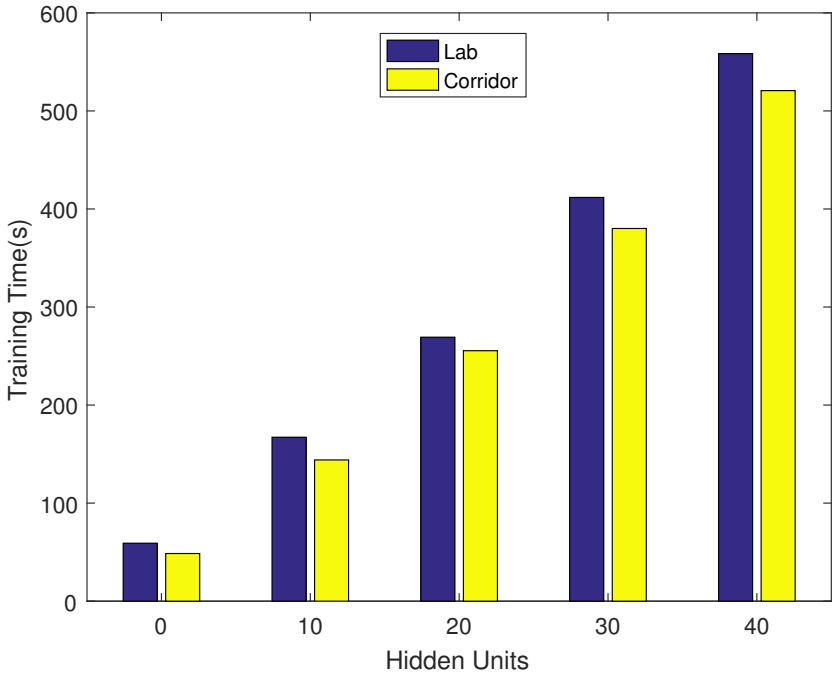

Figure 5.10: The average training time for different number of hidden units.

the greatest mean distance error, 1.39 meters for lab and 1.93 for the corridor, are accessed with 5 hidden units. Generally, the results of the corridor are worse than the indoors and the accuracy gap between two scenarios are 0.5 meter roughly. The reason is probably that the lab

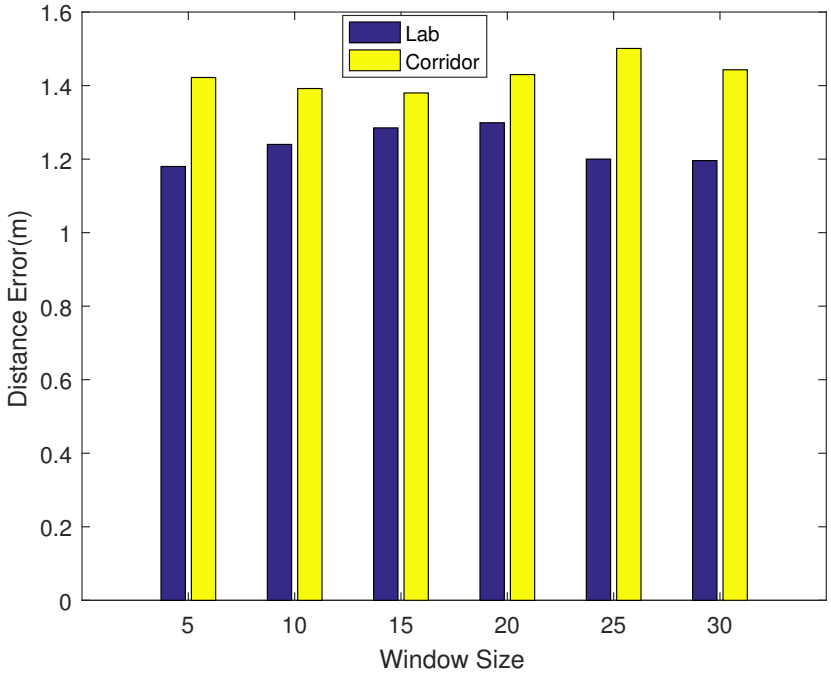

Figure 5.11: The average distance error for different window size.

environment is more complicated so that the light intensity and magnetic field data have more diversity than the corridor.

As we can see in Fig. 5.10, the training time is positively correlated with the number of hidden units. Although the longest training time has reached 520.72 seconds, we choose the dataset with 40 hidden units for training because of the best positioning accuracy.

5.3.4 Impact of Window Size

To further investigate the parameters impact on localization performance, we also deploy six datasets with different numbers of window size. The definition of window size is the length of a sliding window of a time sequence of data. We can understand it as the sentences length when you are reading a book. Figure 5.11 shows similar results under different window size. Although there are fluctuation of mean distance error with six different window size, the difference between the best and worst localization accuracy are only 0.108 meter for the corridor and 0.103 meter for the lab. Considering the computer laboratory and corridor are 6×12 m^2 and 2.4×20 m^2 respectively, 0.1 meter-level different is acceptable. We can still tell that the window size does not disturb the localization performance of DeepML. Figure. 5.12 exhibits

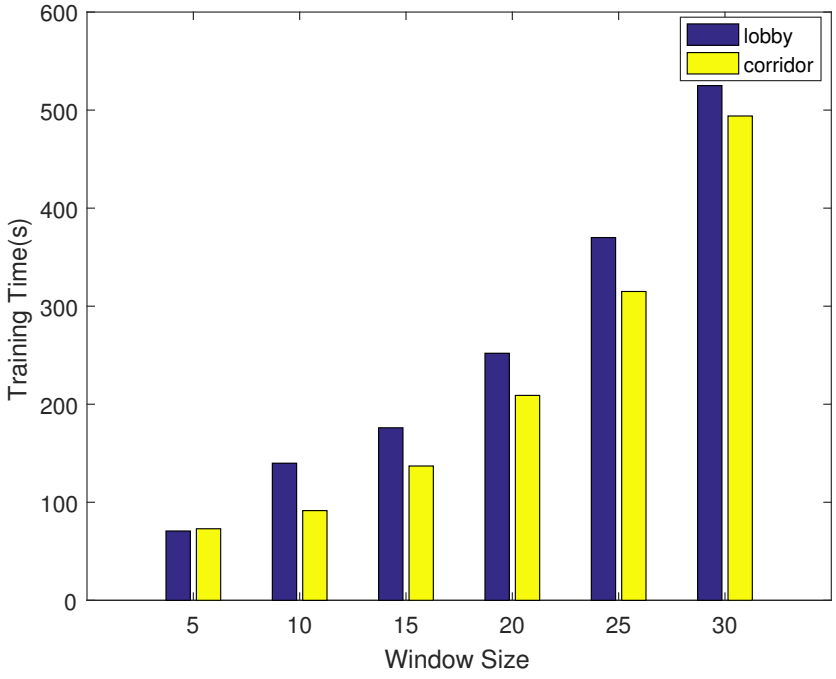

Figure 5.12: The average training time for different window size.

the relations between network training time and window size. Actually, the training time does not decrease user experience much at practical test because the training process belongs to on-line stage. Therefore, we usually consider the positioning accuracy firstly. The dataset with window size of 5 are adopted for each training point finally by reason that the lowest mean distance error of localization are obtained at both scenarios.

5.3.5 Impact of Batch Size

Batch size denotes the subset size of the training sample (e.g. 100 out of 1000) which is going to be used in order to train the network during its learning process. We explore the impact of batch size on localization accuracy under two environments. Figure 5.13 illustrates the mean distance errors for increasing batch size in the lab and lobby scenarios. As we can see, the mean distance error increases with the growth of batch size when larger than 1500. This result indicates that the batch size is related to localization accuracy. The greatest localization performance, 1.08 meters for the lobby and 1.46 meters for the corridor, is achieved when batch size of 1500.

Figure 5.14 presents the training time for different batch size. According to the Fig 5.14, we know the longest training times are 124.276 seconds and 128.64 seconds in the lab and the

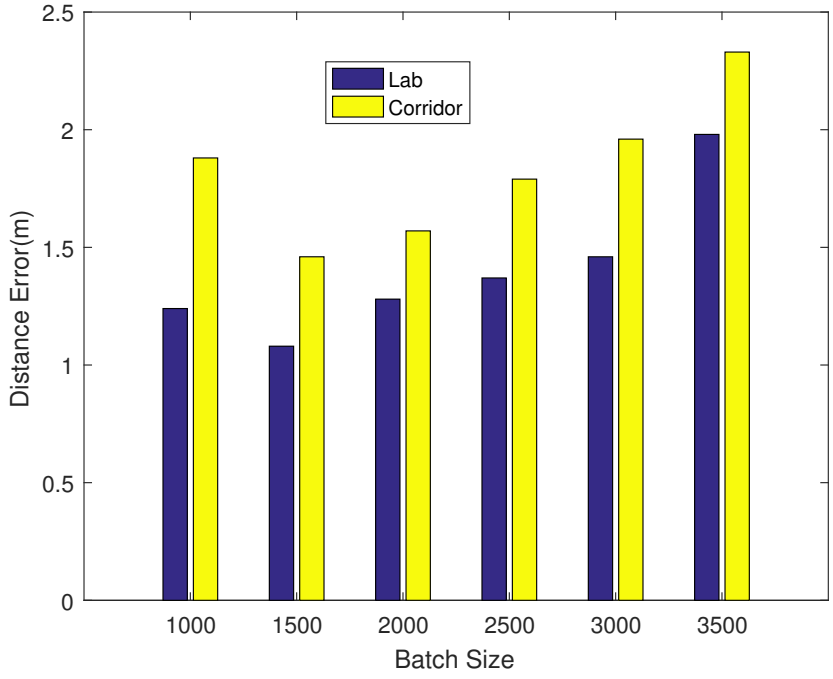

Figure 5.13: The average distance error for different batch size.

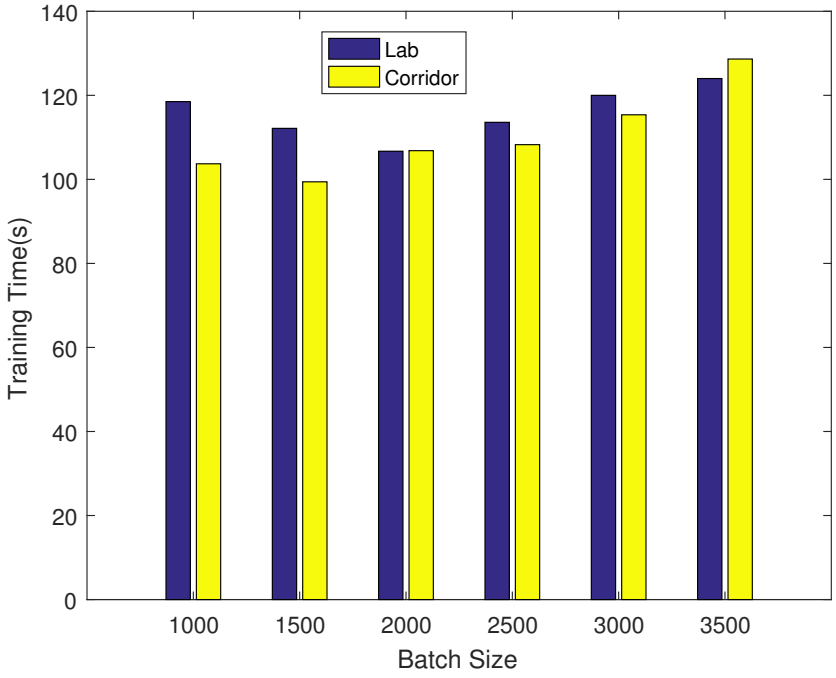

Figure 5.14: The average training time for different batch size.

corridor, and the shortest training times are 106.70 seconds and 99.42 seconds in the lab and the corridor, which means the batch size only has small influence on the network's training time.

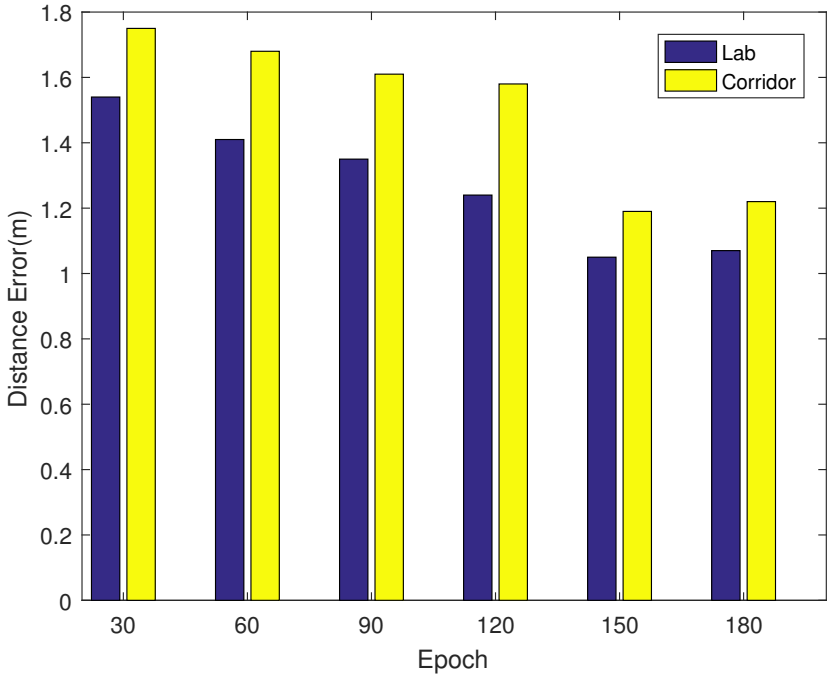

Figure 5.15: The average distance error for different epoch.

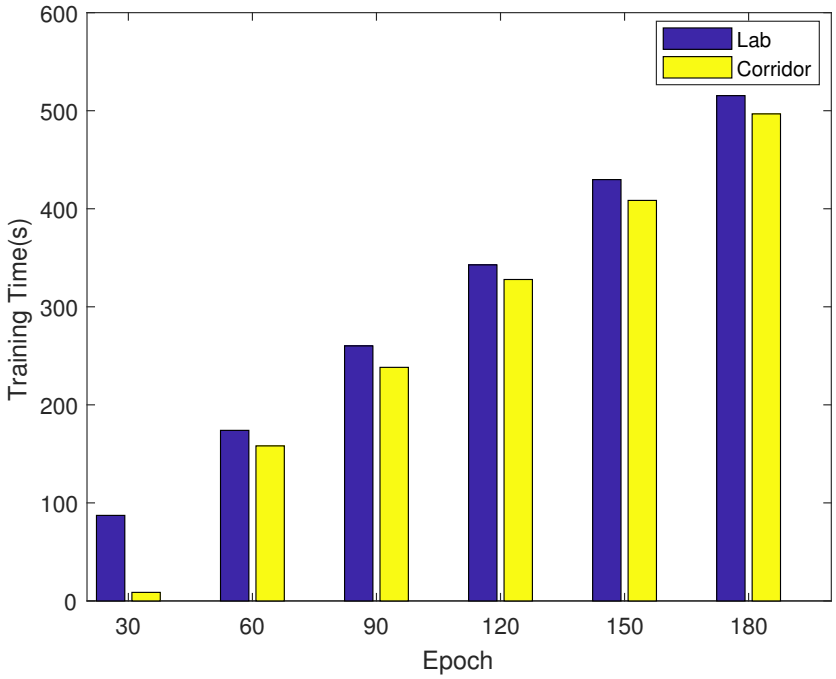

Figure 5.16: The average training time for different epoch.

5.3.6 Impact of Batch Epoch

In training neural network, one epoch means one pass of the full training set. Specifically, a epoch contains one forward pass and one backward pass. We adjusted the value of epoch

and prepared six datasets with different numbers of echo to improve our positioning system accuracy. Figure 5.15 shows the impact of epoch on the mean distance error of positioning accuracy. Obviously, The highest distance error is achieved when the value of epoch is 30 in both lab and corridor indoor environment. Moreover, the mean distance error keeps decreasing along with the growth of the value of epoch. And we can see the error verges to be steady when the value of epoch is over 150. The mean distance error can remain at a relatively stable level wait until the network converges. We notice that the lowest error is obtained at the 150 epochs from the Fig 5.15

Figure 5.16 illustrate the relations between the training time and the value of epoch. As is shown, the training time increases as the value of the epoch increases in both scenarios. The result is reasonable that more epoch loops leads to more time consumed. To reach the lowest distance error, we spend about 429.73 seconds and 408.51 seconds to train the network in the lab and the corridor respectively.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

In this thesis, we presented DeepML, a deep LSTM based system for indoor localization using the magnetic and light sensors in smartphones. We first experimentally verified the feasibility of using the magnetic-light bimodal data from magnetic and light sensors for indoor localization. We then presented the DeepML design, with its data preprocessing, deep LSTM network, and probabilistic location estimation modules, where data preprocessing was implemented to obtain bimodal image data, and two layers LSTM network was then trained. For online test phase, the newly received magnetic field and light data can be used for estimating the location of the mobile device. Our experiments under two representative indoor environments demonstrated the effectiveness of the proposed DeepML system.

6.2 Future Work

Even though DeepML is a high accuracy indoor localization using commercial off-the-shelf smartphone and existing ceiling-implemented lights, there are still many interesting problems that deserve further study. In the future, we plan to enhance the system performance along the following three directions.

We conducted several experiments in the corridor and lab with several windows at noon to evaluate the performance of DeepML. However, the result shows DeepML is sensitive to the effect from the sun light. The sunlight disturbed the light intensity value collected by light

sensors of the smartphone which leads to mismatching in the online test. Therefore, an adequate sunlight elimination and neutralization algorithm is indispensable in our future study.

The DeepML system manifests the feasibilty of using the bimodal dataset. For more accurate indoor positioning, we plan to add in WiFi signals to make trimodal data for network training. In consideration of good performance in other indoor localization systems, it is possible that our system will achieve better positioning accuracy. Moreover, collecting WiFi signals using smartphones is really convenient and fast.

We spent a lot of time on collecting enormous fingerprint data for training the LSTMs network in both lab and corridor scenarios. We can easily predict this would be a much more difficult and time-consuming task in a larger environment, such as shopping malls and airports. We conjecture that generative adversarial nets (GAN) could be a suitable deep learning method for generating datasets and fingerprints to reduce the data collection effort. We will investigate the application of GAN in our future work.

References

[1] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 13–32, Jan. 2009.

[2] X. Wang, X. Wang, and S. Mao, "Rf sensing in the internet of things: A general deep learning framework," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 62–67, 2018.

[3] X. Wang, Z. Yu, and S. Mao, "DeepML: Deep LSTM for indoor localization with smartphone magnetic and light sensors," in *Proc. IEEE ICC 2017*, Kansas City, MO, May 2018.

[4] M. Gong, B. Hart, and S. Mao, "Advanced wireless LAN technologies: IEEE 802.11ac and beyond," *ACM Mobile Computing and Communications Review (MC2R)*, vol. 18, no. 4, pp. 48–52, Oct. 2014.

[5] P. Bahl and V. N. Padmanabhan, "Radar: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM'00*, Tel Aviv, Israel, Mar. 2000, pp. 775–784.

[6] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," in *Proc. ACM MobiSys'05*, Seattle, WA, June 2005, pp. 205–218.

[7] X. Wang, X. Wang, S. Mao, J. Zhang, S. Periaswamy, and J. Patton, "DeepMap: Deep Gaussian Process for indoor radio map construction and location estimation," in *Proc. IEEE GLOBECOM 2018*, Abu Dhabi, United Arab Emirates, Dec. 2018.

[8] J. Xiao, K. Wu., Y. Yi, and L. Ni, "FIFS: Fine-grained indoor fingerprinting system," in *Proc. IEEE ICCCN'12*, Munich, Germany, Aug. 2012, pp. 1–7.

[9] X. Wang, L. Gao, and S. Mao, "CSI phase fingerprinting for indoor localization with a deep learning approach," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1113–1123, Dec. 2016.

[10] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Trans Veh. Technol.*, vol. 66, no. 1, pp. 763–776, Jan. 2017.

[11] X. Wang, L. Gao, and S. Mao, "BiLoc: Bi-modality deep learning for indoor localization with 5GHz commodity Wi-Fi," *IEEE Access Journal*, vol. 5, no. 1, pp. 4209–4220, Mar. 2017.

[12] X. Wang, X. Wang, and S. Mao, "CiFi: Deep convolutional neural networks for indoor localization with 5GHz Wi-Fi," in *Proc. IEEE ICC 2017*, Paris, France, May 2017, pp. 1–6.

[13] X. Wang, L. Gao, S. Mao, and S. Pandey, "DeepFi: Deep learning for indoor fingerprinting using channel state information," in *Proc. WCNC'15*, New Orleans, LA, Mar. 2015, pp. 1666–1671.

[14] X. Wang, X. Wang, and S. Mao, "ResLoc: Deep residual sharing learning for indoor localization with CSI tensors," in *Proc. IEEE PIMRC 2017*, Montreal, Canada, Oct. 2017.

[15] X. Wang, C. Yang, and S. Mao, "Phasebeat: Exploiting csi phase data for vital sign monitoring with commodity wifi devices," in *Distributed Computing Systems (ICDCS), 2017 IEEE 37th International Conference on*. IEEE, 2017, pp. 1230–1239.

[16] ——, "Resbeat: Resilient breathing beats monitoring with realtime bimodal csi data," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[17] ——, "Tensorbeat: Tensor decomposition for monitoring multi-person breathing beats with commodity WiFi," *ACM Transactions on Intelligent Systems and Technology*, vol. 9, no. 1, pp. 8:1–8:27, Sept. 2017.

[18] X. Wang, X. Wang, and S. Mao, "Deep convolutional neural networks for indoor localization with csi images," *IEEE Transactions on Network Science and Engineering*, 2018.

[19] J. Chung, M. Donahoe, C. Schmandt, I.-J. Kim, P. Razavai, and M. Wiseman, "Indoor location sensing using geo-magnetism," in *Proc. ACM MobiSys'11*, Bethesda, MD, June/July 2011, pp. 141–154.

[20] W. Storms, J. Shockley, and J. Raquet, "Magnetic field navigation in an indoor environment," in *Proc. IEEE UPINLBS'10*. Kirkkonummi, Finland: IEEE, Oct. 2010, pp. 1–10.

[21] B. Gozick, K. P. Subbu, R. Dantu, and T. Maeshiro, "Magnetic maps for indoor navigation," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 12, pp. 3883–3891, Dec. 2011.

[22] Y. Shu, C. Bo, G. Shen, C. Zhao, L. Li, and F. Zhao, "Magicol: Indoor localization using pervasive magnetic field and opportunistic WiFi sensing," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1443–1457, July 2015.

[23] Y. Ma, Z. Dou, Q. Jiang, and Z. Hou, "Basmag: An optimized HMM-based localization system using backward sequences matching algorithm exploiting geomagnetic information," *IEEE Sensors J.*, vol. 16, no. 20, pp. 7472–7482, Oct. 2016.

[24] Z. Yang, Z. Wang, J. Zhang, C. Huang, and Q. Zhang, "Wearables can afford: Lightweight indoor positioning with visible light," in *Proc. ACM. MobiSys'15*, Florence, Italy, May 2015, pp. 317–330.

[25] Y.-S. Kuo, P. Pannuto, K.-J. Hsiao, and P. Dutta, "Luxapose: Indoor positioning with mobile phones and visible light," in *Proc. ACM MobiCom'14*, Maui, HI, Sept. 2014, pp. 447–458.

[26] C. Zhang and X. Zhang, "LiTell: Robust indoor localization using unmodified light fixtures," in *Proc. ACM MobiCom'16*, New York City, NY, Oct. 2016, pp. 230–242.

[27] Q. Xu, R. Zheng, and S. Hranilovic, "IDyLL: Indoor localization using inertial and light sensors on smartphones," in *Proc. ACM UbiComp'15*. Osaka, Japan: ACM, Sept. 2015, pp. 307–318.

[28] Z. Zhao, J. Wang, X. Zhao, C. Peng, Q. Guo, and B. Wu, "NaviLight: Indoor localization and navigation under arbitrary lights," in *Proc. IEEE INFOCOM'17*. Atlanta, GA: IEEE, May 2017, pp. 1–9.

[29] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," *J. Neural Comput.*, vol. 12, no. 10, pp. 2451–2471, Oct. 2000.

[30] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "Lstm: A search space odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, 2017.

[31] A. Graves, N. Jaitly, and A.-r. Mohamed, "Hybrid speech recognition with deep bidirectional LSTM," in *Proc. IEEE ASRU'13*, Olomouc, Czech Republic, Dec. 2013, pp. 273–278.

[32] F. J. Ordonez and D. Roggen, "Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition," *MDPI Sensors*, vol. 16, no. 1, p. 115, Jan. 2016.

[33] M. Chen, V. C. M. Leung, S. Mao, and T. Kwon, "Receiver-oriented load-balancing and reliable routing in wireless sensor networks," *Wiley Wireless Communications and Mobile Computing Journal*, vol. 9, no. 3, pp. 405–416, Mar. 2009.

[34] M. Chen, V. C. M. Leung, S. Mao, and M. Li, "Cross-layer and priority path scheduling for real-time video communications over wireless sensor networks," in *Proc. IEEE VTC 2008-Spring*, Marina Bay, Singapore, May 2008, pp. 2873–2877.

[35] M. Feng, T. Jiang, D. Chen, and S. Mao, "Cooperative small cell networks: High capacity for hotspots with interference mitigation," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 108–116, Dec. 2014.

[36] M. Feng, S. Mao, and T. Jiang, "Joint duplex mode selection, channel allocation, and power control for full-duplex cognitive femtocell networks," *Elsevier Digital Communications and Networks Journal*, vol. 1, no. 1, pp. 30–44, Feb. 2015.

[37] D. Hu and S. Mao, "On medium grain scalable video streaming over femtocell cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 641–651, Apr. 2012.

[38] D. Hu, S. Mao, and J. H. Reed, "On video multicast in cognitive radio networks," in *Proc. IEEE INFOCOM 2009*, Rio de Janeiro, Brazil, Apr. 2009, pp. 2222–2230.

[39] X. Wang, R. Huang, and S. Mao, "Sonarbeat: Sonar phase for breathing beat monitoring with smartphones," in *Computer Communication and Networks (ICCCN), 2017 26th International Conference on*. IEEE, 2017, pp. 1–8.

[40] Y. Wang, S. Mao, and R. Nelms, "Distributed online algorithm for optimal real-time energy distribution in the smart grid," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 70–80, Feb. 2014.

[41] Y. Xu and S. Mao, "User association in massive MIMO HetNets," *IEEE Systems Journal*, vol. 11, no. 1, pp. 7–19, Mar. 2017.

[42] M. X. Gong, S. F. Midkiff, and S. Mao, "A cross-layer approach to channel assignment in wireless ad hoc networks," *ACM/Springer Mobile Networks and Applications Journal*, vol. 12, no. 1, pp. 43–56, Feb. 2007.

[43] M. Chen, V. C. M. Leung, S. Mao, and M. Li, "Energy-efficient itinerary planning for mobile agents in wireless sensor networks," in *Proc. IEEE ICC 2009*, Dresden, Germany, June 2009, pp. 1–5.

[44] Y. Wang, S. Mao, and R. Nelms, "On hierarchical power scheduling for the macrogrid and cooperative mi-crogrids," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1574–1584, Dec. 2015.

[45] D. Hu and S. Mao, "Streaming scalable videos over multi-hop cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3501–3511, Nov. 2010.

[46] M. Chen, T. Kwon, S. Mao, Y. Yuan, and V. C. M. Leung, "Reliable and energy-efficient routing protocol in dense wireless sensor networks," *International Journal of Sensor Networks*, vol. 4, no. 1/2, pp. 104–117, July 2008.

[47] M. Chen, Y. Qian, S. Mao, W. Tang, and X. Yang, "Software-defined mobile networks security," *ACM/Springer Mobile Networks and Applications (MONET) Journal*, vol. 21, no. 5, pp. 729–743, Oct. 2016.

[48] M. X. Gong, R. J. Stacey, D. Akhmetov, and S. Mao, "A directional CSMA/CA protocol for mmWave wireless PANs," in *Proc. IEEE WCNC 2010*, Sydney, Australia, Apr. 2010, pp. 1–6.

[49] S. Mao and S. S. Panwar, "A survey of envelope processes and their applications in quality of service provisioning," *IEEE Communications Surveys & Tutorials*, vol. 8, no. 3, pp. 2–20, 2006.

[50] S. Mao, S. S. Panwar, and Y. T. Hou, "On minimizing end-to-end delay with optimal traffic partitioning," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 2, pp. 681–690, Mar. 2006.

[51] M. Gong, S. F. Midkiff, and S. Mao, "Design principles for distributed channel assignment in wireless ad hoc networks," in *Proc. IEEE ICC 2005*, Seoul, Korea, May 2005, pp. 3401–3406.

[52] I. K. Son and S. Mao, "Design and optimization of a tiered wireless access network," in *Proc. IEEE INFOCOM 2010*, San Diego, CA, Mar. 2010, pp. 1–9.

[53] W. Yang, X. Wang, S. Cao, H. Wang, and S. Mao, "Multi-class wheat moisture detection with 5GHz Wi-Fi: A deep LSTM approach," in *Proc. ICCCN 2018*, Hangzhou, China, July/Aug. 2018.

[54] M. Chen, Y. Zhang, Y. Li, S. Mao, and V. Leung, "EMC: Emotion-aware mobile cloud computing in 5G," *IEEE Network*, vol. 29, no. 2, pp. 32–38, Mar./Apr. 2015.

[55] X. Wang, S. Mao, and M. X. Gong, "An overview of 3gpp cellular vehicle-to-everything standards," *GetMobile: Mobile Computing and Communications*, vol. 21, no. 3, pp. 19–25, 2017.

[56] I.-K. Son, S. Mao, Y. Li, M. Chen, M. Gong, and T. Rappaport, "Frame-based medium access control for 5G wireless networks," *Springer MONET Journal*, vol. 20, no. 6, pp. 763–772, Dec. 2015.

[57] M. Chen, J. Yang, Y. Hao, S. Mao, and K. Hwang, "A 5G cognitive system for healthcare," *MDPI Big Data and Cognitive Computing Journal*, vol. 1, no. 1, pp. 2–16, Mar. 2017.

[58] M. Feng, S. Mao, and T. Jiang, "Base station ON-OFF switching in 5G wireless systems: Approaches and challenges," *IEEE Wireless Communications*, vol. 24, no. 4, pp. 46–54, Aug. 2017.

[59] Z. Chang, L. Lei, Z. Zhou, S. Mao, and T. Ristaniemi, "Learn to cache: Machine learning for network edge caching in the big data era," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 28–35, June 2018.

[60] Y. Wang, L. Tao, X. Huang, J. Shi, and N. Chi, "8-gb/s rgby led-based wdm vlc system employing high-order cap modulation and hybrid post equalizer," *IEEE Photonics Journal*, vol. 7, no. 6, pp. 1–7, 2015.

[61] A. Ganick and D. Ryan, "Light positioning system using digital pulse recognition," Aug. 21 2012, uS Patent 8,248,467.

[62] C. Yang, X. Wang, and S. Mao, "AutoTag: Recurrent vibrational autoencoder for unsupervised apnea detection with RFID tags," in *Proc. IEEE GLOBECOM 2018*, Abu Dhabi, United Arab Emirates, Dec. 2018.

[63] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," 2018.

[64] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[65] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[66] T.-H. Do and M. Yoo, "An in-depth survey of visible light communication based positioning systems," *MPDI Sensors*, vol. 16, no. 5, p. 678, May 2016.

[67] J. Randall, O. Amft, J. Bohn, and M. Burri, "LuxTrace: Indoor positioning using building illumination," *Springer Person. ubi. comput.*, vol. 11, no. 6, pp. 417–428, Aug. 2007.

[68] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014, [online] Available: https://arxiv.org/abs/1412.6980.