Bypass Analysis Based on Convolutional Neural Networks

Jingwen Huang *

Department of Physics, North China University of Technology, College of Artificial Intelligence and Computer Science, Beijing, 111000, China

* Corresponding Author Email: hellowen0927@outlook.com

Abstract. Side-channel analysis, as a key branch of cryptanalysis, poses a serious threat to the security of embedded smart devices. With the development of artificial intelligence technology, side-channel cryptanalysis methods based on convolutional neural networks (CNNSCA) have become a research hotspot. This paper focuses on the modeling and analysis of CNNSCA, exploring its application in side-channel cryptographic attacks. In traditional side-channel analysis techniques, non-modeling methods such as Differential Power Analysis (DPA) and Correlation Power Analysis (CPA) require domain knowledge to construct guess models; modeling methods like Template Attacks (TA) involve template construction and matching phases. CNNSCA leverages the advantages of deep learning to automatically extract data features. Experimental results demonstrate that CNNSCA outperforms traditional TA when attacking ideal low-noise side-channel signals, and shows greater advantage in attacking distorted side-channel signals due to noise and countermeasures, providing new avenues for improving the efficiency of side-channel cryptographic attacks and countering protective strategies, which is of significant importance for ensuring the security of cryptographic systems.

Keywords: Convolutional Neural Networks, Side-Channel Analysis, VGGNet, AlexNet.

1. Introduction

Side Channel Analysis (SCA) refers to bypassing the cumbersome analysis of encryption algorithms by utilizing the information leaked during computations from the hardware implementation of cryptographic algorithms, and quickly cracking the cryptographic system integrated with statistical theory. The corresponding attack methods are referred to as side channel attacks.

Convolutional Neural Networks (CNN) learn to optimize self-composed neurons [1]. Each neuron receives inputs and performs operations (such as scalar multiplication followed by a non-linear function). From the raw image vector input to the final output score for categories, the entire network expresses a single perceptual score function (weights). The last layer contains the loss function associated with the categories. The basic architecture of CNN includes the input layer, convolutional layer, activation function layer, pooling layer, fully connected layer, and output layer.

Channel Cryptanalysis methods (SCA) are divided into modeling methods and non-modeling methods [2]. Modeling methods include template attacks, side-channel attacks based on multilayer perceptrons, and side-channel attacks based on Convolutional Neural Networks (CNN SCA).

Deep learning technologies have made significant advancements in recent years, with various advanced CNN model structures continuously emerging. Cryptographic technology is widely used in critical fields such as finance, power, and transportation, which have higher security requirements for cryptographic devices. At the same time, attackers are continually seeking more effective means to obtain sensitive information. This paper introduces mature CNN model structures from the deep learning field into the side channel analysis domain and optimizes them to tackle the non-stationary and high noise characteristics of side channel signals (such as power consumption, electromagnetic radiation, etc.), breaking the reliance on manual feature extraction of traditional side channel analysis techniques and addressing limitations such as low analysis efficiency and weak interference resistance. This cross-domain technology integration not only expands the application boundaries of deep learning technology but also allows advanced neural network models to play a practical role in

cryptographic security assessment scenarios, providing critical technical support for the intelligent development of side channel analysis technology.

2. Classification of Side-Channel Analysis Methods Based on CNN

2.1. Classification Based on Network Structure

Side-channel analysis methods can be categorized based on network structures such as AlexNet and VGGNet. The Alex-CNNSCA network model is a variant based on AlexNet, capable of attacking unprotected encryptions, but it is not very effective in obtaining keys from protected targets. The VGG-CNNSCA model, based on VGGNet, shows certain effectiveness in key retrieval for unprotected targets with various parameters; however, the results for key retrieval in experiments targeted at datasets with first-order masking and random delay protection are also not very promising.

2.2. Classification Based on Analysis Process

The analysis processes can be divided into modeling-based and non-modeling-based side-channel analysis methods. Modeling-based methods, such as template attacks, involve constructing a template before carrying out a matching attack. Non-modeling methods, such as Differential Power Analysis (DPA) and Correlation Power Analysis (CPA), exploit the correlation between the power consumption of cryptographic devices during operation and the values of the processed data [3]. They extract key information through statistical differential analysis without the need to pre-construct device-specific power consumption templates.

2.3. Classification Based on Signal Type

Based on signal types, side-channel analysis can be categorized into power analysis and electromagnetic analysis. Power signals are common subjects of study in side-channel analysis. When cryptographic devices perform computations, they exhibit varying power consumption dependent on data processing. Convolutional Neural Network (CNN) models, with their collaborative multilevel architecture, can capture these subtle power fluctuation characteristics. Electromagnetic radiation similarly contains rich information about cryptographic operations. While a cryptographic device is in operation, it generates electromagnetic signals of varying frequencies and intensities. CNN models can perform in-depth analysis of these electromagnetic signals, with convolutional layers identifying specific frequency components or intensity variation characteristics caused by cryptographic operations.

2.4. Comparative Analysis and Discussion

From the experimental comparison data in Table 1, it can be observed that various CNN methods used for side-channel attacks share significant commonalities in core parameters: the activation function is predominantly RELU, with ResNet under electromagnetic analysis additionally using softmax. This indicates that RELU effectively avoids the vanishing gradient problem when extracting features from side-channel signals (power consumption, electromagnetic), adapting well to the feature learning requirements in attack scenarios. Furthermore, most methods have a convolutional layer count focused at 5 layers (VGG-CNNSCA, Alex-CNNSCA, ResNet), while the custom lightweight architecture based solely on non-modeling DPA has only 3 layers, reflecting that a 5-layer convolutional structure is more effective in balancing 'feature capturing capability' and 'computational efficiency' in side-channel signal feature extraction.

Table 1. Experimental Comparison Analysis

Method	CNN	Activation function	Convolutional layer	Learning rate	Title
Classification based on network structure	VGG-CNNSCA	RELU	5{64,128,256,512,512}	10^{-3}	VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION [4].
	Alex-CNNSCA	RELU	5{96,256,384,384,256}	10^{-2}	A new method for template attacks on encrypted chips based on the AlexNet convolutional neural network [5].
Based on the classification of analytical processes	based on modeling	RELU	Each CNN model's convolutional layer is the same as the one mentioned above.	10^{-3}	Research progress and analysis of modeling methods for side-channel password attack [2].
	Based on non-modeling methods DPA.	RELU	3 kernels, with sizes (3, 3), (5, 5), (3, 3)	10^{-4}	Differential Power Analysis[3].
classification based on signal type	Power consumption analysis,	VGGCNNSCA, AlexCNNSCA, custom lightweight CNN architecture	Each CNN model's convolutional layer is the same as the one mentioned above.	10^{-4}	S-Box Power Consumption Randomization Side-Channel Attack Based on CNN-BPR [6].
	electromagnetic analysis	ResNet	5 pieces, kernel size 3×9	10^{-3} -10^{-2}	Research on Electromagnetic Attacks on AES Cryptographic Chips Based on Deep Residual Neural Networks [7].

3. Security Assessment Indications for Cryptographics attacks Based on Convolutional Neural Networks

Guessing Entropy (GE) is a quantitative measure of 'key uncertainty' in side-channel attacks, used to quantify the efficiency of attack methods [8]. A smaller value indicates that the attack can successfully locate the correct key with fewer traces, reflecting the effectiveness of the attack method. In side-channel attacks, the attacker sorts all possible key candidates based on the acquired side-channel signals (such as power consumption, electromagnetic radiation, etc.) in terms of probability, placing the most likely candidate for the correct key at the top, the next most likely in second place, and so on.

Attack success rate. The attack success rate is a core indicator measuring the "effectiveness of attack results". It specifically refers to the proportion of successful experimental attempts in which an attacker, through the analysis of side-channel signals, correctly infers the complete key (rather than just a portion of the key bytes) under fixed experimental conditions (such as fixed key length, fixed side-channel signal type, fixed signal acquisition parameters, and fixed attack algorithms) to the total number of experimental attempts.

The required number of samples (also known as 'the number of attack trajectories required') is a key indicator for measuring 'attack cost efficiency.' Specifically, it refers to the number of side channel signals (i.e., sample size) that the attacker needs to collect to achieve a predetermined attack objective (such as reaching a 90% attack success rate or reducing the guessing entropy to 5).

4. Characteristics and Analysis of the Existing Models of CNNSCA

I With the rapid development of deep learning technologies, CNNs have significantly improved the efficiency and success rate of side-channel attacks due to their strong feature extraction and pattern recognition capabilities [2] [9]. To adapt to different side-channel analysis scenarios and data characteristics, researchers have developed various structures of CNN-SCA models.

In the ILSVRC-2014 competition, the VGG proposed by Karen et al. had a significant impact on the field of image recognition. Drawing on the advantages demonstrated by VGGnet in image classification, its network structure was adapted to the side-channel analysis scenario. Utilizing the feature extraction capabilities of convolutional neural networks, VGG-CNNSCA emerged as a deep learning-based side-channel attack method. Its structure is centered around "convolution blocks and fully connected layers", providing a practical foundation of deep networks for side-channel cryptanalysis. However, while this model ensures the integrity of feature extraction, it incurs substantial computational overhead. Therefore, future research needs to find a balance between "effectiveness" and "cost"[4].

Guo et al. changed the structure of the neural network by adjusting the hyperparameters and designed Alex-CNNSCA based on the AlexNet model, providing a new paradigm for side-channel attacks on cryptographic chips, especially suitable for high-dimensional and high-noise side-channel signal analysis [5].

Combining the advantages of the VGG-CNNSCA model and the Alex-CNNSCA model, Yun Lin Liu and others designed CNNSCAbase [2]. They embedded the SE module for the first time in this model, effectively alleviating the gradient dispersion problem in the backpropagation of deep networks, enhancing the model's classification performance and reducing training time.

Based on the CNNSCAbase, the final optimized model obtained through hyperparameter optimization has been designed as the CNNSCAnew model. Its optimization includes reducing the number of convolution kernels (from 64 to 32) and the number of channels in the fully connected layer (from 4096 to 1024), thus lowering the consumption of computing resources.

In a study by Ryad Benadjila and others [10,11], ResNet-50 was compared with VGG-16 and Inception-v3. Through cross-validation, it was found that VGG-16's vulnerability performance was significantly better than the other two networks. This brought important insights: a more complex model does not necessarily yield better results. The characteristics of side-channel signals are 'relatively low in dimensions and concentrated on sensitive features.' The residual connections of ResNet-50 and the multi-branch structure of Inception are better suited for handling high-dimensional and dispersed feature data like images, which can be redundant in side-channel signal analysis; whereas VGG-16's simple stack of convolution blocks can focus more on extracting key sensitive features from the signal, leading to superior attack effectiveness. This also reminds researchers that the design of CNN-based side-channel attack models should not blindly pursue 'advanced architectures' but should be tailored to the characteristics of side-channel signals.

5. Limitations That Exist

The training accuracy is relatively low. For instance, models like Alex-CNNSCA and VGG-CNNSCA can perform cryptanalysis, but their training accuracy is not high. The models' ability to learn and extract password features needs improvement, which results in a certain limit on their attack effectiveness. In the future, a design that combines 'model and algorithm' could be experimented with: by adding step labels of the encryption algorithm to the input layer of the CNN, the model can more

specifically extract key features of that step, thereby improving training accuracy and breaking through the limits of attack effectiveness.

The consumption of computing resources is high. Due to its complex network structure and a large number of computation operations, training and running the CNNSCA model requires a significant amount of computing resources, which limits its range of applications.

Data dependency is strong. CNNSCA usually requires a large amount of side-channel trace data for training in order to learn sufficient cryptographic feature patterns. If the data volume is insufficient, the model is prone to overfitting, resulting in poor generalization ability in real attack scenarios, making it unable to effectively cope with cryptographic devices or algorithms under different conditions.

6. Summary In Conclusion

Convolutional neural networks (CNNs) demonstrate significant advantages in side-channel attack analysis due to their unique architecture and powerful feature extraction capabilities. This paper introduces several CNNSCA methods that have successfully broken cryptographic codes so far. The evolution of existing models revolves around the "balance between effectiveness and cost," while future development needs to resolve the "trilemma of accuracy, resources, and data." From the perspective of analysis methods, the application of univariate analysis, multivariate analysis, and adversarial masking techniques allows CNNs to adapt to different scenarios, with existing CNNSCA models showing a trend of continuous optimization. However, their limitations also indicate that future efforts should continue exploring areas such as model structure optimization, computational efficiency improvement, and enhancement of small sample learning capabilities to better meet the actual needs of cryptographic system security assessments, which also has certain academic significance for information security and cryptographic protection.

References

[1] O'Shea K, Nash R. An introduction to convolutional neural networks. arXiv preprint arXiv: 1511.08458, 2015.

[2] Fangfang. Research on power load forecasting based on improved BP neural network. Harbin Institute of Technology, 2011.

[3] Liu Linyun, Chen Kaiyan, Li Xiongwei, Zhang Yang, Xie Fangfang. Overview of side channel analysis based on convolutional neural network. Computer Science, 2022, 49 (5): 296 - 302.

[4] Kocher P, Jaffe J, Jun B. Differential power analysis. In: Annual International Cryptology Conference, 1999: 388 - 397. Berlin, Heidelberg: Springer Berlin Heidelberg.

[5] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Computer Science, 2014.

[6] Guo Dongxin, Chen Kaiyan, Zhang Yang, et al. new template attack method for encryption chip based on VGGNet convolutional neural network. Application Research of Computers, 2019, 36 (9): 2809 - 2812.

[7] Cao Jiahua, Wu Zhen, Wang Yi, et al. S-Box power randomization side channel attack based on CNN-BPR. Journal of Chengdu University of Information Technology, 2022, 37 (1): 5.

[8] Luo Man, Zhang Hongxin. Research on electromagnetic attacks of AES cryptographic chips based on deep residual neural networks. Journal of Electromagnetic Waves and Applications, 2019, 34 (4): 5.

[9] Xiao Chong, Tang Ming. A review of side channel analysis based on deep learning. Journal of Computer Research and Development, 2025 (3).

[10] Maghrebi H, Portigliatti T, Prouff E. Breaking cryptographic implementations using deep learning techniques. In: Carlet C, Hasan M, Saraswat V, eds. Security, Privacy, and Applied Cryptography Engineering. SPACE 2016. Lecture Notes in Computer Science, vol 10076. Springer, Cham, 2016.

[11] Benadjila R, Prouff E, Strullu R, et al. Deep learning for side-channel analysis and introduction to ASCAD database. Journal of Cryptographic Engineering, 2020, 10: 163 – 188.