



PAPER • OPEN ACCESS

Non-intrusive Load Monitoring Based on Convolutional Neural Network Mixed Residual Unit

To cite this article: Yunxin Li *et al* 2019 *J. Phys.: Conf. Ser.* **1176** 022052

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices
to create your essential collection of books in STEM research.

Start exploring the **collection** - download the first chapter of
every title for free.

Non-intrusive Load Monitoring Based on Convolutional Neural Network Mixed Residual Unit

Yunxin Li¹, Bo Yin^{1,*}, Peidong Wang², Rui Zhang³

¹College of information science and engineering, Ocean University of China, Qingdao, China

²Qingdao Product Quality Supervision and Inspection Institute, Qingdao, China

³Qingdao Haier Smart Technology R&D Co., Ltd., Qingdao, China

*Corresponding author e-mail: ybfirst@126.com

Abstract. Non-intrusive Load Monitoring (NILM) is designed to determine the type of electrical appliance after decomposing the total power into each appliance in order to analyze how much electricity was consumed. In the traditional pattern recognition, relying on artificial extraction features, there is a large amount of unreliability, and then the use of convolutional neural networks to automatically extract features, but in the traditional CNN structure, when the network level is deepened, training error and test error will increase. Therefore, this paper proposes a network structure with residual unit added to the traditional convolutional neural network, in order to reduce the probability of gradient explosion. The dataset of the paper is the data collected from 22 different electrical appliances, and its current and voltage trajectory map is used to identify the accuracy rate of 97.9%.

1. Introduction

Due to the rapid development of China's smart-grid, a large number of and a variety of electrical appliances have entered life, and thus the demand for energy has also increased, which has triggered energy consumption. Energy management has also become one of the mainstream issues now. Effective energy management helps power companies make better use of power generation resources by decreasing consumers' energy requirement in peak hour. In order to solve this problem, it is crucial to be able to clearly and accurately determine the power consumption of each appliance due to a reasonable planning of the power usage. Traditional load monitoring is performed by installing power sensor at each power point and this approach is called intrusive load monitoring. However this method is uneconomical because it requires installation of many power sensors and that it may affect system reliability. To overcome such problem, an approach to integrate all the load profile information by using a smart meter is proposed and it is called as NILM. NILM provides an efficient and accurate method to reduce the trouble of monitoring the consumption of each appliance. Firstly, it should track the aggregated electricity quantity, and then the monitored total electricity state is disaggregated using machine learning techniques and disaggregated into each individual electric appliance.

This experiment's focus and dedication is to create a convolutional neural network that incorporate residual units. The experiment procedure is guided by machine learning disciplines which is key process, to obtain the optimal implementation for load identification. Process includes: (1) Data pre-treatment: acquiring data of current and voltage, then plot them in same coordinate system becomes VI trajectory. (2) model structure: constructing convolutional neural network that



incorporate residual unit. The first unit includes convolutional neural network and pooling layer and the second unit is residual unit, and then the last unit is convolutional neural network and pooling layer. Finally, adding two fully connected layers which including output layer. (3) training: using the data that acquired from step (1) to train model that from step (2). (4) classification: selecting validation set to classify appliances.

The structure of this paper is shown below. In the second section, we present the problem statement and solution of the related work. Section 3 describes theory used in this paper. Including VI-trajectory image, convolutional neural network and residual unit. In section 4, we conduct experiment that is about validating the expected effects of the dataset and learning architecture and is about measuring the overall performance comparing with the existing models. In the end, the work of this paper is summarized and the conclusion is given.

2. Related Work

Hart come from MIT came up with the original concept of disaggregating residential electrical load information in 1980. He showed how different electrical appliances produce different power signals, including their active power, current and voltage. The on-off events he shows are enough to describe the use of certain appliances. Even if this method is good, it cannot cover up its limitations. For instance, small appliances, appliances that are set as normally open and appliances which have non-discrete changes in power[2]. There have many researches were studied for this concept, and optimize the NILM model.

There are many features that are ideal for classifying kitchen appliances such as root mean square(RMS) which were proposed by Najmeddine and Figueiredo et al. They used the current I and the voltage V as the signature[3,4]. But there is a disadvantage, it cannot classify the multistate appliances.

A new idea were presented by Cole[5,6] is to extract data and identify a steady-state load for NILM. When one or more appliances are opened or closed, it can implement load switching between appliances by Cole's algorithm. But there is a disadvantage about this algorithm, it takes a long time to accumulate real power(P) and active power(Q). In addition, this algorithm do not have ability to classify any appliances that power consumption does not change[6].

Recently, many researches proposed some improved algorithms. In order to solve low identification accuracy for NILM in the case of low sampling rate, a neoteric algorithm presented by Liang Kong is based on maximum likelihood. The average results of the identification accuracy are more than 85%[7]. Seyed Mostafa Tabatabaei in 2017 proposed to use a multi-label meta-classification framework (RAkEL), and a bespoke multi-label classification algorithm (MLkNN), employing both time-domain and wavelet-domain feature sets[8]. Steven R. Shaw described a system that can decide which operating schedule would be chose and search parameters of physical models, including transient event classification scheme, system identification, and implementation using in NILM[9].

In [10], Leen De Bates proposed images which possess weighted pixel of the voltage and current trajectory were used as input for deep learning method, then constructed CNN model including 2 convolutional layers that have ability to extract nuclear features automatically for appliance classification. For PLAID dataset, the accuracy is up to 77.6% and the WHITED dataset is 75.46%. The number of appliance for the PLAID is 11. This paper updates traditional convolutional neural network model by adding residual unit. The number of appliance for our dataset is up to 21. And this paper also uses VI images. Finally, the accuracy of our model is 97.9%.

3. Theory

3.1. VI image

When the appliance is turned on, it starts to detect the current and voltage values of each appliance, and divides the current and voltage data in one cycle and plots them in the same coordinate system to form a continuous VI trajectory image. Features that facilitate model training can be obtained from the VI trajectory image, such as the closed area, the slope in the middle, etc. There are pixels in each

trajectory image, such as a 64×64 image, which will have 64×64 pixel values, which are the data forms that make up the picture. When the model reads the data, it reads the pixel value of the image and uses the principle of image recognition to classify the appliance.

Past pattern recognition typically uses raw data, and today transforms data into weighted pixelated images with continuous values rather than binary values, maximizing the extraction of features contained in the VI trajectory image.

3.2. Convolutional neural network

The premise of classifying appliance is to extract features and some existing load identification algorithms are based on transient and steady-state electrical characteristics which relies excessively on feature extraction. Subsequently, AlexNet, which was born in 2012, led to the rapid development of convolutional neural networks. Convolutional neural network can automatically extract features, unlike traditional neural network, which requires manual feature extraction, nor feature extraction algorithm such as SIFTCompletion of feature extraction and abstraction greatly reduces the difficulty of applying image recognition. Compared with general neural networks, CNN is closer in structure and spatial image structure, and CNN convolution connection method and human visual nerve processing optical signal, the way is similar, so CNN is highly suitable for the classification of images.

The network structure of common neural networks is mostly consistent. There are several hidden layers between the input and output. Each layer is composed of multiple neurons, and there is no connection between them. The layers are fully connected. The convolutional neural network is a special structure of the common neural network. On the one hand, the connections between its neurons are Sparse Connectivity. For another, the weights between some units in the same layer are shared, and the network structure of non-full connectivity and weight sharing makes it like the biological neural networks, reducing the complexity of the network model and reducing the number of weights.

The most basic three layers of the convolutional neural network are the convolutional layer, the pooling layer, and the fully connected layer. The above three layers constitute a complete CNN. Each node of the convolutional layer is multiplied with an input matrix called a convolution kernel, which is then weighted and summed. This layer can extract features. After the convolutional layer is generally pooling layer of the dimension to reduce the dimension, pooling will reduce the total amount of parameters. The operation of lowering the dimension is to use the sliding window of $a \times a$ to pick the maximum value of the current window in the input map.

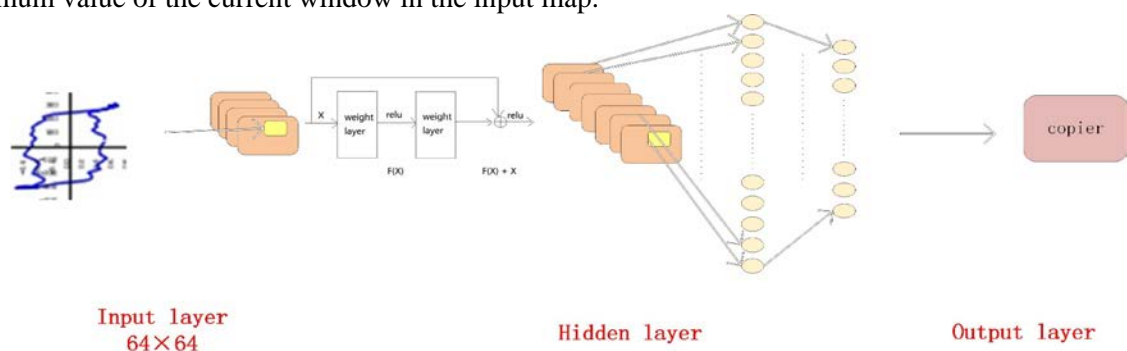


Figure 1. Model Structure

3.3. Residual unit

During the training of convolutional neural networks, it is found that the depth of the neural network is crucial for the classification effect. If you want to further improve the load recognition, you need to deepen the network level. When the network layer is deepened, the gradient will be dispersed. The recognition rate will rise first and then decrease rapidly. In order to avoid this problem, the residual network (ResNet) appears.

There are some characteristics of residual unit as follows:

1. Reducing the network layers in order to control parameters' number.

2. There are obvious levels, and the number of feature maps is progressively layer by layer to ensure the output feature expression ability.
3. Without dropout, regularization with BN and global average pooling speeds up training.
4. When the number of layers is high, the number of 3×3 convolutions is reduced, and the number of input and output feature maps of 3×3 convolution is controlled by $1 \times$ convolution. This structure is called “bottleneck”.

Rather than having every few stacked layers fit directly into the desired underlying mapping, we explicitly make these layers fit into the residual mapping. If the layers behind the deep network are identity maps, the model degenerates into a shallow network. The problem to be solved now is to learn the identity mapping function $H(x) = x$, which may be the reason why deep networks are difficult to train. However, if the network is designed as $H(x) = F(x) + x$, as shown below. We can convert to learn a residual function $F(x) = H(x) - x$. As long as $F(x) = 0$, we form an identity map $H(x) = x$.

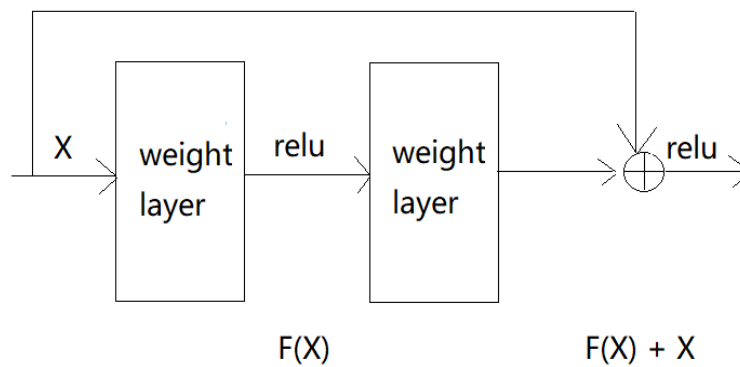


Figure 2.Residual unit structure

4. Experience

In order to verify that the addition of residual units in the traditional convolutional neural network structure is effective, we conducted related experiments.



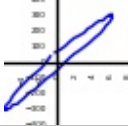


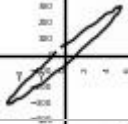
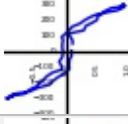
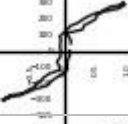
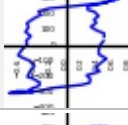
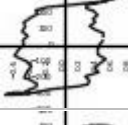


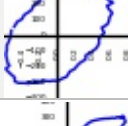


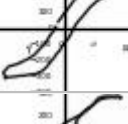
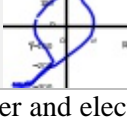
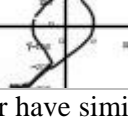
4.1. Data pre-treatment

First, we need a large number of data sets to train the model. In order to extract the effective features better, we use two signals of voltage and current, which are combined into a VI trajectory image. The sampling rate of the sub-meters V and I of the data is 4 kHz, that is, 80 data points are acquired within 0.02s. The experimental data includes 21 electrical appliances and transient data in no-load state, each of which is turned on and off at least 100 times. Then, from the transient data of each appliance, the data is divided into one cycle, and the current and voltage data in each cycle are plotted in the same coordinate system. We get 1000 VI trajectory image for each appliance. As an input, the image has a size of 64×64 , of which 800 images are used as training sets and 200 images are used as test sets.

The appliances including: 1+ mobile phone, rice cooker, refrigerator, filter, oxygen machine, vacuum cleaner, hair dryer, copier, mini fridge, microwave oven, printer, hanging machine, kettle, air conditioner, electric heater, computer, TV, notebook, induction cooker, Nokia tablet, drinking fountain, fan.

The table 1 turn images into grayscale because colour will become a feature that is useless for identifying.

Table 1.Artwork and its grayscale images

Appliances	Artwork	Grayscale
1+mobile phone		
Hair Dryer		
Electric heater		
TV		
Copier		
Oxygen Machine		
Mini Fridge		
Vacuum Cleaner		
Microwave Oven		

From the table 1, we can see that hair dryer and electric heater have similar current waveform. At this time, The axes can be used as a reference to distinguish between the two appliances.

4.2. Model structure

The network structure proposed in this paper is as follows: VI track map with size 64×64 is taken as input. The hidden layer is: one convolutional layer that kernel is 3×3 , one pooling layer, residual unit, one pooling layer, one convolutional layer that kernel is 3×3 , one pooling layer, one fully connected layer of 1024 units, finally an output layer of 22 units. Batch is 50 at a time. Output's units depend on the types of appliances.

Table 2.Parameters of each layer of the model

Layer	Type	Input Size	Kernel Size	Feature map/Hidden Layer	Output Size
1	Convolution	64×64	3×3	64	64×64
2	Pooling	64×64	2×2	64	32×32
3	Residual unit	32×32	2×2	64	32×32
4	Pooling	32×32	2×2	64	16×16
5	Convolution	16×16	3×3	128	16×16
6	Pooling	16×16	2×2	128	8×8
7	Fully connected layers	$8 \times 8 \times 128$	--	1024	1024
8	Fully connected layers	512	--	22	55

4.3. Training

The model is built. After the model is initialized, CNN learns the classification of electrical appliances. In this process, various types of training data are needed. These training data are the input VI track map $X=(X_1, X_2, \dots, X_n)$, each of which corresponds to a type label $Y=(Y_1, Y_2, \dots, Y_n)$, where the label Y_i is a group One-hot encoding. The purpose of training the model is to optimize the weight and offset values within the model, the cost function reaches a minimum.

The cost function:

$$-\sum_{i=1}^N \sum_{k=1}^K t_{i,k} \log(y_{i,k}) \quad (1)$$

The cost function uses the AdamOptimizer descent method to update the minimum value, update and calculate the network parameters that affect the model training and model output, so that it approximates or reaches the optimal value, thus minimizing the loss function. This algorithm uses the gradient of each parameter. Value to minimize or maximize the loss function.

4.4. Result

4.4.1. Evaluation Standard. There are more than one label of a picture in a multi-label image classification task, so the evaluation cannot use the standard of ordinary single-label image classification---- mean accuracy. This task uses a similar method to information retrieval—mAP (mean Average Precision)----average the AP value that is an average AP value for multiple verification sets. As shown below:

mAP can be understood in three parts: P, AP, MAP:

$$\text{Presion}_c = \frac{N(\text{TruePositives})_c}{N(\text{TotalObjections})_c} \quad (2)$$

$$\text{AveragePresion}_c = \frac{\sum \text{presion}_c}{N(\text{TotalImages})_c} \quad (3)$$

$$\text{MeanAveragePresion} = \frac{\sum \text{AveragePresion}_c}{N(\text{classes})} \quad (4)$$

4.4.2. Confusion matrix. Figure 3 is the result on dataset represented by confusion matrix. The Y-axis is the actual label of the appliances, and the X-axis is the label predicted by the model. From this figure, we can observe the probability that the appliance is misjudged into other label.

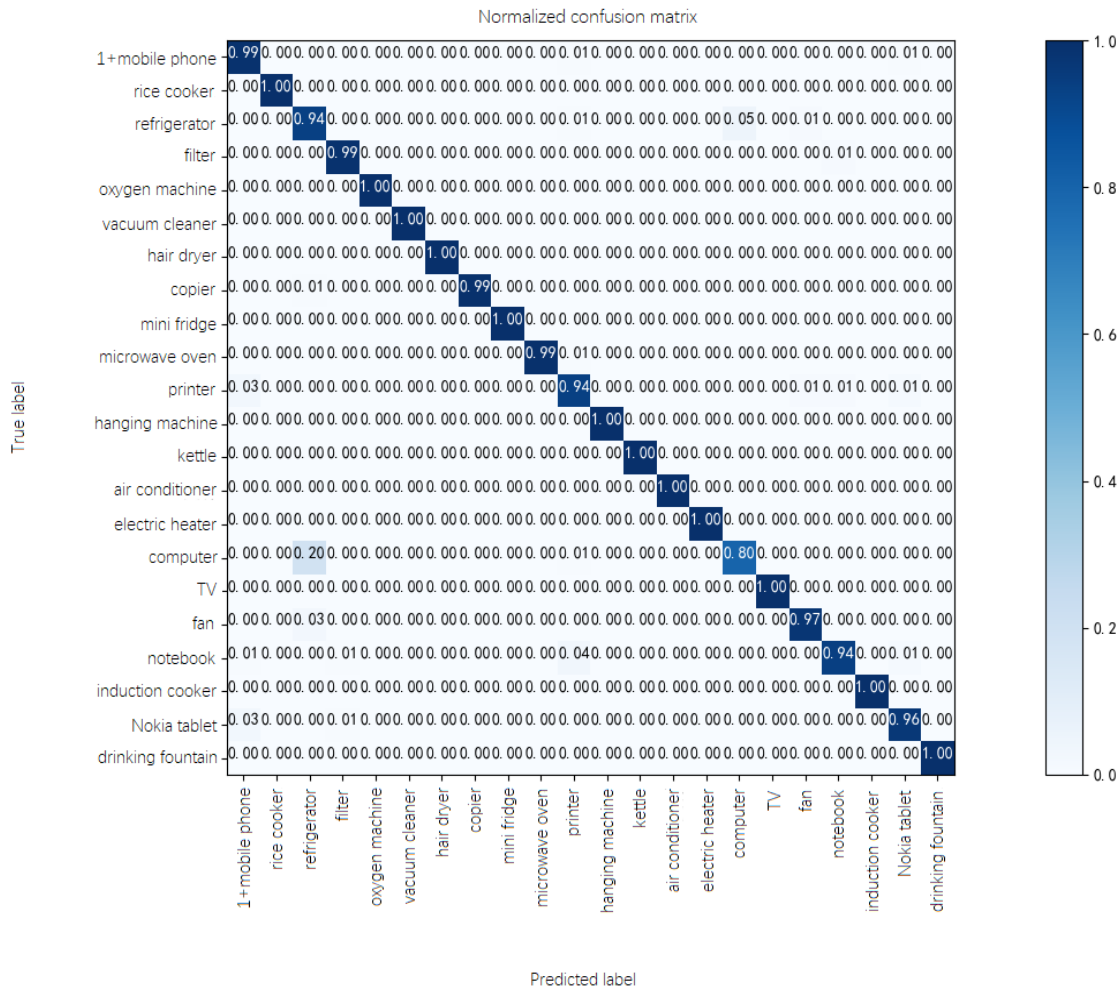


Figure 3. Confusion matrix

5. Conclusion

In this paper, a novel method is used to input the VI trajectory map as a model, and the residual unit is added to the traditional network to eliminate the gradient dispersion. Applying this method to the dataset of 22 electrical appliances, the result reached 97.9%. For a large number of electrical appliances, there is a good classification effect.

Acknowledgments

This work was supported in part by China International Scientific and Technological Cooperation Special (2015DFR10490), Qingdao national laboratory for marine science and technology Aoshan science and technology innovation project (2016ASKJ07-4) and Qingdao innovation and entrepreneurship leading talent project (13-cx-2).

References

- [1] Kim J, Le T T, Kim H. Nonintrusive Load Monitoring Based on Advanced Deep Learning and Novel Signature[J]. Computational Intelligence & Neuroscience, 2017, 2017(2-3):4216281.W.
- [2] Chang H H. Non-Intrusive Demand Monitoring and Load Identification for Energy Management Systems Based on Transient Feature Analyses[J]. Energies, 2012, 5(11):4569-4589.
- [3] H.Najmeddine,K.ElKhamlichiDrissi,C.Pasquieretal.,“State of art on load monitoring methods,”in Proceedings of the 2008 IEEE 2nd International Power and Energy Conference, PECon 2008,pp.1256–1258,mys,December2008.

- [4] M. B. Figueiredo, A. De Almeida, and B. Ribeiro, “An experimental study on electrical signature identification of non-intrusive load monitoring (nilm) systems,”*International Conference on Adaptive and Natural Computing Algorithms*, pp.31–40, 2011.
- [5] Cole, A.I.; Albicki, A. Data extraction for effective non-intrusive identification of residential power loads. In *Proceedings of the IEEE Instrumentation and Measurement Technology Conference (IMTE/98)*, St. Paul, MN, USA, 18–21 May 1998; pp. 812–815.
- [6] Cole, A.I.; Albicki, A. Algorithm for non-intrusive identification of residential appliances. In *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS’98)*, Monterey Conference Center, Monterey, CA, USA, 31 May–3 June 1998; pp. 338 – 341.
- [7] Kong L, Yang D, Luo Y. Non-Intrusive Load Monitoring and Identification Based on Maximum Likelihood Method[C]// *IEEE International Conference on Energy Internet*. IEEE, 2017:268-272.
- [8] Tabatabaei S M, Dick S, Xu W. Towards Non-Intrusive Load Monitoring via Multi-Label Classification[J]. *IEEE Transactions on Smart Grid*, 2017, PP(99):1-1.
- [9] Shaw S R, Leeb S B, Norford L K, et al. Nonintrusive Load Monitoring and Diagnostics in Power Systems[J]. *IEEE Transactions on Instrumentation & Measurement*, 2008, 57(7):1445-1454.
- [10] Baets L D, Ruysinck J, Develder C, et al. Appliance classification using VI trajectories and convolutional neural networks[J]. *Energy & Buildings*, 2018, 158.