A DNN-based Image Retrieval Approach for Detection of Defective Area in Carbon Fiber Reinforced Polymers through LDV Data

Erfan Basiri Dept. of electrical engineering, University of Guilan Rasht, Iran erfan.basirirasht@gmail.com Reza PR Hasanzadeh Dept. of electrical engineering, University of Guilan Rasht, Iran <u>hasanzadehpak@guilan.ac.ir</u>

Abstract— Carbon fiber reinforced polymer (CFRP) materials, due to their specific strength and high consistency against erosion and corrosion, are widely used in industrial applications and high-tech engineering structures. However, there are also disadvantages: e.g. they are prone to different kinds of internal defects which could jeopardize the structural integrity of the CFRP material and therefore early detection of such defects can be an important task. Recently, local defect resonance (LDR), which is a subcategory of ultrasonic nondestructive testing, has been successfully used to solve this issue. However, the drawback of utilizing this technique is that the frequency at which the LDR occurs must be known. Further, the LDR-based technique has difficulty in assessing deep defects. In this paper, deep neural network (DNN) methodology is employed to remove this limitation and to acquire a better defect image retrieval process and also to achieve a model for the approximate depth estimation of such defects. In this regards, two types of defects called flat bottom holes (FBH) and barely visible impact damage (BVID) which are made in two CFRP coupons are used to evaluate the ability of the proposed method. Then, these two CFRPs are excited with a piezoelectric patch, and their corresponding laser Doppler vibrometry (LDV) response is collected through a scanning laser Doppler vibrometer (SLDV). Eventually, the superiority of our DNN-based approach is evaluated in comparison with other well-known classification methodologies.

Keywords— Deep neural network, defect image retrieval, laser Doppler vibrometry, nondestructive testing, carbon fiber reinforced polymer.

I. INTRODUCTION

Composite materials including Carbon fiber reinforced polymer (CFRP) due to many advantages including low weight, low thermal expansion, and high physical and chemical resistance have turned to the fully used ones in industry and advanced structures such as aerospace industry and automotive engineering in which the low weight of construction while maintaining high resistance is vital. One of the disadvantages of using these materials is the possibility of occurring damage inside them and their high sensitivity to this damage; therefore, recognition and localization of such defects are essential to prevent them from being spread. Since defects are often Saman Hadi Dept. of electrical engineering, University of Guilan Rasht, Iran <u>saman.hadi72@gmail.com</u> Mathias Kersemans Mechanics of materials and structures Ghent University Ghent, Belgium mathias.kersemans@ugent.be

invisible, nondestructive testing (NDT) is used for identification of defects that investigates the properties of a specimen without damaging it. NDT techniques include a variety of methods, one of which is ultrasonic. One of the newest approaches in ultrasonic NDT methods is making use of local defect resonance (LDR) to recognize different defects such as flat bottom holes (FBH), barely visible impact damage (BVID), etc. [1]. The behavior of LDR is very similar to the usual resonance behavior of stable structures. When a various range of frequencies of ultrasonic excitation waves applied to the specimen, some of them match with the resonance frequency of local defects, and therefore the vibration amplitude of defects will be increased considerably in 3 dimensions (3D)s relative to the other parts of the specimen. There is an approximate relation between the resonance frequency and the size of the defect. For a circular FBH and a square FBH, the following equations can be raised to determine the resonance frequency [2, 3]

$$f_0^c \approx \left(1.6.\frac{H}{R^2}\right) . \sqrt{E/12\rho(1-v^2)}$$
 (1)

$$f_0^s \approx \left(4\pi \cdot \frac{H}{3s^2}\right) \cdot \sqrt{E/6\rho(1-v^2)} , \qquad (2)$$

where f_0^c and f_0^s are resonance frequency of circular and square defects in order. Also, the material parameters E, ρ , and v are Youngs's modulus, density, and Poisson number respectively. Also, H, R, and S are residual thickness, radius, and size of FBH defects, respectively. Fig. 1 depicts two circular and square FBH with different H, R, and S. In the traditional LDR technique, for predicting LDR frequencies of a FBH defect in the specimen, it is essential to have some approximate information about the shape of a defect such as H, R, and S.

Due to the fact that we usually have no a priori information about the shape of defects, we use broadband vibrations for exciting the specimen. In this regards, a piezoelectric actuator is used to excite CFRPs. Then the locations of the defects are obtained from measuring the vibrational response by 3D infrared scanning Laser Doppler Vibrometer (SLDV). The data gathered in this way is called laser Doppler vibrometry (LDV) which includes both resonance and non-resonance frequencies of defects.



Fig. 1. (a) Illustration of a shallow square FBH defect (defect₁) and a deep circular FBH defect (defect₂). (b) Cross-sectional view of defect₁ and defect₂. H_1 and H_2 are the remaining thickness of these defects. S and R represent the size of defect₁ and the radius of defect₂.

Automated retrieval of location and depth of deep defects as well as complex damages is an arduous task in CFRP materials due to their particular structure. Therefore, the recently used signal processing technique such as Frequency Band Data (FBD) [4] may have problem to detect these kinds of defects with an acceptable accuracy.

FBD is a signal processing method often used in data compression, while it could also be employed individually as a technique for defect localization. In brief, it computes the average amplitude of vibration for each point at location (x, y)over the entire measurement bandwidth. The FBD of a scan point is computed as follows:

$$FBD(x_i, y_i, f_1, f_2) = (f_{res}/f_2 - f_1) \cdot \sum_{f=f_1}^{f_2} V_z(x_i, y_i, f) / U_{exc}(f) \quad (3)$$

In (3), f_{res} is the frequency resolution of the Fast Fourier Transform (FFT). f_1 and f_2 are the lowest and highest frequencies of the frequency bandwidth which have been selected from a frequency response of excitation signal. V_z and U_{exc} are the representation of the out-of-plane (z-plane) velocity amplitude and the voltage amplitude of the excitation signal, respectively [4].

In this study, we firstly use Deep Neural Network (DNN) [5] for automated retrieval of the location of deep FBHs and BVID in CFRP_{FBH,5} and CFRP_{BVID}, respectively. Then another DNN is used to Estimate the depth of FBH defects. It should be noted that the FBD method cannot detect the deepest FBH at all. Ultimately, we compare the performance of our proposed models with FBD and some other well-known classification methodologies namely Support Vector Machine (SVM) [6], AdaBoost [7, 8], and Shallow Neural Network (SNN) [5].

The paper is organized as follows: in Section II, we explain the formation of datasets. In Section III, DNN is introduced, and the proposed DNN method is described in section IV. Then the experimental results are presented in section V, and finally we conclude the paper in Section VI.

II. DATASET

As depicted in Fig. 2, we use two datasets in this study where each of them is obtained from a different specimen. As shown in Fig. 2-a, the first sample is made from unidirectional CFRP laminate based on layup $[(45/0/-45/90)]_{3s}$. The size is $150 \times 90 \times 5.52 \text{ mm}^3$, and it contains 5 FBHs with different remaining thickness *H*. The second sample is made from unidirectional CFRP laminate based on layup $[(0/90)]_{6s}$, and the



Fig.2. (a) CFRP plate containing 5 FBHs (**CFRP**_{FBH.5}) with different *H* as shown below the defects. The plate are divided into 6 regions by red dashed lines. The first four regions are used in training phase and the others are used in testing phase. (b) Back side of CFRP plate containing 6.3J BVID (**CFRP**_{BVID}). The approximate defective area is diamond-shaped due to the shape of impactor-tip used in making this damage.



Fig. 3. Schematic diagram of LDV setup. The size of LDV data is $M \times N \times K$

size is $150 \times 100 \times 5.52$ mm³, which is shown in Fig. 2-b. This CFRP plate contains BVID created by a 7.72 kg drop weight with 16 mm impactor-tip, based on the ASTM D7136 standard test method, whose measured impact energy is 6.3 J.

A burst chirp signal from 5 kHz to 80 KHz is applied to the CFRP samples through a low power piezoelectric patch (type EPZ-20MS64W from Ekulit, with a diameter of 12 mm) which was bonded to the surface. Finally, the out-of-plane vibrational response is obtained using a 3D infrared scanning laser Doppler vibrometer (Polytec PSV-500-3D XTRA) and the FFT of them are computed using Hanning windows to avoid spectral leakage. Because of using a SLDV with $f_{res} = 25$ Hz and a burst chirp signal with BW = 75 KHz, the out of plane (Z-axis) vibrational response is measured in K = 3001 different frequency at each scan point. Thus two 3D LDV datasets of sizes ($106 \times 194 \times 3001$) and ($122 \times 148 \times 3001$) are achieved for the CFRP_{FBH,5} and CFRP_{BVID} respectively. The setup of this experiment is shown in Fig. 3.

III. DEEP NEURAL NETWORK (DNN)

Nowadays, Artificial Intelligence (AI) is one of the main causes of evolution in industries; a fast developing section of which is called deep learning, an approach of machine learning that enhances computer systems using experience and data. Deep learning architectures such as DNN, Deep Belief Networks (DBN) [9], Recurrent Neural Networks (RNN) [10], and Convolutional Neural Network (CNN) [10] have been utilized in vast fields including computer vision, speech and



Fig. 4. Schematic diagram of N-1 hidden layers fully connected DNN [11].

audio processing, natural language processing, robotics, material properties, etc. It has immensely employed human knowledge experience and also the information about the human brain, especially in computer-aided design. Over the last few years, it has had broad advancement as a result of having more powerful computers, larger datasets, and techniques to train deeper networks [10].

A DNN is formed from multiple layers between the input and output layers named hidden layers, and it finds a proper mathematical linear or non-linear relationship between the inputs and outputs. DNNs have a similar structure with SNN, but deeper (DNN has three hidden layers or more while SNN has just one or two), and it has a more noticeable hierarchy structure. Hence DNN is more robust than SNN according to the discovery of the hidden relationships between inputs and outputs [12]. The general structure of a DNN with N-1 hidden layers and n_v outputs is demonstrated in Fig.4.

The inputs (x) and outputs (y) of the network are vectors of dimension n_x and n_{y_y} , respectively. The circles are the representation of neurons, and the lines between them show their connections. Each hidden layer consists of the N_j neurons. In fact, each neuron consists two steps of computation. First, it calculates the weighted sum of its inputs, then it commonly applies a nonlinear transformation activation function (e.g. ReLU, sigmoid, etc.) to the weighted sum and sends the result out of the neuron. The parameters W_j (weights matrix) and b_j (bias vectors) control the behavior of the DNN models by being adjusted repeatedly until the cost function reaches its minimum value [11].

Training DNN is a challenging task, and the problems of that include vanishing gradients, sticking in poor local optima, overfitting, and plateaus, all exist with adding more hidden layers which can make learning pretty slow. This is where that regularization methods such as l_1 regularization [13], l_2 regularization [13], or dropout regularization [14], and new optimization algorithms such as momentum [15], RMSprop [10], or Adam [16] can significantly contribute the learning algorithm.

TABLE I. DNN STRUCTURE FOR DETECTING THE LOCATION OF FBH AND BVID DEFECTS

Hidden layer number	Layer (type)	Output shape	Parameters
1	Dense layer	64	192128
1	Leaky ReLU	64	0
2	Dense layer	64	4160
2	Leaky ReLU	64	0
3	Dense layer	32	2080
	Leaky ReLU	32	0
	Dense layer	32	1056
4	Leaky ReLU	32	0
	Dropout (30%)	32	0
5	Dense layer	32	1056
	Leaky ReLU	32	0
	Dropout (30%)	32	0
Output layer	Dense layer	1	33

TABLE II. DNN STRUCTURE FOR ESTIMATING THE DEPTH OF FBH DEFECTS

Hidden layer number	Layer(type)	Output shape	parameters
1	Dense layer	128	384256
1	Leaky ReLU	128	0
2	Dense layer	128	15612
2	Leaky ReLU	128	0
2	Dense layer	64	8256
3	Leaky ReLU	64	0
Output layer	Dense layer	1	65

IV. PROPOSED METHOD

A. Dataset division

As mentioned before, the first aim of this study is the automated identification and retrieval of the location of 2 deepest defects in $CFRP_{FBH,5}$. However, to prove that our model can properly detect another type of defects, we also try to localize the location of BVID defect in $CFRP_{PBVID}$ with the same model. Therefore, we split our datasets as follows:

1) The first 4 regions shown in Fig. 2-a, including LDV data of the FBH defects 1, 2, 3, and their surrounding sound areas are chosen for the **train set**.

2) Regions 4 and 5 shown in Fig. 2-a, including LDV data of the FBH defects 4, 5, and their surrounding sound areas are chosen for our **first test set**.

3) The corresponding LDV data of $CFRP_{BVID}$ defects shown in Fig. 2-b is chosen as our **second test set**.

Furthermore, as the second aim, we use FBH defects for depth estimation. In this regards, we form our train/test set as follows:

1) 60% of vibrational response data of first dataset is used to train the second DNN.



Fig. 5. Top row and bottom row show the results for $CFRP_{FBH,5}$ and $CFRP_{BVID}$ datasets, respectively. From left to right: the corresponding output results of (a) FBD, (b) SVM, (c) AdaBoost, (d) SNN, (e) DNN with 3 hidden layers, (f) DNN with 5 hidden layers. The actual locations of defective areas in our test sets are shown by red dashed lines.

2) Then we test our model on the remaning 40%.

B. Proposed models

In the first model used for detection the location of 2 deepest FBHs, the network consists of 5 hidden layers, the first and second hidden layer has 64 neurons, but the last 3 hidden layers have 32 neurons. In the proposed model, neurons are densely connected as each neuron receives input from all the neurons in the previous layer. A modified version of Rectified Linear Unit (ReLU), known as Leaky ReLU [17]¹, is used as the activation function for each hidden layer except for the output layer, in which sigmoid function is used. Xavier initialization [18] is used for initializing weights matrix, and bias vectors are initialized with zeros. In the case of Xavier initialization, the weights of each layer are initialized as random draws from a truncated normal distribution centered on 0 and variance as

$$var = 2/n_{in} + n_{out},\tag{4}$$

where n_{in} and n_{out} are respectively the number of input and output units in the weight tensor. After initializing, we train DNN with Adam optimizer with constant learning rate of 10^{-4} and exponential decay rates for the moment estimates of $\beta_1 = 0.99$ and $\beta_2 = 0.999$ [16]. Adam optimizer is planned for training deep networks, and in fact, it is a combination of RMSprop [10] and stochastic gradient descent with momentum [15]. Dropout regularization [14] is used in the last 2 hidden layers, which is a popular technique that allows us to train on the bigger and deeper network by reducing overfitting possibility. It randomly selects some neurons and puts them aside of a layer during training. For loss function, binary crossentropy is used [10].

For the second aim, a regression model is used for estimating the depth of defects. It is almost similar to the first model, although it consists of 3 hidden layers, the first and second hidden layer has 128 neurons and the last hidden layer has 64 neurons. Also, linear activation function is used in the output layer and Mean Absolute Error (MAE) is employed as the loss function [10]. Tables I and II summarize our models. The parameters mentioned in the tables refer to the number of weights used in each layer of the networks. In training phase, each model is trained with 10 different random seeds, and the maximum training epoch number is set to 1000 for each of them. We use keras [19] with tensorflow backend [20] for all calculations in this work. Finally, we employ hyperparameter optimization with GridSearchCV tool using sciket-learn [21] (python package) for both DNNs.

V. RESULTS

After training, we test first DNN model, designed for defect detection, on LDV scan points of regions 4 and 5 of CFRP_{FBH,5} and also CFRP_{BVID}. For comparing the proposed model with other classification methodologies including AdaBoost, SVM, SNN (with 1 hidden layer), and DNN (with 3 hidden layers), 4 types of evaluation metrics, including, *accuracy*, *recall*, *precision*, and *f1_score* are used. These evaluation metrics are given as below:

$$Accuracy = \frac{TP+TN}{TN+TP+FP+FN} (\%)$$
(5)

$$Recall = \frac{TP}{TP + FN}(\%) \tag{6}$$

$$Precision = \frac{TP}{TP + FP} (\%) \tag{7}$$

$$f1_score = 2 \times \frac{\frac{Precision \times Recall}{Precision + Recall}}{(\%)}.$$
(8)

The mentioned classification methodologies which are used for the evaluation of our first DNN model are designed as follows: AdaBoost algorithm with SAMME.R [9] version is used, and SVM [7] with Radial Basis Function (RBF) is utilized. Hyperparameters of AdaBoost including n estimators and learning rate, and also hyperparameters of SVM including C and Gamma are fine-tuned with grid search algorithm. In addition, in order to show that our deep proposed model has better performance than shallower networks, we compare the proposed model with a 1-hidden layer SNN consisting 64 neurons and a 3-hidden layer DNN including 64 neurons in the first and second hidden layer and 32 neurons in the last hidden layer. Also for comparing our models with FBD according to the above mentioned evaluation metrics, we need to apply a threshold to the results of FBD. This threshold is chosen in such a way as to achieve best results for FBD method. The estimated shapes of defects obtained by the mentioned methods are shown in Fig.5. Tables III and IV show the quantitative results of the proposed DNNs, FBD technique, and some well-known clas-

¹ Leaky ReLU has two advantages over ReLU: 1) It solves dying ReLU problem due to having none zero slope in negative domain, 2) It increases the speed of training due to having mean close to zero.

sification methodologies for CFRP_{FBH,5} and CFRP_{BVID} datasets, respectively.

TABLE III: PROPOSED MODEL FOR DETECTING FBH DEFECT IN COMPARISON WITH FBD AND SOME CLASSIFICATION METHODOLOGIES. ALL VALUES ARE IN PERCENTAGE. THE BOLDED VALUES SHOW THE BEST RESULTS. (* DNN3 MEANS WITH 3 HIDDEN LAYERS, ** DNN5 MEANS WITH 5 HIDDEN LAYERS)

Evaluation	Classification methodologies and FBD					
metrics	FBD	SVM	AdaBoost	SNN	*DNN3	**DNN5
Precision	100	100	86.0	83.1	89.9	71.9
Recall	15.6	19.9	29.3	29.0	44.7	56.2
Accuracy	89.5	89.8	90.0	89.8	92.0	91.3
F1_score	27.0	33.2	43.7	42.9	59.7	63.1

TABLE IV: PROPOSED MODEL FOR DETECTING BVID DEFECT IN COMPARISON WITH FBD AND CLASSIFICATION METHODOLOGIES. ALL VALUES ARE IN PERCENTAGE. THE BOLDED VALUES SHOW THE BEST RESULTS. (* DNN3 MEANS WITH 3 HIDDEN LAYERS, ** DNN5 MEANS WITH 5 HIDDEN LAYERS)

Evaluation	Classification methodologies and FBD					
metrics	FBD	SVM	AdaBoost	SNN	*DNN3	**DNN5
Precision	99.7	99.6	100	99.5	99.6	99.4
Recall	68.5	47.3	5.50	67.5	68.2	75.3
Accuracy	98.7	97.9	96.4	98.7	98.8	99.0
F1_score	81.2	64.2	10.4	80.4	81.0	85.7

As can be observed in Figs. 5(f-1) and 5(f-2), defects 4 and 5 are estimated sufficiently with the proposed DNN model, while the FBD method is not able to detect these defects acceptably, especially defect 5 which is deepest. Besides, DNN model shows significant performance on CFRP_{BVID} as well as CFRP_{FBH,5}, while they are two different types of defects. The two most important factors in Tables III and IV, which is important to show the strength of the model in defect detection, are *f1_score* and *recall*. The results as highlighted in the tables for *f1_score* and *recall* show DNN5 (with five hidden layers) is superior in detecting defects in comparison with other methods.

TABLE V: ESTIMATED DEPTHS FOR 5 FBHS DEFECTS .

	Defects	actual depth	Estimated depth	
	Defect 1	1.02 mm	1.01 mm	
	Defect 2	1.84 mm	1.83 mm	
	Defect 3	2.98 mm	3.00 mm	
	Defect 4	3.81 mm	3.93 mm	
	Defect 5	4.69 mm	4.79 mm	
		Shallowest	<u>···</u>	Shallowest
•	0	Deepest Sound	• • • • • • •	Deepest Sound
	а		b	

Fig. 6. (a) Estimitaed shapes for five FBHs by the second proposed DNN model. (b) Estimated shapes for five FBHs by the FBD method.

Finally, as shown in Table V, using the DNN model designed and developed for depth estimation, depth of all FBH defects is estimated acceptably. It is worthwhile to mention that the FBD method could not detect defects' depth at all. FBD could just classify the first 3 defects while our second DNN could also classify 5 FBHs based on their depth properly as can be observed in Fig. 6. Also, we compare the performance of our regression model quantitatively with a 1-hidden layer SNN (with 128 neurons) and a 2-hidden layer SNN (including 128 neurons in the first and 128 neurons in the second hidden layer) based on Mean Square Error (MSE) criterion. This factor is computed as follows:

$$MSE = (1/_{M \times N}) \cdot \sum_{i=1}^{M \times N} (P_i - O_i)^2 , \qquad (9)$$

where $M \times N$ is the number of scan points. P_i and O_i are actual and estimated depth of scan points, respectively. Based on the results listed in Table VI, it can be concluded that DNN with 3 hidden layers has better performance than SNNs in estimating depth of defects due to having less MSE.

TABLE VI: PROPOSED MODEL FOR ESTIMATING FBH DEFECT IN COMPARISON WITH REGRESSION METHODOLOGIES. THE BOLDED VALUE SHOW THE BEST RESULT. (*DNN3 MEANS WITH 3 HIDDEN LAYERS, *SNN2 MEANS WITH 2 HIDDEN LAYERS, *** SNN1 MEANS WITH 1 HIDDEN LAYER)

Regression methodologies	MSE
*DNN3	0.12
**SNN2	0.15
***SNN1	0.26

VI. CONCLUSION

In this study, DNN method was used to localize the existed defects and then estimate the depth of defects in CFRP composites. Two CFRP coupons containing different types of damage were used as our datasets. Firstly, a DNN was designed for defects image segmentation of CFRP_{FBH,5} and CFRP_{BVID}. The subjective and objective results indicated that the proposed DNN model outperforms the other classification methodologies specifically in detecting deep defects. Secondly, another DNN model was proposed for FBH defects depth estimation. Results showed that the obtained DNN model is able to approximate depth of FBH defects promisingly. Thus, DNN can be employed as a proper method for detecting different kinds of damage especially deep defects and complicated damages in CFRP composites.

Further study aims to detect the mentioned damages with the combination of FBD and DNN which could be considered as an unsupervised method. Also, due to the challenging size of LDV data, using the convolutional neural network to detect and classify FBH defects can be promising.

Refrences

- I. Solodov, J. Bai, S. Bekgulyan and G. Busse, "A local defect resonance to enhance acoustic wave-defect interaction in ultrasonic nondestructive evaluation," Applied Physics Letters, vol. 99, no. 21, 2011.
- [2] I. Solodov, J. Bai and G. Busse, "Resonant ultrasound spectroscopy of defects: Case study of flat-bottomed holes," Journal of Applied Physics, 2013.
- [3] I. Solodov, "Resonant acoustic nonlinearity for defect-selective imaging and NDT," In Proceedings of the AIP conference, 2015.

- [4] J.Segers, et al "automated extraction of local defect resonance for efficient non-destructive testing of composites," 9th european workshop on health monitoring, July, 2018.
- [5] Y. Bengio, "Learning Deep Architectures for AI, "Foundations and Trends in Machine Learning," vol. 2, no. 1, pp. 1–127, 2009.
- [6] C. Cortes and V. Vapnik, "Support-vector networks," Machin learning 20, pp. 273–297, 1995.
- [7] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and Application to Boosting," Journal of Computer and System Sciences, vol. 55, no. 1, pp. 119–139, Agust, 1997.
- [8] J. Zhu, H. Zou, S. Rosset, and T. Hastie, "Multi-class adaboost," Stat. Interface, vol. 2, no. 3, pp. 349–360, 2009.
- [9] G. E. Hinton, "Deep belief networks," Scholarpedia, 2009.
- [10] I. Goodfellow, Y. Bengio, and A. Courville. Deep Learning. 2016. MIT Press.
- [11] A. Lozano-Diez, R. Zazo, D. T. Toledano and J. Gonzalez-Rodriguez, "An analysis of the influence of deep neural network (DNN) topology in bottleneck feature based language recognition," PLoS ONE, vol. 12, no. 8, August, 2017.
- [12] S. Feng, H. Zhou and H. Dong, "Using deep neural network with small dataset to predict material defects," Materials and Design, vol. 62, pp. 300–310, 2019.

- [13] A. Y. NG, "Feature selection, 11 vs. 12 regularization, and rotational invariance," In ICML, 2004.
- [14] N. Srivastava, G. Hinton, A. Krizhevsky, L. Sutskever and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," J. Mach. Learn. Res., vol. 15, no.1, pp. 1929-1958, 2014.
- [15] N. Qian, "On the momentum term in gradient descent learning algorithms," Neural Netw., vol. 12, no. 1, pp. 145–151, 1999.
- [16] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," Int. Conf. Learn. Represent., pp 1–13, 2015.
- [17] A. L. Mass, A. Y. Hannun and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," Proc. ICML, vol. 30, 2013.
- [18] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," Proc. Int. Conf. Artif. Intell. Stat., pp. 249–256, May, 2010.
- [19] F. Chollet et al., "Keras", https://keras.io, 2015.
- [20] M. Abadi et al., "Tensorflow: a system for largescale machine learning," in OSDI, vol. 16, pp. 265–283, 2016.
- [21] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," JMLR, vol. 12, pp. 2825–2830, 2011.