CapsNets algorithm

Ping Zhang1, Ping Wei1* and Shuhuan Han1

1 School of electrical engineering and information engineering, Lanzhou University of Technology, Lanzhou, Gansu, 730050, China

* Corresponding author’s e-mail: 912751386@qq.com

Abstract: Capsule networks (CapsNets) is a commonly used neural network structure, especially in computer vision research, which may have a significant impact on deep learning. Convolutional neural networks (CNNs) have reached superhuman level in various computer classification tasks such as classification, parsing, object detection, semantic segmentation, and instance segmentation. Traditional deep learning tends to deteriorate when the relevance of detection objects changes. For this kind of situation, this paper improves a new network, namely the capsule network (CapsNets), the network can achieve a higher detection rate in the hostile environment.

1. Introduction

CNN has many problems: for example, the direction of the component and the relative relationship in space are not important to it, it only cares about the characteristics [1]. In addition, CNN has a problem, that is, the pooling layer. Hinton said the performance of the maximum pooling layer was such a huge error that it was a disaster. From a network design perspective, Hinton is right, the pooling layer not only reduces the parameters, but also avoids over-fitting. However, it does discard some information, such as location information. Despite the different angles, your brain can easily identify them as the same object, which CNN doesn't. It can only achieve similar functions by increasing the amount of training data [2].

2. CapsNet Principle

Geoffrey Hinton et al first proposed that CapsNets is in the document "Transformation Encoder" [3] in 2011. But in 2017, Sara Sabour et al. published an article called "Dynamic Routing between Capsules" [4] that introduced a new CapsNets structure. The best performance was achieved on MNIST [5] (a well-known data set for handwritten digital pictures), and a lot better performance were obtained on MultiMNIST (a deformed handwritten digital image dataset of different numbers versus overlap) than CNNs [6].
Figure 1. MultiMNIST image (white) and CapsNets reconstructed image (red + green) [7]

Figure 1 MultiMNIST image (white) and CapsNets reconstructed image (red + green). "R" stands for refactoring and "L" [8] stands for label. For example: this example predicts (upper left) is correct, and the reconstructed image is also correct. But the fifth example prediction is wrong, (5,0) is predicted to be (5,7). Therefore, 5 is correctly reconstructed, but 0 is not [9].

3. Step
A capsule is a vector, which can contain any number of values, and each value represents a feature of an object (such as a picture) to be identified. Combined with the study of traditional CNN earlier, [10] we already know that convolution layer of each value is a convolution of the region and the convolution kernel, this is the result of the linear weighted sum, and there is only one value, and it is also a scalar. In our capsule network, each value is a vector that represents not only the characteristics of the object, but also its direction and state, etc [11]. Now we assume that there are three lower levels of capsules required to deliver a higher level of four capsules [12].

Like a fully connected neural network, each connection to the capsule network is also weighted. In figure 2, W represents weighting. What we need to pay attention to is: [13] C is not a weight. It is called a coupling coefficient. Next, I will explain in detail. The weight now only concerns W. In a neural network, we know that each neuron is a scalar, that is, there is only one digital value, [14] so each weight is both a scalar and a numeric value [15].

But in the capsule network, the vector is represented by each capsule neuron, which means that it contains multiple values (such as [x1, x2, x3, ..., xn], [16] the specific number n is obtained according
to the network design), [17] so the weight $W$ of each capsule neuron should also be a vector. $W$ is still updated based on backpropagation.

The input of the fully connected neural network is a linear weighted summation [18]. This is similar to the capsule network, but the coupling coefficient $C$ is increased during the linear summation phase. The input $S$ of the capsule network is obtained by the following formula.

$$S_j = \sum c_{ij} \hat{u}_{ji}, \quad \hat{u}_{ji} = W_{ij} \mathbf{u}_i$$

Where $\mathbf{u}$ is the output value of the network of the upper layer. $W$ is the weight to be multiplied by each output [19]. $C$ is calculated by the following formula [20]:

$$C_{ij} = \frac{\exp(b_{ij})}{\sum_i \exp(b_{ij})}$$

$C$ is called the coupling coefficient. In order to find $C$[21], we have to figure out $b$ first, $b$ according to the following formula[22]:

$$b_i \leftarrow b_i + \hat{u}_{ji} \cdot \mathbf{v}_j$$

$b$ initial value is 0. Therefore, in the process of forward propagation $S$, [23] we design $W$ as a random value, $b$ initialized to 0 to get $C$, the output of the capsule network in the previous layer is $\mathbf{u}$, with three values, we can get the next layer $S$ [24].

In a fully connected neural network, the activation functions we choose are usually [25]: sigmoid, tanh, etc. But in the capsule network, Hinton constructs a new activation function Squashing, so the output $V$ is calculated as shown below [26]:

$$V_j = \frac{||S_j||}{1+||S||} \frac{S_j}{||S||}$$

4. Simulation

The three pictures on the left are the detection effects under the CapsNet network, [27] and the three pictures on the right are detected under the R-CNN network. [28] It can be seen that the CapsNet network is much more robust [29].

5. Conclusion
Compared to other advanced technologies, CapsNet has the highest success rate for capsule networks with MNIST data sets [30]. Using a smaller data set will be more successful. (By forcing the model to learn the characteristic variables in the capsule, it can more effectively infer possible variables with less training data). The routing-by-agreement algorithm allows us to distinguish among objects with overlapping images [31]. It is easier to understand the image with the activation vector. The capsule network retains information such as homomorphism, hue, pose, albedo, texture, deformation, speed and object position [32].

Acknowledgments
We gratefully acknowledge the assistance of Hong Zhang in revising and improving the paper. We also thank Peng Shen for their excellent comments on the paper.

References
[1] Zhang, X., Zhao, S. G. (2019). Fluorescence microscopy image classification of 2D HeLa cells based on the CapsNet neural network. Medical & biological engineering & computing. 57(6): 1187-1198.
[2] Yang, Y., Ni, W., Sun, Q., Wen, H., Teng, Z. (2013). Improved Cole parameter extraction based on the least absolute deviation method. Physiological measurement. 34(10): 1239.
[3] Wilson, M. G., Beres, A., Baird, R., Laberge, J. M., Skarsgard, E. D., Puligandla, P. S. (2013). Congenital diaphragmatic hernia (CDH) mortality without surgical repair? A plea to clarify surgical ineligibility. Journal of pediatric surgery. 48(5): 924-929.
[4] Nagdewate, A. B., Paraskar, S. R. (2016). Discrimination between magnetizing inrush and Interturn fault current in transformer: Hilbert transform-ANN approach. In 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC). IEEE. pp. 466-469.
[5] Moon, R. R., Dhatrak, R. K. (2014). Extraction of features to discriminate and detect transformer inrush current and faulty condition using ANN. In 2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD). IEEE. pp. 1-5.
[6] Nasr, A., Langer, J. C., & Network, C. P. S. (2012). Influence of location of delivery on outcome in neonates with gastroschisis. Journal of pediatric surgery. 47(11): 2022-2025.
[7] Wang, X., Tan, K., Chen, Y. (2018). CapsNet and Triple-GANs Towards Hyper spectral Classification. In 2018 Fifth International Workshop on Earth Observation and Remote Sensing Applications (EORSA). IEEE. pp. 1-4.
[8] Algamdi, A. M., Sanchez, V., Li, C. T. (2019). Learning Temporal Information from Spatial Information Using CapsNets for Human Action Recognition. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. pp. 3867-3871. IEEE.
[9] Guo, J. L., Fang, F., Wang, W., Ren, F. (2018). EEG Emotion Recognition Based on Granger Causality and CapsNet Neural Network. In 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS). IEEE. pp. 47-52.
[10] Yin, J., Li, S., Zhu, H., Luo, X. (2019). Hyper spectral Image Classification Using CapsNet With Well-Initialized Shallow Layers. IEEE Geoscience and Remote Sensing Letters.
[11] Wang, X., Tan, K., Du, Q., Chen, Y., Du, P. (2019). Caps-TripleGAN: GAN-Assisted CapsNet for Hyper spectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing.
[12] Tobing, E., Murtaza, A., Han, K., Mun, Y. Y. (2018). [Regular Paper] EP-CapsNet: Extending Capsule Network with Inception Module for Electrophoresis Binary Classification. In 2018 IEEE 18th International Conference on Bioinformatics and Bioengineering (BIBE). IEEE. pp. 327-333.
[13] Wang, Q., Xu, C., Zhou, Y., Ruan, T., Gao, D., He, P. (2018). An attention-based Bi-GRU-CapsNet model for hypernymy detection between compound entities. In 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE. pp. 1031-1035.
[14] Lin, A., Li, J., Ma, Z. (2019). On Learning and Learned Data Representation by Capsule Networks. IEEE Access. 7: 50808-50822.

[15] Xiang, C., Zhang, L., Tang, Y., Zou, W., Xu, C. (2018). MS-CapsNet: A novel multi-scale capsule network. IEEE Signal Processing Letters. 25(12): 1850-1854.

[16] Ahmad, A., Kakillioglu, B., Velipasalar, S. (2018). 3D Capsule Networks for Object Classification from 3D Model Data. In 2018 52nd Asilomar Conference on Signals, Systems, and Computers. IEEE. pp. 2225-2229.

[17] Ha, K. W., Jeong, J. W. (2019). Decoding Two-Class Motor Imagery EEG with Capsule Networks. In 2019 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE. pp. 1-4.

[18] Huang, Y., Hu, R., Zeng, Z. (2018). Three-Dimensional Memristor-Based Crossbar Architecture for Capsule Network Implementation. In 2018 Eighth International Conference on Information Science and Technology (ICIST). IEEE. pp. 170-175.

[19] Afshar, P., Mohammadi, A., Plataniotis, K. N. (2018, October). Brain tumor type classification via capsule networks. In 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE. pp. 3129-3133.

[20] Kakillioglu, B., Ahmad, A., Velipasalar, S. (2018). Object Classification from 3D Volumetric Data with 3D Capsule Networks. In 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP). IEEE. pp. 385-389.

[21] Vesperini, F., Gabrielli, L., Principi, E., Squartini, S. (2019). Polyphonic Sound Event Detection by Using Capsule Neural Networks. IEEE Journal of Selected Topics in Signal Processing. 13(2): 310-322.

[22] Wang, S., Liu, G., Li, Z., Xuan, S., Yan, C., Jiang, C. (2018). Credit Card Fraud Detection Using Capsule Network. In 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE. pp. 3679-3684.

[23] Nasr, A., Wayne, C., Bass, J., Ryan, G., Langer, J. C. (2013). Effect of delivery approach on outcomes in fetuses with gastroschisis. Journal of pediatric surgery. 48(11): 2251-2255.

[24] Shahroudnejad, A., Afshar, P., Plataniotis, K. N., Mohammadi, A. (2018, November). Improved explainability of capsule networks: Relevance path by agreement. In 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP). IEEE. pp. 549-553.

[25] Sultana, F., Sufian, A., Dutta, P. (2018). Advancements in Image Classification using Convolutional Neural Network. In 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN). IEEE. pp. 122-129.

[26] Paoletti, M. E., Haut, J. M., Fernandez-Beltran, R., Plaza, J., Plaza, A., Li, J., Pla, F. (2018). Capsule networks for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing. 57(4): 2145-2160.

[27] Afshar, P., Plataniotis, K. N., Mohammadi, A. (2019). Capsule Networks for Brain Tumor Classification Based on MRI Images and Coarse Tumor Boundaries. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. pp. 1368-1372.

[28] do Rosario, V. M., Borin, E., Breternitz, M. (2019). The Multi-Lane Capsule Network. IEEE Signal Processing Letters. 26(7): 1006-1010.

[29] Li, W., Liu, H., Wang, Y., Li, Z., Jia, Y., Gui, G. (2019). Deep Learning-Based Classification Methods for Remote Sensing Images in Urban Built-Up Areas. IEEE Access. 7: 36274-36284.

[30] Mobiny, A., Lu, H., Nguyen, H. V., Roysam, B., Varadarajan, N. (2019). Automated Classification of Apoptosis in Phase Contrast Microscopy Using Capsule Network. IEEE transactions on medical imaging.

[31] Wu, X., Liu, S., Cao, Y., Li, X., Yu, J., Dai, D., Meng, H. (2019). Speech Emotion Recognition Using Capsule Networks. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. pp. 6695-6699.
[32] Wang, W. Y., Li, H. C., Pan, L., Yang, G., Du, Q. (2018). Hyperspectral Image Classification Based on Capsule Network. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE. pp. 3571-3574.