

Median Filtering Detection Based on Quaternion Convolutional Neural Network

Jinwei Wang^{1,2,3,4}, Qiye Ni³, Yang Zhang³, Xiangyang Luo^{2,*}, Yunqing Shi⁵,
Jiangtao Zhai³ and Sunil Kr Jha³

Abstract: Median filtering is a nonlinear signal processing technique and has an advantage in the field of image anti-forensics. Therefore, more attention has been paid to the forensics research of median filtering. In this paper, a median filtering forensics method based on quaternion convolutional neural network (QCNN) is proposed. The median filtering residuals (MFR) are used to preprocess the images. Then the output of MFR is expanded to four channels and used as the input of QCNN. In QCNN, quaternion convolution is designed that can better mix the information of different channels than traditional methods. The quaternion pooling layer is designed to evaluate the result of quaternion convolution. QCNN is proposed to features well combine the three-channel information of color image and fully extract forensics features. Experiments show that the proposed method has higher accuracy and shorter training time than the traditional convolutional neural network with the same convolution depth.

Keywords: Median filtering forensics, quaternion convolution layer, quaternion pooling layer, color image.

1 Introduction

With the rapid development of multimedia technology, the video [Pan, Lei, Zhang et al. (2018); Pan, Yi and Chen (2018)] and digital image has become one of the indispensable information carriers in people's daily life. There are many ways to obtain images, such as mobile phones, cameras and so on. Nowadays, with the popularity of the Internet, images are able to spread rapidly around the world. At the same time, with the appearing of many

¹ Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Nanjing University of Information Science & Technology, Nanjing, 210044, China.

² State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, 450001, China.

³ Department of Computer and Software, Nanjing University of Information Science & Technology, Nanjing, 210044, China.

⁴ State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, 100093, China.

⁵ Department of Electrical Computer Engineering, New Jersey Institute of Technology, New Jersey, NJ07102, USA.

*Corresponding Author: Xiangyang Luo. Email: xiangyangluo@126.com.

Received: 08 March 2019; Accepted: 28 March 2019.

image editing software, it becomes easier to tamper the content of digital images. However, it is difficult for the human to distinguish whether the image has been modified by the image editing software. The altered images often lead to misunderstandings and give the lawbreaker an opportunity to take advantage of it. Therefore, digital image forensics technology is particularly important today. In recent years, image manipulation and forensics mainly focus on JPEG compression [Luo, Huang and Qiu (2010)], geometric transformation [Popescu and Farid (2005)], contrast enhancement [Stamm and Liu (2010)], sharpening [Cao, Zhao, Ni et al. (2011)] and median filtering [Cao, Zhao, Ni et al. (2010); Yuan (2011)]. Median filtering is a statistical filter commonly used in digital image processing. It has good properties of filtering random noise and preserving edge details. Tamperers can easily use median filtering for anti-forensics operations. For example, the nonlinearity of median filtering is used to destroy the periodicity between the adjacent pixels of resampled images so that the technique based on periodic detection geometric transformation is ineffective [Kirchner and Bohme (2008)]. Due to the role played by median filtering in digital image anti-forensics technology and information hiding, the detection and forensics of median filtering have attracted more and more attention.

In recent years, the convolutional neural network has achieved great success in the computer vision field. It can automatically learn digital image features and classify them. The development of neural networks has gone through many booms and valleys over the years. After Yann LeCun successfully used LeNet-5 in handwritten digit recognition in 1998, deep learning has once again experienced rapid development. Among them, there are many landmark structures, such as AlexNet [Krizhevsky, Sutskever and Hinton (2012)], GoogleNet [Szegedy, Liu, Jia et al. (2015)], ResNet [He, Zhang, Ren et al. (2016)], DenseNet [Huang, Liu, Van Der Maaten et al. (2017)], and DCGAN [Radford, Metz and Chintala (2015)], etc. These models have achieved remarkable success in computer vision and other related fields. Fang et al. [Fang, Zhang, Sheng et al. (2018)] used DCGAN to improve image recognition based on CNN and get a big ascension. The use of deep learning for remote sensing image object detection is also very effective [Cheng, Zhou and Han (2016), Han, Zhang, Cheng et al. (2015)].

Quaternion and quaternion transform are widely used in image compression, texture classification, face recognition, digital watermarking and other color image processing fields. Soulard et al. [Soulard and Carré (2010), Soulard and Carré (2011); Carré and Denis (2006)] used color quaternion wavelet transform (CQWT) to encode color images the first time. Compared with discrete wavelet transform (DWT), CQWT has the characteristics of a large compression ratio and a small distortion. Subsequently, the amplitude and angle coefficients of CQWT are used to classify color images. For example, Wang et al. [Wang, Li, Luo et al. (2018)] used CQWT to identify computer-generated images. Due to the continuous multi-scale analysis characteristics, CQWT has achieved an improvement in classification accuracy than DWT and continuous wavelet transform (CWT) under the same characteristics. It is verified that CQWT is superior in performance. Schauert et al. [Schauerte and Stiefelhagen (2012)] detected faces using quaternion discrete cosine transform (QDCT) and joint focus model. Rizo et al. [Rizo and Ziou (2015)] used quaternion-based local binary patterns (LBP) and illumination invariance to study face recognition. This method simply defines LBP in the form of quaternion, and solves the problem of information redundancy between adjacent points in

traditional LBP. Many researchers combine quaternion with neural networks and get good results [Parcollet, Morchid, Bousquet et al. (2016), Parcollet, Zhang, Morchide et al. (2018)]. It can be seen from the above that quaternion has obvious advantages in processing color digital images.

Therefore, it is a feasible method to combine quaternion with neural network and apply it to color image [Shi and Funt (2007)], but most applications of quaternion only use a quaternion to preprocess color images and then takes their output as the input of neural networks [Rishiyur (2006); Kendall, Grimes and Cipolla (2015)]. In this paper, a median filtering detection based on the quaternion convolutional neural network is proposed. The contribution of the scheme is summarized as follows:

- (1) A quaternion is used to combine the information of three channels of color images.
- (2) The quaternion convolution layer and quaternion pooling layer are designed.
- (3) Quaternion convolutional neural network has better detection performance for median filtering than CNN with the same depth.

The remaining part of this paper is organized as follows. Section 2 briefly introduces some related work. Then the proposed scheme is introduced in Section 3. The performance of the experiment and the results are provided in Section 4. Finally, Section 5 provides the conclusion.

2 Related works

Median filtering was proposed by Turkey in 1971. The principle is to replace the value of a point in the sequence with the median value of each point in the neighborhood of the point. Let X and Y represent an image and its image after median filtering respectively. The median filtering operation with $w \times w$ as the neighborhood is performed according to Eq. (1):

$$Y_{i,j} = \text{median} \{ X_{i+h,j+v} \mid h, v \in \{(1-w)/2\}, \dots, (w-1)/2 \} \quad (1)$$

In the above formula, $X_{i+h,j+v}$ and $Y_{i,j}$ respectively represent the pixel values of X and Y at points $(i+h, j+v)$ and (i, j) . Typical filtering window sizes are 3×3 and 5×5 .

2.1 Median filtering residual

In Chen et al. [Chen, Kang, Liu et al. (2015)], the convolutional neural network is used to extract the features of the median filtered image. Unlike conventional convolutional neural networks, the filter layer is designed before the first convolution layer. For median filtering forensics, directly input the color images to convolutional neural networks is not very good. The filter layer is used to get the median filtering residual (MFR) of an image. The output of the filter layer is then used as the input of the convolutional neural network. A $w \times w$ sliding window is used to convolve the input image $x(i, j)$ to get the output $y(i, j)$. MFR is designed as Eq. (2):

$$\text{MFR} = \text{med}_w(x(i, j)) - x(i, j) = y(i, j) - x(i, j) \quad (2)$$

The filter layer is followed by five convolution layers, two fully connected layers. The above method does not take into account the relationship between the RGB channels of the color image, and only directly convolution color image to extract features. But it often loses important structural features.

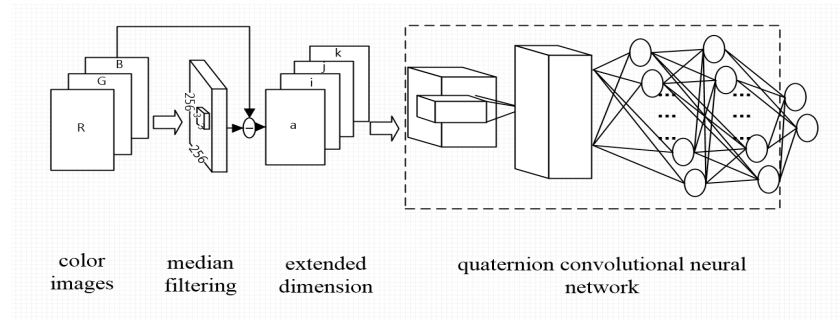


Figure 1: An overall framework of the proposed method. RGB is the channels of color image, and it will be filtered by the median filter with 3×3 size. Next, the original going to be subtracted from the result and get MFR. After that, the MFR is expanded to four dimensions and used as input to the quaternion neural network

2.2 Quaternion convolution

In document [Zhu, Xu, Xu et al. (2018)], the traditional convolutional neural network adds a different channel of input through convolution operation has some inherent deficiencies [Zhu, Xu, Xu et al. (2018)], like:

- (1) Simple addition ignores the complex connections between channels.
- (2) Some important structural information may be lost.
- (3) The added way gives the convolution kernel a very large degree of freedom, which easily leads to overfitting.

Therefore, quaternion network is proposed and the quaternion convolution layer is designed. The elements of the quaternion convolution kernel can be designed according to Eq. (3):

$$w_{\mu} = s_{\mu} \left(\cos \frac{\theta_{\mu}}{2} + \sin \frac{\theta_{\mu}}{2} \mu \right) \quad (3)$$

θ_{μ} is the angle of rotation, s_{μ} is the scaling factor, the rotation axis is μ .

Traditional CNN only supports scaling transformation, although document [Zhu, Xu, Xu et al. (2018)] combines quaternion with neural network to support the rotation of color space (which provides a more reasonable representation of color). But quaternion convolution in document [Zhu, Xu, Xu et al. (2018)] has a high degree of complexity, which is not conducive to practical use.

3 Proposed model

The overall structure is shown in Fig. 1. In this section, quaternion is first introduced and then the details of the proposed model are explained.

3.1 Definition of quaternion

The concept of quaternion was introduced the concept of quaternion in 1843. The representation of quaternion hyper-complex q is as follows:

$$q = a + b \cdot i + c \cdot j + d \cdot k \quad (4)$$

where $a, b, c, d \in R$, i, j, k are imaginary numbers, and be defined as follows:

$$\begin{aligned} i^2 &= j^2 = k^2 = -1 \\ i \cdot j &= -j \cdot i = k \\ j \cdot k &= -k \cdot j = i \\ k \cdot i &= -i \cdot k = j \end{aligned} \quad (5)$$

Quaternions are extensions of real and complex numbers, which are generally seen as a combination of real and imaginary parts. a is the real part of quaternion, which can also be expressed as q_r . b , c and d are the imaginary parts of quaternion, which can also be expressed as $q_i \cdot i + q_j \cdot j + q_k \cdot k$. If q_r is equal to 0, quaternion q is known as the pure quaternion. Quaternion q can also be expressed as the complex form $q = A + B \cdot j$, where A and B are also the complex form, and $A = q_r + q_i \cdot i$, $B = q_j + q_k \cdot i$, putting A and B into q , then

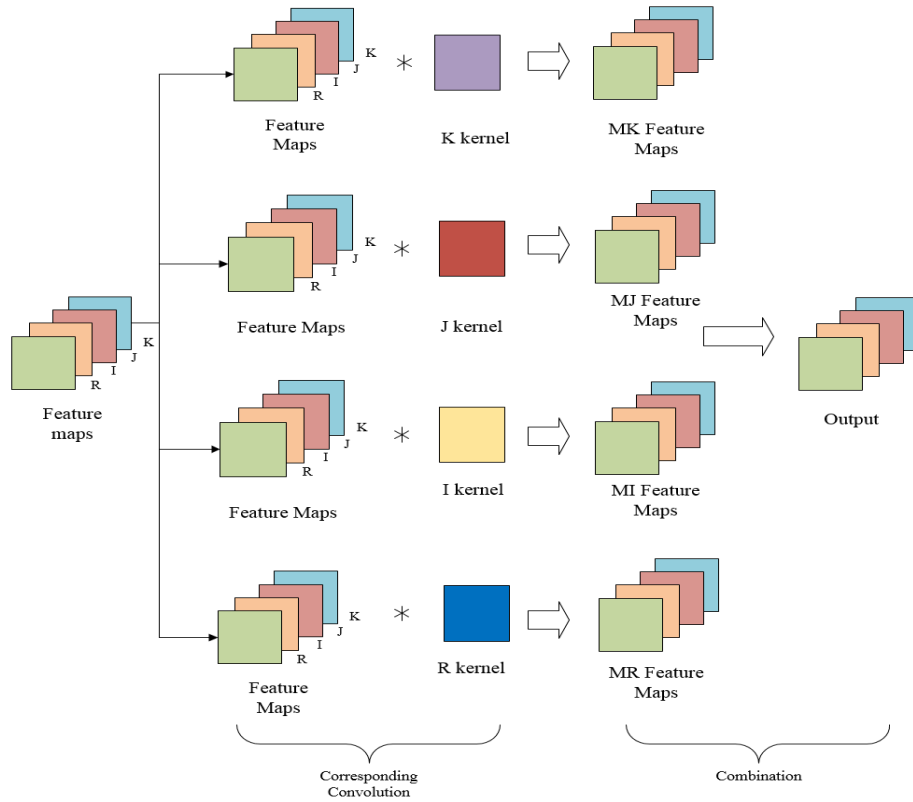
$$\begin{aligned} q &= A + B \cdot j \\ &= q_r + q_i \cdot i + (q_j + q_k \cdot i) \cdot j \\ &= q_r + q_i \cdot i + q_j \cdot j + q_k \cdot ij \\ &= q_r + q_i \cdot i + q_j \cdot j + q_k \cdot k \end{aligned} \quad (6)$$

The operation rules for quaternion, for quaternion q_1 and q_2 , satisfy:

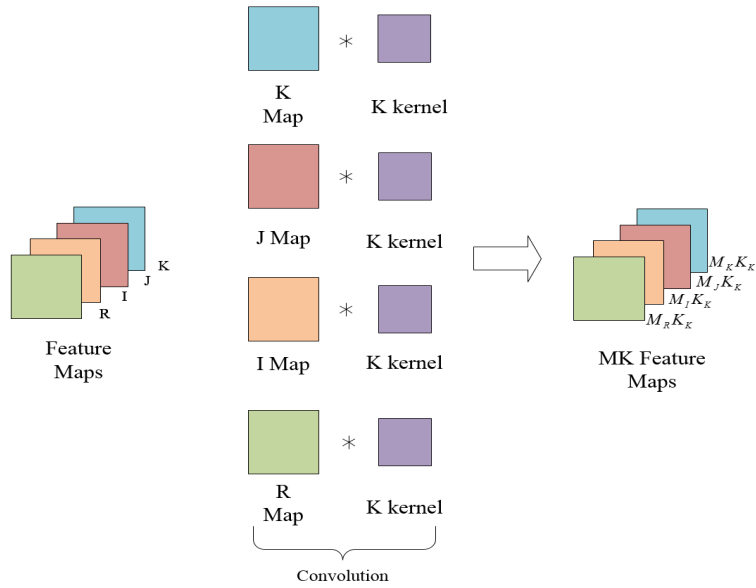
$$q_1 = q_{1r} + q_{1i} \cdot i + q_{1j} \cdot j + q_{1k} \cdot k \quad (7)$$

$$q_2 = q_{2r} + q_{2i} \cdot i + q_{2j} \cdot j + q_{2k} \cdot k \quad (8)$$

Then the addition, subtraction and multiplication operations of quaternion q_1 and q_2 are respectively defined as follows:



(a)



(b)

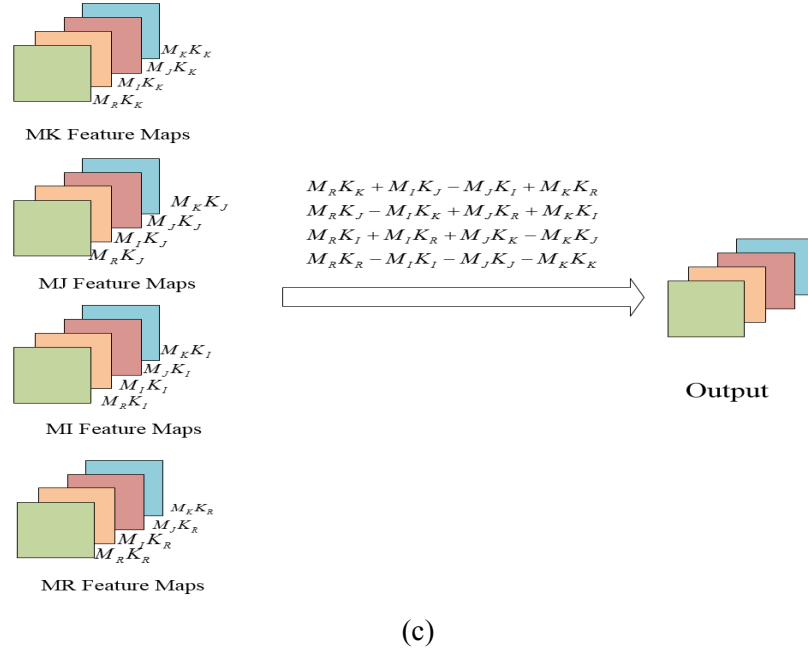


Figure 2: Quaternion convolution layer. (a) shows the process of quaternion convolution, including corresponding convolution and mutual combination. (b) is the operation corresponding convolution. (c) is a way of combining (a) and (b)

$$q_1 + q_2 = (q_{1r} + q_{2r}) + (q_{1i} + q_{2i}) \cdot i + (q_{1j} + q_{2j}) \cdot j + (q_{1k} + q_{2k}) \cdot k \quad (9)$$

$$q_1 - q_2 = (q_{1r} - q_{2r}) + (q_{1i} - q_{2i}) \cdot i + (q_{1j} - q_{2j}) \cdot j + (q_{1k} - q_{2k}) \cdot k \quad (10)$$

$$\begin{aligned} q_1 q_2 = & (q_{1r} q_{2r} - q_{1i} q_{2i} - q_{1j} q_{2j} - q_{1k} q_{2k}) \\ & + (q_{1r} q_{2i} + q_{1i} q_{2r} + q_{1j} q_{2k} - q_{1k} q_{2j}) \cdot i \\ & + (q_{1r} q_{2j} + q_{1j} q_{2r} + q_{1k} q_{2i} - q_{1i} q_{2k}) \cdot j \\ & + (q_{1r} q_{2k} + q_{1k} q_{2r} + q_{1i} q_{2j} - q_{1j} q_{2i}) \cdot k \end{aligned} \quad (11)$$

3.2 Proposed scheme

3.2.1 Preprocess

Median filtering is very effective for the removal of speckle noise and salt and pepper noise on the image. However, usually the noise contained in the actual image is relatively insignificant. The features of the images extracted from convolutional neural network are mainly texture and color, etc. But, there are relatively few high-frequency features that can reflect the difference between the original image and the image after median filtering. At the same time, a large number of features such as textures and colors extracted may even have some interference in the discrimination. Therefore, in some previous studies based on the detection of median filtering using deep learning, it was found that if the

original image was taken as the input of the network directly, the final result was not very ideal. So the image was preprocessed. First, the filter layer in Chen et al. [Chen, Kang, Liu et al. (2015)] is used, which uses the median filtering window with 3×3 size as shown in Fig. 1 to filter the original image $x(i, j)$ and then use the output $y(i, j)$ to subtract the original image and extract fingerprints of median filtering operations. Finally, the median filtering residual (MFR) is obtained. Such processing can highlight the characteristics of the high-frequency part, thus improving the accuracy of classification.

Quaternions are generalization of complex numbers. In the present research results, it is considered to represent an image with quaternion by encoding the RGB channels of color image as the imaginary part of quaternion. The above method is also used in this paper. In this way, quaternion can process their pixels as a whole vector in multidimensional space and consider color association. At the same time, many classical tools previously used for grayscale image processing have been successfully extended to color image processing by means of quaternion representation, such as Fourier transform, wavelet transform, and neural network and so on. However, compared with the above methods, the quaternion convolution layer designed in this study needs to correspond to four channels, and color images have only three channels, in order to match the two dimensions, the color image is added a channel to form a four-channel image. The fourth dimension of the expansion is filled with zero.

3.2.2 Quaternion convolution layer

The local pixels of the image are closely related, while the distant ones are weakly correlated. Therefore, it is not necessary for each neuron to perceive the global image, but only the local image, and then the global information can be obtained by integrating the local information at a higher level. This is called the local perceptual field. It shows the importance of convolution layer in the neural network. Therefore, the key to complete the quaternion neural network is the design of the convolution layer of a quaternion. If two quaternions are multiplied, each element in the matrix is combined with the other elements. For example, there is a quaternion convolution kernel matrix $F = a + b \cdot i + c \cdot j + d \cdot k$, a quaternion matrix $M = w + x \cdot i + y \cdot j + z \cdot k$. If the quaternion convolution kernel F is used to convolve the quaternion matrix M , it can be expressed by the following formula:

$$\begin{aligned} F \times M = & (a \times w - b \times x - c \times y - d \times z) + \\ & i(a \times x + b \times w + c \times z - d \times y) + \\ & j(a \times y - b \times z + c \times w + d \times x) + \\ & k(a \times z + b \times y - c \times x + d \times w) \end{aligned} \quad (12)$$

Through the quaternion convolution, three channels of a color image can be effectively merged. The quaternion convolution process of the image is shown in Fig. 2. Fig. 2(a) describes the overall process of quaternion convolution. First, the input feature maps (RIJK) are convolved four times: same input, different convolution kernel (all four convolutions kernel are initialized by random method). The convolution method designed in this study is different from the traditional convolutional neural network. The conventional convolution

method generates the number of channels of the convolution kernel according to the number of channels of the input feature map. For example, for the size of a convolution kernel is $F \times F$ and the number of channels of the input feature map is d , then the convolution kernel of size $F \times F \times d$ can be obtained. The convolution kernel can obtain a channel of the output feature map by convolving the input feature map. The convolution method in this study is to convolve the channels of the input feature map using the same convolution kernel. Thus, output channels are generated with the same number of the channels that are entered in the feature map. These output channels are connected to form a new output feature map. As shown in Fig. 2(b), M in MK Feature Maps represents a feature map, and K represents convolution kernel and the process is called corresponding convolution. The feature map obtained by corresponding convolution has the characteristics of the corresponding channel of the input feature map. Four feature maps (MK, MJ, MI, MR) can be obtained after four corresponding convolutions. The channels at different positions of each output feature map are then combined according to the formula shown in Fig. 2(c). in this way the final output of the quaternion convolution layer is obtained. Each channel of the output feature map obtained by the above method contains the feature information of the input feature map. Finally, color image information is combined and makes feature more sufficient extraction.

3.2.3 Quaternion pooling layer

The use of pooling layer brings great improvement to a convolutional neural network. The pooling layer introduces the displacement invariance and pays more attention to the existence of the silent feature rather than the specific location of the feature. The use of the pooling layer reduces the input size of the next layer, reduces the amount of calculation and the number of parameters, which is the most intuitive function. At the same time, the use of pooling layer can also prevent overfitting to some extent. Therefore, pooling layer is a very important part of a convolutional neural network. A quaternion pooling layer is specially designed for quaternion convolutional neural networks. The quaternion pooling is based on the output of the quaternion convolution layer. The amplitude of the convolution result of a quaternion is evaluated, and the maximum value is obtained. The formula to obtain amplitude is as follows:

$$D = \sqrt{a_{ij}^2 + b_{ij}^2 + c_{ij}^2 + d_{ij}^2} \quad (13)$$

where ij is the position in the channel in which the element is located, a, b, c and d represent the element values of the four channels corresponding to the same position.

3.2.4 Other layers

The other layers select the ones used in the traditional convolutional neural network. For the activation layer, sigmoid, tanh and RELU are the most commonly used. Since the sigmoid and tanh gradients are very gentle in the saturated region and close to zero, it is easy to cause the vanishing gradient problem and slow down the convergence speed, while ReLU is easier to learn and optimize. Because of its piecewise linearity nature, the derivation of its forward propagation and backward propagation are both piecewise linear,

which is more convenient. Therefore, RELU is selected as the activation function of this structure. Because the activation input value of deep neural network before nonlinear transformation gradually deviates or changes with the deepening of network depth or in the training process. So this leads to the disappearance of the gradient of the low-level neural network in the case of backpropagation.

Through a certain standardization method, forcibly pulls the distribution of the input value of any neuron in each layer of neural network back to the standard normal distribution with a mean of zero variance, so that the gradient becomes larger and the gradient disappearance problem is avoided. Moreover, the gradient becomes larger, which means that the learning convergence speed is fast, and the training speed can be greatly accelerated. Therefore, the BatchNorm layer will be added to this structure, and the BatchNorm function is selected from TensorFlow.

4 Experiment

In this section, the experiments performed are presented to demonstrate the effects of median filtering forensics using a quaternion convolutional neural network. We also used the conventional convolutional neural network for comparison, which showed the improvement of our scheme.

4.1 Introduction of datasets

To run the experiments, we first built a dataset using the CASIA TIDE v2.0 and UCID datasets. CASIA TIDE v2.0 contains 12,614 color images, which have a different size from 240×160 to 900×600 pixels and three kind of formats: JPEG, BMP and TIFF. The UCID dataset has 886 color images of the two kinds of size of 384×512 and 512×384 respectively.

Next, we performed median filtering on the original image by using the 3×3 sliding window to generate a set of tampered images. Finally, we got 13500 tampered images. Now we had 27000 images in total, and we then used a ratio of 2:1 to separate training and testing datasets, i.e., 18,000 images for training dataset, and the remaining 9,000 images as testing dataset. In addition, the images are all cut to the size of 256×256 .

4.2 Experimental environment and parameter settings

Two experiments are conducted on the dataset, respectively using the conventional convolution neural network and the proposed quaternion convolution neural network. Both experiments consist of three convolution layers, three pooling layers, three fully connected layers and followed by one softmax layer. In addition to the fact that convolution and pooling layers are different from the conventional ones, quaternion convolution neural network has the same structure as conventional convolution neural network in both fully connected and softmax layer, which also includes the operation of Flatten before the fully connected layer. ReLU is selected as the activation function in both experiments, which make the network nonlinear. Both networks use cross entropy loss as loss function. For the detection task of median filtering, it is actually the task of classifying the tampered images from the original ones. The method used to judge the

model is to check the accuracy of classifying the detected image to the correct label. Both experiments are performed on a NVIDIA GeForce GTX 1080Ti GPU.

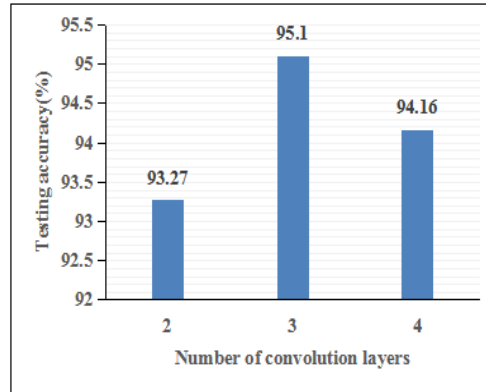


Figure 3: Testing accuracy versus number of convolution layers

When training the proposed model, an optimization algorithm called stochastic gradient descent is employed, with the parameters set as follows: momentum=0.9, decay=5e-4, and a learning rate $\epsilon=1\text{e-}6$ that decreases every 5 epochs by a factor $f=0.5$. The model in experiment are both trained for 50 epochs.

4.3 Performance of the proposed scheme

In order to obtain a better model for median filtering forensic detection, a self-comparison experiment was designed based on the scheme. Our original architecture involves two quaternion convolution layers, two quaternion pooling layers to learn the association across quaternion features and three fully connected layers used for the classification. However, the number of quaternion convolution layers cannot be determined, and different depths must be tried to find the appropriate number of convolution layers. Thus, a set of experiments are conducted to find the best depth of quaternion convolution layers.

The performance of each depth is shown in Fig. 3. It can be noticed that the overall detection rate can be improved while we increase the number of quaternion convolution layers until reaching the best depth.

It can be observed that the best performance is achieved when we use the architecture mentioned, i.e., three quaternion convolution layers each followed by a quaternion pooling layer and ended with three fully connected layers. Additionally, it can be seen that when four quaternion convolution layers are used, the final detection rate has declined by 0.94% compared to the best one. Because when the number of convolution layers is 4, compared with 3 layers, it is easier to overfit, resulting in a decrease in accuracy. It can show us that it's not the deeper the quaternion convolution network is, the better detection result it has in the field of color image manipulation forensics.

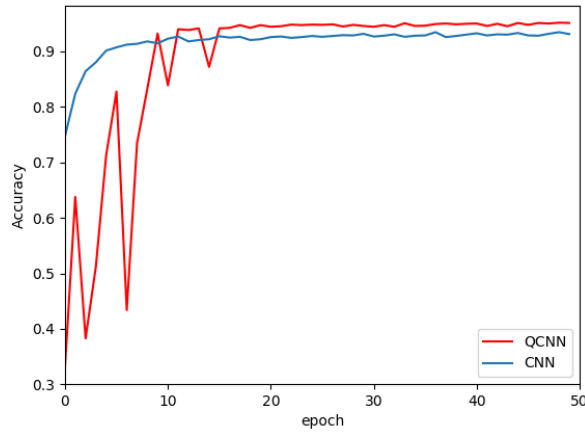


Figure 4: The accurate trend of different epoch on different models

4.4 Comparison with other methods

The experiments were conducted on the earlier mentioned datasets. We collected the test accuracy for each of the two comparative experiments and put them together in Fig. 4 for comparison.

Fig. 4 depicts the curves of the image manipulation detection accuracy rate versus the number of training epochs for each network, i.e., the proposed QCNN and the conventional CNN. It can be observed that the QCNN method is superior to the conventional method in the training period, and the detection rate of the best implementation is 95.10% and 93.46% that are shown in Tab. 1, respectively. The convergence speed of the conventional CNN is very fast, but the proposed scheme is better after about 15 epochs.

Table 1: Testing accuracy on different models

	Test_acc
CNN	0.936
QCNN	0.951

The conventional approach achieves a lower detection rate since it is a suboptimal solution for a trained network with not fully utilizing the correlation information among three color channels. Thus, the proposed QCNN can capture color channel features that may not be captured using real-value convolution. And due to its better applicability to color information, the proposed QCNN can get richer features than the conventional CNN, leading to higher detection accuracy.

5 Conclusion

In this paper, a quaternion - based neural network is proposed and used for median filtering detection. In the scheme, the color image has only three channels, while

quaternion contains four channels. The method to match quaternion to dimensions of color images becomes very important. After consulting relevant literature and research, the method of extending one dimension to color image is finally adopted. In the end, the result is better than the direct input method. The quaternion convolution layer perfectly combines the convolutional neural network with quaternion. Considering the importance of the pooling layer to the neural network, the quaternion pooling layer is specially designed. Of course, there are still many problems that have not been solved so far, such as how to fill the fourth dimension of color images. We will continue to explore these issues in the following work.

Funding Statement: This work was supported in part by the Natural Science Foundation of China under Grants (Nos. 61702235, 61772281, U1636219, U1636117, 61702235, 61502241, 61272421, 61232016, 61402235 and 61572258), in part by the National Key R\&D Program of China(Grant Nos. 2016YFB0801303 and 2016QY 01W0105), in part by the plan for Scientific Talent of Henan Province (Grant No. 2018JR0018), in part by the Natural Science Foundation of Jiangsu Province, China under Grant BK20141006, and in part by the Natural Science Foundation of the Universities in Jiangsu Province under Grant 14KJB520024, the PAPD fund and the CICAET fund.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- Cao, G.; Zhao, Y.; Ni, R.; Kot, A. C.** (2011): Unsharp masking sharpening detection via overshoot artifacts analysis. *IEEE Signal Processing Letters*, vol. 18, no. 10, pp. 603-606.
- Cao, G.; Zhao, Y.; Ni, R.; Yu, L.; Tian, H.** (2010): Forensic detection of median filtering in digital images. *IEEE International Conference on Multimedia and Expo*, pp. 89-94.
- Carré, P.; Denis, P.** (2006): Quaternionic wavelet transform for colour images. *Wavelet Applications in Industrial Processing IV, International Society for Optics and Photonics*, vol. 6383, 638301.
- Chen, J.; Kang, X.; Liu, Y.; Wang, Z. J.** (2015): Median filtering forensics based on convolutional neural networks. *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 1849-1853.
- Cheng, G.; Zhou, P.; Han, J.** (2016): Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7405-7415.
- Fang, W.; Zhang, F.; Sheng, V. S.; Ding, Y.** (2018): A method for improving CNN-based image recognition using DCGAN. *Computers, Materials & Continua*, vol. 57, pp. 167-178.
- Han, J.; Zhang, D.; Cheng, G.; Guo, L.; Ren, J.** (2015): Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp.

3325-3337.

He, K.; Zhang, X.; Ren, S.; Sun, J. (2016): Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778.

Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K. Q. (2017): Densely connected convolutional networks. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700-4708.

Kendall, A.; Grimes, M.; Cipolla, R. (2015): PoseNet: a convolutional network for real-time 6-dof camera relocalization. *IEEE International Conference on Computer Vision*, pp. 2938-2946.

Kirchner, M.; Bohme, R. (2008): Hiding traces of resampling in digital images. *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 4, pp. 582-592.

Krizhevsky, A.; Sutskever, I.; Hinton, G. E. (2012): Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, pp. 1097-1105.

Luo, W.; Huang, J.; Qiu, G. (2010): JPEG error analysis and its applications to digital image forensics. *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 480-491.

Pan, Z. Q.; Lei, J. J.; Zhang, Y. J.; Wang, F. L. (2018): Adaptive fractional-pixel motion estimation skipped algorithm for efficient HEVC motion estimation. *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 14, no. 1, pp. 12-19.

Pan, Z. Q.; Yi, X. K.; Chen, L. M. (2018): Motion and disparity vectors early determination for texture video in 3D-HEVC. *Multimedia Tools and Applications*, pp. 1-18.

Parcollet, T.; Morchid, M.; Bousquet, P. M.; Dufour, R.; Linarès, G. et al. (2016): Quaternion neural networks for spoken language understanding. *IEEE Spoken Language Technology Workshop*, pp. 362-368.

Parcollet, T.; Zhang, Y.; Morchid, M.; Trabelsi, C.; Linarès, G. et al. (2018): Quaternion convolutional neural networks for end-to-end automatic speech recognition. arXiv: 1806.07789.

Popescu, A. C.; Farid, H. (2005): Exposing digital forgeries by detecting traces of resampling. *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 758-767.

Radford, A.; Metz, L.; Chintala, S. (2015): Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv: 1511.06434.

Rishiyur, A. (2006): Neural networks with complex and quaternion inputs. arXiv: cs/0607090.

Rizo-Rodriguez, D.; Ziou, D. (2015). Quaternion-based local binary patterns for illumination invariant face recognition. *International Conference on Biometrics*, pp. 257-263.

Schauerte, B.; Stiefelhagen, R. (2012): Predicting human gaze using quaternion DCT image signature saliency and face detection. *IEEE Workshop on The Applications of Computer Vision*, pp. 137-144.

- Shi, L.; Funt, B.** (2007): Quaternion color texture segmentation. *Computer Vision and Image Understanding*, vol. 107, no. 1, pp. 88-96.
- Soulard, R.; Carré, P.** (2010): Quaternionic wavelets for image coding. *18th European Signal Processing Conference*, pp. 125-129.
- Soulard, R.; Carré, P.** (2011): Quaternionic wavelets for texture classification. *Pattern Recognition Letters*, vol. 32, no. 13, pp. 1669-1678.
- Stamm, M. C.; Liu, K. R.** (2010): Forensic estimation and reconstruction of a contrast enhancement mapping. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1698-1701.
- Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S. et al.** (2015): Going deeper with convolutions. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-9.
- Wang, J.; Li, T.; Luo, X.; Shi, Y. Q.; Liu, R. et al.** (2018): Identifying computer generated images based on quaternion central moments in color quaternion wavelet domain. *IEEE Transactions on Circuits and Systems for Video Technology*.
- Yuan, H. D.** (2011): Blind forensics of median filtering in digital images. *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 4, pp. 1335-1345.
- Zhu, X.; Xu, Y.; Xu, H.; Chen, C.** (2018): Quaternion convolutional neural networks. *European Conference on Computer Vision*.