Data Augmentation for Object Detection: A Review

Parvinder Kaur
Department of CSE
SUSCET, Tangori
Mohali, India
parvinderkaur1984@gmail.com

Dr. Baljit Singh Khehra
Department of CSE
BBSBEC, Fatehgarh Sahib
Fatehgarh Sahib, India
baljitkhehra@rediffmail.com

Er. Bhupinder Singh Mavi
Head of Engineering
HNC Virtual Solutions
Wixom, MI, USA
bmavi@outlook.com

2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS) | 978-1-6654-2461-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/MWSCAS47672.2021.9531849

***Abstract*— Deep learning has been a game changer in the field of object detection in the last decade. But all the deep learning models for computer vision depend upon large amount of data for consistent results. For real life problems especially for medical imaging, availability of enough amounts of data is not always possible. Data augmentation is a collection of techniques that can be used to extend the dataset size and improve the quality of images in the dataset by a required amount. Logically it is used to make the deep learning model independent of the counterfeit features of the data space. In this paper a comprehensive review of data augmentation techniques for object detection is done. Problem of class imbalance is also outlined with possible solutions. In addition to train time augmentation techniques an overview of test time augmentations is also presented.**

Keywords—convolution neural network, deep learning, overfitting, normalization.

I. Introduction

Deep learning gave a gigantic increase to the already speedily developing arena of computer vision. With deep convolution neural networks, a vast variety of new applications of computer vision have been announced and are playing a significant role in our daily lives. These include object detection and recognition, face recognition, image content retrieval, self-driving cars, image search, image style transfer, image colorization, semantic segmentation, robotics, medical image analysis and disease detection, online shopping and many more[1] [2] [3]. The latest gush of curiosity in deep learning techniques is due to the reason that they have been beating previous state-of-the-art techniques in many domains. The problem of data insufficiency is very crucial since data are at the core of any deep learning model. The lack of sufficient training examples is every so often answerable for deprived performances in deep learning solutions to problems [4] [5]. Having a large amount of data for training is vital for good performance of any deep learning model. Image data augmentation is a procedure that can be used to extend the size of dataset required by deep learning model for training by artificially generating variations of existing actual images in the dataset. Training a deep learning model includes, tweaking its parameters in order to map a specific input to some defined output. Goal is to hunt the spot where model's loss is minimal and it can only be achieved by fine-tuning the parameters in a right direction.

The remainder of this paper is organized into five sections. In Section II the main challenge associated with deep learning models is discussed. Section III represents various augmentation techniques used in literature. Section IV presents various image data augmentation techniques with illustrative examples. Section V represents additional design decisions that should be considered while applying augmentation techniques to image data. Finally the whole paper is concluded in section VI along with significant findings.

II. Challanges In Deep Learning

There is no distrust that deep convolution neural networks are outperforming conventional methods for many computer vision applications. But, these cannot be considered as a rule of thumb that can be applied to all kind of problems and dominate all conventional technologies. There are some limitations and challenges that hinder their performance to beat the human vision system [6]. The biggest challenge is improving generalization performance of deep learning model. Large amount of labelled training data is the key to the success of any deep learning model. But it is not possible to give every possible labelled example related to a problem domain in the real world scenario. Medical imaging is such a domain where we always lack the real world training samples. A good deep learning model should respond to the new input data in the same way as it responds to previously seen data. This ability of a model is known as generalization. Whenever the trained model is given the unseen image it is not able to sometimes classify the image correctly. It may be due to the problem of underfitting and overfitting [7]. Overfitted model is not able to capture the patterns in the image in proper way. The problem of underfitting can be due to two reasons. First reason is when there the lack of availability of sufficient data for training. Other reason can be an attempt to make a linear model from non-linear data. The problem of overfitting often arises because the model has too many parameters to learn; hence the system memorizes the dataset instead of learning the underlying concept in the training dataset. So there is a need to improve the generalization performance. Generalization is the ability to classify correctly the newly seen patterns by the experience of previously seen patterns. The substantial dependence on accurate and plenty of data also makes deep learning models susceptible to fooling. Deep learning systems are pretty good at some huge portion of a given problem domain, but still can be easily fooled. Class imbalance is another major issue in deep learning [8]. The classes with abundant of exemplars are called minority classes and the classes with shortage of exemplars are known as minority classes [9]. Image data augmentation is one method that can be used successfully to improve the generalization ability of a model and addresses the problem of class imbalance. Both are the main attention of this paper. Some other methods also provided in literature and are outlined here for comparison purpose:

A. Reduce the Number of Parameters

Overfitting is directly proportional to number of parameters. Fewer the no. of parameters, fewer will be the no. of layers resulting in fewer activation functions and fewer will be the chances of overfitting. More activation functions lead to more non-linearity [10]. Matching of Convolution Neural Network Complexity with dataset complexity is the key to success but there is no magic bullet of doing this.

B. Dropout Neurons

The concept of dropout neurons makes the neurons in the model more independent of the neighboring neurons and thus making the model more agile to learn [11]. At every instant of time while training there are some neurons that are dropped and thus making the model less complex than the original in terms of weights. The process is equivalent to "ensembling" in which many different neural networks are trained and their weights are combined to give better performance. But adding dropout neurons makes it possible only by the use of a single convolution neural network. At every point of time the neural network has a different configuration as the neurons are dropped in random fashion. But once the training is started, dropout will not give good results. To get better performance it is required to train till final convergence.

C. Weight Regularization

As more and more training examples are provided to model the weights go on increasing and become unnecessarily large. Increased input output variance is the consequence of these large weights. The result is that the network will not be able to distinguish between the similar inputs. All this leads to unpredictable results on the testing data hence poorly generalized system. Moreover larger weights make the model more complex leading to overfitted system. Two possible solutions to overcome this problem are L1 and L2 weight regularizations [12].

D. Transfer Learning

Transfer learning techniques can be a choice when adequate training data is not available, and the pretrained and target models have some similar problem domains but are not exactly alike. Transfer learning [13] uses the weights from a pretrained model trained for the similar task on the larger dataset. The new network attains the totally new structures but it uses the weights of already trained model. A dissimilar task solved with the help of quite similar problem can act as a starting point to extract learned features and tweak the model, decreasing the separation paradigm of having two different problems [14]. But the problem of negative transfer and overfitting again hinder the performance sometimes. It is very perplexing to find the solutions to decreased performance due to negative transfer learning [15].

E. Batch Normalization

Batch normalization increases learning independence among hidden layers by reducing the covariance shift. It is done to speed up the learning process and to reduce the use of dropout neurons in the network. Only using batch normalization will not improve the results, so it should be used in combination with regularization and dropout to reduce overfitting.

III. Related Work

Before beginning with the augmentation techniques a brief review of literature is must. Data augmentation is a technique that solves the main cause of overfitting or poor generalization by adding more training examples. All this saves time and cost as capturing a large number of new images for any domain is cumbersome. Jonas Nilsson *et al*. in 2014 generated augmented images by mapping virtual pedestrians onto real image backgrounds and used it for pedestrian detection [16]. The results were improved by a good factor with augmented data combined with real data as compared to only real data. Sebastiaen C. Wong et al. in 2016 used data warping and synthetic data sampling, the two new approaches for data augmentation [17]. Data warping created more exemplars in data space and synthetic data sampling created more exemplars in the feature space. For checking new data sets potential they used convolution back propagation neural network. They concluded that data warping provides better results than synthetic data sampling if reasonable transforms for the data can be identified. Joseph Lemle *et al.* in 2017 used the concept of smart augmentation. They automatically combined the images in a way that improved regularization [18]. To learn the best suitable augmentation technique for a given class of input data, smart augmentation learns to combine two or more examples in one class. Jia Shijie et al. in 2017 evaluated the imapact of various data augmentation techniques on image classification tasks. They used flipping, rotating, PCA jittering, color jittering, noise, Generative Adversial Networks (GANs) etc [19]. They concluded that cropping, flipping, WGAN and rotation outperperformed the other methods. Kosaku Fozita *et al.* in 2018 proposed a new data augmentation strategy based on the idea of image transformation [20]. They added some noise to the images and generated new samples in order to reduce scarcity in a given feature space. The model trained with augmented dataset outperformed the model trained with original dataset. Medical image analysis always faces the problem of data scarcity. Real images of every disease are not always available in abundance. Abdulaziz Namozov *et al.* in 2018 used augmentation techniques to computed tomography scan images dataset and showed that it boosted up the model's classification ability [21].Cheng Lei *et al.* evaluated the impact of some parameters like augmentation method, size of basic dataset per label, augmentation rate etc. on the performance of deep learning model [22]. They concluded that geometric transformations always do not improve the performance. Blending two geometric transformations tainted the generalization. Combining geometric transformation with photometric transformation attained better results. Cherry khosla et *al.* in 2020 did a survey on various data augmentation techniques based on data warping and oversampling. They also addressed the problem of overfitting in their survey [23].

IV. Data Augmentation Techniques

Data augmentation techniques can be broadly divided into following categories: techniques based on geometric transformations, techniques based on color transformations, techniques based on random occlusion and techniques based on deep learning. Figure 1 depicts the classification of data augmentation techniques.

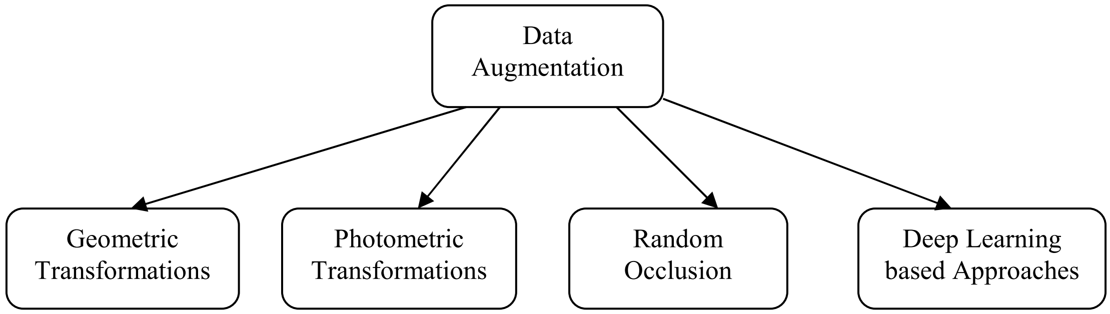

Fig.1. Taxonomy of data augmentation techniques

A. Techniques based on Geometric Transformations

Geometric transformations tend to change the pixel position of the image in the training dataset. These include rotation, scaling, flipping, cropping, padding, translation, affine transformation etc.

Interpolation is used in scaling and rotation transformations in order to adjust the gap introduced in between the original pixels. Replication of pixels can also be done depending upon the situation. Figure 2 depicts various geometric and photometric transformations.

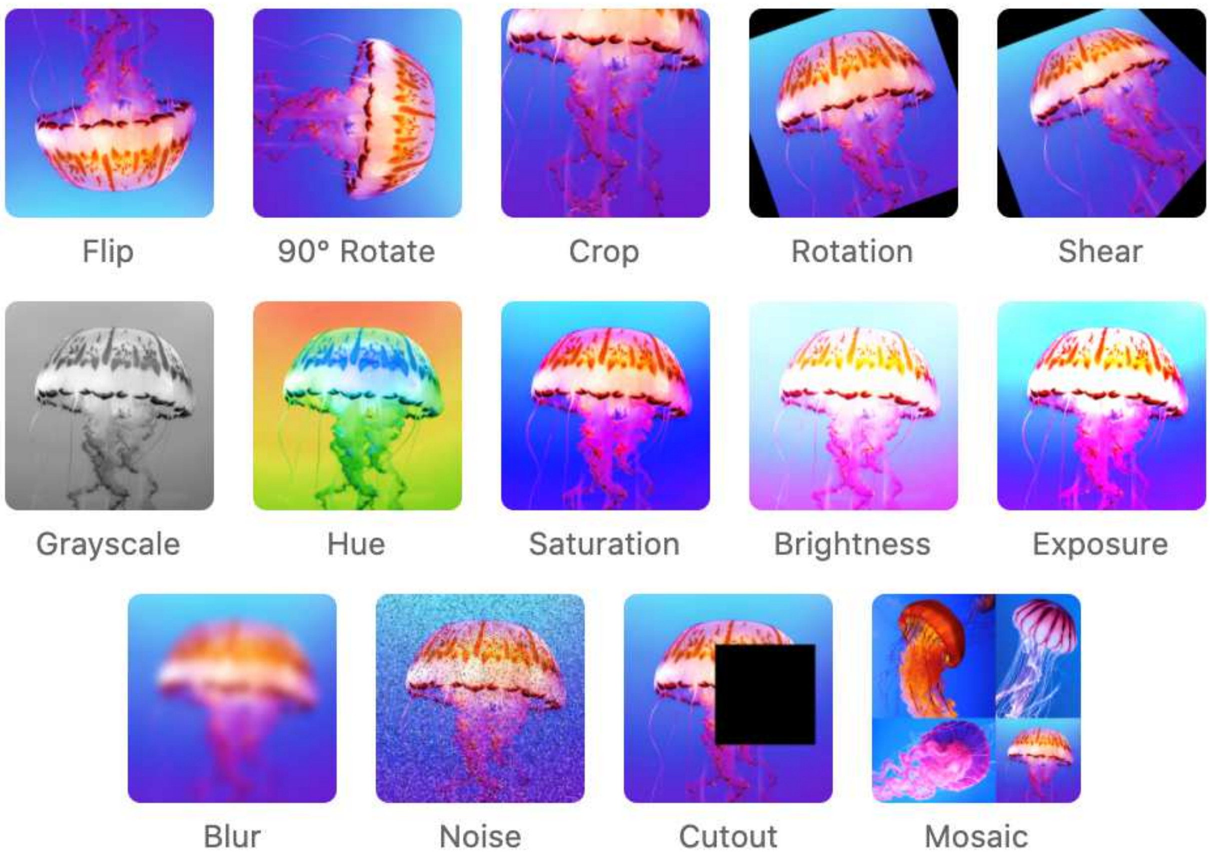

Fig. 2. Some geometric and photometric transformations examples

- Scaling changes the size of a given image by a given amount. Height and width of the image both can be changed. It is helpful to make the network familiar with objects at different scales.
- Cropping selects a portion of the image and discards the remaining image. Cropping is very useful technique to add variety to training examples. But it should not end up with the loss of object of interest.
- Flipping creates mirrored copies of images. Creating many variations of images by flipping them at different orientations presents the model with more information. Best example of flipping is when you take your selfie with front camera versus the photograph taken with back camera of your phone. In real world application like face recognition both versions may be of equal importance. Horizontal and vertical are the most common type of flip transformations.
- Padding adds some value at all the sides of an image.
- Affine transformations map intensity value at some point in image to some other new point while preserving the points and straight lines. Scaling, translation, shearing, rotation are special cases of affine transformations.

Alexnet [24] used random cropping, flipping and rotation as its main augmentation technique and increased the training data size by a factor of 2048. VGG16 [25] applied random multiple scaling to each training image. The range of rescaling was from 256 to 512. ResNet [26] also used the same idea of random rescaling and then a crop was taken after flipping and color augmentation.

B. Techniques based on Photometric Transformations

Photometric augmentations tend to change the pixel or intensity values instead of pixel positions. Basically they are based on the change in color components of an image. Color augmentations include changing the brightness, contrast, hue, saturation, noise in an image. Color is very important component of an image. It is sometimes helpful to recover the intrinsic physical properties of the scene.

- PCA based Color Augmentation computes PCA components on all the images in the data set. Then based on these computed PCA components an intensity value is calculated and added to each channel of RGB image. All this changes the color balance of an image by changing color illumination and intensity. AlexNet [24] uses this type of augmentation technique and achieves error rate reduction by 1 percent.
- . VGG16 [25] also applied color augmentation by simply computing mean RGB value on the training set and then subtracting it from each pixel in the image. These operations are done on individual RGB channels.

C. Techniques based on Random Occlusion

A good object detection or image recognition model should be robust to image occlusions. In other words it should be able to correctly detect objects even if some part of the object or image cannot be seen or occluded. Random occlusion augmentation techniques make the occluded images available for training and thus providing better information or experience to the model. Basic random occlusion techniques are random erase, cutout, hide and seek, grid mask, cutmix and mosaic augmentation etc. Authors and Affiliations

1) Random Erasing [27]: interchanges some portions of the image by some random values or mean pixel values of the training dataset. It can be thought of a regularization technique and hence avoids memorization of training dataset. Figure 3 depicts the concept of random erasing.

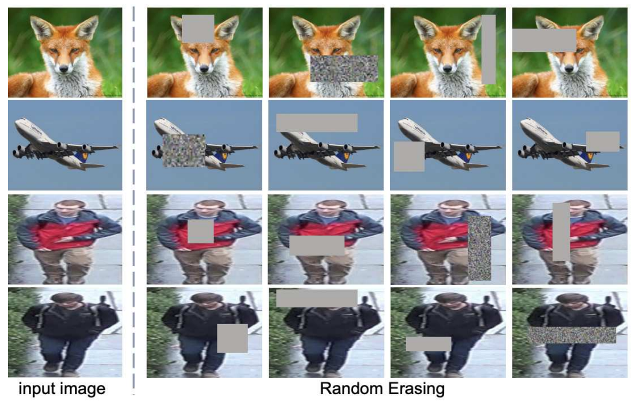

Fig. 3.Random erasing [27]

2) Cutout: is same as random erase except that the cutout portions are not shown to first layer of the CNN. It also helps to improve generalization and overcome overfitting. Figure 4 shows the mixup, cutout and cutmix transformation examples:

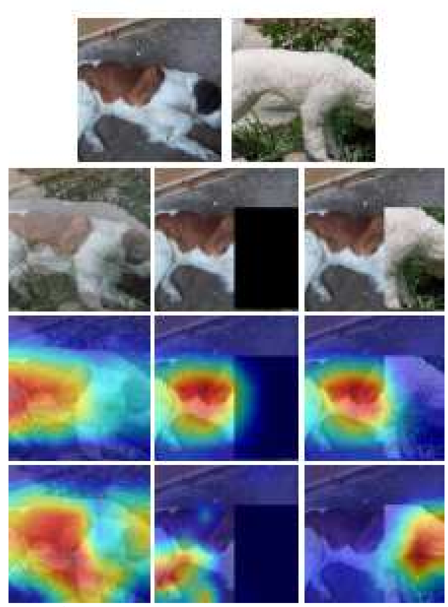

Fig. 4. Examples of photometric augmentations [28]

3) Cutmix: removed the drawback of all the above random occlusion techniques. All the techniques discussed above lead to loss of information. Cutmix attains the good properties of these techniques and on the same time removes the drawback of important information loss. Random regions of the image are cut and pasted in the manner shown in figure 4 above.

4) Hide and seek: tends to hide arbitrarily some parts of the object to be detected and forces the model to seek the other pertinent features when the most important features are not available for an object to be distinguished from another.

5) Grid mask [29]: overcomes the drawbacks of random erasing and cutout by covering the whole image. Thus the chance of losing the whole object during cutout or random erase is totally removed. The image is divided into many uniformly distributes square tiles and some of them are randomly erased. Concept is shown in figure 5 given below:

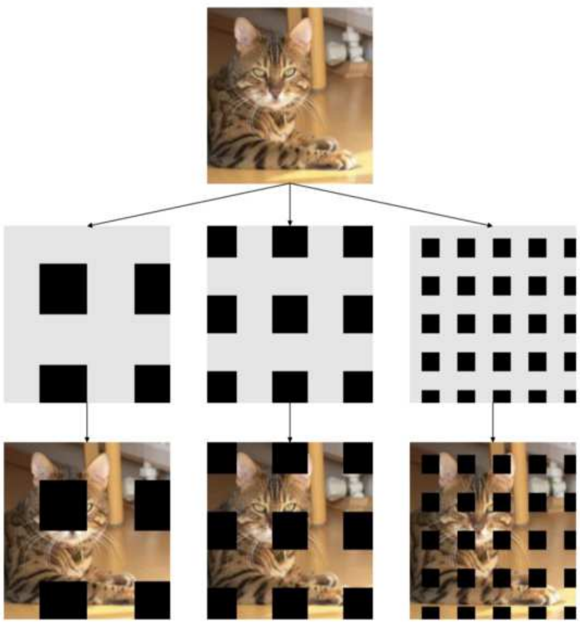

Fig. 5. Concept of Grid Mask [29]

6) Mosaic Augmentation: joins four images into one single image. As a result the objects in the joined image appear at a smaller scale than the original image. This kind of augmentation is beneficial in improving the detection of small objects in images and is used first time by YoloV4 [30]. Figure 6 shows the result of joining four images using mosaic augmentation.

Fig. 6. Mosaic augmentation

D. Techniques based on Deep Learning

All the three kind of augmentations discussed above do not always capture all the disparities in the environments. Also there are chances of lost information or features of the original dataset because these techniques try to change the geometry or lighting conditions of the images. Now-a-days deep learning based methods for data augmentation are providing more convincing transformations [31]. Some of such techniques are Neural Style Transfer (NST), Adversial training and Generative Adversial Networks (GAN).

1) Neural Style Transfer: makes style transformations in an image. One image's style is transferred to another image's content. All this is done by training a deep neural network [32]. Some samples generated using NST are shown below in figure 7:

Fig. 7. Training samples generated from neural style transfer

2) Adversial Training: generates some filters that are applied to training images to produce more examples [33]. These newly generated images are able to deceive a deep learning model in such a way that the model misclassifies them. These images are given as a part of training dataset and model is then experienced enough to classify them correctly.

3) Generative adversial network: [34] is made up of two components- generator and discriminator. Both these components are neural networks that are trained concurrently. Generator generates forged images from the underlying space and the discriminator extricates the synthetic forged images from the real images. Claus Aranha et al. in 2019 [35] used GAN to generate absolutely artificial data for the problem of cancer detection and found outstanding results as compared with real limited dataset.

V. CLASS IMBALANCE

Class imbalance is another major issue that should be addressed to get better performance from a deep learning model. The problem occurs when you have lesser number of training samples in one of many classes you want to classify. In medical imaging this is a major problem. For example in cancer detection the images belonging to cancer category will be few as compared to non-cancer images. It is very difficult to get good results for the imbalanced class. Moreover the validation of such classes is difficult to be tested due to the lack of samples. The problem can be solved using resampling methods discussed below:

1) Undersampling: deletes the training samples from the class that has abundance of samples. Random undersampling is a common method in which the samples are randomly picked and deleted from the majority class. But major problem with this approach is that sometimes the important information may be deleted. So some heuristic function can be used that can identify the redundant data in the majority class.

2) Oversampling: randomly copies the samples in the minority class to increase the number of samples in the training dataset. But oversampling sometimes leads to overfitting. Undersampling and oversampling both are applied together.

3) Synthetic minority oversampling technique (SMOTE) [ ]: generates new images from the existing images in the minority classes. Firstly a random image m from the minority class is chosen and then k nearest neighbours of that image are identified and then The fake image is then created by randomly selecting one of the k nearest neighbors n and joining a and b to make a line in the feature space. The fake samples are generated as a curved blend of the two selected images m and n.

VI. TIME AUGMENTATION

All the methods covered in the above sections are applied to training datasets and are known as train time augmentations. Augmentations can also be applied to testing and validation datasets. Test time augmentations change the way to test the model on the testing dataset although methods for creating synthesized data are same as in train time augmentation. AlexNet [24] applied cropping and horizontal flip and created ten samples from each image. These images are then averaged to get the final result. GoogLeNet [36] used scaling and rescaled the test images at four different scales and then applied cropping on these images. Each test image resulted in 144 crops that are averaged to give the final synthesized image. VGG16 [25] used the same scale jitter approach that was used during the training time. Test-time augmentation gives a fit model each chance to predict vigorously.

VII. CONCLUSION

Many data augmentation techniques have been proposed in literature for image classification, image recognition and object detection. Many studies and experiments are also conducted to publically available datasets like CIFAR-10, CIFAR-100, Caltech-101, Caltech-256, MNIST (Modified National Institute of Standards and Technology), ImageNet to validate the usefulness of these techniques. But there is no magic bullet that can solve all the problems. Cherry-picking the suitable set of augmentation techniques for a custom problem is still an experiment. What kind of augmentations should be applied depends upon the type of dataset, problem domain and no. of training samples available in each class. Test-time augmentation focuses on taking organized crops of the input images to guarantee features present in the input images are perceived. Geometric and photometric transformations are common due their simplicity. On the other hand different versions of GAN show promising results but it is very elusive to train GAN. Transfer learning and low shot learning are the most encouraging environments for using GAN for image data augmentation.

REFERENCES

[1] Q. Wu, Y. Liu, Q. Li, S. Jin and F. Li, "The application of deep learning in computer vision," Chinese Automation Congress (CAC), Jinan, China, pp. 6522-6527, 2017, doi: 10.1109/CAC.2017.8243952.

[2] X. Zhou, W. Gong, W. Fu and F. Du, "Application of deep learning in object detection," 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), Wuhan, 2017, pp. 631-634, doi: 10.1109/ICIS.2017.7960069.

[3] U. Shah and A. Harpale, "A Review of Deep Learning Models for Computer Vision," 2018 IEEE Punecon, Pune, India, 2018, pp. 1-6, doi: 10.1109/PUNECON.2018.8745417.

[4] C. Khosla and B. S. Saini, "Enhancing Performance of Deep Learning Models with different Data Augmentation Techniques: A Survey," 2020 International Conference on Intelligent Engineering and Management (ICIEM), London, UK, 2020, pp. 79-85, doi: 10.1109/ICIEM48762.2020.9160048.

[5] C. Shorten, T. M. Khosgoftaar, "A survey on image data augmentation for deep learning" Journal of Big Data **6,** 60, 2019. https://doi.org/10.1186/s40537-019-0197-0

[6] O. Sharma, "Deep Challenges Associated with Deep Learning," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, 2019, pp. 72-75, doi: 10.1109/COMITCon.2019.8862453.

[7] X. Chen and X. Lin, "Big Data Deep Learning: Challenges and Perspectives," in IEEE Access, vol. 2, pp. 514-525, 2014, doi: 10.1109/ACCESS.2014.2325029.

[8] P. Shukla and K. Bhowmick, "To improve classification of imbalanced datasets," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2017, pp. 1-5, doi: 10.1109/ICIIECS.2017.8276044.

[9] X. Guo, Y. Yin, C. Dong, G. Yang and G. Zhou, "On the Class Imbalance Problem," 2008 Fourth International Conference on Natural Computation, Jinan, China, 2008, pp. 192-201, doi: 10.1109/ICNC.2008.871.

[10] H. Chung, E. Chung, J. G. Park and H. Jung, "Parameter Reduction For Deep Neural Network Based Acoustic Models Using Sparsity Regularized Factorization Neurons," 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 2019, pp. 1-5, doi: 10.1109/IJCNN.2019.8852021.

[11] N. Srivastava, G. Hinton, A. Krizhevsk, I. Sutskever and Ruslan Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting, " Journal of Machine Learning Research, Vol.15, No. 56, pp. 1929-1958,2014, url: http://jmlr.org/papers/v15/srivastava14a.html

[12] W. Yan and Z. Limimg, "The effect of initial weight, learning rate and regularization on generalization performance and efficiency," 6th International Conference on Signal Processing, 2002., Beijing, China, 2002, pp. 1191-1194 vol.2, doi: 10.1109/ICOSP.2002.1180003.

[13] K. Wang, X. Gao, Y. Zhao, X. Li, D. Dou and C. Z. Xu, "Pay Attention to Features, Transfer Learn Faster CNNs", 8th International Conference on Learning Representations, (ICLR) 2020, url: https://openreview.net/forum?id=ryxyCeHtPB.

[14] R. Ribani and M. Marengoni, "A Survey of Transfer Learning for Convolutional Neural Networks," 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T), Rio de Janeiro, Brazil, 2019, pp. 47-57, doi: 10.1109/SIBGRAPI-T.2019.00010.

[15] Z. Wang, Z. Dai, B. Póczos and J. Carbonell, "Characterizing and Avoiding Negative Transfer," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 11285-11294, doi: 10.1109/CVPR.2019.01155.

[16] J. Nilsson, P. Andersson, I. Y. -H. Gu and J. Fredriksson, "Pedestrian Detection Using Augmented Training Data," 2014 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 2014, pp. 4548-4553, doi: 10.1109/ICPR.2014.778.

[17] S. C. Wong, A. Gatt, V. Stamatescu and M. D. McDonnell, "Understanding Data Augmentation for Classification: When to Warp?," 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Gold Coast, QLD, Australia, 2016, pp. 1-6, doi: 10.1109/DICTA.2016.7797091.

[18] J. Lemley, S. Bazrafkan and P. Corcoran, "Smart Augmentation Learning an Optimal Data Augmentation Strategy," in IEEE Access, vol. 5, pp. 5858-5869, 2017, doi: 10.1109/ACCESS.2017.2696121.

[19] J. Shijie, W. Ping, J. Peiyi and H. Siping, "Research on data augmentation for image classification based on convolution neural networks," 2017 Chinese Automation Congress (CAC), Jinan, China, 2017, pp. 4165-4170, doi: 10.1109/CAC.2017.8243510.

[20] K. Fujita, M. Kobayashi and T. Nagao, "Data Augmentation using Evolutionary Image Processing," 2018 Digital Image Computing: Techniques and Applications (DICTA), Canberra, ACT, Australia, 2018, pp. 1-6, doi: 10.1109/DICTA.2018.8615799.

[21] A. Namozov and Y. I. Cho, "An Improvement for Medical Image Analysis Using Data Enhancement Techniques in Deep Learning," 2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT), Busan, Korea (South), 2018, pp. 1-3, doi: 10.1109/ICT-ROBOT.2018.8549917.

[22] C. Lei, B. Hu, D. Wang, S. Zhang, Z. Chen, "A Preliminary Study on Data Augmentation of Deep Learning for Image Classification," Proceedings of the 11th Asia-Pacific Symposium on Internetware 2019 Pages 1–6https://doi.org/10.1145/3361242.3361259.

[23] C. Khosla and B. S. Saini, "Enhancing Performance of Deep Learning Models with different Data Augmentation Techniques: A Survey," 2020 International Conference on Intelligent Engineering and Management (ICIEM), London, UK, 2020, pp. 79-85, doi: 10.1109/ICIEM48762.2020.9160048.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Proceedings of the 25th International Conference on Neural Information Processing Systems – vol. 1, pp. 1097–1105, 2012.

[25] K. Simonyan & A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition, " ICLR 2015.

[26] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016.

[27] Z. Zhong, L. Zheng, G. Kang, S. Li, Y. Yang, "Random Erasing Data Augmentation," Proceedings of AAAI Conference on Artificial Intelligence, Vol. 34, No. 07, 2020, doi: https://doi.org/10.1609/aaai.v34i07.7000

[28] S. Yun, D. Han, S. Chun, S. J. Oh, Y. Yoo and J. Choe, "CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, pp. 6022-6031, doi: 10.1109/ICCV.2019.00612.

[29] P. Chen, S. Liu, H. Zhao, J. Jiya, "GridMask Data Augmentation," url: https://arxiv.org/abs/2001.04086.

[30] A. Bockokovskiy, C. Wang, H. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection, " 2020, url: https://arxiv.org/abs/2004.10934

[31] H. Chen and P. Cao, "Deep Learning Based Data Augmentation and Classification for Limited Medical Data Learning," 2019 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), Shenyang, China, 2019, pp. 300-303, doi: 10.1109/ICPICS47731.2019.8942411.

[32] B. Georgievski, "Image Augmentation with Neural Style Transfer," In: Gievska S., Madjarov G. (eds) ICT Innovations 2019. Big Data Processing and Mining. ICT Innovations 2019. Communications in Computer and Information Science, vol 1110. Springer, Cham, https://doi.org/10.1007/978-3-030-33110-8_18.

[33] A. Bagheri, O. Simeone and B. Rajendran, "Adversarial Training for Probabilistic Spiking Neural Networks," 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Kalamata, Greece, 2018, pp. 1-5, doi: 10.1109/SPAWC.2018.8446003.

[34] H. Mansourifar, L. Chen and W. Shi, "Virtual Big Data for GAN Based Data Augmentation," 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 1478-1487, doi: 10.1109/BigData47090.2019.9006268.

[35] F. H. K. S. Tanaka and C. Aranha, "Data Augmentation Using GANs," ArXiv, 2019, url: https://arxiv.org/abs/1904.09135.

[36] C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions,", 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) , pp. 1-9, doi: https://doi.org/10.1109%2fCVPR.2015.7298594.