International Journal of Science and Engineering Applications
Volume 15-Issue 05, 104 – 106, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1505.1018

Research and Application of Siamese Neural Network

Si Yang Pei
Yangtze University
Jingzhou Hubei, China

Abstract: A Siamese Neural Network (SNN) is a specialized deep learning framework characterized by weight sharing and similarity learning. Compared with traditional neural networks, it overcomes the limitations of heavy dependence on large-scale labeled datasets and poor performance in similarity comparison tasks. Owing to their strong capability in feature representation and metric learning, SNNs have been widely applied in pattern recognition, object tracking, medical diagnosis, and few-shot learning. This paper introduces the Siamese Neural Networks from four aspects: fundamental concepts and network structures, improvement strategies and key technologies, major application fields, and future development trends. Furthermore, the advantages and current limitations of SNNs are summarized, providing a reference for subsequent theoretical research and engineering applications.

Keywords: Siamese Neural Network; similarity learning; contrastive loss; deep learning

1. Introduction

With the rapid development of deep learning technology, neural networks have been widely used in computer vision, natural language processing, pattern recognition and other fields.Traditional neural networks mostly adopt a single-input structure for classification tasks. However, they have obvious limitations in similarity judgment: they require large amounts of labeled data and struggle to learn deep similarity features between samples, making it difficult to meet practical demands such as signature verification, face matching and object tracking. To address this issue, Bromley et al. first proposed the Siamese Neural Network in 1993 and applied it to handwritten signature verification[1].By constructing two subnetworks with identical structures and shared weights, the model can effectively measure the similarity between two input samples. This architecture provides a novel solution for similarity-based learning tasks and lays the foundation for subsequent metric learning research.

Following this work, researchers have continuously improved Siamese Neural Networks. Chopra et al. applied SNNs to face verification and proposed a discriminative similarity metric learning strategy[2].Hadsell et al. further improved the robustness of feature mapping and proposed invariant metric learning methods [3] . In 2015, Koch et al. successfully introduced Siamese Neural Networks into few-shot learning tasks, demonstrating their effectiveness under limited labeled data conditions[4] . In the same year, Hoffer et al. proposed the Triplet Network, which significantly improved feature discrimination capability through triplet-based metric learning[5].

In recent years, Siamese Neural Networks have evolved rapidly with the integration of advanced deep learning techniques such as deep convolutional networks, attention mechanisms, and Transformer architectures. Their application scope has expanded from traditional verification tasks to object tracking, image retrieval, medical diagnosis, and autonomous driving. Therefore, conducting a systematic review of Siamese Neural Networks is of great significance for understanding their development trends and promoting further applications.

2. Concept and Structure of Siamese Neural Network

2.1 Concept of Siamese Neural Network

A Siamese Neural Network is a specialized deep learning architecture designed for similarity measurement tasks. Unlike conventional classification networks, SNNs focus on learning the relationship between two samples rather than directly predicting category labels.

The core idea of SNNs is to use two subnetworks with identical structures and shared parameters to extract feature representations from paired inputs. After mapping the input samples into a common feature space, the network evaluates their similarity through a distance metric. During training, similar samples are pulled closer together, while dissimilar samples are pushed farther apart. Owing to this architecture, Siamese Neural Networks demonstrate excellent performance in verification, matching, and few-shot learning tasks.

2.2 Network Structure

This paper mainly focuses on the narrow definition of the Siamese neural network, whose structure is illustrated in Figure 1.

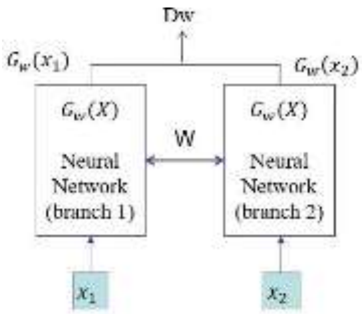

Figure 1

As shown in Figure 1, the Siamese neural network consists of two identical subnetworks, also called branch networks, which share the same weights . Each branch network is responsible for extracting feature representations from the input samples, denoted as () . The similarity between two samples is determined by calculating the distance between their corresponding feature vectors, represented as . Common distance metrics include the Manhattan distance (L1 norm) and Euclidean distance (L2 norm).

During training, a loss function is used to evaluate the performance of the network, while the backpropagation algorithm updates the network parameters. Unlike traditional classification networks, the main objective of a Siamese neural network is to distinguish between similar and dissimilar samples rather than directly classify them. Therefore, the network is commonly trained using the contrastive loss function, which measures the ability of the model to discriminate between pairs of input samples.

The contrastive loss function can be expressed as:

$$o() = (1 -) \frac{1}{2}{}^{2} + \frac{1}{2} \max(0, -)^{2}$$

where $= 0$indicates that the two samples are similar, and the loss decreases as the distance becomes smaller.

Conversely, $=1$ indicates that the samples are dissimilar, and represents the predefined margin parameter. When the distance between dissimilar samples approaches the margin , the loss tends toward zero, indicating effective discrimination between different samples.

3. Improvements of Siamese Neural Network

Although the original Siamese Neural Network achieved promising results in similarity learning, it still suffers from several limitations, including insufficient feature extraction capability, inaccurate similarity measurement in complex scenarios, and poor robustness under imbalanced data distributions. To address these problems, researchers have proposed various improvement strategies focusing on network structure, feature extraction, similarity measurement, and loss function optimization.

3.1 Improvement of Feature Extraction Sub-network

The feature extraction capability directly determines the performance of Siamese Neural Networks. Early Siamese models mainly adopted shallow CNN structures, which had limited representation ability for complex features.

To improve feature learning performance, researchers introduced deeper convolutional architectures such as ResNet[9]. Residual connections effectively alleviate gradient vanishing problems and enable networks to learn richer semantic information. Integrating ResNet into Siamese architectures significantly improves feature discrimination capability, especially in object tracking and face recognition tasks.

In addition, feature fusion strategies have been widely adopted to combine shallow local features with deep semantic information. Attention mechanisms further enhance important feature regions while suppressing redundant information, thereby improving robustness under complex backgrounds and occlusion conditions.

These improvements indicate that Siamese Neural Networks are gradually evolving from shallow similarity comparison models toward deep semantic representation learning framework.

3.2 Improvement of Similarity Measurement Layer

Traditional similarity measurement methods, such as Euclidean distance and cosine similarity, mainly capture linear relationships between features. However, they often fail to represent complex nonlinear relationships in high-dimensional feature spaces.

To solve this problem, Chopra et al. proposed a discriminative metric learning strategy that jointly optimizes feature extraction and similarity measurement. Hadsell et al. further introduced invariant feature learning methods to improve robustness against transformations such as translation, rotation, and scaling .

By learning task-specific similarity metrics, improved Siamese Neural Networks achieve stronger generalization ability and better adaptability to complex scenarios.

3.3 Improvement of Loss Function

Loss functions play a crucial role in Siamese Network optimization. Traditional contrastive loss often suffers from inefficient convergence and sample imbalance issues.

To address these limitations, Hoffer et al. proposed the Triplet Network architecture. Instead of using paired samples, the network adopts triplet samples consisting of anchor, positive, and negative examples, enabling the model to learn more discriminative feature representations.

FaceNet further introduced online hard example mining strategies to dynamically select challenging training samples[8]. This significantly improves training efficiency and enhances the discriminative capability of the embedding space.

3.4 Other Improvement Directions

In addition to the above improvements, researchers have explored enhancing network efficiency and flexibility. For instance, the Fully-Convolutional Siamese Network (SiamFC) removes fully connected layers to achieve faster inference speed, making it suitable for real-time applications. The Siamese Region Proposal Network (SiamRPN) further improves localization accuracy by integrating region proposal mechanisms.

Moreover, Siamese Neural Networks have been successfully combined with few-shot learning techniques, enabling effective learning from limited labeled data. This significantly expands their applicability to domains such as medical diagnosis and rare object recognition, where labeled data are scarce.

4. Applications of Siamese Neural Network

With continuous technological development, Siamese Neural Networks have been widely applied in various fields due to their excellent similarity learning capability.

4.1 Pattern Recognition Field

Pattern recognition is the earliest application field of Siamese Neural Networks, including signature verification, face verification, fingerprint recognition and other tasks. The core is to realize accurate identity recognition by comparing the features of samples.

Signature verification is the first application scenario of the siamese neural network. In 1993, Bromley et al. first applied the siamese network to handwritten signature verification, and realized the recognition of forged signatures by extracting the temporal and morphological features of signatures and comparing the similarity of two signatures. This research provides a brand-new technical solution for signature verification, which has higher accuracy and robustness compared with the traditional manual feature extraction method.

Face verification is one of the most widely used scenarios of siamese neural networks. In 2005, Chopra et al. applied the siamese network to face verification, and realized identity verification by extracting the deep features of human faces and comparing the similarity of two human faces. FaceNet proposed by Schroff et al., based on the improved siamese network architecture, adopts the triplet loss and online hard example mining strategy, realizing high-precision face verification and clustering, and is widely used in scenarios such as access control systems and mobile phone unlocking. In addition, the siamese network is also applied to biometric identification tasks such as fingerprint recognition and iris recognition, realizing accurate identity authentication by comparing the similarity of biometric features.

4.2 Object Tracking Field

Object tracking is one of the most important application areas of Siamese neural networks in computer vision. By comparing the similarity between a target template and candidate regions in video frames, Siamese networks can accurately locate and track moving objects. Compared with traditional tracking methods, Siamese-based approaches exhibit stronger robustness against target occlusion, background interference, and appearance changes.

Representative models such as **SiamFC** and **SiamRPN** further improve real-time performance and target localization accuracy by integrating fully convolutional structures and region proposal mechanisms. Owing to their high efficiency and robustness, Siamese-based tracking algorithms have been widely applied in intelligent surveillance, autonomous driving, and UAV systems.

4.3 Few-shot Learning Field

Few-shot learning is an important research direction in the field of deep learning, whose core is to realize the recognition and classification of new categories by using a small number of labeled samples, applicable to scenarios with scarce labeled data (such as medical diagnosis and rare species recognition). Relying on the characteristic of weight sharing, the siamese neural network can effectively use a small number of samples for training, becoming one of the core architectures of few-shot learning.

In 2015, Koch et al. applied the siamese neural network to few-shot image recognition, and realized the recognition of new categories by training the network to learn sample similarity measurement. This study shows that the siamese neural network can achieve a high recognition accuracy with only a small number of labeled samples, and has stronger data adaptability compared with traditional neural networks.

The triplet network proposed by Hoffer et al. further optimizes the performance of the siamese network in few-shot learning. Through triplet sample training, it improves the network's ability to capture sample feature differences, achieving good results in tasks such as few-shot image classification and text classification. At present, the siamese neural network has become one of the mainstream architectures of few-shot learning, providing an important support for research in fields with scarce labeled data.

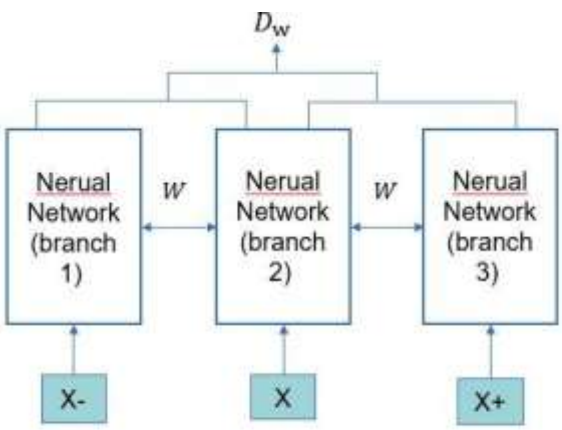

Figure 2

4.4 Other Application Fields

In addition to the above fields, the siamese neural network is also applied to many fields such as text matching, medical diagnosis and image retrieval. In the field of text matching, by converting texts into vectors and using the siamese network to compare the similarity of text vectors, tasks such as text duplicate checking and semantic matching are realized; in the field of medical diagnosis, by comparing the similarity of medical images (such as CT images and MRI images), doctors are assisted in disease diagnosis, improving the accuracy of diagnosis; in the field of image retrieval, by extracting image features and using the siamese network to compare the similarity of images, efficient image retrieval is realized.

5. Conclusion

This paper systematically reviews the concepts, structures, improvement strategies, and application fields of Siamese Neural Networks. As a representative framework for similarity learning and metric learning, Siamese Neural Networks effectively overcome the limitations of traditional neural networks in similarity comparison tasks through weight sharing and feature embedding mechanisms.

Recent studies have significantly improved SNN performance by introducing deep feature extraction architectures, attention mechanisms, feature fusion strategies, and optimized loss functions. These improvements have expanded their applications from traditional verification tasks to object tracking, few-shot learning, medical diagnosis, and autonomous driving.

Nevertheless, several challenges still remain. First, Siamese Neural Networks still face performance degradation under complex scenarios such as severe occlusion and highly imbalanced datasets. Second, deep architectures and complex loss functions often increase computational costs and training complexity. Third, multimodal similarity learning still requires further optimization.

Future research directions mainly include lightweight network design, multimodal feature fusion, Transformer-based Siamese architectures, and the integration of reinforcement learning techniques. With continuous advancements in deep learning, Siamese Neural Networks are expected to play an increasingly important role in intelligent systems and practical engineering applications.

6. References

[1] Bromley J., Bentz J., Bottou L., et al. Signature verification using a “Siamese” time delay neural network. *International Journal of Pattern Recognition and Artificial Intelligence*, 1993.

[2] Chopra S., Hadsell R., LeCun Y. Learning a similarity metric discriminatively, with application to face verification. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.

[3] Hadsell R., Chopra S., LeCun Y. Dimensionality reduction by learning an invariant mapping. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2006.

[4] Koch G., Zemel R., Salakhutdinov R. Siamese neural networks for one-shot image recognition. *ICML Deep Learning Workshop*, 2015.

[5] Hoffer E., Ailon N. Deep metric learning using triplet network. *International Workshop on Similarity-Based Pattern Recognition*, 2015.

[6] Bertinetto L., Valmadre J., Henriques J., Vedaldi A., Torr P. Fully-convolutional Siamese networks for object tracking. *European Conference on Computer Vision (ECCV)*, 2016.

[7] Li B., Yan J., Wu W., Zhu Z., Hu X. High performance visual tracking with Siamese region proposal network. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[8] Schroff F., Kalenichenko D., Philbin J. FaceNet: A unified embedding for face recognition and clustering. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[9] He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[10] Liu X., Cui Y. Research on improved Siamese neural network method. *Computer Knowledge and Technology*, 2021.