



HAL
open science

Deep Complex Convolutional Neural Networks for Remote Sensing Image Classification

Lingling Li, Yukai Sun, Fuhai Ma, Jingjing Ma, Licheng Jiao, Fang Liu

► **To cite this version:**

Lingling Li, Yukai Sun, Fuhai Ma, Jingjing Ma, Licheng Jiao, et al.. Deep Complex Convolutional Neural Networks for Remote Sensing Image Classification. 5th International Conference on Intelligence Science (ICIS), Oct 2022, Xi'an, China. pp.252-259, 10.1007/978-3-031-14903-0_27. hal-04666442

HAL Id: hal-04666442

<https://hal.science/hal-04666442v1>

Submitted on 1 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Deep Complex Convolutional Neural Networks for Remote Sensing Image Classification ^{*}

Lingling Li, Yukai Sun, Fuhai Ma, Jingjing Ma, Licheng Jiao, and Fang Liu

School of Artificial Intelligent, Xidian University, Xian,710071, P.R. China.
llli@xidian.edu.cn

Abstract. At present, the neural network is often based on the real field of operation, research shows that, compared with the real field, the complex has incomparable advantages in the field of image processing, such as the complex represents more information, such as the phase information and modulus value, which play a great role in some fields. To take full advantage of complex data, this paper mainly studies CNN network, and through complex value processing, and get Complex Convolutional Neural Networks (CCN), complete the construction of complex convolutional neural network. In order to study complex neural network, we start from two aspects, one is convolution operation, the other is network construction. In this paper, we use ENet as the basic structure of the model, replace the convolutional structure, pooling structure, and BatchNorm structure with the complex form, use it in the Flevoland dataset, and get a good test results.

Keywords: Polarimetric SAR · Complex Convolution Neural Network
· Image Classification

1 INTRODUCTION

In recent years, polarimetric synthetic aperture radar (PolSAR) is one of the most important research directions in the field of remote sensing. PolSAR image classification is a basic and effective way to interpret PolSAR images, and has extensive research and application value [1, 2, 14].

Based on the convolutional neural network, this paper studies the operation mode of the real number in the neural network, summarizes and analyzes the data of the convolutional neural network which is widely used at present. Further introduce the application of complex numbers in neural networks, the advantages of complex numbers in processing polarimetric SAR data, and whether they can be implemented in neural networks based on complex number domains. Take lessons from the popular CNN networks and improve them. This paper introduces the convolutional neural network which changes the input from the

^{*} This work was supported in part by the National Natural Science Foundation of China (No.61906150, 62076192), the State Key Program of National Natural Science of China (No.61836009), the Major Research Plan of the National Natural Science Foundation of China (No.91438201, 91438103).

real number domain to the complex number domain. The main research contents of this paper are summarized as the following three points:

1. Research on complex number-based convolutional neural networks It mainly studies how to perform convolution operations on complex numbers, how to input complex numbers, and how to train. In this part, we simulate the convolution operation of the real part and the imaginary part of the complex number to complete.

2. Research on deep convolutional neural networks based on complex numbers

In this part, we refer to the ENet network, because the real part, imaginary part, and modulus value of complex numbers in the ENet network can be used as input for convolution operations, and the ENet network has a significant effect on image segmentation.

3. Analysis of experimental results based on real number field and complex number field

2 RELATED WORK

2.1 Convolutional neural network

Over the past few years major computer vision research efforts have focused on convolutional neural networks, commonly referred to as ConvNets or CNNs. These efforts have resulted in new state-of-the-art performance on a wide range of classification [3, 4] and regression [9, 10] tasks.

Convolutional Neural Networks (CNNs) are a class of neural networks that are particularly suitable for computer vision applications because of their ability to abstract representations hierarchically using local operations. There are two key design ideas that drive the success of convolutional architectures in computer vision. First, CNNs exploit the 2D structure of images and pixels within adjacent regions are often highly correlated. Therefore, instead of using one-to-one connections between all pixel units (as most neural networks do), CNNs can use grouped local connections. Second, the CNN architecture relies on feature sharing, so each channel (ie, the output feature map) is generated by convolution with the same filter at all locations as depicted in figure.1.

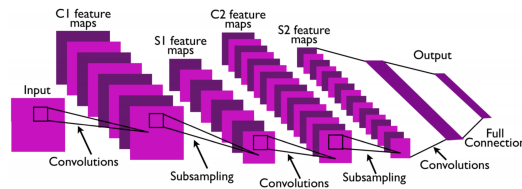


Fig. 1. Illustration of the structure of a standard Convolutional Network. Figure reproduced from [6]

2.2 ENet

ENet is an extension of the CNN network. It has been very popular in the field of image segmentation in recent years. It is an encoding and decoding network. Convolutional Neural Networks are designed to perform classification encoders. A decoder formed by an upsampling network that segments the original image[7].

Each block in ENet contains three convolutional layers like Figure.2: a 1x1 map, the main convolutional layer, and a 1x1 dilation. Interspersed with BN layers and PReLU layers, this structural combination is defined as a bottleneck model. If bottleneck is downsampling, a max pooling layer is added to the main branch. At the same time, the first 1x1 mapping in bottleneck is replaced by a convolution of size 2x2 and stride of 2. The types are ordinary convolution, hole convolutions, and transposed convolutions, etc., are sometimes replaced with 1x5 or 5x1 asymmetric convolutions.

The advantages of ENet are as follow: The ENet network structure is asymmetric and consists of a large encoding layer and a small decoding network. And The model replaces the convolutional layer in bottleneck with atrous convolution and concatenates it, which increases the receptive field and improves the IOU of segmentation [8].Then, the paper replaces all ReLUs with PReLUs, adding an extra parameter for each featuremap.

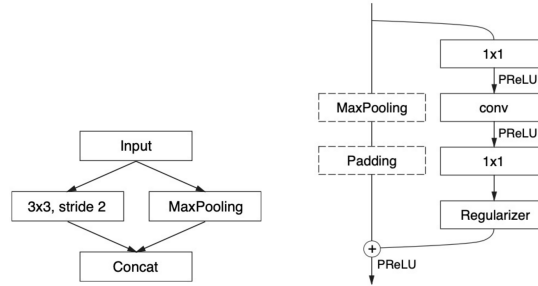


Fig. 2. The structure of ENet

3 METHODOLOGY

3.1 Complex convolution network

The method we use here is to simulate the convolution of two real numbers by the real and imaginary parts of complex numbers to realize the convolution operation of complex numbers. First, we define the complex filter parameters as $\vec{W} = \vec{A} + i\vec{B}$, the complex input information is $\vec{h} = \vec{x} + i\vec{y}$, through the deep network in the real number domain to simulate the value operation of complex numbers (where A and B are real numbers matrix, \vec{x} and \vec{y} are real vectors),

the complex number is expressed in the form of convolution output, as shown in the formula 3.1; And the real and imaginary parts of the operation are $\mathcal{R}(W \times \vec{h})$ and $\mathcal{J}(W \times \vec{h})$.

$$W \times \vec{h} = (A \times \vec{x} - B \times \vec{y}) + i(B \times \vec{x} + A \times \vec{y}) \quad (3.1)$$

$$\begin{bmatrix} \mathcal{R}(W \times \vec{h}) \\ \mathcal{J}(W \times \vec{h}) \end{bmatrix} = \begin{bmatrix} A & -B \\ A & A \end{bmatrix} \times \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \quad (3.2)$$

The above formula tells us that it can be used not only as a complex convolutional network, but also as a fully connected layer. The operations of the complex convolutional neural network are actually convolution operations and matrix product operations. Next, we try to build complex network structures: complex convolutional networks (CCN) and complex fully connected layers (CFC).

The input of the complex number field is that h with dimension $T_{in}/2$ consists of real part \vec{x} and imaginary part \vec{y} with dimensions $T_{in}/2$. The 3D tensor in the complex domain is the mapping to the input features required in the convolutional layer in the 2D complex domain. Where $W, H, T_{in}/2$ is the input signal, $i \times i$ is the size of the convolution kernel, and the parameter quantity of the four-dimensional complex weight tensor W of the output feature map is $T_{out}/2 \times T_{in}/2 \times i \times i$.

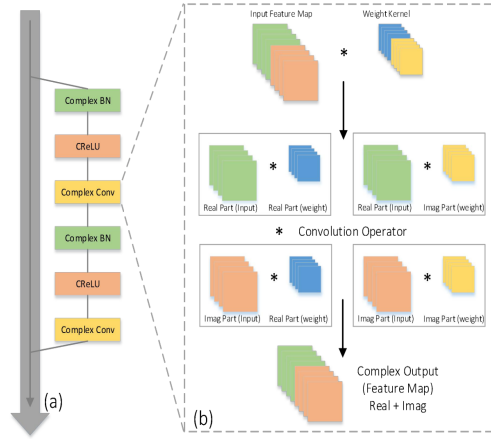


Fig. 3. Basic block of CCN

As shown in figure.3, the convolution kernel and feature map are the objects of convolution in the complex domain. We simulate the convolution operation using the real and imaginary parts of complex numbers as complex feature maps and complex convolution kernels. , perform multiplication and accumulation operations, and obtain the convolution result.

3.2 Network parameters analysis

Compared with the CNN, CCN require a larger amount of parameters and operations. Set the input of the CNN and the complex number convolutional neural network as a three-dimensional tensor similar to $W \times H \times T_{in}$, and the size of the convolution kernel of size $i \times i$, T_{out} as the output channel of the convolutional layer. For real network input, the number of input channels is T_{in} . The dimension of the weight tensor required by the model is $i \times i \times T_{in} \times T_{out}$.

For the complex convolutional neural network, the input feature of the channel number T_{in} represents the $T_{in}/2$ -dimensional complex feature, and the number of channels the model needs to generate is a complex number of $T_{out}/2$ output features. At this time, each channel needs a complex convolution kernel of size $i \times i$. Then $i \times i$ is the number of parameters of the real and imaginary parts of the convolution kernel, so the number of parameters of a convolution kernel required by CCN is $2 \times i \times i$, and the required dimension of the complex convolutional neural network at this time is $i \times i \times T_{in} \times T_{out}/2$.

Table 1. parameters for CNN and CCN

	Input	Kernel Size	Output	Parameters
CNN	T_{in}	$i \times i$	T_{out}	$i \times i \times T_{in} \times T_{out}$
CCN	$T_{in}/2$	$2 \times i \times i$	$T_{out}/2$	$i \times i \times T_{in} \times T_{out}/2$

3.3 Complex Batch Norm

The covariance matrix V of the complex input is given by the formula.3.3:

$$V = \begin{pmatrix} V_{rr} & V_{ri} \\ V_{ir} & V_{ii} \end{pmatrix} = \begin{pmatrix} Conv(R(x), R(x)) & Conv(R(x), i(x)) \\ Conv(i(x), R(x)) & Conv(i(x), i(x)) \end{pmatrix} \quad (3.3)$$

To scale the complex data x , the complex data needs to be centered and multiplied by the reciprocal of the quadratic root of a covariance matrix of size 2×2 . The expression for x obtained after complex normalization is as follows:

$$\tilde{x} = (V)^{(-\frac{1}{2})}(x - E[x]) \quad (3.4)$$

Analogous to real numbers, in the complex domain β^c is the complex shift parameter and γ^c is the scaling parameter equivalent to the covariance matrix V .

$$\gamma^c = \begin{pmatrix} \gamma_{rr}^c & \gamma_{ri}^c \\ \gamma_{ir}^c & \gamma_{ii}^c \end{pmatrix} \quad (3.5)$$

Therefore, the mathematical formula for complex BN is as follows:

$$BN(\tilde{x}) = \gamma^c \tilde{x} + \beta^c \quad (3.6)$$

3.4 Complex Activate Function

The expression of the activation function ReLU function used in the convolutional neural network is show in formula.3.7

$$\text{ReLU}(h) = \begin{cases} h & \text{if } h \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.7)$$

The formula.3.7 is the activation function of the real number domain. However, when we consider the calculation method of complex numbers, when h is a complex number variable[13], the size of the real part and the imaginary part is uncertain, and the relationship with 0 is also uncertain. Therefore, the ReLU function is in the complex number domain. cannot be used in complex arithmetic, and needs to be adjusted accordingly.

Here we use the phase of the complex domain as the inflection point of the relu function:

$$\text{ZReLU}(h) = \begin{cases} h & \text{if } \theta_h \in [0, \frac{\pi}{2}] \\ 0 & \text{otherwise} \end{cases} \quad (3.8)$$

4 EXPERIMENT

The cropland dataset in the Flevoland region of the Netherlands is the L-band image data of the Flevoland region of the Netherlands acquired by TASA/JPL ARISAR. This is a four-view full-polarization image with a size of 750×1024 and a resolution of $12m \times 5m$, including 15 different types of ground objects (without background), namely potatoes, dried beans, rapeseed, sugar beets, alfalfa, and barley, grass, three kinds of wheat, peas, bare land, forest, water, built-up area. Unmarked areas are marked in white on the real class plot. After removing the background class, the total number of 15 different features is 157712[12].

We use the AUG function for data enhancement, using point-by-point segmentation and random segmentation methods to separate 30 points in height and width, generate 30 points in height and 30 points in width, and obtain 900 random interpolations. The data is reshaped, and then pickled and persisted with the Numpy library to form a dataset that can be read by Pytorch. The labels come from BMP images to identify specific pixels. Then 1200 pieces of data of size 100×100 were formed, 1100 for training and 100 for testing.

According to the figure.4,it is obvious that the accuracy rate is still relatively low in the 30 epochs, the difference between the predicted image and the original image is still very large, and only two color blocks are distinguished,the accuracy is only 82.41%; after the 90 epochs, it can be seen that the same as the original image The difference is no longer large, and the classification effect has been significantly improved,the accuracy is increase to 89.19%. After 300 epochs, the accuracy rate can be maintained at 97.45%, and the training loss is maintained at a low level, achieving optimal classification results.

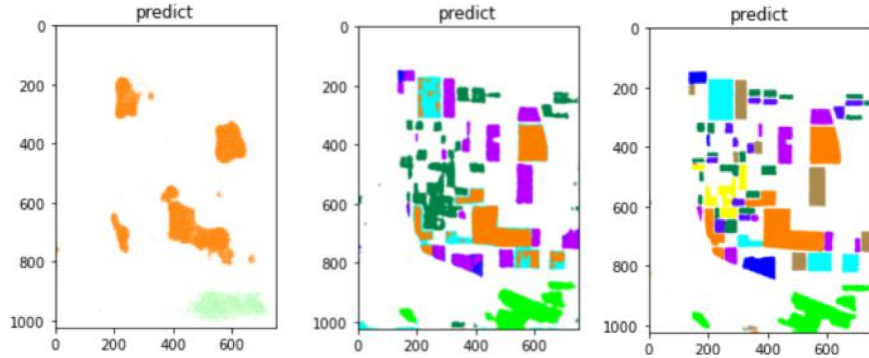


Fig. 4. Visualization of test results at epochs 30,90 and 300

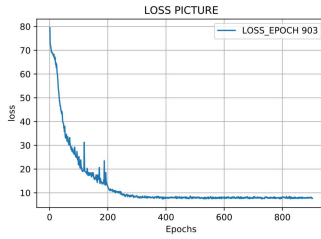


Fig. 5. Train loss

Table 2. parameters for CNN and CCN

Epochs	30	90	120	300
Accuracy	82.4%	89.2%	93.1%	97.4%

Table.2 is a table of the output accuracy. From the table, it can be intuitively seen that our highest accuracy is maintained at 97.5%, and its position is exactly between 200-300. The subsequent test accuracy remained stable at about 96%, and the accuracy results were better.

Through the adjustment of experimental parameters, a loss map is output. Corresponding to the loss graph.5 and the test accuracy table.1, it can be seen that when the loss value is at a high position, the test results are relatively poor. When the value of loss is close to stable, the model converges and the optimal result is obtained.

5 CONCLUSION

This paper mainly studies and analyzes deep complex neural network and complex residual neural network, summarizes and analyzes from the two, and finally

realizes the task of remote sensing image classification based on deep complex convolutional neural network. The differences in the number of parameters and network depth between the real-domain convolutional network and the complex-domain convolutional network are compared, and it is found that the complex-domain convolutional neural network has advantages in parameters and network depth. In order to realize the complex domain data processing based on the deep residual network, how to combine the ENet network to construct the convolution layer, pooling layer and activation layer of the complex domain is analyzed in detail. The four modules are organized according to the network organization form of the real number domain, and finally the complex convolutional network of this paper is formed. In the follow-up research, we need to simulate the method of real number neural network to analyze the complex convolutional neural network, and explore the potential of more complex convolutional networks, which can be applied to different scenarios.

References

1. Harris, R A. Envisat : ASAR science and applications[J]. 1998.
2. Ulaby F T, Elachi C. Radar polarimetry for geoscience applications[J]. 1990, 5(3):38-38.
3. Khan A , Sohail A , Zahoora U , et al. A survey of the recent architectures of deep convolutional neural networks[J]. Artificial Intelligence Review, 2020.
4. Tran D , Bourdev L , Fergus R , et al. Learning Spatiotemporal Features with 3D Convolutional Networks[J]. 2014.
5. F Liu, Shen C , Lin G . Deep Convolutional Neural Fields for Depth Estimation from a Single Image[J]. IEEE, 2014.
6. Lecun Y , Kavukcuoglu K , Farabet C . Convolutional Networks and Applications in Vision[C]// Proceedings of 2010 IEEE International Symposium on Circuits and Systems. IEEE, 2010.
7. Paszke A , Chaurasia A , Kim S , et al. ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation[J]. 2016.
8. F Yu, Koltun V . Multi-Scale Context Aggregation by Dilated Convolutions[J]. 2016.
9. Eigen D , Fergus R . Predicting Depth, Surface Normals and Semantic Labels with a Common Multi-Scale Convolutional Architecture[J]. 2015 IEEE International Conference on Computer Vision (ICCV), 2014.
10. F Liu, Shen C , Lin G . Deep Convolutional Neural Fields for Depth Estimation from a Single Image[J]. IEEE, 2014.
11. Hadji I , Wildes R P . What Do We Understand About Convolutional Networks?[J]. 2018.
12. Zhou Y , Wang H , Xu F , et al. Polarimetric SAR Image Classification Using Deep Convolutional Neural Networks[J]. IEEE Geoscience Remote Sensing Letters, 2017, 13(12):1935-1939.
13. Trabelsi C , Bilaniuk O , Zhang Y , et al. Deep Complex Networks[J]. 2017.
14. Liu F , Jiao L , Tang X . Task-Oriented GAN for PolSAR Image Classification and Clustering[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019:1-13.