Noname manuscript No.
(will be inserted by the editor)

An Improved Deep Convolutional Neural Network to Predict Airfoil Lift Coefficient

Boping Yu · Liang Xie · Fuxin Wang

Received: date / Accepted: date

Abstract Recently, significant developments in deep learning have made many possibilities in the field of fluid mechanics. This paper presents a new method of using improved convolutional neural network to learn airfoil lift coefficient calculated by OpenFOAM simulation tool. We propose a "feature-enhanced-image" data preprocessing method to prepare the training and testing data set. A novel convolutional neural network is designed which uses deeper convolution and pooling layers coupled with batch normalization technique. In addition, before linear regression, in fully-connected layers, we use dropout method to reduce the risk of over-fitting. Mini-batch stochastic gradient descent(SGD) optimization algorithm is chosen and mean square error(MSE) is used to do the model evaluation when training and testing the model. It is demonstrated that this improved deep convolutional neural network(IDCNN) provides more accurate lift coefficient prediction compared to other state-of-the-art neural networks. We also test the effect of batch size and full batch normalization implementation on the performance of the whole convolutional neural network. Finally, it is concluded that the best predicting performance is achieved in the condition of 10 batch size and the mean square error of blind test can reach 3.1×10^{-4}. Furthermore, the "feature-enhanced-image" method we proposed can achieve 85.2% decreasing of testing MSE.

Keywords IDCNN · lift coefficient · batch normalization · dropout · mini-batch SGD

Boping Yu
Shanghai Jiao Tong University
E-mail: boping_yu@sjtu.edu.cn

Liang Xie
Shanghai Jiao Tong University
E-mail: xl2016@sjtu.edu.cn

Fuxin Wang
Shanghai Jiao Tong University
E-mail: fuxinwang@sjtu.edu.cn

1 Introduction

The calculation of aerodynamic coefficients of airfoil has always been an important process of airfoil optimization and aircraft design[2]. At present, the main methods to obtain aerodynamic coefficients are wind tunnel test and computational fluid dynamics(CFD) simulation. In addition, there exists many effective surrogate models[6,9,8] to deal with these high dimensional computing problems and many researchers also utilize artificial neural network(ANN) technique to perform the aerodynamic coefficients approximation[4,7]. Recently, great advances in neural networks and machine learning have been made due to the booming development of computational power. The development of deep neural network(DNN) technique has opened the new doors to improve the robustness, generalization and precision of the aerodynamic coefficients approximation model.

More recently, more and more researchers have applied the deep learning techniques in the field of aerospace and achieve some positive results. Julia Ling et al[12] proposed a novel deep neural network architecture which contains a multiplicative layer with embedded invariance for the Reynolds stress anisotropy prediction and this model shows improved prediction precision compared with a commonly used neural network. S.Suresh et al[17] used a RNN method to predict lift coefficient at high angle of attack and it demonstrates RNN can has high predicted performance than memory neuron networks(MNN). Yao Zhang et al[18] used a CNN approach to predict airfoil aerodynamic coefficients for unseen airfoil shapes. However, there exists a MSE gap between the testing data set and training data set, which means the model they proposed has a risk of over-fitting.

Inspired by the work by Yao Zhang et al[18], we design an improved deep convolutional neural network to do the aerodynamic coefficients approximation and this paper will demonstrate that the prediction model can provide more accurate and general coefficients approximation. For a CNN model, the most important thing is the preprocessing of the training data. Enough and extensive data can contribute to a good learning ability and generalization performance model[10]. Within our known knowledge, we made the following innovations in data preparation.

- We use user-defined OpenFOAM simulation tool to guarantee the accuracy of training data.
- We propose a "feature-enhanced-image" method to prepare the training data, thus broadening the width of the airfoil data image and changing the brightness of the image to do the data augmentation.
- Feed some bad shapes of airfoil data image into the IDCNN model we proposed, in oder to enhance its generalization ability.

2 Improved Deep CNN for Lift Coefficient Prediction

Yao Zhang et al[18] used a highly similar architecture to LeNet-5[11], to do the airfoil lift coefficients approximation. It has demonstrated that CNN technique has been a useful perspective in engineering meta-modeling task, but the CNN model they proposed did not get a satisfactory accuracy. We harness a deeper CNN architecture to predict lift coefficient, including three convolution layers, three pooling layers and four fully-connected layers.

2.1 IDCNN Architecture for Lift Coefficient Prediction Model

The improved CNN architecture we proposed is based on the traditional LeNet-5 architecture, but we have made the following changes and improvements. We deepening the whole neural network, including the convolutional layer, pooling layer and fully-connected layer. Besides, batch normalization[5] is used after each pooling operation and it can accelerate training speed and reduce internal covariate shift between previous pooling layer and next convolutional layer. In order to avoid the over-fitting problems effectively which may affect the prediction accuracy, we adopt the dropout[16] technique and set the dropout rate to 0.7. ReLU activation function and mini-batch SGD[1] optimized method are chosen in the training of the CNN model.

The data input layer is a mini-batch of airfoil images and the resolution of the image is 100×100. The size of batch can affect the training results of model, and we will compare testing accuracy with different given batch sizes. In convolutional layer 1, we used 40 convolution kernels which have the size of 5×5. After feeding the output data into pooling layer 1, we add a BN layer and the same operation occurs between every convolutional layer and pooling layer. In pooling layer 1, we select max-pooling method and the pooling filter size is 2×2. In addition, zero-padding method and moving step size which is equal to filter size in each direction are chosen. In convolutional layer 2, 60 convolutional filters are used during the convolution operation and the filter size is 5×5. The settings of pooling 2 are the same as pooling layer 1 and a BN layer is added before the pooling 2. The last convolutional layer, we use 100 kernels which have the same size as previous kernels. Then the max-pooling method is chosen in the pooling layer 3. After the convolutional and pooling layers, four fully-connected layers are used to do the linear regression, in other words, airfoil lift coefficient prediction. The details of the number of nodes in each fully-connected layer are shown in figure 1.

To summarize, we have constructed an improved deep CNN model that contains 13 hidden layers, including three convolutional layers, three pooling layers, three BN layers and four fully-connected layers. A visual illustration architecture of the IDCNN we proposed is shown in figure 1.

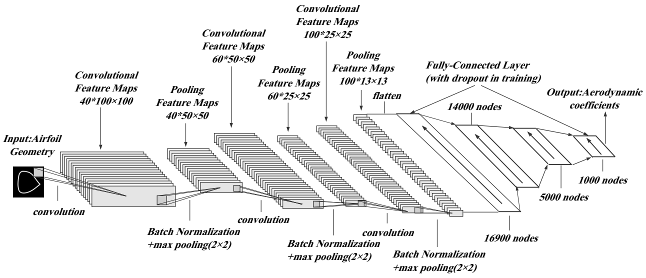

Fig. 1 Improved Deep CNN Architecture for Lift Coefficient Prediction Model

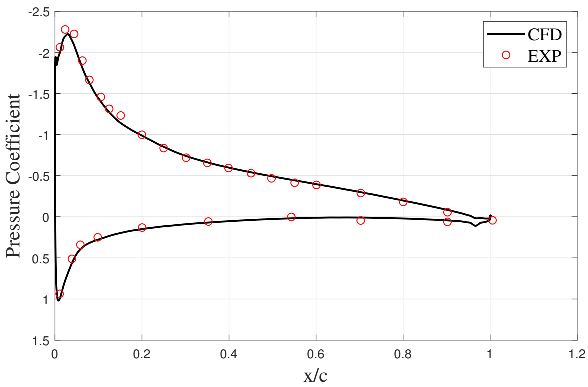

Fig. 2 Comparing Result of Experimental and CFD Data

2.2 Data Preparation

The CNN model we proposed is trained and tested on a database of airfoil data images and lift coefficients. According to the goal of airfoil optimization in future, SC1095 airfoil is chosen, and Hicks-Henne fitting method[3] is used to generate the training and testing airfoil datasets. In this context, the user-defined OpenFOAM solver is chosen as the simulation tool to obtain the aerodynamic coefficients. The simulating mach number is 0.4, reynolds number is 3.63×10^6 and the angle of attack is $6.13°$. Spalart-Allmaras turbulence model[15] and SIMPLE algorithm[13] are selected. In order to check the accuracy of CFD results, in this paper, we compare the CFD result with experimental data[14]. Figure 2 shows the comparing result and it shows that the CFD result is highly consistent with experimental data. In other words, figure 2 illustrates that the user-defined OpenFOAM solver and computing settings are feasible in this situation.

Data preprocessing is crucial to the training and testing of CNN model. Besides, we expect to get more training samples with less computational cost. In view of the above two situations, we propose a novel data preprocessing

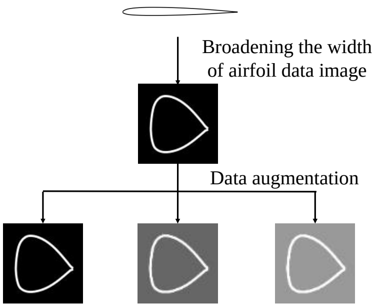

Fig. 3 "feature-enhanced-image" Method

method "feature-enhanced-image" for the airfoil data image and the description of this method is given below. A initial SC1095 airfoil data image is shown at the top of figure 3. The maximum thickness of the SC1095 airfoil is only 9.50% of the chord length. So a small changes of control variables could not lead to a visible shape changes in a low resolution data image. In order to maximize the shape changes of the given airfoil at low resolution, we broaden the width of airfoil data image by multiplying the thickness of airfoil by 10 times. Then we do the data augmentation by changing the brightness of the data image into 3 different types. In other words, three different types of airfoil data image have the same lift coefficient label. We can get massive training and testing samples using only one third of CFD computation cost through this method. Changing brightness means changing the number of airfoil data matrix, and the variance of all the data in the matrix can keep the same. Thus the information of the airfoil shape can be preserved and the cost of training data preparation can be reduced at the same time. The process of applying the "feature-enhanced-image" method in the IDCNN model is shown in figure 3.

In addition to the above data preprocessing methods, we also add some bad airfoil data images into the training data set. Some bad airfoil image examples are shown in figure 4. Bad images means some airfoils that can not be regarded as a proper airfoil shape. We can consider this operation as the noise adding. Through changing the control variables of Hicks-Henne fitting function, 11550 pairs of normal input/output(I/O) training data set and 1050 pairs of I/O testing data set are generated. Then we use the "feature-enhanced-image" method to do the data augmentation, thus we get 34650 pairs of I/O training data set for the following hyper parameter tunning.

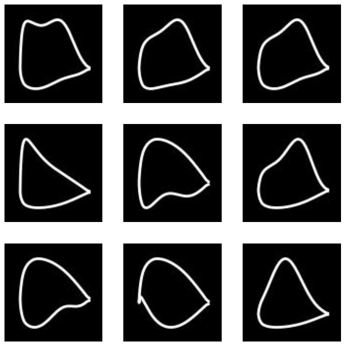

Fig. 4 Examples of Bad Airfoil Images

3 Results and Discussion

In a linear regression model, MSE is the most popular used evaluation criteria. The MSE calculated by summing over all of the squared differences between the predicted and true values and then divided by the number of the batch size. In equation 1, m_{batch} represents the number of batch size, t_i is the true value and y_i is the predicted value.

$$MSE = \sum_{i=1}^{m_{batch}} \frac{(t_i - y_i)^2}{m_{batch}} \tag{1}$$

In the training of the IDCNN model we proposed. We utilize the early-stopping method to save the trained model at a specific epoch to get the minimum testing MSE.

3.1 "feature-enhanced-image" Method Influence on the Testing MSE

We select 11550 training data randomly from the 34650 training data set to check the influence of the "feature-enhanced-image" method on the testing MSE compared with the normal training data set mentioned above. The other parameters in the CNN model are the same and the comparing results of testing MSE are showing in figure 5. In these two CNN model, dropout technique is used to fight the over-fitting problem. The testing MSE of the CNN with normal training data set can only reach 2.5×10^{-3}. However, the testing MSE of the randomly selected training data that using the "feature-enhanced-image" method can reach 3.7×10^{-4}. The decreasing rate of the testing MSE is 85.2% which shows the "feature-enhanced-image" is a very effective method for reducing the testing MSE. We can also figure out that the amplitude of the MSE fluctuation has increased after using the "feature-enhanced-image" method. This is because we have added some noise airfoil data images into the training data set.

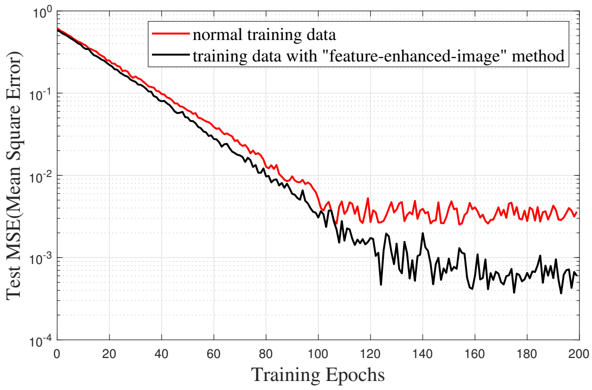

Fig. 5 "feature-enhanced-image" Method Influence on the Testing MSE

3.2 Comparing Results of Empty and Full BN layers Implementation

Batch normalization is always be implemented after the data input layer, but multiple BN layers are added after each convolutional layer in order to reduce internal covariate shift in this paper. The comparing result is shown in the figure 6, it is easy to conclude that full BN layers implementation can dramatically reduce the training time. One training epoch means feeding the whole training data to the CNN model. When we use a normal CNN architecture without full BN layers implementation, the MSE does not reach a acceptable level until at about 120 epochs. However, when applying full BN layers in the previous CNN architecture, the testing MSE reaches 1×10^{-3} at about 2 epochs. Moreover, the fluctuation of MSE during the model training has also been affected by the full BN layers implementation. In figure 6, the amplitude of the MSE fluctuation is enclosed by two black lines for easily comparing. When applying full BN layers in the CNN model, the amplitude of the MSE fluctuation reduce a little bit than that of CNN model without full BN layers applied.

The given comparing result shows that full BN implementation can highly reduce the training time, thus being good to the following hyper-parameters tunning and model constructing. Smaller amplitude of the MSE fluctuation can lead to easily early-stopping at a appropriate epoch to save the CNN model.

3.3 Batch Size Influence on the Testing Results

Mini-batch SGD optimized algorithm is used in the neural network training. Batch size is a highly influential factor on the performance of the CNN model and an intuitive judgement is that the larger the batch size, the slower the convergence. From the CNN model tunning experience, using about 1% of the

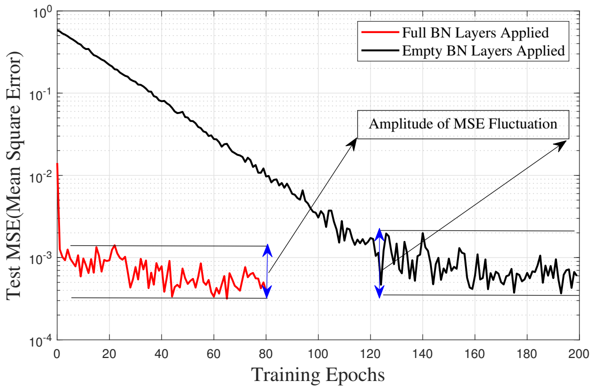

Fig. 6 Comparing Results of Empty and Full BN layers Implementation

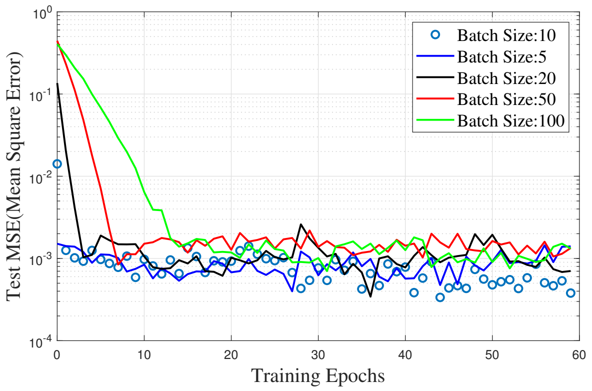

Fig. 7 Comparing Result of batch size on the performance of model

training data set as the size of mini-batch is a proper choice. According to the goal of faster training and best MSE performance, we set the top-level of comparing batch sizes to be 100. Then five different batch sizes are compared when we do the mini-batch SGD optimization, and the results are presented in the figure 7. It is easy to figure out that 10 batch size is an optimal choice for the improved CNN model. Because under this condition, both fast convergence and minimum MSE are guaranteed.

3.4 Over-fitting Checking and MSE Quantitative Analysis

A satisfactory testing accuracy is not the end of training a neural network. Over-fitting is the most likely occur during neural network training. For over-fitting phenomenon, a simple judgement method is checking the difference between the training and testing accuracy of the neural network. In the im-

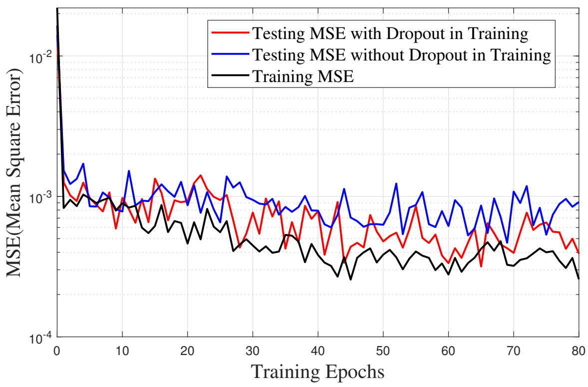

Fig. 8 Training and Testing History

Table 1 Testing MSE Quantitative Analysis of Different Neural Networks

Network Type	MSE	Remark
CNN	0.005	Reference[18]
RMLP	0.0009	Reference[17]
IDCNN	0.0003	Our model

proved CNN airfoil lift coefficient prediction model, dropout technique is used to avoid this problem and the comparing result is shown in the figure 8. The testing MSE of CNN without dropout is higher than the that of CNN with dropout in most training epochs, except for a few epochs in the beginning of the training. It is shown that dropout implementation can reduce the MSE gap between the training and testing procedure.

Three different neural networks have been compared with the same evaluation criteria MSE and the comparing results are shown in table 1. The list two other neural networks, CNN and recurrent multilayer perceptron(RMLP), are also used to predict the lift coefficients. The improved deep CNN model we proposed have a much smaller MSE than other two neural networks. It demonstrates that the improved techniques we employed, including BN, dropout and mini-batch SGD, are feasible and of great value in predicting the airfoil lift coefficient.

4 Conclusion

An improved deep CNN model for lift coefficient prediction is presented which uses full BN layers implementation, dropout and mini-batch SGD techniques. In the preprocessing of the training and testing data, the "feature-enhanced-image" method we proposed is utilized and it can gain a effective 85.2% declining of testing MSE in the model training. Besides, we also add some bad airfoil data images into the model training data to enhance the model generalization ability.

As for the improvements of the "feature-enhanced-image", full BN and dropout methods on the initial CNN model, the comparing results are given to show that these three techniques are effective in reducing the training time and increasing the prediction accuracy. The training and testing history is presented to show that the model we established has a good generalization ability than other state-of-the-art neural networks. Furthermore, the IDCNN model we constructed presents a more accurate prediction than other two neural networks. In conclusion, the IDCNN model we proposed can guarantee a much less training time and a more precise testing accuracy at the same time.

References

1. Bottou, L.: Large-scale machine learning with stochastic gradient descent. In: Proceedings of COMPSTAT'2010, pp. 177–186. Springer (2010)
2. Buckley, H.P., Zhou, B.Y., Zingg, D.W.: Airfoil optimization using practical aerodynamic design requirements. Journal of Aircraft **47**(5), 1707–1719 (2009)
3. Hicks, R.M., Henne, P.A.: Wing design by numerical optimization. Journal of Aircraft **15**(7), 407–412 (1978)
4. Ignatyev, D.I., Khrabrov, A.N.: Neural network modeling of unsteady aerodynamic characteristics at high angles of attack. Aerospace Science and Technology **41**, 106–115 (2015)
5. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
6. Iuliano, E.: Global optimization of benchmark aerodynamic cases using physics-based surrogate models. Aerospace Science and Technology **67**, 273–286 (2017)
7. J.Linse, D., F.Stenge, R.: Identification of aerodynamic coefficients using computational neural networks. Journal of Guidance Control & Dynamics **16**(6), 1018–1025 (1993)
8. Jouhaud, J.C., Sagaut, P., Montagnac, M., Laurenceau, J.: A surrogate-model based multi-disciplinary shape optimization method with application to a 2d subsonic airfoil. Computers & Fluids **36**(3), 520–529 (2007)
9. Laurenceau, J., Sagaut, P.: Building efficient response surfaces of aerodynamic functions with kriging and cokriging. Aiaa Journal **46**(2), 498–507 (2008)
10. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. nature **521**(7553), 436 (2015)
11. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998)
12. Ling, J., Andrew, K., Jeremy, T.: Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. Journal of Fluid Mechanics **807**, 155–166 (2016)
13. Patankar, S.: Numerical heat transfer and fluid flow. CRC press (1980)
14. Qing, W., Qi-jun, Z.: Synthectical optimization design of rotor airfoil by genetic algorithm. Journal of Aerospace Power **31**(6), 1486–1495 (2016)
15. Spalart, P., Allmaras, S.: A one-equation turbulence model for aerodynamic flows. In: 30th aerospace sciences meeting and exhibit, p. 439 (1992)
16. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research **15**(1), 1929–1958 (2014)
17. Suresh, S., Omkar, S.N., Mani, V., Prakash, T.N.G.: Lift coefficient prediction at high angle of attack using recurrent neural network. Aerospace Science and Technology **7**(8), 595–602 (2003)
18. Zhang, Y., Sung, W.J., Mavris, D.N.: Application of convolutional neural network to predict airfoil lift coefficient. In: 2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, p. 1903 (2018)