Learning-Based Load Identification and Forecasting using Low-Frequency Measurements

Amanie Azzam

A Thesis

In the Department of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of

Master of Applied Science (Electrical and Computer Engineering)

at Concordia University

Montréal, Québec, Canada

December 2023

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Amanie Azzam

Entitled: Learning-Based Load Identification and Forecasting using Low-Frequency Measurements

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Electrical and Computer Engineering)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

______________________________ Chair
Dr. Chunyan Lai

______________________________ Examiner
Dr. Wenfang Xie

______________________________ Supervisor
Dr. Amir G. Aghdam

Approved by ______________________________
Yousef R. Shayan, Chair
Department of Electrical and Computer Engineering

____________ 2024

____
Mourad Debbabi, Dean
Faculty of Engineering and Computer Science

Abstract

Learning-Based Load Identification and Forecasting using Low-Frequency Measurements

Amanie Azzam

Non-intrusive load monitoring (NILM) is a technique used for effective and cost-efficient electricity consumption management. This thesis presents two different NILM methods. One of them employs a hybrid model of convolutional neural network (CNN) and bidirectional long short-term memory (BiLSTM) for low-frequency data disaggregation, and the oher one utilizes a graph neural network (GNN) along with a long short-term memory (LSTM) network for load prediction, both combined with attention mechanism. The first study is adept at extracting temporal and spatial features from low-frequency power data, enhanced by an attention mechanism for event detection and load disaggregation. We conduct simulations using the publicly available low-frequency REDD dataset to assess our model's performance. The proposed approach exhibits superior accuracy and computational efficiency compared to existing methods. The second study explores NILM load prediction, integrating a GNN to represent complex time correlations between appliances, forming a graph-based foundation for feature extraction. The outcome is coupled with LSTM for temporal pattern capturing and attention processes for focusing on key information. The results confirm the effectiveness of this approach in predicting load and uncovering hidden power consumption patterns. Both studies contribute significantly to the field of NILM, offering advanced methodologies for energy management in smart homes.

Acknowledgments

I would like to express my deepest gratitude to my advisor, Dr. Amir G. Aghdam, for his invaluable guidance, patience, and support throughout the course of this research. His insights and expertise have been instrumental in shaping both the direction and success of my work.

My sincere thanks go to Saba Sanami for her significant contributions to this project. Her perceptive advice and continuous support were essential in bringing this thesis to completion.

I am also thankful to my family and my friends, who have provided me with endless love, encouragement, and understanding. Their contributions and support have been vital to the completion of this thesis.

Contribution of Authors

This dissertation is the research work of Amanie Azzam and it is submitted under the supervision of Dr. Amir G. Aghdam for the degree of Master of Applied Science at Concordia University. The dissertation is original.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Literature Review

Non-intrusive load monitoring (NILM) has emerged as a critical technology in the field of energy consumption analysis and management. It enables one to quantify individual home appliances' energy consumption without requiring additional instrumentation or sensor deployment. As society continues to grapple with the challenges of rising energy demands and environmental concerns, NILM stands as a promising solution to promote energy efficiency, reduce wastage, and enable informed decision-making for homeowners and utilities alike.

The development of NILM is timely and pertinent, as our society faces escalating challenges in managing increasing energy demands alongside pressing environmental concerns. The global push towards sustainability highlights the urgent need for efficient energy use, making technologies like NILM not only relevant but essential. NILM's ability to disaggregate total energy consumption into specific appliance usage offers a detailed insight into energy usage patterns. This level of granularity is crucial for identifying areas where energy efficiency can be improved, thus playing a vital role in reducing overall energy consumption.

The benefits of NILM extend to both consumers and energy providers. For consumers, NILM empowers homeowners with insights into their energy consumption habits. This awareness encourages energy-efficient behavior, resulting in reduced utility bills and a lower carbon footprint. It also provides real-time information about appliance operation, helping users detect malfunctioning or inefficient devices [1]. This, in turn, facilitates timely maintenance and replacement decisions. Integrating NILM with smart home systems allows for intelligent energy management [2], [3]. Appliances can be automatically controlled based on user preferences and real-time energy consumption data. For energy providers, NILM facilitates demand response programs by enabling utilities to communicate with consumers and control appliances during peak demand periods, thereby balancing the grid's load. Accurate appliance-level consumption data aids in load forecasting and peak demand management [4]. This enables utilities to allocate resources effectively and prevent grid overloads [5]. Detailed appliance-level data also helps utilities strategically plan infrastructure upgrades and expansions, improving grid reliability and efficiency. NILM allows both consumers and energy providers to identify which appliances consume the most energy, facilitating targeted energy-saving measures. For instance, replacing or adjusting the usage of high-energy-consuming appliances can substantially reduce electricity bills. Additionally, it enables the early detection of defects in appliances. By monitoring and analyzing energy consumption patterns, NILM systems can identify anomalies that may indicate a malfunctioning or inefficient appliance. Early detection of such issues not only prevents further energy wastage but also allows for timely maintenance or replacement, which can extend the lifetime of appliances and avert costly repairs. Moreover, integrating NILM with smart home systems enhances overall energy management, automatically controlling appliances based on real-time data to optimize energy usage, thus further driving down costs.

NILM was first introduced by Hart [6] during the early 1900s. The concept encompasses four distinct components: event detection, data processing, load decomposition, and load identification. Load identification is a fundamental task in NILM among these components. It consists of sophisticated algorithms which categorize loads using input features either provided by human input or extracted from electrical signal attributes. The early investigations into load identification effectiveness [7], [8], [9] often faced constraints arising from emerging hardware and software capacities.

Traditional NILM methods focused on statistical analysis and pattern recognition [6], [10]. Rule-based methods and statistical analyses constitute prevalent non-learning techniques, which rely on predefined heuristics and pattern recognition algorithms like hidden markov models (HMMs) [6], and clustering methods [11]. These techniques often struggle to capture the complicated variability and complexities of modern appliances' behaviors, particularly in cases of overlapping power signatures or subtle differences.

NILM can be classified into two primary groups: load identification and load forecasting. Each of these categories addresses specific aspects of energy monitoring and management, contributing uniquely to the overall effectiveness of NILM systems. Load Identification focuses on discerning the individual appliances contributing to the overall energy consumption within a household or building [33]. This process is foundational to NILM, as it allows for the detailed analysis of which appliances are using energy and when. The significance of load identification lies in its ability to provide granular data about energy usage patterns. By identifying which appliances are active and their consumption profiles, homeowners and energy managers can gain a clearer understanding of where and how energy is used. This insight is crucial for implementing targeted energy-saving strategies, such as replacing inefficient appliances or adjusting usage patterns. The methodologies employed in load identification range from simple signature recognition techniques, where appliances are identified based on unique electrical signatures, to more complex machine learning algorithms that can discern appliances in real time, even in environments with overlapping signals and variable usage patterns.

While load identification provides a snapshot of current or past appliance usage, load forecasting extends this analysis into the future. It involves predicting the future energy consumption of different appliances within a house or building. This predictive aspect of NILM is increasingly important in the context of dynamic energy markets and smart grid technologies. Effective load forecasting enables better energy planning and decision-making. For instance, it can inform homeowners when it is most cost-effective to run certain appliances based on predicted energy prices or guide energy providers in managing supply and demand more effectively. Advanced forecasting models leverage a variety of data inputs, including historical consumption patterns, weather data, and user behaviors, to make accurate predictions. The challenge in load forecasting lies in the variability of these factors and the need for models to adapt to changing conditions.

Both load identification and load forecasting are integral to the development of intelligent energy systems. Load identification lays the groundwork by providing detailed data on current usage patterns, while load forecasting builds on this foundation to anticipate future energy needs. Together, they form a comprehensive approach to energy monitoring and management, enabling more efficient and sustainable energy use. Moreover, the integration of these two aspects of NILM can lead to more sophisticated energy management solutions. For example, data derived from load identification can be used to train and refine load forecasting models, enhancing their accuracy and reliability. In a broader context, these capabilities are critical for the advancement of smart grid technologies, where understanding and predicting energy usage at a granular level is key to optimizing grid performance and integrating renewable energy sources.

The introduction of deep learning in the field of NILM marked a significant turning point, particularly in the realm of load disaggregation research. This shift, underscored in studies [11], [12], [13], represented a fundamental change in approach, addressing many of the limitations inherent in non-learning-based methodologies. Deep learning's ability to process and analyze large datasets, discern complex patterns, and adapt to new data types has made it particularly well-suited for the nuanced challenges presented by NILM. The transition to deep learning in NILM research was initially spearheaded by the work referenced in [15], which marked a notable departure from traditional methods. This pioneering study introduced a novel composite architecture that synergistically combined long short-term memory (LSTM) networks with denoising autoencoders (DAE). The LSTM networks, renowned for their efficiency in handling time-series data, were adept at capturing temporal dependencies in energy consumption patterns. This feature was crucial for NILM, where understanding usage patterns over time is essential. The integration of DAEs, on the other hand, addressed the issue of noise in the data – a common challenge in NILM due to the variability of appliance signatures and the presence of external disturbances. The combination of these two elements into a single architecture marked a significant advancement in the field, enabling more accurate and reliable load disaggregation. Further extending the capabilities of deep learning in NILM, Davies et al. presented an innovative approach to feature learning in [16]. This work tackled the problem of dealing with high-frequency electrical data, which, while rich in information, presents significant processing challenges. By downsampling this data into four distinct channels, the authors were able to simplify the data while retaining critical information. This pre-processing step was instrumental in enabling the subsequent application of a five-layer convolutional neural network (CNN). CNNs, with their ability to extract features from spatial data, proved highly effective in classifying the transformed data, thereby achieving a more effective and efficient load disaggregation process. The advancements in NILM were further augmented by the introduction of a fusion of sequence-to-point, spatial, and channel attention mechanisms, as detailed in [17]. This development led to the formulation of a convolutional block attention model. The significance of this model lies in its ability to focus on specific features of the input data, akin to the human attention mechanism. By selectively concentrating on certain aspects of the data while ignoring others, this model could learn the distinctive characteristics of target devices with greater accuracy and detail. This attention-based approach represented a step forward in the quest to enhance the interpretability and effectiveness of NILM systems.

The collective impact of these deep learning-based advancements on NILM has been profound. By enabling more precise and efficient analysis of energy consumption data [34], [35], these technologies have opened up new possibilities for energy management. They have paved the way for smarter, more responsive, and more sustainable energy usage in both residential and commercial settings. The ability to accurately disaggregate and forecast energy consumption at the appliance level has implications for energy conservation, operational efficiency, and the optimization of power grids. Furthermore, the integration of these deep learning models into practical applications promises to revolutionize the way energy consumption is monitored and managed. From smart homes that can optimize their energy use in real-time to utility providers better equipped to manage demand and supply, the potential applications of these advancements are vast. Additionally, load forecasting can facilitate demand response programs that can adjust the electricity prices and loads according to the supply and demand, to lower the peak demand and greenhouse gas emissions [34], [35]. Moreover, it can be used for predictive maintenance by identifying the faults and anomalies in the appliances, and hence, improving their performance and durability [34], [36]. It can also support smart home automation, tailoring the living environment to the occupants' activities and preferences, to enhance their comfort and convenience [34], [37]. As the field continues to evolve, further research is likely to yield even more sophisticated models and algorithms, driving the future of energy management towards greater sustainability and efficiency. As renewable energy sources become more prevalent, the ability to accurately predict and manage load becomes even more crucial. Load forecasting can help integrate these variable energy sources more effectively into the grid, thereby supporting the transition to a cleaner energy mix.

The main challenge in load forecasting lies in its complexity, as the electricity demand is influenced by a multitude of factors, including weather conditions, seasonal variations, time-of-day patterns, and user behavior. These elements introduce a level of unpredictability and variability that makes accurate forecasting a demanding task. Weather conditions, for instance, can significantly affect energy usage; cold or hot weather may lead to increased heating or cooling demands. Seasonal changes impact not only the temperature but also the length of daylight, affecting lighting and heating requirements. Time-of-day patterns reflect typical usage cycles, with peak demands often occurring in the mornings and evenings. Furthermore, individual user behavior, which can vary greatly among different households and change over time, adds another layer of complexity to forecasting.

To address these challenges, numerous advanced methods have been proposed in the literature, each aiming to improve the accuracy and reliability of load forecasting. A notable contribution in this field is the work [44], which presents a short-term load forecasting approach that integrates empirical mode decomposition (EMD), bi-directional long short-term memory (BiLSTM) networks, and an attention mechanism (EMD-BiLSTM-ATT). This approach is particularly innovative in how it combines EMD with advanced neural network architectures. EMD is effective in decomposing nonlinear and non-stationary time series data, making it easier for the LSTM network to capture the underlying patterns in energy usage. The bi-directional nature of LSTM allows the model to learn from both past and future data points, providing a more comprehensive understanding of energy consumption patterns. The addition of an attention mechanism further enhances the model's ability to focus on the most relevant features of the data, thereby improving forecasting accuracy.

This integrated approach has been shown to outperform traditional state-of-the-art methods in terms of prediction accuracy and adaptability. Another significant advancement is outlined in [43], which proposes a real-time prediction model for smart grids based on the combination of an attention mechanism with a bi-directional LSTM (CNN-ATT-BiLSTM).

The authors in [42] introduce an enhanced graph neural network (GNN) based on the bi-directional LSTM framework (EnGAT-BiLSTM). This model represents a significant leap forward in load forecasting, employing a dynamic load knowledge graph to represent both the load data and external influencing factors. The use of a knowledge graph allows the model to incorporate a wide range of data. The graph attention mechanism enables the model to effectively extract high-quality load spatiotemporal features, focusing on the most relevant parts of the graph to improve forecasting accuracy. The methodologies and findings presented in the papers [44], [43], and [42] collectively demonstrate the effectiveness of employing advanced neural network architectures, such as GNNs and LSTMs, combined with attention mechanisms, for load forecasting using low-frequency data. However, these studies do not fully address the time correlation of different appliances within a household, which could further enhance forecasting accuracy. The simultaneous operation of multiple appliances and their time correlation can have a significant impact on overall energy consumption patterns. By incorporating these time correlations between different appliances into the forecasting models, one can achieve higher levels of accuracy and reliability. Moreover, these models do not account for the uncertainty and variability in human behavior and preferences, which can change over time, influencing appliance usage patterns. Human behavior is often unpredictable and can vary due to a multitude of factors, such as changes in lifestyle, occupancy patterns, and personal preferences.

Both load identification and load forecasting stem from the ability to recognize unique power signatures of appliances, which are generally consistent across different households.

While the time correlation between appliance usage varies based on individual human behaviors and household-specific patterns, the fundamental power signatures of appliances remain largely invariant. This characteristic of NILM allows for effective identification and monitoring of individual appliances independent of the household. Each appliance (refrigerator, washing machine, etc.) has a distinctive electrical signature that NILM systems can detect and analyze. This capability ensures that NILM is not limited by the variability in household usage patterns. Instead, it can reliably disaggregate total energy consumption into specific appliance usage based on the power signatures. This universal applicability of power signatures across different environments underlines the method's efficiency, making NILM a robust and versatile tool for energy monitoring and management in diverse residential settings. Consequently, NILM provides valuable insights into energy consumption, ensuring its effectiveness and broad application.

Utilizing low-frequency data, as adopted in this thesis, presents distinct advantages over high-frequency data, making it a preferable choice for several applications. Low-frequency data significantly enhances computational efficiency. This offers a significant advantage, especially when dealing with large datasets or operating with limited computational resources. The reduced volume of data due to longer time intervals allows for more efficient processing and analysis, leading to quicker insights into energy consumption patterns. This efficiency is particularly beneficial in real-time or near-real-time monitoring systems, where speed is essential. The cost-effectiveness of low-frequency data processing cannot be overstated, as it requires lower storage capacity and processing power, making it an economically viable option for widespread deployment, especially in residential or small-scale commercial settings. This cost advantage extends to implementing and maintaining NILM systems, making energy monitoring more accessible and sustainable. Furthermore, while high-frequency data provides greater accuracy in identifying specific appliance signatures, its level of detail may not always be necessary for effective energy management. Low-frequency data, although less granular, still provides sufficient accuracy for most practical real-world applications. It enables the identification of major energy consumption patterns and recognizing key appliances, often the primary focus in energy efficiency and management strategies. Moreover, the simplicity of low-frequency data analysis makes it more adaptable and less prone to overfitting, a common challenge with more complex high-frequency data models. This adaptability is essential in diverse environments where energy usage patterns vary significantly.

1.2 Thesis Organization and Contributions

The rest of this thesis is organized as follows. Chapter 2 introduces CNN-BiLSTM model with an attention mechanism using low-frequency power data to develop a NILM algorithm. The proposed approach offers a cost-effective advantage, as low-frequency data collection requires less expensive measurement equipment compared to its high-frequency counterpart. The CNN-BiLSTM model combines two distinct neural network architectures, to leverage the strengths of each for enhanced NILM performance. The CNN part specializes in detecting spatial features and patterns, which is crucial for distinguishing different appliances based on their distinct energy consumption signatures. The BiLSTM component, on the other hand, excels at capturing temporal patterns and dependencies within sequential data, making it adept for load disaggregation tasks. The integration of an attention mechanism further refines the model's focus on relevant information, improving its ability to accurately identify appliance activation and deactivation. This work has been submitted to a Conference. Chapter 3 introduces the second proposed method called GNN-LSTM model combined with an attention mechanism to predict the power consumption of home appliances using low-frequency power data. The proposed model leverages the strengths of the LSTM network to capture temporal dependencies in electricity consumption while integrating a GNN to explore the time correlation between appliances. The attention mechanism allows the model to dynamically weigh the significance of various factors in the input sequence. By combining the three components mentioned above, the model captures short-term fluctuations and long-term trends in load data. This work has been submitted to the 18th annual IEEE International Systems Conference (SYSCON2024). Finally, Chapter 4 provides conclusions and possible future research directions.

Chapter 2

Low-Frequency Load Identification using CNN-BiLSTM Attention Mechanism

2.1 Problem Statement

Consider the energy consumption data for a household, available as an aggregate signal. It is desired to decompose the overall aggregated power waveform into its constituent appliance-level components as depicted in Figure 2.1. The main challenges for power disaggregation include: (i) handling appliances' power signature variations [18]; (ii) identifying appliances whose power signature overlap partially or completely [20], and (iii) achieving the desired accuracy. Previous research efforts have tackled these issues by utilizing high-frequency data, as they contain the complete signal and can be used to extract maximum information. However, the significant drawback lies in the cost of gathering such high-frequency data. It demands dedicated infrastructure, which incurs costs in terms of installing hardware to collect this type of data. Conversely, low-frequency features come

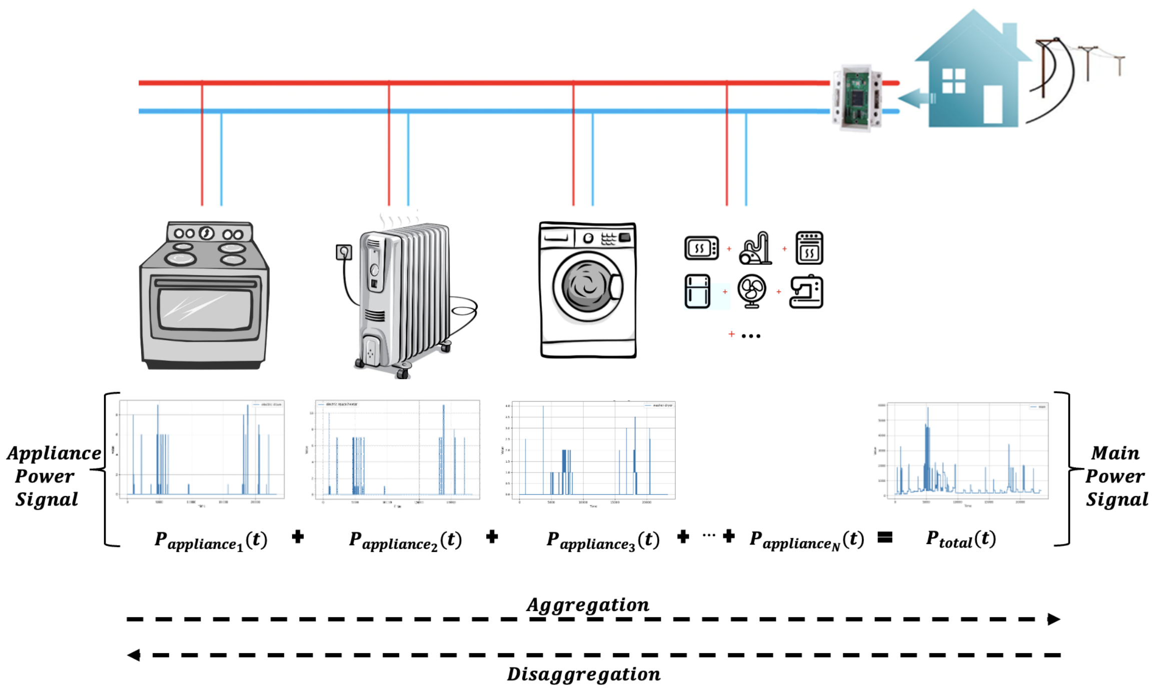

Figure 2.1: Household appliances power consumption

with information loss but offer the advantage of being easily collectible [21]. The forward challenge revolves around the disaggregation of power consumption data, characterized by a low frequency of 0.1 Hz. Mathematically, the disaggregation is formulated in [6] as follows:

$$P_{total}(t) = \sum_{i=1}^{N} P_{appliance_i}(t) + P_{noise}(t), \tag{2.1}$$

where $P_{total}(t)$ represents the total power consumption at time t, $P_{appliance_i}(t)$ is the power consumption of the i-th appliance, N is the number of appliances, and $P_{noise}(t)$ signi-fies the noise component in the signal. The objective is to estimate $P_{appliance_i}(t)$ for each appliance, despite challenges posed by low-frequency measurements and signal noise.

CNNs are utilized for the classification of load types, leveraging spatial patterns within the data [26], [27]. The CNN architecture consists of convolutional layers followed by pooling layers. Each convolutional layer detects progressively more complex features. It can learn to distinguish the energy consumption patterns of various appliances, e.g., the distinct start-up and shut-down sequences of a dishwasher and an electric stove. Pooling layers reduce spatial dimensions, preserving important information. This process is particularly essential for our application, as it helps condense the information while preserving the critical spatial characteristics that distinguish appliances. The convolution operation utilized in [29] is:

$$C_{ij} = \sum_{m=1}^{M} \sum_{n=1}^{N} I_{i+m,j+n} K_{m,n}, \tag{2.2}$$

where I is the input matrix, K is the convolutional kernel, and M and N are the kernel dimensions. Pooling (max-pooling) extracts the maximum value from a region of the input, reducing spatial dimensions while retaining significant features denoted by P_{ij}, i.e.:

$$P_{ij} = \max(I_{2i,2j}, I_{2i,2j+1}, I_{2i+1,2j}, I_{2i+1,2j+1}). \tag{2.3}$$

A CNN is an essential part of our load identification and appliance classification technique because it is a powerful tool for capturing the spatial patterns in energy consumption data. This enables our model to effectively distinguish between appliances based on their unique energy consumption signatures.

2.2 Long Short Term Memory (LSTM)

Long short-term memory (LSTM) is a type of recurrent neural network (RNN) designed to process sequential data and remember long-term dependencies [22]. It is widely used in deep learning for sequence prediction tasks such as natural language processing [23], speech recognition [24], and time series forecasting [25]. The core idea behind LSTM is the use of specialized memory cells that can store and retrieve information over long sequences.

2.3 Bidirectional Long Short Term Memory (BiLSTM)

The BiLSTM unit plays a crucial role in the proposed disaggregation process. Its bidirectional processing capability helps our model comprehend sequential data and identify the complex temporal features in energy consumption data. Home appliances have unique energy consumption patterns that can be observed (and learned) over time. These patterns have nuanced transitions and interactions. The BiLSTM is a powerful tool for capturing these subtleties. It consists of two LSTM layers, one processing the input sequence in the forward direction and the other in reverse, enabling it to capture more context (compared to the LSTM) from both past and future inputs. The method can recognize the temporal sequences associated with appliance operation, such as the recurring cycles of a refrigerator's compressor or the periodic fluctuations in power consumption caused by a washing machine's agitator. Mathematically, the forward LSTM computes hidden states as follows [28]:

$$h_t^f = LSTM_f(x_t, h_{t-1}^f, c_{t-1}^f), \tag{2.4}$$

where x_t is the input at time t, h_{t-1}^f is the forward hidden state at time $t-1$, and c_{t-1}^f is the forward cell state at time $t-1$. The backward LSTM computes hidden states in a similar manner i.e.[28]:

$$h_t^b = LSTM_b(x_t, h_{t+1}^b, c_{t+1}^b), \tag{2.5}$$

where h_{t+1}^b is the forward hidden state at time $t+1$, and c_{t+1}^f is the forward cell state at time $t+1$. The final hidden state is a concatenation of the forward and backward hidden states:

$$h_t = [h_t^f, h_t^b]. \tag{2.6}$$

The BiLSTM is a powerful tool that helps the proposed model distinguish between different appliances based on their temporal intricacies. It can accurately break down total power consumption into individual appliance-level components by effectively modeling temporal patterns and dependencies.

2.4 Attention Mechanism

The attention mechanism is integrated into the BiLSTM model to enhance the model's capability to focus on specific time steps crucial for load identification. Attention mechanisms have been widely adopted in sequence-to-sequence tasks, and in our context, they dynamically weigh the importance of different time steps within the input sequence. This dynamic weighting is essential for accurately identifying appliance patterns, as certain time steps may contain more relevant information than others. For better identification of appliances, one can also consider the turning on and off time instants, especially in relation to other appliances, as features to improve the accuracy of the method. The model uses the attention mechanism to assign weights to each moment in the input sequence. Therefore, the moments when significant changes occur can be used in the identification task. This leads to a more accurate disaggregation of the power consumption data into the individual components.

The attention score α_t for a time step t is calculated using the hidden state h_t and a context vector v according to [30]:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{T} \exp(e_j)}. \tag{2.7}$$

e_t in the above equation is the attention energy at time t, computed as:

$$e_t = v^T \tanh(W_h h_t + b_h), \tag{2.8}$$

where W_h and b_h are weight and bias learnable parameters for attention calculation, and T

is the total number of time steps. The context vector c is obtained as the weighted sum of hidden states according to [30]:

$$c = \sum_{t=1}^{T} \alpha_t h_t. \tag{2.9}$$

The final prediction y in the regression task, according to [30], is then obtained by passing c through a linear layer as follows:

$$y = W_y c + b_y, \tag{2.10}$$

where W_y and b_y are weight and bias learnable parameters for linear transformation in regression.

The attention mechanism is a mechanism for improving load identification performance by offering a more focused and selective understanding of the data. This allows the model to better identify important temporal events that define appliance behaviour, ultimately leading to more accurate and reliable load disaggregation. Therefore, integrating an attention mechanism is crucial for improving the model's ability to recognize patterns of appliance activation and deactivation.

2.5 Proposed Model

In our study, we propose a hybrid model for the precise identification of home appliances' energy consumption patterns. Our model seamlessly integrates three key components: a CNNs, a BiLSTM network, and an attention mechanism for pinpointing critical time steps within energy data sequences. The CNNs leverage spatial patterns in the data to categorize the identified appliances effectively. The BiLSTM, on the other hand, captures intricate temporal dependencies in energy consumption data for enabling accurate appliance identification. Finally, the attention mechanism enhances the model to focus on

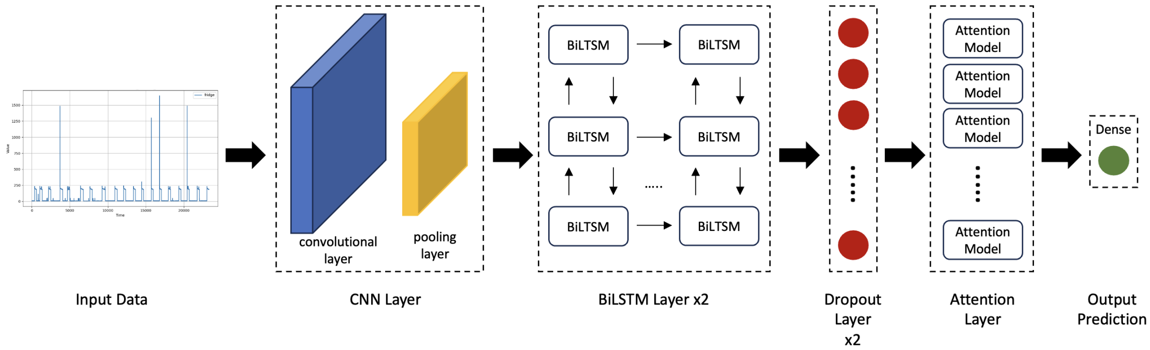

Figure 2.2: Proposed model architecture

more informative time steps, crucial for accurate load estimation. This hybrid approach, shown in Figure 2.2, capitalizes on the strengths of each component to create a robust and comprehensive solution for load identification using low-frequency power data.

2.6 Simulations

The REDD dataset is a well-established benchmark in the field of load identification and energy disaggregation. This dataset comprises a rich set of electrical load measurements recorded from various sensors and appliances within a residential setting. We will use the data for six appliances: dishwasher, electric space heater, electric stove, refrigerator, microwave, and washer dryer. It is publicly available and was collected by the authors of [18]. The dataset is characterized by a low monitoring frequency of 1 Hz. We reduced the sampling frequency of the REDD dataset from 1 Hz to 0.1 Hz by downsampling in time. Downsampling to 0.1 Hz effectively reduces the data volume by a factor of ten, significantly alleviating these challenges. This downsampled frequency provides an appropriate balance between maintaining sufficient detail for effective energy consumption analysis and reducing the computational burden. At 0.1 Hz, the dataset still retains enough granularity to capture the essential characteristics of appliance energy usage patterns, which is vital for accurate load disaggregation and monitoring. Note that if the downsampling rate is too

Table 2.1: The proposed model's hyperparameters used in the simulations

hyperparameters	method
No. of epochs	20
Loss function	MSE
Optimizer	Adam
CNN layer number	1
BiLSTM layer number	2
Attention mechanism layer number	1
Dropout layer number	2
Dense layer number	1

low, it could risk losing important information, while a rate that is too high might not sufficiently alleviate the computational challenges. Therefore, 0.1 Hz represents a pragmatic middle ground, ensuring the dataset remains manageable and the analysis remains feasible without significantly compromising the quality and utility of the information derived from the data. This tailored approach reflects a nuanced understanding of the trade-offs involved in data sampling for NILM. It demonstrates a thoughtful consideration of the practical aspects of data handling and analysis in this research. To prepare the data for our experiments, we perform standard preprocessing steps as part of the proposed model shown in Figure 2.3, including data cleaning, normalization, and sequence splitting into training and testing sets. Additionally, we transform the data into a format suitable for both the BiLSTM model (for regression) and the CNN model (for classification). Table 2.1 provides an overview of the proposed model's hyperparameters: epochs, cost function, optimization method, and the configuration of various neural network layers, used in the simulations. It is to be noted that 80% of data was used for training and 20% was used for testing. The proposed method was trained over a span of 20 epochs. The normalized data for appliances and total household energy consumption are shown in Figure 2.4 and 2.5 respectively. The training and validation loss values are shown for each epoch as depicted in Figure 2.6. This figure demonstrates that the model successfully converges, with a loss value lower than 0.00025. Despite the depth and complexity of the proposed model, it is noteworthy that the

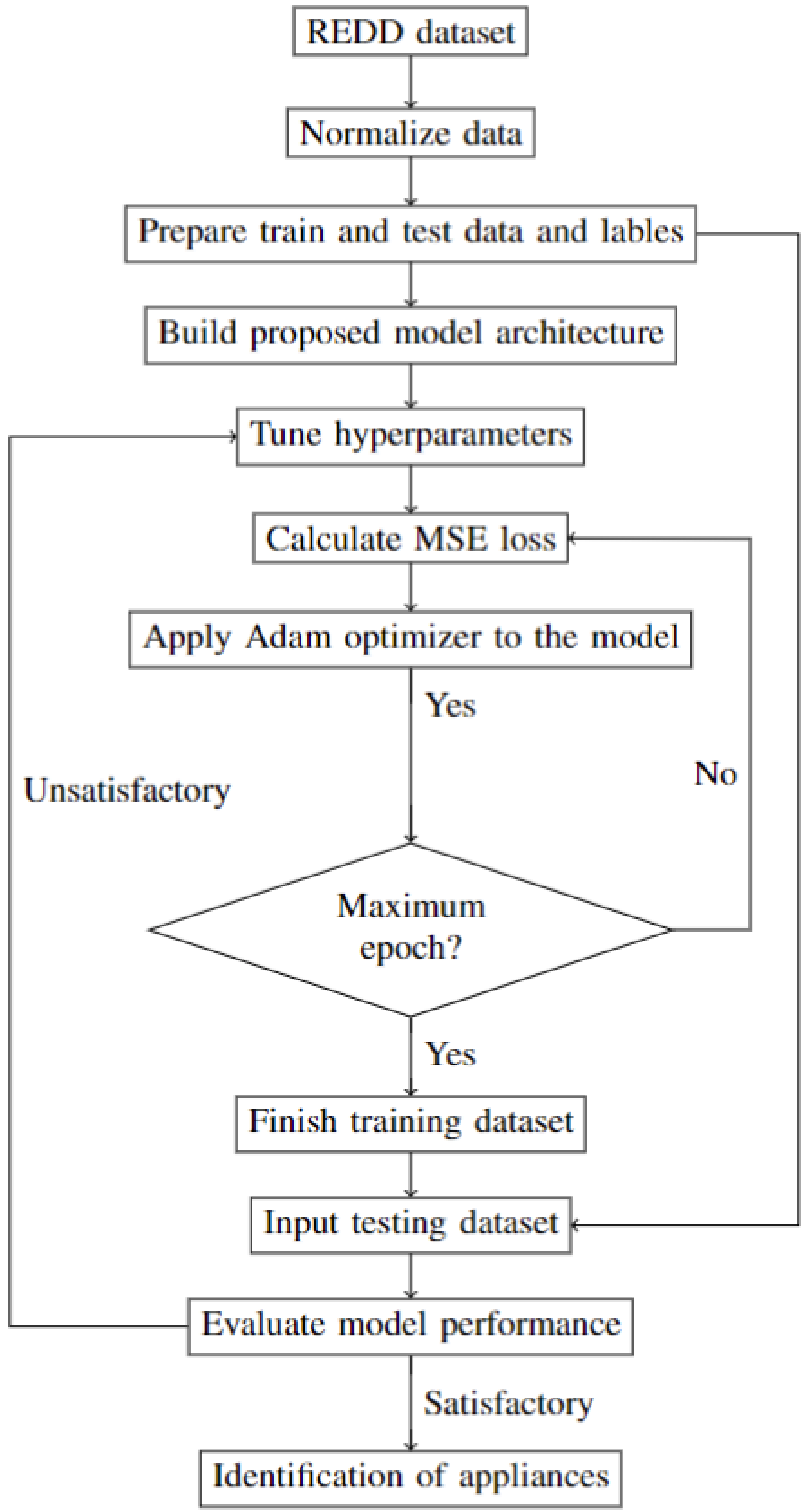

Figure 2.3: Flowchart of the proposed model

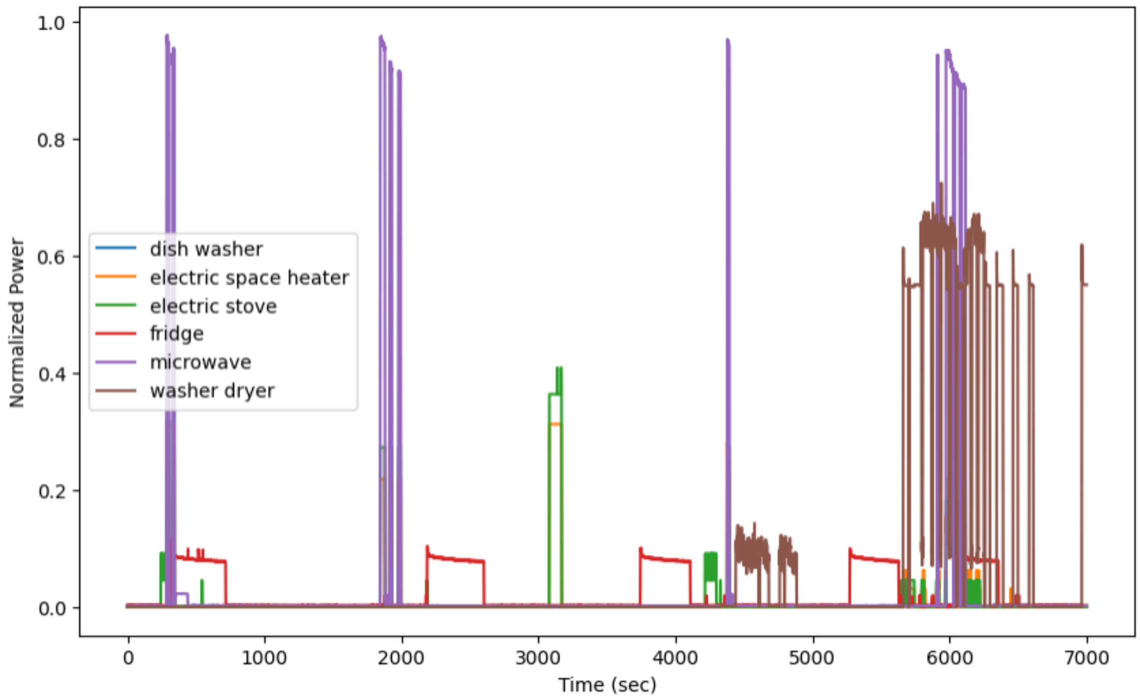

Figure 2.4: Normalized appliance data

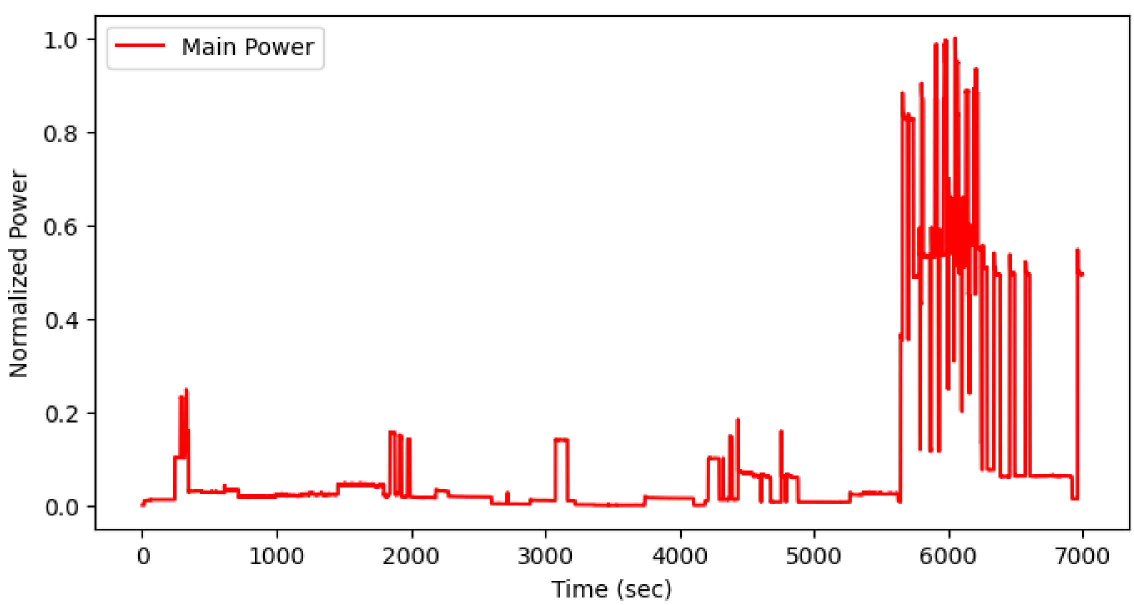

Figure 2.5: Normalized total household energy consumption data

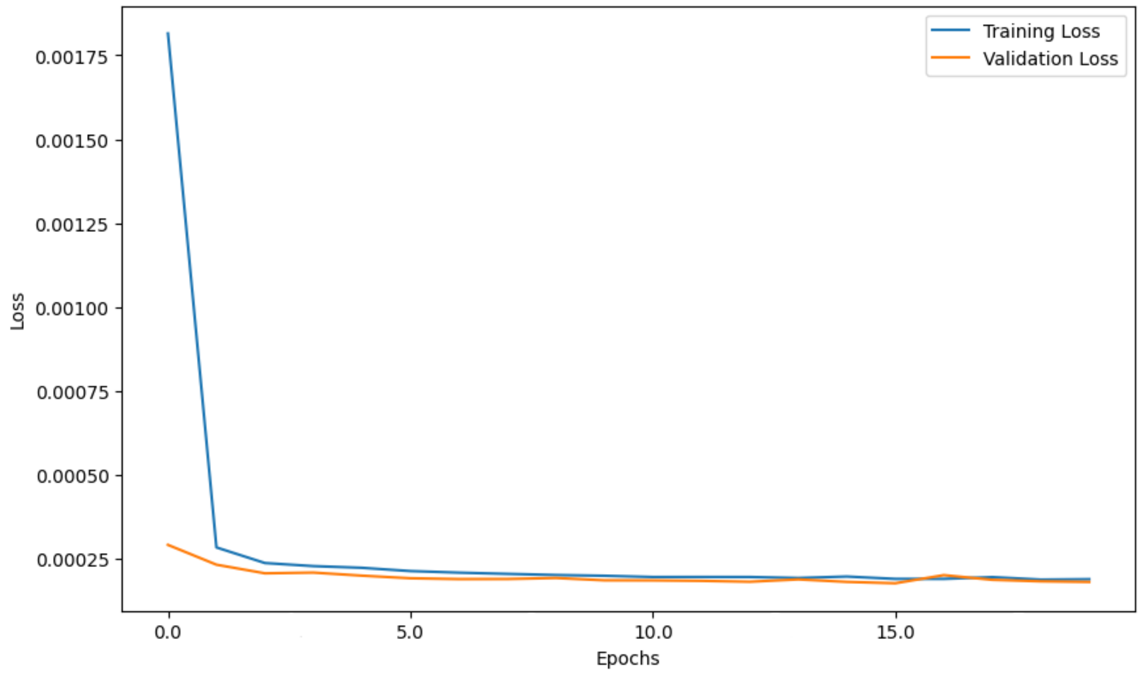

Figure 2.6: Training and validation loss

runtime (computation time) was notably fast (around 31 s 19ms/step). This computational efficiency is essential for real-time or near-real-time applications, where timely load identification is paramount. The power consumption of the refrigerator, microwave, electric space heater, and electric stove are shown in Figures 2.7, 2.8, 2.9 and 2.10, where each appliance is properly identified by the proposed method. In assessing the performance of our load identification model, we utilize three essential metrics: precision, recall, and F1-score. These metrics are fundamental for evaluating the model's classification accuracy. $Precision$ quantifies the accuracy of the model's positive identification of a particular load type as described by [31]:

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives} \tag{2.11}$$

$Recall$, also known as the true positive rate, measures the model's ability to correctly identify instances of a specific load type among all actual instances of that type, according to

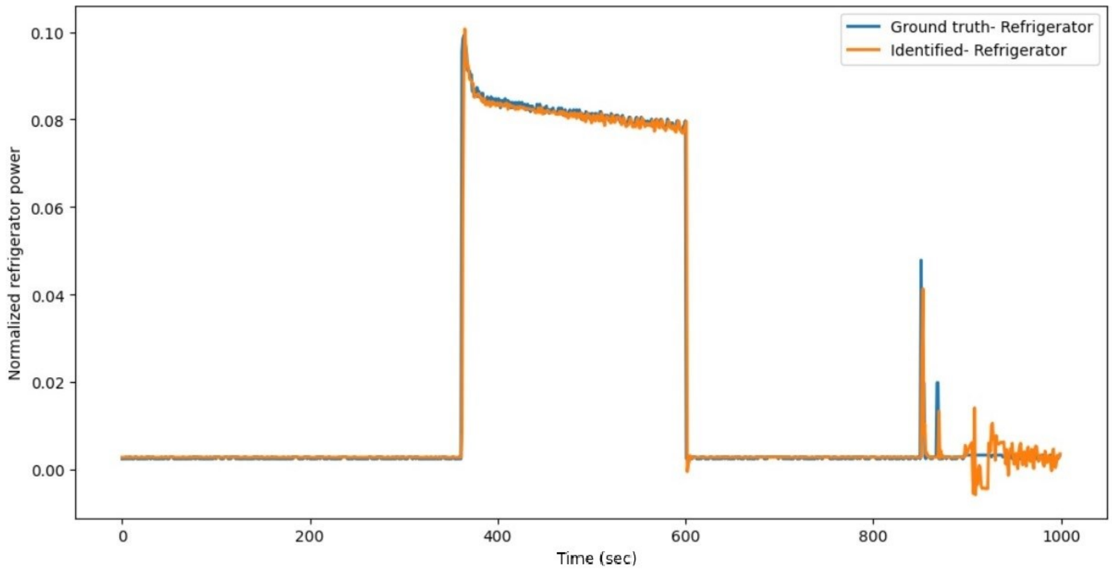

Figure 2.7: Identified vs. ground truth power consumption for a refrigerator

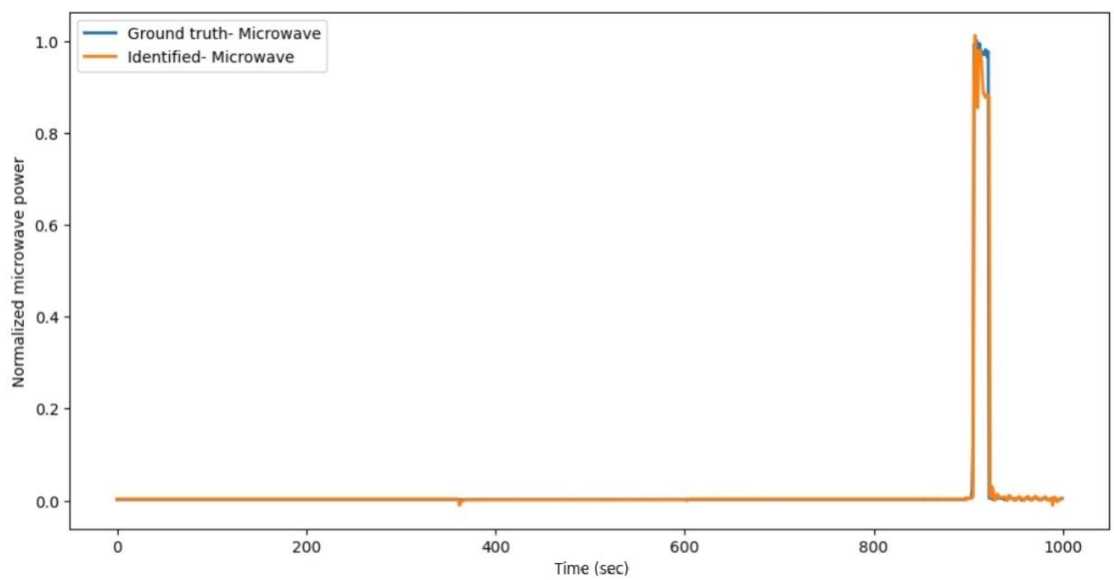

Figure 2.8: Identified vs. ground truth power consumption for a microwave

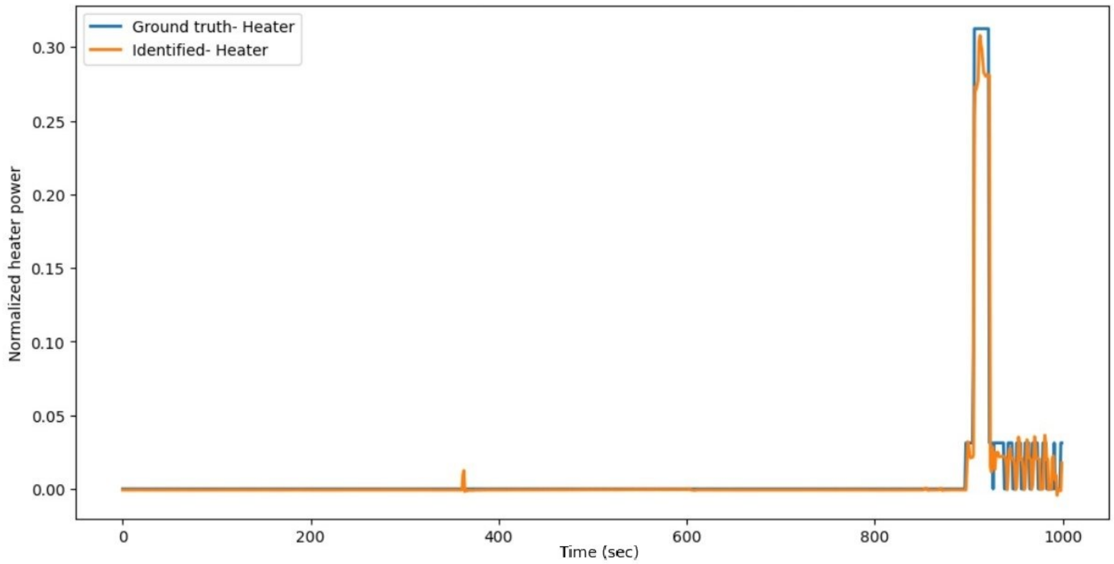

Figure 2.9: Identified vs. ground truth power consumption for an electric space heater

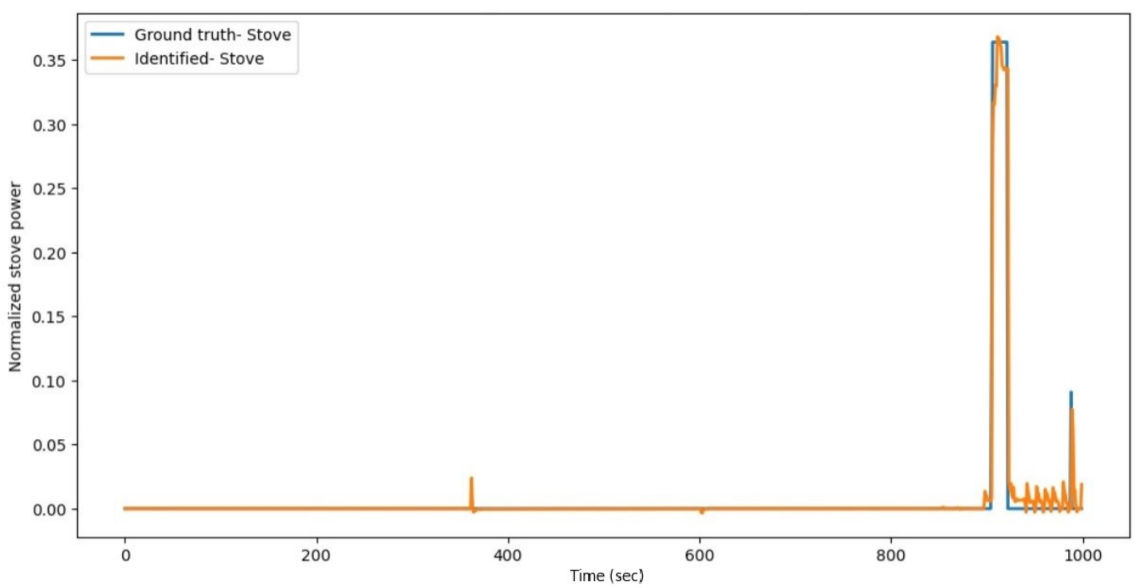

Figure 2.10: Identified vs. ground truth power consumption for an electric stove

Table 2.2: Model performance

Appliances	Precision	Recall	F1
dish washer	1.0000	0.9524	0.9756
electric space heater	1.0000	1.0000	1.0000
electric stove fridge	1.0000	1.0000	1.0000
microwave	1.0000	0.9890	0.9940
washer dryer	0.9810	1.0000	0.9901
	0.9823	0.9901	0.9860

the following formula [31]:

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives} \tag{2.12}$$

$F1 - score$ is defined as the harmonic mean of precision and recall, balancing the trade-off between false positives and false negatives, as follows [31]:

$$F1 = \frac{2.Precision.Recall}{Precision + Recall} \tag{2.13}$$

Table 2.2 provides a summary of these metrics for test data, illustrating our model's remarkable performance in accurately classifying load types.

Chapter 3

Load Forecasting using GNN-LSTM Attention Mechanism with Low-Frequency

3.1 Problem Statement

In the load forecasting problem, it is desired to predict the electricity demand or consumption in a given time interval. Load forecasting is essential for planning and operating power systems, as well as for optimizing energy efficiency and reducing costs. It is a challenging problem due to several reasons, such as:

- the diversity and complexity of electrical loads with potentially large fluctuations, overlapping patterns, or nonlinear behaviors [38];
- the lack of standardized and representative datasets, metrics, and procedures for assessing the efficacy of NILM algorithms [39], [40], and
- the trade-off between the accuracy, scalability, and practicability of the existing NILM techniques may necessitate high-resolution data, complex models, or user feedback

[38].

These challenges necessitate novel and robust approaches capable of dealing with the uncertainty of load data, as well as user demands and expectations. A fully connected graph structure is employed to represent the time correlation of appliances. An adjacency matrix is constructed with a controlled level of sparsity, introducing a realistic representation of the correlation between appliances. This sparsity mimics the real-world scenario, where not all appliances' operation times depend on each other. The diagonal elements of the matrix are set to zero, neglecting the self-dependency of the appliances. The objective is to predict each appliance's load based on its time correlation with other appliances despite the challenges mentioned earlier.

3.2 Graph Neural Network (GNN)

In the context of our energy consumption prediction model, GNN plays a pivotal role in capturing and leveraging the intricate time correlation among household appliances. The sparse graph structure is defined by an adjacency matrix, incorporating a controlled level of sparsity to realistically model the time correlation between appliances. A commonly used equation in GNN-based load forecasting is the message-passing formula given in [42] as follows:

$$h_v^{(l+1)} = \sigma\left(\sum_{u \in N(v)} W^{(l)} h_u^{(l)}\right), \tag{3.1}$$

where $h_v^{(l+1)}$ is the updated value of node v at layer $(l+1)$ and represents the hidden state of the node after information aggregation at layer $(l+1)$; σ is the activation function; $W^{(l)}$ is the weight matrix at layer l used to transform and combine node features and messages from neighbors, which is specific to each layer of the GNN and is learned during training; $h_u^{(l)}$ is the value of neighboring node u at layer l and represents the hidden state of neighboring nodes at layer l. The equation describes how the hidden state of a node v in a graph is updated in the next layer $(l+1)$ by aggregating information from its neighboring nodes, applying a weight matrix transformation $W^{(l)}$, and then passing the result through a nonlinear activation function σ. This process is repeated for each node in the graph over multiple layers, allowing the network to capture complex time correlations and dependencies between appliances within the graph structure.

3.3 Long Short Term Memory (LSTM)

LSTM networks have become a pivotal tool in load forecasting due to their effectiveness in capturing temporal dependencies in electricity consumption patterns. LSTM is a type of recurrent neural network (RNN) that excels at modeling time series data. In the context of load forecasting, LSTM cells are used to retain information over extended time periods and remember past consumption patterns. This enables the network to grasp not only immediate dependencies but also longer-term trends in appliance usage. The core equation of an LSTM cell for load forecasting given in [46] is:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \tag{3.2}$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \tag{3.3}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{3.4}$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \tag{3.5}$$

$$h_t = o_t \cdot \tanh(c_t), \tag{3.6}$$

where x_t represents the input at time t, which can include information like historical load values and external factors; h_t is the hidden state at time t, denoting the learned representation at that moment; c_t is the cell state at time t, capturing memory and information retention; i_t, f_t, and o_t are the input, forget, and output gates, respectively, and they control the flow of information in and out of the cell; σ is the sigmoid activation function, and tanh is the hyperbolic tangent activation function; W and b are, respectively, the weight matrix and bias term learned during training. LSTM networks demonstrate remarkable performance in capturing complex load patterns and are crucial tools for accurate load forecasting in the energy domain.

3.4 Attention Mechanism

Attention mechanisms have revolutionized load forecasting by allowing models to focus on specific aspects of the input data that are most relevant for prediction. These mechanisms enable the model to assign varying levels of importance to different time steps, factors, or features, adapting dynamically to the input sequence. In load forecasting, attention mechanisms prove valuable in capturing seasonality and identifying crucial patterns in energy consumption. The core equation for a simple attention mechanism is the weighted sum of input features, described in [48] as:

$$\text{Attention}(Q, K, V) = \sum_{i=1}^{N} \frac{\exp(QK_i)}{\sum_{j=1}^{N} \exp(QK_j)} \cdot V_i, \tag{3.7}$$

where Q represents the query vector used to weigh the importance of different keys; K represents the keys, which are the elements attended to, and V represents the values associated with the keys. The attention mechanism computes a weighted sum of the values, where the weights are determined by the similarity between the query and the keys. This is an equation according to which the model pays attention to specific time steps or features most

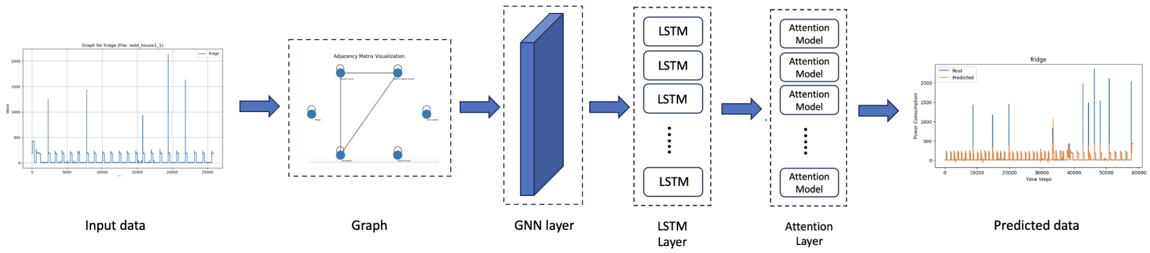

Figure 3.1: Proposed model architecture

relevant to making load forecasts and is a powerful tool for enhancing prediction accuracy in load forecasting applications.

3.5 Proposed Model

In our study, we propose a hybrid model for predicting the energy consumption patterns of home appliances. Our model integrates three key components: a GNN, an LSTM network, and an attention mechanism for pinpointing critical time steps within energy data sequences. The GNN captures complex time correlations and dependencies between home appliances, represented as nodes in the graph. The LSTM, on the other hand, effectively models temporal dependencies and sequences, allowing the model to remember past consumption patterns. Finally, the attention mechanism dynamically assigns varying levels of importance to different time steps within energy data sequences, facilitating the identification of crucial patterns for energy consumption predictions. This hybrid approach, shown in Figure 3.1, capitalizes on the strengths of each component to create a robust and comprehensive solution for load forecasting using low-frequency power data.

3.6 Simulations

The reference energy disaggregation data set (REDD) [47] is a benchmark data collection in load prediction and energy disaggregation research. This dataset comprises a rich

Table 3.1: The proposed model's hyperparameters used in the simulations

hyperparameters	method
No. of epochs	5
Loss function	MAE
Optimizer	Adam
GNN layer number	1
LSTM layer number	1
LSTM layer units number	30
Attention mechanism layer number	1
Dense layer number	1

set of electrical load measurements recorded from various sensors and appliances within a residential setting. We will use this dataset for six appliances: dishwasher, electric space heater, electric stove, refrigerator, microwave, and washer-dryer. The dataset is characterized by a low monitoring frequency. To prepare the data for our simulations, we perform standard preprocessing steps as part of the proposed model, including data normalization and sequence splitting into training and testing sets. Additionally, we transform the data into a format suitable for both the GNN and LSTM model. Table 3.1 provides an overview of the proposed model's hyperparameters: epochs, loss function, optimization method, and the configuration of various neural network layers used in the simulations. It is to be noted that 80% of the data was used for training and 20% for testing. The proposed method was trained over a span of 5 epochs, as the model will be over-fitted after five iterations. The training and validation loss values are shown for each epoch as depicted in Figure 3.2. This figure demonstrates that the model successfully converges with a loss value lower than 16.81, less than the existing results, as shown in Table 3.2. Despite the depth and complexity of the proposed model, it is remarkable that the runtime (computation time) was notably fast. This computational efficiency is essential for real-time or near-real-time applications, where timely load identification is paramount. The power consumptions of the refrigerator, microwave, dishwasher, and washer-dryer are shown in Figures 3.3, 3.4, 3.5 and 3.6. These figures demonstrate satisfactory load identification performance when

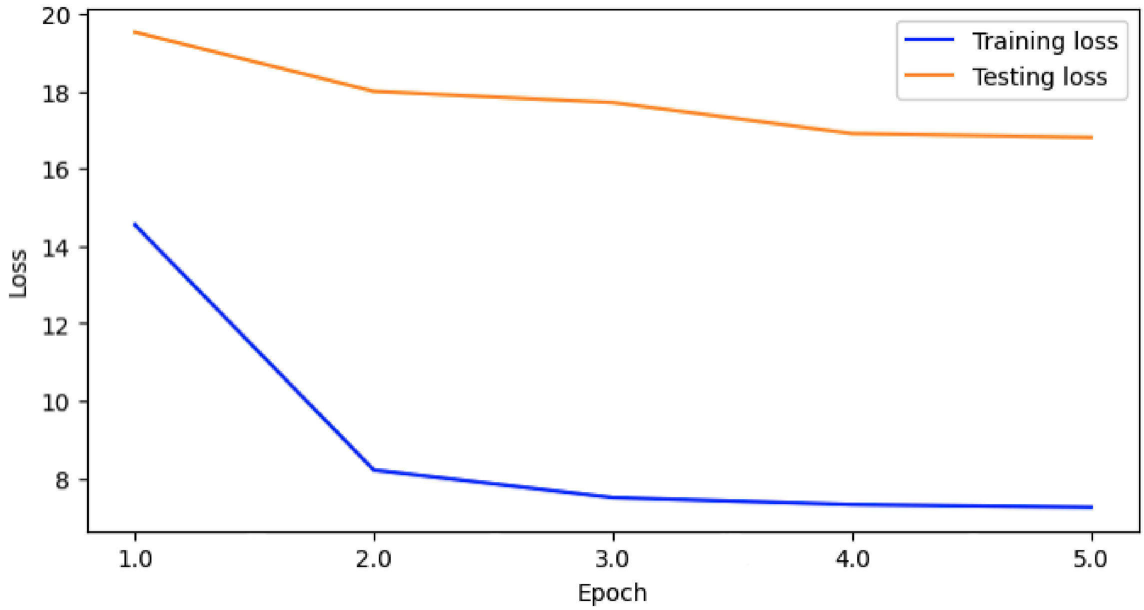

Figure 3.2: Training and validation loss

Table 3.2: Performance comparison of different models

Model	MAE
GNN-LSTM-ATT	16.81
GNN4NILM	18.87
BERT	26.35
LSTM	30.80
GRU	28.73
CNN	28.92

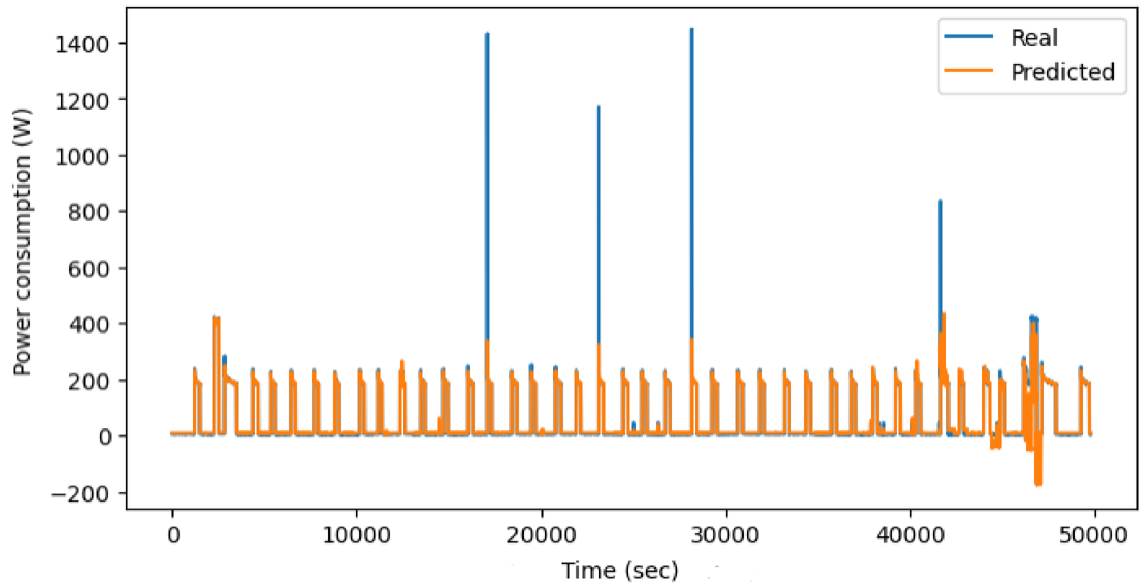

Figure 3.3: Predicted power consumption of a fridge compared to the ground truth

compared with the ground truth.

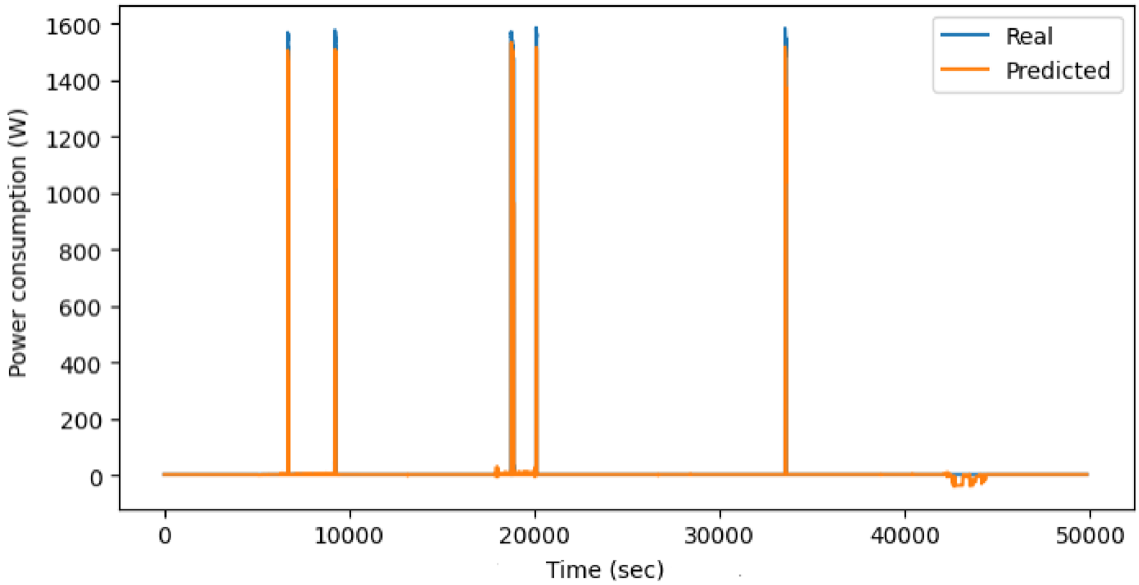

Figure 3.4: Predicted power consumption of a microwave compared to the ground truth

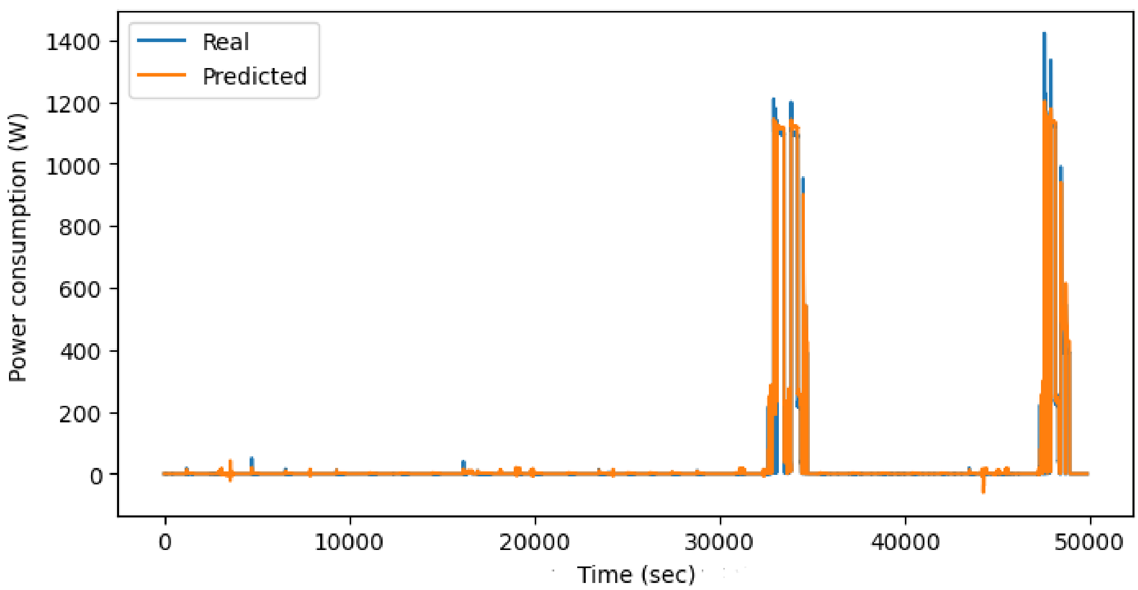

Figure 3.5: Predicted power consumption of a dishwasher compared to the ground truth

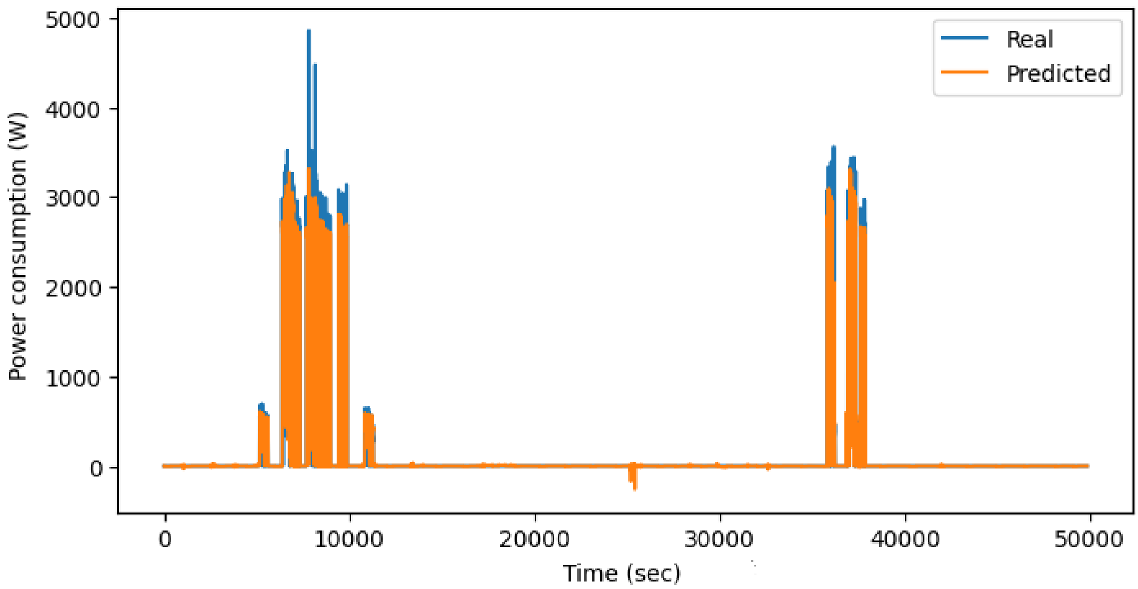

Figure 3.6: Predicted power consumption of a washer-dryer compared to the ground truth

Chapter 4

Conclusions and Future Work

In this thesis, we present two models within the non-intrusive load monitoring (NILM) framework using low-frequency data: the CNN-BILSTM model and the GNN-LSTM model, both enhanced by an attention mechanism. The CNN-BILSTM model stands out for its computational efficiency and accuracy, surpassing existing methods in solving NILM problems. This model's cost-effectiveness is a significant advantage, particularly considering its use of low-frequency data. Its efficiency and accuracy make it highly suitable for real-time applications, offering a new horizon for energy monitoring in smart homes.

The CNN-BILSTM model not only demonstrates superior performance in terms of computational resources but also sets a benchmark in appliance identification within the NILM context. The model's ability to operate with low-frequency data without sacrificing accuracy is a notable breakthrough, making it an ideal solution for real-world applications where high-frequency data may not be readily available. This aspect of the model shows great promise for widespread implementation in various residential and commercial settings, promoting energy efficiency and cost savings.

Furthermore, this thesis explores the potential of the GNN-LSTM model in forecasting future energy consumption patterns of individual appliances. This model's approach, which integrates controlled graph representation with an attention mechanism, enhances interpretability and prediction accuracy. Such capabilities are crucial in developing smart energy systems that can adapt and respond to changing consumption patterns, thereby optimizing energy usage.

Our simulation results demonstrate the efficacy of both models, highlighting their potential to transform energy management in residential environments. The findings suggest that these models can play a pivotal role in advancing proactive energy management, contributing significantly to the evolution of intelligent and sustainable smart homes. By enabling more precise and efficient energy usage, these models support efforts in energy conservation and play a crucial role in mitigating the environmental impact of energy consumption.

The impact of this research extends beyond just its technical successes. It helps achieve the larger aim of sustainable living by improving how we monitor and predict energy use. Using these models in smart homes could significantly lower unnecessary energy use, supporting worldwide initiatives to cut down on carbon emissions and fight against climate change. The results lay the groundwork for upcoming studies in NILM and smart energy management. The development of the CNN-BILSTM and GNN-LSTM models marks a major advancement in the quest for more efficient and intelligent energy use in domestic applications. As the global community shifts towards sustainable energy solutions, the methodologies introduced in this research provide essential contributions to this critical area.

Nevertheless, despite the encouraging outcomes, there are limitations that should be acknowledged. For instance, the effectiveness of these models across various environmental settings and with different types of appliances requires more study. Future research could concentrate on improving the robustness of these models against diverse data sets and real-life situations. Moreover, investigating how these models can be combined with other smart home technologies might pave the way for more comprehensive home energy management systems.

Bibliography

[1] M. Eirinaki, I. Varlamis, J. Dahihande, A. Jaiswal, A. A. Pagar and A. Thakare, ”Real-time recommendations for energy-efficient appliance usage in households,” Frontiers in Big Data, vol. 5, article 972206, 2022.

[2] Q. Yan, W. Xudong and W. Zun, ”Applications of NILM in the optimization management of intelligent home energy management system,” IEEE Sustainable Power and Energy Conference, Chengdu, China, pp. 2300-2305, 2020.

[3] T. Hussain, F. U. Min Ullah, K. Muhammad, S. Rho, A. Ullah, E. Hwang, J. Moon and S. W. Baik, ”Smart and intelligent energy monitoring systems: A comprehensive literature survey and future research guidelines,” International Journal of Energy Research, vol. 45, no. 3, pp. 3590-3614, 2021.

[4] P. Ma, S. Cui, M. Chen, S. Zhou and K. Wang, ”Review of family-level short-term load forecasting and its application in household energy management system,” Energies, vol. 16, no. 15, article 5809, 2023.

[5] Y. Himeur, A. Alsalemi, F. Bensaali, A. Amira and A. Al-Kababji, ”Recent trends of smart nonintrusive load monitoring in buildings: A review, open challenges, and future directions,” International Journal of Intelligent Systems, vol. 37, no. 10, pp. 7124-7179, 2022.

[6] G. W. Hart, ”Nonintrusive appliance load monitoring,” IEEE, vol. 80, no. 12, pp. 1870-1891, 1992.

[7] J. Z. Kolter and T. Jaakkola, ”Approximate inference in additive factorial hmms with application to energy disaggregation,” the 15th International Conference on Artificial Intelligence and Statistics, La Palma, Canary Islands, Spain, 21–23, 2012.

[8] P. R. Z. Taveira, C. H. V. De Moraes and G. Lambert-Torres, ”Non-intrusive identification of loads by random forest and fireworks optimization,” IEEE Access, vol. 8, pp. 75060-75072, 2020.

[9] H. Y. Lam, G. S. K. Fung and W. K. Lee, ”A novel method to construct taxonomy electrical appliances based on load signatures,” IEEE Transactions on Consumer Electronics, vol. 53, pp. 653-660, 2007.

[10] O. Parson, S. Ghosh, M. Weal and A. Rogers, ”Non-intrusive load monitoring using prior models of general appliance types,” the AAAI Conference on Artificial Intelligence, vol. 26. no. 1. 2012.

[11] J. Kelly and W. Knottenbelt, ”Neural NILM: Deep neural networks applied to energy disaggregation,” the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, 2015.

[12] C. Zhang, M. Zhong, Z. Wang, N. Goddard and C. Sutton, ”Sequence-to-point learning with neural networks for non-intrusive load monitoring,” the AAAI Conference on Artificial Intelligence, vol. 32. no. 1. 2018.

[13] J. Song, H. Wang, M. Du, L. Peng, S. Zhang and G. Xu, ”Non-intrusive load identification method based on improved long short term memory network,” Energies, 14(3), p.684, 2021.

[14] H. Kim, M. Marwah, M. Arlitt, G. Lyon and J. Han, ”Unsupervised disaggregation of low-frequency power measurements,” Energy and Buildings, vol. 43, no. 11, pp. 3480-3487, 2011.

[15] J. Kelly and W. Knottenbelt, ”Neural NILM: Deep neural networks applied to energy disaggregation,” the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, Seoul, Republic of Korea, 4–5, pp. 55–64, 2015.

[16] P. Davies, J. Dennis, J. Hansom, W. Martin, A. Stankevicius and L. Ward, ”Deep neural networks for appliance transient classification,” the IEEE International Conference on Acoustics, Speech and Signal Processing, Brighton, UK, 12–17, pp. 8320–8324, 2019.

[17] H. Liu, L. Li, G. Ding and Q. Zhang, ”Nonintrusive load disaggregation combining with external attention mechanism and Seq2Piont,” the 7th International Conference on Power and Renewable Energy, 2022.

[18] J. Z. Kolter and M. J. Johnson, ”REDD: A public data set for energy disaggregation research,” Workshop on Data Mining Applications in Sustainability, 2011.

[19] N. Batra, J. Kelly, O. Parson, H. Dutta, W. Knottenbelt, A. Rogers, A. Singh and M. Srivastava, ”NILMTK: An open-source toolkit for non-intrusive load monitoring,” the 5th International Conference on Future Energy Systems, 2014.

[20] A. Zoha, A. Gluhak, M. A. Imran and S. Rajasegarar, ”Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey,” Sensors, vol. 12, no. 12, pp. 16838-16866, 2012.

[21] P. Huber, A. Calatroni, A. Rumsch and A. Paice, ”Review on deep neural networks applied to low-frequency NILM,” Energies, vol. 14, no. 9, article 2390, 2021.

[22] S. Hochreiter and J. Schmidhuber, ”Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.

[23] A. Graves and J. Schmidhuber, ”Framewise phoneme classification with bidirectional LSTM and other neural network architectures,” Neural Networks, vol. 18, no. 5-6, pp. 602-610, 2005.

[24] A. Graves, A. R. Mohamed and G. Hinton, ”Speech recognition with deep recurrent neural networks,” the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 6645-6649, 2013.

[25] Z. C. Lipton, J. Berkowitz and C. Elkan, ”A critical review of recurrent neural networks for sequence learning,” arXiv:1506.00019, 2015.

[26] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, ”Gradient-based learning applied to document recognition,” IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.

[27] A. Krizhevsky, I. Sutskever and G. E. Hinton, ”ImageNet classification with deep convolutional neural networks,” Advances in Neural Information Processing Systems, 2012.

[28] M. Schuster and K. K. Paliwal, ”Bidirectional recurrent neural networks,” IEEE Transactions on Signal Processing, vol. 45, no. 11, pp. 2673-2681, 1997.

[29] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, ”Gradient-based learning applied to document recognition,” IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.

[30] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin, ”Attention is all you need,” Advances in neural information processing systems, 30, 2017.

[31] R. Yacouby and D. Axman, ”Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models.” The first workshop on evaluation and comparison of NLP systems, 2020.

[32] M. Etezadifar, H. Karimi, A. G. Aghdam and J. Mahseredjian, ”Resilient event detection algorithm for non-intrusive load monitoring under non-ideal conditions using reinforcement learning,” IEEE Transactions on Industry Applications, 2023, to appear.

[33] A. Azzam, S. Sanami, and A. G. Aghdam, ”Low-frequency load identification using CNN-BiLSTM attention mechanism,” arXiv:2311.08536, 2023.

[34] A. L. P. De Ocampo, A. M. M. Baes, and D. G. D. Ronquillo, ”Non-intrusive load monitoring and forecasting for home appliances using artificial intelligence – a review,” International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Bangalore, India, p. 1-7, 2022.

[35] M. Zhuang, M. Shahidehpour, and Z. Li, ”An overview of non-intrusive load monitoring: Approaches, business applications, and challenges,” International Conference on Power System Technology (POWERCON), Guangzhou, China, p. 4291-4299, 2018.

[36] S. Makonin, and F. Popowich, ”Nonintrusive load monitoring (NILM) performance evaluation: A unified approach for accuracy reporting,” Energy Efficiency 8, p. 809-814, 2015.

[37] V. Gkalinikis, N., C. Nalmpantis, and D. Vrakas, ”Torch-NILM: An effective deep learning toolkit for non-intrusive load monitoring in pytorch,” Energies, 15(7), p. 2647, 2022.

[38] X. Zhou, J. Feng, and Y. Li. ”Non-intrusive load decomposition based on CNN–LSTM hybrid deep learning model,” Energy Reports 7, p. 5762-5771, 2021.

[39] M. Zhuang, M. Shahidehpour, and Z. Li, ”An overview of non-intrusive load monitoring: Approaches, business applications, and challenges,” International Conference on Power System Technology (POWERCON), Guangzhou, China, p. 4291-4299, 2018.

[40] C. Nalmpantis, and D. Vrakas, ”Machine learning approaches for non-intrusive load monitoring: from qualitative to quantitative comparation,” Artificial Intelligence Review 52, p. 217-243, 2019.

[41] M. Kaselimi, E. Protopapadakis, A. Voulodimos, N. Doulamis, and A. Doulamis, ”Towards trustworthy energy disaggregation: A review of challenges, methods, and perspectives for non-intrusive load monitoring,” Sensors 22, no. 15, p. 5872, 2022.

[42] Y. Jiang, L. Mingsheng, L. Yangyang, L. Yaping, Z. Jingyun, L. Yifeng, and L. Chunyang, ”Enhanced neighborhood node graph neural networks for load forecasting in smart grid,” International Journal of Machine Learning and Cybernetics, p. 1-20, 2023.

[43] D. Zhang, X. Jin, P. Shi, and X. Chew, ”Real-time load forecasting model for the smart grid using bayesian optimized CNN-BiLSTM,” Frontiers in Energy Research 11, p. 1193662, 2023.

[44] Z. Meng, Y. Xie, and J. Sun. ”Short-term load forecasting using neural attention model based on EMD,” Electrical Engineering, p. 1-10, 2022.

[45] H. Chi, and S. Kang, ”Multiobjective metaheuristic load control algorithm for interaction between smart home and humans,” Mathematical Problems in Engineering, vol. 2022, Article ID 2210159, 2022.

[46] Y. Wang, W. Jiang, C. Wang, Q. Song, T. Zhang, Q. Dong, and X. Li, ”An electricity load forecasting model based on multilayer dilated LSTM network and attention mechanism,” Frontiers in Energy Research 11, p. 1116465, 2023.

[47] J. Z. Kolter, and M. J. Johnson, ”REDD: A public data set for energy disaggregation research,” Workshop on Data Mining Applications in Sustainability, 2011.

[48] X. Yang, ”An overview of the attention mechanisms in computer vision,” Journal of Physics: Conference Series, vol. 1693. No. 1. IOP Publishing, 2020.

[49] R. Yacouby, and D. Axman, ”Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models,” The first workshop on evaluation and comparison of NLP systems, 2020.