IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11390

The Significance of Image Augmentation in Deep Learning: A Review

Dipmala Salunke[1], Prasadu Peddi[2], Ram Joshi[3]

Research Scholar, Department of Computer Engineering, Shri Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan, India[1]

Assistant Professor, Department of Information Technology, JSPM's Rajarshi Shahu College of Engineering Tathawade, Pune, Maharashtra, India[1]

Professor, Department of Computer Engineering, Shri Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan, India[2]

Professor, Department of Information Technology, JSPM's Rajarshi Shahu College of Engineering Tathawade, Pune, Maharashtra, India[3]

Abstract: Many computer vision tasks have shown that deep convolutional neural networks perform exceptionally well. However, in order to avoid overfitting, these networks rely extensively on huge amounts of data. Many disciplines, such as medical image analysis, lack access to massive data sets. Data augmentation approaches enable applications with limited datasets to achieve higher accuracy. The process of creating samples by transforming training data is known as data augmentation, with the goal of increasing classifier accuracy and resilience. Geometric transformations, colour space augmentations, kernel filters, random erasing, feature space augmentation, adversarial training, generative adversarial networks, neural style transfer, and meta-learning are some of the image augmentation technologies explored in this review. Data Augmentation will help readers learn how to improve the performance of their models and expand their limited datasets in order to take advantage of big data.

Keywords: Machine learning, Deep learning, Data augmentation, GAN, Medical imaging.

I. INTRODUCTION

One of the most well-known computer vision problems is image classification. Hundreds of studies have recently been conducted in order to attain excellent performance in image recognition, detection, and classification. The advancement of various deep learning algorithms, particularly Convolutional Neural Networks, has resulted in significant increases in classification outcomes. To learn thousands of different kinds of objects from millions of images, a deep learning model with a large learning capacity is required. Alex et al employed first CNN model that achieved top-1 and top-5 error rates of 37.5 % and 17 % in ImageNet LSVRC-2010 [1,2]. Following that, in natural image datasets, other deep CNN models demonstrated even better classification and detection accuracy[3,4]. Deep neural network also performed well on MNIST dataset [5]. To try to expand Deep Learning for use on smaller datasets, functional techniques including dropout regularization, batch normalization, transfer learning, and pre-training have been developed [6]. To assess medical images, a large range of applications and algorithms have been proposed. Deep learning advances, particularly deep convolution neural networks (CNN), have increased the performance of medical image classification algorithms. However, training a deep CNN using medical images from the scratch is a difficult task that demands a vast amount of annotated data.

One of the most well-known computer vision problems is image classification. Hundreds of studies have recently been conducted in order to attain excellent performance in image recognition, detection, and classification. The advancement of various deep learning algorithms, particularly convolutional neural networks, has resulted in significant increases in classification outcomes. A deep learning model with a large learning capacity is required to learn thousands of different kinds of objects from millions of images. Alex et al. employed the first CNN model that achieved top-1 and top-5 error rates of 37.5% and 17.5% in ImageNet LSVRC-2010 [1,2]. Following that, in natural image datasets, other deep CNN models demonstrated even better classification and detection accuracy [3,4]. Deep neural networks also performed well on the MNIST dataset [5]. To try to expand Deep Learning for use on smaller datasets, functional techniques including dropout regularization, batch normalization, transfer learning, and pre-training have been developed [6]. To assess medical images, a large range of applications and algorithms have been proposed. Deep learning advances, particularly

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⚟ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11390

deep convolution neural networks (CNN), have increased the performance of medical image classification algorithms. However, training a deep CNN using medical images from scratch is a difficult task that demands a vast amount of annotated data.
When your model is overfitted to the training set, it is known as overfitting. The model's ability to generalise to new cases that were not part of the training set gets increasingly difficult. Different methods are available to reduce overfitting, like adding more data, data augmentation, dropout method, regularisation techniques. The figure 1 illustrates the overfitting of the model.

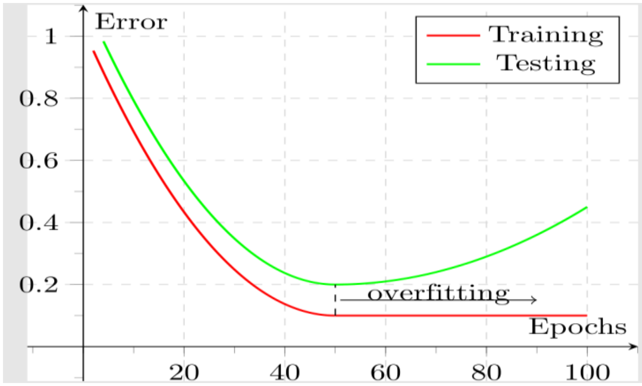

Fig. 1 Overfitting of model (Image Source: Wikipedia)

Image augmentation can be applied before training the model or in real time. In this paper, we will learn different image data augmentation techniques that are based on the notion that by augmenting the original dataset, more information may be recovered to improve the accuracy of deep learning models.

II. IMAGE AUGMENTATION TECHNIQUES

Few recently used techniques for image data augmentation are given below:

A. Geometric transformations

The possibility of keeping the label post-transformation is defined as a data augmentation method's safety. On ImageNet problems like cat against dog, rotations and flips are generally safe, but not on digit identification tasks like 6 versus 9. A non-label-preserving transformation may improve the model's capacity to produce a response, indicating that it is not confident in the forecast.

A.1 Cropping
Cropping an image's focal patch can be utilised as a useful processing step for image data with mixed height and width dimensions. In addition, random cropping can be employed to create a similar effect to translations. This may or may not be a label-preserving transformation depending on the cropping reduction threshold selected.

A.2 Flipping
Flipping the horizontal axis is far more common than vertical axis flipping. This is one of the simplest augmentations to use, and it has been shown to work on datasets like CIFAR-10 and ImageNet.

A.3 Color space transformations
Image-editing programmes can also be used to generate colour space changes. A colour histogram is made up of the pixel values in each RGB colour channel of a picture. This histogram can be used to apply filters that alter an image's colour space features. Changing the colour distribution of images can be a wonderful way to overcome lighting issues while assessing data.

A.4 Translation
To prevent positional bias in data, shifting images left, right, up, or down can be a highly effective adjustment. The leftover space can be filled with a constant value such as 0 s or 255 s, or it can be filled with random or Gaussian noise as the original image is translated in one direction.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⨳ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11390

A.5 Color space

Another method that is quite feasible to apply is doing augmentations in the colour channel space. Isolating a single-color channel, such as R, G, or B, is a very simple colour augmentation. To enhance or reduce the brightness of an image, simple matrix operations can be used to modify the RGB values. Deriving a colour histogram that describes the image allows for more advanced colour augmentations.

A.6 Rotation

The image is rotated right or left on an axis between 1° and 359° for rotation augmentations.

B. GAN based methods

Ian Goodfellow [7] proposed the GAN architecture as a framework for generative modelling using adversarial training. GANs can learn to synthesise data that is unrecognisable from the actual data from a dataset. This method involves feeding a random vector into a generator network, which then maps it to the desired size of the output image. Then a discriminator network identifies the fake image according to the original data. Consider a dataset of 100 cats and 100 dogs. Using the GAN, we can generate 400+ cats and 400+ dogs, giving us 500 of each class. We now apply translational shifts to images generated by GANs and have tripled the size of our dataset. Then we flip the photos, add noise to a sample of them, and rotate them slightly. Following all of these augmentation strategies, it appears that we may be able to 100x the size of our dataset in a manageable manner [7].
The DCGAN architecture was employed by Alec R et al. where they used CNNs for generator and discriminator functions instead of multiplayer perceptron's. The architecture was tested on the LSUN dataset for bedroom images of size 64 x 64 x 3. The goal of DCGAN is to make the generator network more complicated such that it can project the input into a high-dimensional tensor and then add deconvolutional layers to convert the projected tensor to an output image [8].It reduces dimensions from $14 \times 14 \times 6$ to $28 \times 28 \times 1$ whereas the spatial dimensions of a convolutional layer will be reduced, for example, from $14 \times 14 \times 32$ to $7 \times 7 \times 64$.

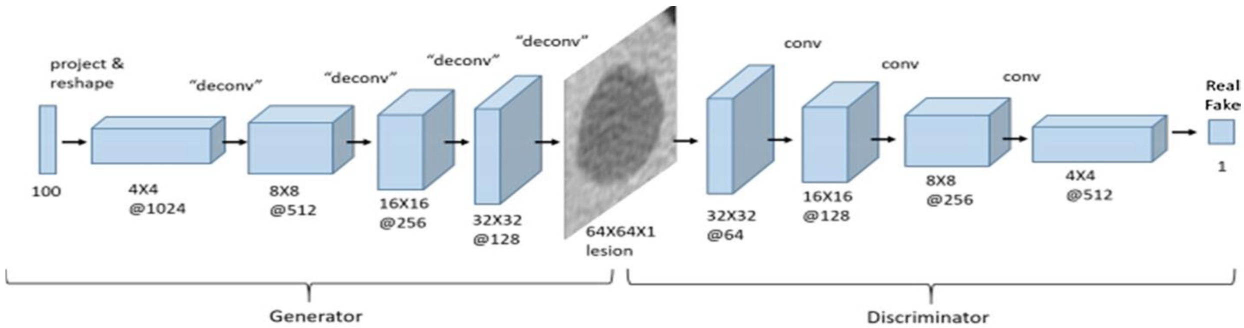

Fig. 2 DCGAN Architecture used to generate liver lesion images by Maayan F-A [9]

Frid-Adar et al. have used DCGAN architecture to generate liver lesion images from CT scans for sizes of $64 \times 64 \times 1$. The size of dataset used was 182 images. After achieving 78.6% sensitivity and 88.4% specificity using traditional augmentations, they saw an increase to 85.7% sensitivity and 92.4% specificity when they added the DCGAN-generated samples.

CycleGAN was introduced by Jun-Yan Z et al. [10] and includes a new cycle-consistency loss function to aid in the stabilisation of GAN training. This is used to convert images from one domain to another domain, for example, from a cat domain to a dog domain. A generator maps the cats' images to dogs' images, and the discriminator is unable to guess that, originally, they were part of a cat or a dog, and then determines whether the retranslated images belong to a cat or a dog.

CycleGAN was also tested for emotion classification problems where images were classified into 7 emotion types where it proved as a method of intelligent oversampling [11]. Progressively growing GANs was employed by [12] which is based on the idea that GANs can take in both images and random vectors as input. As a result, the GANs in this series work by passing samples from a lower resolution GAN to a higher resolution GAN. On facial images, this has yielded incredible results.

The Neural-style transfer method was used as a data augmentation tool to manipulate the representation of images created using CNN. While preserving the original content of an image, neural style transfer is used to transfer it into another type [13].

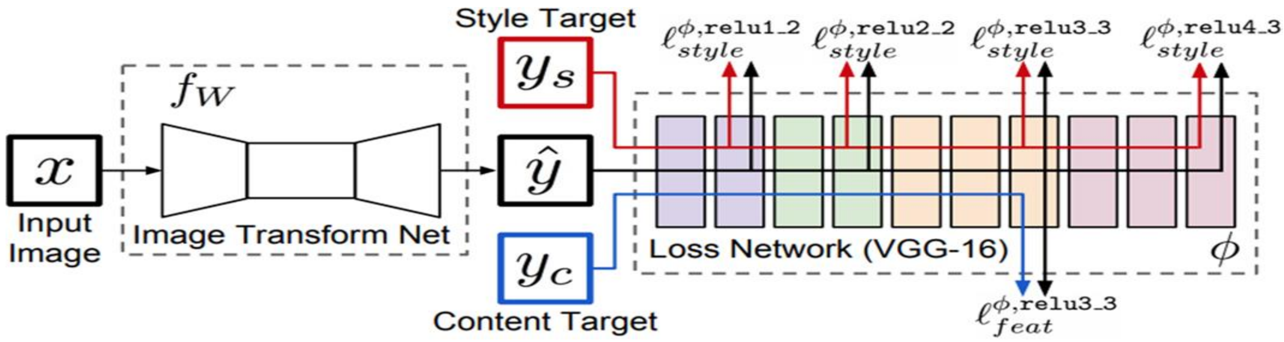

Fig. 3 Fast neural style algorithm by Johnson et al. [14]

III. GEOMETRIC TRANSFORMATION RESULT ON DENTAL RADIOGRAPH

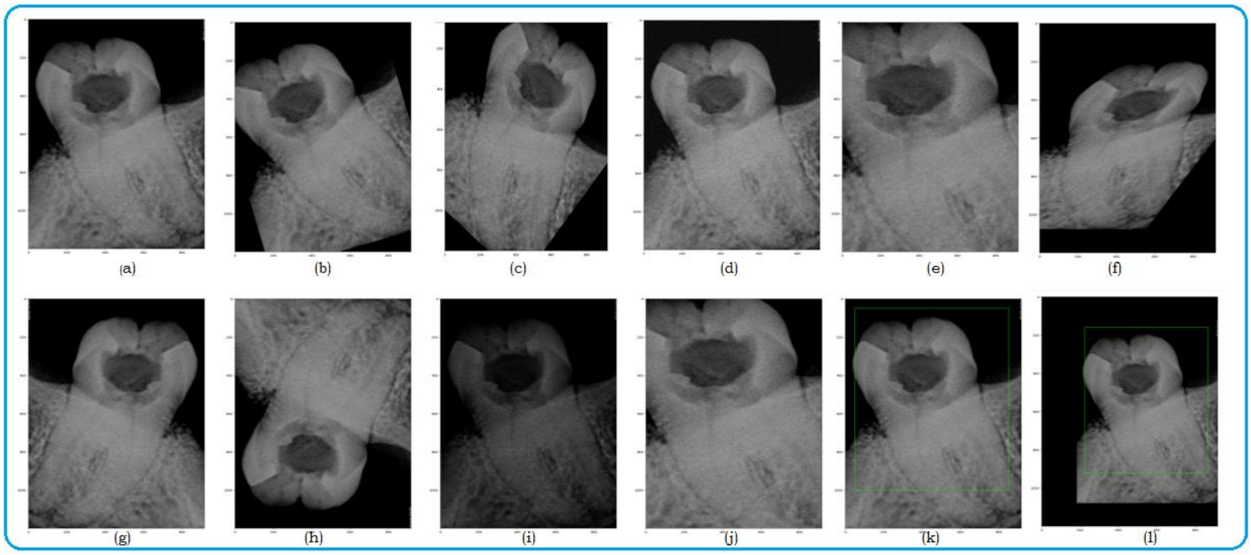

Fig. 4 Result of geometric transformation technique

Figure 4 shows the result of geometric transformation on dental x-ray image. Different types of image augmentations results on dental x-rays are: (a) original image,(b)rotate image from -50 degree to 30 degree, (c) rotate image from 30 degree to 70 degree,(d) adding noise to data,(e) Cropping of image, (f) Shearing the image (g) flipping image horizontally (h) flipping image vertically, (i)Changing the brightness of the image,(j)Scaling the image below to 150% to 80% of the image height/width,(k)bounding box around the original image,(l)displaying the bounding box on top of the original image.

IV. DISCUSSION

The advantage of horizontal flipping or random cropping is straightforward to explain. However, it is unclear why combining pixels or complete images, as in patch shuffle regularisation or sample pairing, is so successful. For GAN-based networks, it is difficult to interpret the representations learned by the neural network. CNN's intermediate layers are useful for the data augmentation process. Traditional geometrical transformation techniques can be improved with the help of neural style transfer and GAN-based methods. Traditional colour space augmentations are far less strong than neural style transfer, although both can be combined. All these augmentation methods are useful if training and testing data are drawn from the same distribution.

V. CONCLUSION

This paper gives an overview of different data augmentation techniques that help solve the problem of overfitting in deep learning applications. In medical imaging generally, we don't get enough data, so the methods discussed here help to increase the size of the dataset. Geometric transformation based and GAN based data augmentation techniques have been

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁙ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11390

explained and can be applied before training the model or in real time. The input layer is where the majority of the augmentations surveyed function. Some, on the other hand, are formed from representations of hidden layers. In the future, we can work on performance benchmarks for various image recognition tasks.

ACKNOWLEDGMENT

The authors would like to express their gratitude to **Dr. Vikas Kashid** (Tirupati Dental and Implant Centre, Pune, Maharshtra, India), and **Dr. Aditya Patakrao** (Adity's Dental Clinic, Pune, India) who supported dental radiographs and dental disease information for this study.

REFERENCES

[1]. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
[2]. Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." In European conference on computer vision, pp. 818-833. Springer, Cham, 2014.
[3]. Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." Cvpr, 2015.
[4]. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778 2016.
[5]. LeCun, Yann, Corinna Cortes, and C. J. Burges. "MNIST handwritten digit database." AT&T Labs [Online]. Available: http://yann.lecun. com/exdb/mnist 2, 2010.
[6]. Jan K, Vladimir G, Daniel C. Regularization for deep learning: a taxonomy arXiv preprint. 2017.
[7]. Ian JG, Jean PA, Mehdi M, Bing X, David WF, Sherjil O, Aaron C, Yoshua B. Generative adversarial nets. NIPS. 2014.
[8]. Alec R, Luke M, Soumith C. Unsupervised representation learning with deep convolutional generative adversarial networks. ICLR, 2016.
[9]. Maayan F-A, Eyal K, Jacob G, Hayit G. GAN-based data augmentation for improved liver lesion classification. arXivpreprint. 2018.
[10]. Jun-Yan Z, Taesung P, Phillip I, Alexei AE. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: International conference on computer vision (ICCV), 2017.
[11]. Goodfellow IJ, Erhan D, Carrier PL, Courville A, Mirza M, Hamner B, Cukierski W, Tang Y, Thaler D, Lee DH, et al. Challenges in representation learning: A report on three machine learning contests. In: NIPS. Berlin: Springer; 2013.p.117–24.
[12]. Tero K, Timo A, Samuli L, Jaakko L. Progressive growing of GANs for improved quality, stability, and variation. CoRR, abs/1710.10196, 2017.
[13]. Leon AG, Alexander SE, Matthias B. A neural algorithm of artistic style. ArXiv. 2015.
[14]. Justin J, Alexandre A, Li FF. Perceptual losses for real-time style transfer and super-resolution. ECCV, 2016;2016:694–711.
[15]. Changhee H, Hideaki H, Leonardo R, Ryosuke A, Wataru S, Shinichi M, Yujiro F, Giancarlo M, Hideki N. GAN-based synthetic brain mr image generation. In: 2018 IEEE 15th International Symposium on biomedical imaging (ISBI 2018). IEEE, 2011. P. 734-8.
[16]. Esteban R, Alok A, Yanping H, Quoc VL. Regularized evolution for image classifier architecture search. arXiv preprint. 2018.
[17]. Q. Zheng, X. Tian, M. Yang, Y. Wu, and J. Su, "PAC-Bayesian framework-based drop-path method for 2D discriminative convolutional network pruning," Multidimensional Systems and Signal Processing, 2019.
[18]. J. Li, M. Yang, Y. Liu et al., "Dynamic hand gesture recognition using multi-direction 3D convolutional neural networks," Engineering Letters, vol. 27, no. 3, pp. 490–500, 2019.
[19]. Dipmala Salunke, Ram Joshi, Prasad Peddi, Deepak Mane, Deep Learning Techniques for Dental Image Diagnostics: A Survey, Springer "Lecture Notes in Social Network, 4th International Conference on Computer Networks, Big Data and IoT, ICCBI 2021.
[20]. Krishna Sheth, Prof. D. T. Salunke,Atul Raj, Astha Kubavat, Dental Disease Classification using Deep Learning: A Survey, GIS Science Journal,2021.volume 8 issue 6,pp.1829-1841.
[21]. S. Zhang, R. Benenson, M. Omran, J. Hosang, and B. Schiele, "Towards reaching human performance in pedestrian detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 973–986, 2018.
[22]. Pradnya Jadhav Dipmala Salunke ,Pooja Sabne, , Hitesh Saini, Vivekanand Shivanagi Handwritten Devanagari Word Recognition using Customized Convolution Neural Network, IEEE-International Conference on Computing, Communication and Green Engineering (CCGE21), 2021

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁑ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11390

[23]. Qinghe Zheng, ,Mingqiang Yang ,Xinyu Tian ,Nan Jiang ,and Deqiang Wang, A Full Stage Data Augmentation Method in Deep Convolutional Neural Network for Natural Image Classification Hindawi Discrete Dynamics in Nature and Society Volume 2020, Article ID 4706576, 11 pages. https://doi.org/10.1155/2020/4706576.