 CMC, vol.57, no.3, pp.521-532, 2018

Symmetric Learning Data Augmentation Model for Underwater Target Noise Data Expansion

Ming He[1, 2], Hongbin Wang[1, *], Lianke Zhou[1], Pengming Wang[3] and Andrew Ju[4]

Abstract: An important issue for deep learning models is the acquisition of training of data. Without abundant data from a real production environment for training, deep learning models would not be as widely used as they are today. However, the cost of obtaining abundant real-world environment is high, especially for underwater environments. It is more straightforward to simulate data that is closed to that from real environment. In this paper, a simple and easy symmetric learning data augmentation model (SLDAM) is proposed for underwater target radiate-noise data expansion and generation. The SLDAM, taking the optimal classifier of an initial dataset as the discriminator, makes use of the structure of the classifier to construct a symmetric generator based on antagonistic generation. It generates data similar to the initial dataset that can be used to supplement training data sets. This model has taken into consideration feature loss and sample loss function in model training, and is able to reduce the dependence of the generation and expansion on the feature set. We verified that the SLDAM is able to data expansion with low calculation complexity. Our results showed that the SLDAM is able to generate new data without compromising data recognition accuracy, for practical application in a production environment.

Keywords: Data augmentation, symmetric learning, data expansion, underwater target noise data.

1 Introduction

Deep learning has made impressive steps forward in various industries due to its capacity for nonlinear representation. However, it is dependent on input data, and model training and parameter optimization rely on abundant real-world data. Model training without data support and mathematical optimization are meaningless. It is expensive to collect abundant data from a real-world environment the collect on process itself typically requires a large amount of human resources, material resources and time. Additionally, more conditions appear, which reflect an increase in the dimensionalities of the data

[1] College of Computer Science and Technology, Harbin Engineering University, Harbin, 150001, China.

[2] College of Computer and Information Engineering, Heilongjiang University of Science and Technology, Harbin, 150022, China.

[3] School of Software and Microelectronics, Harbin University of Science and Technology, Harbin, 150080, China.

[4] Department of Computer Science and Information Systems, University of Limerick, Limerick, Ireland.

[*] Corresponding Author: Hongbin Wang. Email: wanghongbin@hrbeu.edu.cn.

collected. The intersection of multiple dimensions makes the data complexity increase exponentially. Furthermore, for some extreme tasks including tasks on plateau or in deep sea environments, the working condition and equipment condition make it impossible to collect data.

Based on antagonistic generation, we have designed an easy and simple symmetric learning data augmentation model (SLDAM) for underwater target radiate-noise data generation and expansion, It takes the optimal classifier of the initial datasets as the discriminator and makes use of the structure of the classifier to construct a generator, so as to simulate data similar to the initial datasets, in order to expansion the training data set for deep learning algorithms. Here, we describe the model and report verification that it is able to fulfill data expansion quickly with a lower calculation complexity. By analyzing the experimental data, we found that the SLDAM is able to generate large amount of new data while ensuring the data recognition accuracy is not compromised. The output is sufficient to be used with deep modelling application.

This paper is organized as follows: Related technology and development of data expansion are introduced in the second part, the symmetric learning model is introduced in the third part, the analysis of experimental results in found in the fourth part, the fifth part summarizes the defects of the SLDAM and introduces future work, and the sixth part concludes the paper.

2 Related works

There is often a lack of data for training and optimizing deep learning models. It is well-known that training a model with too small a dataset leads to overfitting. Therefore, during model training, researchers are interesting in training datasets or adjusting the model to make up for data deficiency.

Artificial training datasets are mainly generated through geometric transformation of images with shearing, rotation, reflection, flipping, zooming, translation, scaling, contrasting, noise disturbance, color switching and so on or the combination of several of the above methods to enlarge the size of the dataset [Wu (2014)]. However, the geometric transformation of artificially generated data cannot change the pixel value but change the location of the pixel, and the dataset scope is changed by data augmentation [Hobert (2011)], which can be used as the input to expect the invariant feature in network learning. Salehinejad et al. [Salehinejad, Barfett, Valaee et al. (2017); Salehinejad, Valaee, Dowdell et al. (2018)] have proposed data augmentation by polar coordinate transformation in paper, in which the pixel represented by original (x, y) was represented by (r, θ) , and later, it was represented as a two-dimensional picture.

In addition to the artificial generation of data, researchers have tried to solve the model training problem in small dataset by optimizing or adjusting the model for lacking datasets. As a mathematical method, regularization is able to solve the overfitting caused by small dataset [Kolář, Hradiš and Zemčík (2016)]. The introduction of regularization would minimize the training error while maximize the test error, and the adding of regular terms such as L1 or L2 in loss function is able to prevent overfitting. The defect of regularization is that a hyper parameter [Chen, Liu, Tang et al. (2017)] which needs to be set artificially is introduced into loss function. Dropout [Pham, Bluche, Kermorvant et al. (2014); Srivastava, Hinton, Krizhevsky et al. (2014)] is regarded as a regular term used in model adjustment, which is to remove certain neural network units temporally in accordance with certain probability in the training of deep learning network. The temporary here is compared with Batch. In the stochastic gradient descent, each of mini-batch is trained in different networks for they are removed randomly. Unsupervised Pre-training [Erhan, Bengio, Courville et al. (2010); Paine, Khorrami, Han et al. (2014)] is proposed in which the Auto-Encoder or RBM convolution is used to perform unsupervised training tier by tier. Finally, classifier tier is added as Fine-Tuning. However, it must be acknowledged that the effect of data augmentation is limited, and it is a wise and feasible method to choose a good dataset in model training.

The GAN model [Goodfellow, Pouget-Abadie, Mirza et al. (2014); Papernot, Mcdaniel, Goodfellow et al. (2016)] is a generative model proposed by Ian Goodfellow in 2014, which was not so popular at that time for its non-significant effect. Until 2016, with the deepening of deep learning, confrontation generation network was developed again. The combination of GAN with deep learning brings great generation capacity. GAN can be used to generate sample data whose distribution is consistent with the real data. Ledig et al. [Ledig, Theis, Huszár et al. (2017)] in Twitter have used VGG network as discriminator to work with parameterized residual network to represent generator so as to achieve the high score construction of low score pictures. Azad et al. [Azad, Ahmadzadeh and Azad (2015)] have blended simulation images and real images based on confrontation thought to take it as a training set to achieve human eye detection, so as to improve the anti-overfitting capacity of the model. Shrivastava et al. [Shrivastava, Pfister, Tuzel et al. (2016)] have proposed SimGAN to refine simulation images. In the SimGAN, self-regularization term is introduced to represent the classification, which is able to minimize combination error and maintain as many simulation images as possible. Meanwhile, the local confrontation loss function would discriminate each of the local image to make the local information more abundant. In addition, GAN can be also used to process speech and language, for example, to generate dialogue or to generate image with texts, etc. Fischer [Fischer (2017)] and Li et al. [Li, Monroe, Shi et al. (2017)] have used GAN to learn the implicit relevance between dialogues. Zhang et al. [Zhang, Gan, Fan et al. (2017)] have proposed to take CNN and LSTM as discriminator to solve the optimization problem with moment matching; and in the training, different from the traditional ones in which the discriminator parameter should be updated several times and then the generator is updated, it asks several updating of generator and then to update CNN discriminator. Yu et al. [Yu, Zhang, Wang et al. (2017)] have combined GAN with reinforcement learning to propose SeqGAN and to train generator based on policy gradient. The policy gradient is obtained through Monte Carlo search. And it is showed that SeqGAN has a better performance than the traditional methods in terms of speech, poetry and music generation. Reed et al. [Reed, Akata, Lee et al. (2016)] have proposed to generate image based on text description, and the text coding is used as the input of the generator.

3 The SLDAM

Currently, it is necessary to design a delicate network structure for the GAN, and to design an effective algorithm and iterative process for the training process. Generally, design of a GAN model has a process with high training cost. This paper, from the perspective of GAN designation ideas, designs a comprehensive loss function called SC-Loss which takes into consideration of both the sample loss and feature loss based on the sample similarity and feature overlapping. The design of SLDAM is easy and simple. It is showed that the quality of samples generated by SLDAM is better and the robustness in classifier is better, which is of great value in the underwater target recognition data expansion application.

3.1 Model designation and training

Based on the GAN, this paper proposes a method to enlarge the original data with confrontation idea based on current mature classifiers. In this paper, it is called symmetric data augmentation model as showed in Fig. 1. SLDAM is divided into two stages for training. The first is the training for classifiers, in which the current samples are used to train the classifiers to obtain a well characterized sample feature classifier, showed as the learn feature and representation in Fig. 1. The second stage is that for generating/expanding data, in which the effective sample feature obtained from former trained classifier is used as discriminator. And a reverse structure similar to the classifier is used to “translate” random noise into original data feature regarded by the model, and also the data is input to be compared with the sample feature data obtained from classifier. The model would construct a “new” data based on the existed features until it fails to distinguish the two features.

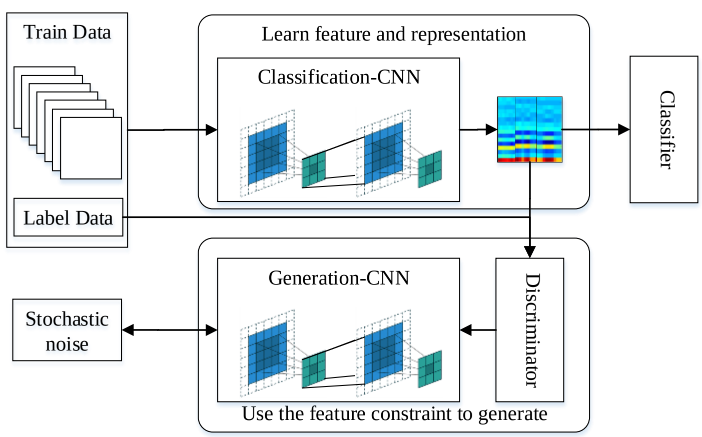


Figure 1: First proposal for the symmetric learning data augmentation model (SLDAM)

In the model’s training stage of classifier, the task of model training is to make use of the existed labeled data to train model effectively, so as to obtain available and qualified characterized sample data feature and classifier. In the stage of data expansion, the obtained qualified feature is used as the discriminating standard of this stage, which can be regarded as the application of automatic encoding machine in the convolutional neural field. It is feasible theoretically, however, there is a serious application problem in practice. The data generated in accordance with the above methods are different from the original sample, or there is lower accuracy when the generated data discriminate the classification than current classifiers. For these problems, the model and loss function are improved in this paper to balance these two problems. The process of SLDAM is shown in Fig. 2.

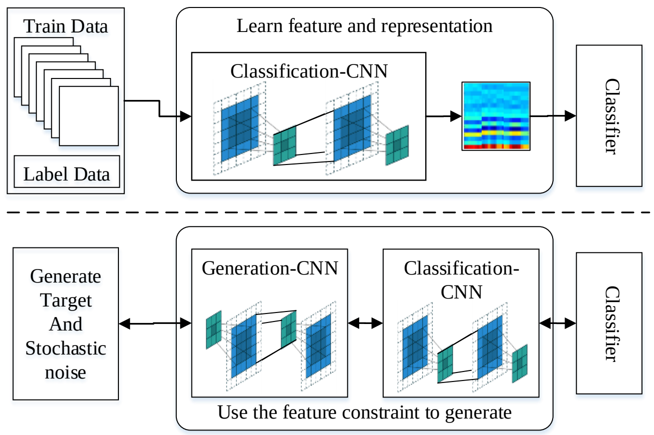


Figure 2: Outline of the final SLDAM used

The feature set used to discriminate generation results are replaced in this paper, and the convolutional neural network trained in classification is used. There are two advantages. The first is that it is difficult to discriminate the feature because the feature in feature set is a multiple-dimension tensor, and there is a great cost in the comparison of model generated feature and feature used as standard either in terms of time complexity or space complexity. Certainly, two feature tensors with same projection can be compared in the metric space with the methods of vector transformation or clustering. However, the introduction of other transformation equals to that of the third-party discriminating standard, which is not fair to the features extracted from the two dimensions and not of great reliability. The second is that discriminating based on the original feature cannot ensure the quality of data generated by the signals. When taking the feature as discriminating standard, its loss function can only be optimized with the feature related dimension gap being loss value, which is able to constrain the generated data to be obtained with feature in classification. It seems that the constraint method is reasonable, however, it is a little loose for the machine learning algorithm and practical sample using. Taking the original feature as discriminating method to discriminate generated data is able to generate data highly accepted by classifier, namely, the extracted data are able to be used to re-characterize data to a great extent. However, this characterization can only ensure the generated data being recognized by the current classifier but fail to ensure the generated data being accepted by other classifiers with a higher classification accuracy, namely, it cannot ensure that the generated data are same as the original sample data.

Oriented by the "similarity" and "same", a loss function to discriminate similarity and feature overlapping ratio is proposed in this paper, which is called Similar Coincidence loss function (SC-Loss).

3.2 Similar-coincidence loss function

Loss function in machine learning is set oriented to single sample, representing the gap between model predicting value and sample's true value. It is used to discriminate the error between generated feature and original feature. It is same as the description above that making use of the error between two groups of features can only ensure the generated data being recognized by the classifier with a higher accuracy, but fail to ensure the similarity between generated data and original input sample, meanwhile, the generated data guided by the loss rule can only show a good accuracy in current classifier but may not in other classifiers. That is to say, if the current classifier is not an authorized classifier, the data generated under its standard can only be used in it and the generated data is of no generality, namely, there may be "overfitting" in the generated data.

From the perspective of loss function designation, the error minimization taking into consideration of all samples is defined as model's experimental risk, showed as Eq. (1). It is to minimize the average of the loss function of all samples in the training set, and the minimization of which is to minimize the equation.

$$R_{emp}(f)=\frac{1}{N}\sum_{i=1}^{N}L(y_i, f(x_i)) \tag{1}$$

Where the *N* in the cumulative sum is the number of samples in training sample, and *i* is used to represent sample order.

When taking into consideration of experimental risk only may lead to overfitting, in which the extreme condition is that the model f(x) is able to predict all the samples in the training set but fail to predict the sample data out of the set. When there are unknown samples, the overfitting can be reduced by introducing structural risk into loss function to minimize the structural risk in the loss function. Structural risk is the comprise between experimental risk and expectation risk. To add a regular term after experimental risk function is the structural risk, and it is showed as the Eq. (2).

$$R_{str}(f)=\frac{1}{N}\sum_{i=1}^{N}L(y_i, f(x_i))+\lambda J(f) \tag{2}$$

Where λ is a coefficient greater than 0 and J(f) is the complexity of model f. For the Eq. (2) is able to observe that the complexity of model decision function is the prerequisite of overfitting, the prerequisite should be eliminated if we want to prevent the overfitting, that is, to reduce the complexity of decision function so as to minimize J(f)J(f). Based on the above consideration, the loss function in this paper is showed as Eq. (3).

$$Loss(y, fs(x), g(x))=\frac{1}{N}\sum_{i=1}^{N}\left[loss_c\left(fs(y_i), fs(x_i)\right)+loss_s(y_i, g(fs(x_i))\right] \tag{3}$$

Where N is the number of input samples, y is the input sample, fs(x) is the extracted feature set, and g(x) is the reverse sample obtained based on feature x. And Lossc and Losss are the feature loss and sample loss respectively.

$$Loss_c(y, fs(x)) = fs(y_i) - \alpha \bullet fs(x_i) \tag{4}$$

The proposed loss function in this paper is composed of two parts, one is to calculate error between features, namely, the Lossc, is showed as Eq. (4). By calculating the error between the feature generated by model and feature extracted by original classifier model as well as minimizing the error, the model is updated reversely. The optimization direction in this part is to regulate the sample generated by models, which is generated based on feature of original sample sets, therefore, it can ensure that the generated data are observable in original dataset, where α is the generating expansion factor. In the following experiments, when its value is in the range of [1-2.2], the accuracy of generated data is higher than the original data. The other part is used as the penalty item of Latin model to generate performance of model in other classifiers, namely, the Losss, is showed as Eq. (5), to constrain the quality of generated sample.

$$Loss_s = y_i - g(fs(x_i) \tag{5}$$

It is verified by the experiments that the loss function proposed in this paper is able to ensure the performance of generated sample in the current classifiers. Meanwhile, the sample data generated in this model have the same performance as the original dataset in other samples.

4 Experiments

The experiment is finished in the underwater target recognition dataset. There are two aims in the experimental verification, the first is to verify the quality of sample data generated by the proposed model; and the second is to verify the capture capacity of the feature generated by the model proposed in this paper, namely, the comparison between the number of generated sample and the recognition accuracy rate of generated sample under current model.

4.1 Underwater target dataset

At present, the feature extraction and recognition of underwater target are finished with such methods as time sequence information structure of radiated noise, power spectrum feature or time-frequency spectrum analysis. In the current experimental condition, the experiment collects data by setting several underwater target noise collecting equipment in real waters. In terms of real observing value, the width of the river is a dynamic value and it is affected by the outside environment greatly. And the target vessel moves in the river along different directions at different speeds in accordance with the experimental requirements. The vessel moves at different speeds and takes different sounding bodies, and they are defined as different targets, and the specific data number is showed in Tab. 1.

Table 1: Total number of experiment data and distribution

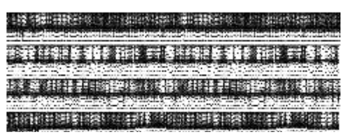
Input Data	Audio parameters				Total
	channels	Digitalizing bit	Sampling frequency	Sampling point	
3K	1	2	48000	4800	4879
5K	1	2	48000	4800	5045
9K	1	2	48000	4800	4821
13K	1	2	48000	4800	4836
15K	1	2	48000	4800	5122

It can be observed from Tab. 1 that the underwater target data adopted in the experiment are multiple, in which the 3 K-15 K data are target noise obtained from self-soundings in real underwater activities. The number of samples after segmentation reaches 18000, which meets the application amount of deep learning model training. In this paper, 3 K, 5 K and 9 K data are used in the experiment.

4.2 Comparison of generated data quality

The aim of this paper is to solve the problem that the underwater target recognition experiment fails to be implemented for the sake of geography or actual difficulties and the experiment cannot be done so that multiple repeated data cannot be obtained from an experiment. For the underwater target data, because of the fact that the generated sample data do not match the category in DCGAN, it cannot be applied in problems asking rigorous classification such as generating underwater target data. The SLDAM method proposed in this paper generates the sound frequency spectrum of underwater target data, and the generated sample of 9K data are showed in Fig. 3, in which list(a) is the frequency spectrum figure of original noise in sample; and the list(b) is that of the sample generated by SLDAM.

(a)



(b)

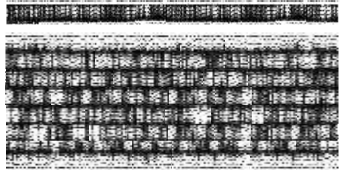


Figure 3: Frequency spectrum of original sample (a) and that of sample generated by SLDAM (b)

In this paper, 2800 samples in 9 K datasets are selected randomly to be compared in the experiment. And the number of samples generated by SLDAM is showed in Tab. 2, where α is the expansion factor.

Table 2: Accuracy of samples generated under different *a*

Order	Original sample number	Accuracy	Expansion factor	Generated sample number	Accuracy of generated sample in original classifiers
1	2800	93.00%	1	2800	95.00%
2	2800	93.00%	1.2	3360	94.70%
3	2800	93.00%	1.7	4760	94.30%
4	2800	93.00%	2	5600	95.10%
5	2800	93.00%	2.2	6160	93.00%

It can be seen from Tab. 2 that the recognition accuracy of original classifier is not reduced and the data are expanded effectively when α ranges from 1-2.2. The accuracy rate of generated data is higher than that of original data in the original classifier, showing that the proposed SLDAM is able to make use of feature captured by model effectively to generate data. Because feature is recognized by the classifier, the generated sample data is easy to be recognized by the classifiers.

4.3 Experiments in other classifiers

Tab. 3 showed the accuracy of different generated samples in underwater target recognition dataset. And the original data model is LeNet-5, which is iterated for 10,000 times to reduce loss to 0.0026598 with the accuracy rate being 99.11%. The accuracy of original sample data expanded with RT which are classified with LeNet-5 is decreased to 26%, reflecting that the data expansion of RT does not fit the LeNet-5 model. Similarly, though the data generated by DCGAN model is close to the original data, its recognition accuracy does not reach the recognition effect of original dataset. Similar to the above reason, the main reason is that the classification of sample data generated by DCGAN is random, while its accuracy in SVM classifier is obviously higher than that in original classifier. It also reflects indirectly that the quality of sample data generated by DCGAN is higher. SLDAM and its simplified SLDAM-s are used in this paper. The latter mainly takes the feature loss in generating sample data process as discriminating standard but overlooks the sample loss. The SLDAM proposed in this paper adopts the symmetric structure, whose classifier is the classification structure of LeNet-5, therefore, the classification recognition accuracy of the sample data generated by the two models in original model are higher, with no much relevance with the original data. However, there is a greater difference between them in classification recognition rate in SVM. Because there is no constraint in the sample loss in SLDAM-s, though the generated sample data have a higher classification accuracy in original classification model, it overfits the original feature instead of the original data for its data are composed of features captured by original model, therefore, a better performance cannot be obtained in classifiers that are not discriminated by original feature sets. Certainly, it has not been predicted that the accuracy rate would be 3%. Looking back to SLDAM, the generated sample data perform better in SVM for taking into consideration of sample loss.

Table 3: Discriminating generated sample with other classifiers

Order	Model	Accuracy of original model	Accuracy of SVM
1	RT	73.0%	67.4.0%
2	DCGAN	89.0%	80.2%
3	DCGAN-ordered	97.6%	94.6%
4	SLDAM-s	98.9%	93.0%
5	SLDAM	98.7%	95.0%

5 Further work

In terms of experimental results, the SLDAM proposed in this paper is feasible and easy to be realized for its structure, which can be applied in underwater target data expansion and generation. Symmetric augmentation model is feasible in application because it is able to construct network and generate data fast, however, it depends too much on the classifier, and when the symmetric generation network is constructed in a rather mature classifier, the qualified feature captured by the symmetric generation model with classifier is able to generate qualified and reliable data, so as to enrich the original dataset to expand the data effectively. However, when the capacity of classifier to characterize dataset is weaker, the quality of sample data generated by symmetric generation model is decreased quickly with a general performance. Meanwhile, the SLDAM proposed in this paper fails to solve the condition in which there is few generated certain target data, meanwhile, it fails to generate and expand data of nonexistent target in dataset, namely, problem of few-shot and zero-shot, which is also the problem to be improved and studied for this paper.

6 Conclusion

To cope with situations where it is impossible to recurrently obtain experimental underwater target recognition data under the same condition due to natural conditions easy and simple SLDAM. It takes into consideration of sample loss and feature loss to ensure the quality of generated data and its reliability in classifier. Symmetric augmentation model provides a feasible data expansion and generation method for underwater target noise based on confrontation generation thought. It was able to construct a model and generate data from an existing dataset, Simulations showed that a better classification performance could be obtained with the symmetric augmentation model than the original data, which reflects that the symmetric generation model is able to capture the features of the original model effectively and they work well together. However, there are defects in the symmetric generation model, since it depends heavily on the classifier. While it is able to characterize data with the features found by the classifier, when the classifier has a lower feature extraction capacity in original dataset, the symmetric augmentation model also has a lower performance. The symmetric augmentation model fails to solve problems like zero-sample problem-this will be the subject of future studies.

Acknowledgement: This work was funded by the National Natural Science Foundation of China under Grant (No. 61772152 and No. 61502037), the Basic Research Project (No. JCKY2016206B001, JCKY2014206C002 and JCKY2017604C010) and the Technical Foundation Project (No. JSQB2017206C002).

References

Azad, R.; Ahmadzadeh, E.; Azad, B. (2015): Real-time human face detection in noisy images based on skin color fusion model and eye detection. *International Journal of Intelligent Systems*, vol. 309, no. 1, pp. 435-447.

Chen, G.; Liu, T.; Tang, Y.; Jian, Y.; Jie, Y. et al. (2017): A regularization approach for instance-based superset label learning. *IEEE Transactions on Cybernetics*, vol. 99, no. 99, pp. 1-12.

Erhan, D.; Bengio, Y.; Courville, A.; Manzagol, P. A.; Vincent, P. et al. (2010): Why does unsupervised pre-training help deep learning? *Journal of Machine Learning Research*, vol. 11, no. 3, pp. 625-660.

Fischer, J. V. (2017): A case study on the relevance of the competence assumption for implicature calculation in dialogue systems. *International Conference of the German Society for Computational Linguistics and Language Technology*, pp. 276-283.

Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D. et al. (2014): Generative adversarial nets. *International Conference on Neural Information Processing Systems*, pp. 2672-2680.

Hobert, J. P. (2011) The data augmentation algorithm: Theory and methodology. *Handbook of Markov Chain Monte Carlo*, pp. 253-293.

Kolář, M.; Hradiš, M.; Zemčík, P. (2016): Deep learning on small datasets using online image search. *Spring Conference*, pp. 87-93.

Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A. et al. (2017): Photo-realistic single image super-resolution using a generative adversarial network. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 105-114.

Li, J.; Monroe, W.; Shi, T.; Jean, S. B.; Ritter, A. et al. (2017): Adversarial learning for neural dialogue generation. *Computation and Language*.

Paine, T. L.; Khorrami, P.; Han, W.; Huang, T. S. (2014): An analysis of unsupervised pre-training in light of recent advances. *Computer Science*.

Papernot, N.; Mcdaniel, P.; Goodfellow, I.; Jha, S.; Celik, Z. B. et al. (2016): Practical black-box attacks against machine learning. *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, pp. 506-519.

Pham, V.; Bluche, T.; Kermorvant, C.; Louradour, J. (2014): Dropout improves recurrent neural networks for handwriting recognition. *International Conference on Frontiers in Handwriting Recognition*, pp. 285-290.

Reed, S.; Akata, Z.; Lee, H.; Schiele, B. (2016): Learning deep representations of fine-grained visual descriptions. *IEEE Conference on Computer Vision & Pattern Recognition*, pp. 49-58.

Salehinejad, H.; Barfett, J.; Valaee, S.; Dowdell, T. (2017): Training neural networks with very little data-A draft. *Computer Vision and Pattern Recognition*.

Salehinejad, H.; Valaee, S.; Dowdell, T.; Barfett, J. (2018): Image augmentation using radial transform for training Deep Neural Networks. *Computer Vision and Pattern Recognition*.

Shrivastava, A.; Pfister, T.; Tuzel, O.; Susskind, J.; Wang, W. (2016): Learning from simulated and unsupervised images through adversarial training. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2242-2251.

Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. (2014): Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929-1958.

Wu, Y. N. (2014): Data augmentation. *Computer Vision*, pp. 165-166.

Yu, L.; Zhang, W.; Wang, J.; Yu, Y. (2017): Seqgan: Sequence generative adversarial nets with policy gradient. *AAAI Conference on Artificial Intelligence*, pp. 2852-2858.

Zhang, Y.; Gan, Z.; Fan, K.; Chen, Z.; Henao, R. et al. (2017): Adversarial feature matching for text generation. *Machine Learning*.