

A Study of Full Stage Data Augmentation Method

Amit Kumar Pandey^{1*} Dr. P. K. Bharti² Dr. Prashant Singh³

¹ PhD Scholar, Department of Computer Science and Engineering, Shri Venkateshwara University Gajraula, Amroha, Uttar Pradesh

² Vice Chancellor, Shri Venkateshwara University, Gajraula, Amroha, Uttar Pradesh

³ Department of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology and Management, Delhi

Abstract – Many computer vision-related tasks may now be accomplished using deep learning, which relies heavily on the use of huge data. A full-stage data augmentation approach for deep learning is proposed in this, which may also serve as an implicit model ensemble without incurring additional model training expenses. A data-space solution to the problem of limited data can be achieved by simultaneously augmenting training and testing data. Deep Learning models may be developed utilizing larger and more accurate training datasets that have been enhanced by data augmentation approaches. Geometric transformations, color space augmentations, kernel filters, picture mixing, random erasing, feature space augmentation, adversarial training, generative adversarial networks, neural style transfer, and meta-learning are among the image augmentation methods explored in this review. Deep Learning models may be developed utilizing larger and more accurate training datasets thanks to Data Augmentation approaches. Improve model performance and increase datasets to take use of big data's possibilities through data augmentation.

Keywords – Full Stage, Data, Augmentation Method, Deep Learning, etc.

-----X-----

INTRODUCTION

Computer vision is the earliest and most widespread use of deep learning technology. The use of deep convolution neural networks in computer vision has increased dramatically since Alex Net's arrival, with solutions for a wide range of computer vision challenges. The ability of over parameterized deep learning models to perform better based on their very nonlinear fitting skills has grown as more and more data and computing power become available [1]. A wide range of deep learning models have been developed and refined, combining different architectures and connections based on training methodologies. It's still a work in progress when it comes to deep learning, and the theory behind it is far from complete. Because of their complexity, deep learning models are challenging to improve in a targeted manner. In research, it's common to combine optimization with generalization. As a result of the "over fitting" problem, deep CNNs trained on enormous datasets nevertheless fall short when applied to fresh datasets that have not been trained. Larger models tend to perform better, but this comes at the cost of making trade-offs between accuracy and reasoning speed. A photograph's natural picture must be cleaned up of any noise if it is to retain its

expressive features [2]. A lack of training data and the need for excellent real-time performance in automated driving systems have hampered its application in several instances, such as medical diagnostics.

Photographs are better at conveying information because they show the situation exactly as it is. An photograph may reveal a wealth of information about a subject. Visual data accounts for more than 90% of the information sent to the brain [2-3]. Visual data is processed and responded to by the human brain 60,000 times better than any other type of data. Image classification has become a dynamic field thanks to breakthroughs in science and technology, such as the development of high-tech instruments like digital cameras, sensors, scanners, and image processing techniques. The photographs must be in digital form to be processed by image processing systems. An picture in digital form may be thought of as a two-dimensional array of numbers, where the numbers represent the reflectance values of the objects in the image at different spatial positions. When conducting mathematical operations like addition, subtraction, multiplication, division, etc., pixels (picture components) can achieve spatial coherence. It is

vital to note that statistical alterations of pixel values are used in a wide range of image processing techniques. A computer or machine system does image categorization. Relative positioning of pixels is used to identify objects in photographs [3].

IMAGE CLASSIFICATION

An essential part of digital image analysis is the categorization of images. Humans, by their very nature, are able to see the world around them, comprehend its meaning, and organize its items into categories. The use of a computer-based categorization system, on the other hand, is a useful addition to the manual procedure. In order to classify a picture, a label is assigned to it depending on the attributes that have been calculated. Classifying the photos into two or more categories can do this. In general, there are four processes in the image classification process: picture capture, preprocessing, feature extraction, and classification, all of which take place simultaneously. Steps in the image categorization process are depicted in this flow chart. Digital cameras, thermal imaging systems, low-level, mid-level, and high-level ultrasonic scanning, as well as fundus fluoroscopy are all examples of image acquisition. The first stage in every image processing system is preprocessing. Color images are converted to grayscale at this step, and then resized, cleaned up of noise, and enhanced for contrast. Feature extraction is the next step after preprocessing [4]. An image's significant and meaningful characteristics are taken from the preprocessed version at the step of feature extraction. Finally, a picture is assigned a class based on its properties using a classifier.

DATA AUGMENTATION

To train a deep neural network, it is necessary to have a large amount of training data. Since each image in the training set needs to be manually annotated by an expert, the process of collecting labeled training data is both time-consuming and expensive. A lack of training data might lead to over fitting, since the neural network is more likely to memorize particular characteristics of the training set [4-5]. It is standard practice to use data augmentation in order to prevent this form of over fitting from occurring. By altering input training data, more training data may be generated, resulting in more accurate results. If data augmentation is done online, then the input picture (or input image mini-batch for SGD) is changed directly before feeding into the deep neural network. Medical image segmentation has been demonstrated to benefit most from data augmentation. It has been proven that data augmentation may be useful for highly deep architectures, even for big datasets like Image Net. In addition, data augmentation makes it simple to include existing information about previously viewed data into the analysis process. For example, if the brightness of the test pictures varies, adding random

intensity shifts might be useful to account for the variable brightness of the test data. A wide range of data augmentation techniques are available, from basic visual alterations like adjusting the intensity of a color to more complex geometric transformations like scaling or rotating an image. Various data augmentation strategies are discussed here [5], as well as their probable applications and implications on training deep neural networks.

- Additive Noise
- Intensity Scaling
- Intensity Shift
- Random Flipping
- Random Translation
- Random Rotation
- Random Cropping
- Elastic Deformation
- Synthetic Data Augmentation

MANUAL DATA AUGMENTATION

When manually augmenting data, photometric and geometric alterations are commonly used. A popular method of photometric variation is to change an image's color palette, which includes rotations, translation, flipping, scaling and shearing. In training data, affine displacements can be used to estimate geometric variations. To determine the new target position of each pixel (x' , y'), the old position is taken into account (x, y) [6]. We can shift the picture to the left by a factor of one, for example. In this case, the new target position for each pixel is computed using the formula: $x' = (x-1-y)$. Displacement fields $x_0 = y$, $y_0 = 0$, would cause the picture to be rotated 90 degrees clockwise. Due to the fact that scaling values might be non-integers, interpolation is required to compute new values for each pixel when scaling is applied [7].

FRAMING THE HIGH-LEVEL CONTEXT OF DATA AUGMENTATION AND DEEP LEARNING

Dropout: Regularization is achieved by reducing the activity levels of randomly chosen neurons during training. Rather of relying on a small number of neurons for all of the network's prediction capabilities, this constraint forces the network to develop more resilient properties.

Batch normalization: a layer's activations can be normalized using this method. In order to normalize an experiment, the batch mean is subtracted from

each activation and the batch standard deviation is divided by the standard deviation.

Transfer learning: is a novel approach to avoid over fitting. To use Transfer Learning, you first train a network on a large dataset like ImageNet, and then apply those weights to a new classification problem.

Pretraining: is extremely comparable to the transfer of knowledge. A large dataset like ImageNet is used to define and train the network architecture in Pertaining. Transfer learning, on the other hand, does not necessitate the transfer of network architectures like VGG-16 or ResNet, but only the weights.

One-shot and Zero-shot learning: Algorithms may be used to generate models with very little data. Facial recognition software frequently makes advantage of one-shot learning. Siamese networks that learn a distance function such that image classification may be done even if the network has been trained on one or a few examples are an example one-shot learning. Data Augmentation, on the other hand, focuses on over fitting from the base of the problem, which is the training dataset. This is based on the premise that augmentations can extract more information from the original dataset. Data warping and oversampling are used to artificially expand the training dataset size [8].

IMAGE DATA AUGMENTATION TECHNIQUES

Data Augmentations were first shown in basic operations including horizontal flipping, color space augmentations, and random cropping. It is possible to convey many of the in variances stated earlier in picture recognition tasks by transforming the images. GAN-based augmentation and neural style transfer are also included in the list of enhancements in this survey. Geometric and color space transformations are also included in the list. Kernel filters are also included in the list of augmentations. Detailed explanations of how each augmentation method works, as well as experimental findings, will be included in this section. Based on simple picture modifications, these data augmentations [9]:

Geometric transformations

The many image processing routines and geometric adjustments that go into the various enhancements. The augmentations detailed below fall into a category that may be classified as simple to execute. A thorough understanding of these changes will serve as a beneficial foundation for further research into Data Augmentation approaches.

Flipping

It is far more usual to flip the horizontal axis than to flip the vertical axis. CIFAR-10 and Image Net datasets have shown that this augmentation is easy to apply and useful. Using this technique on datasets involving text recognition, such as MNIST or SVHN, does not preserve the labels.

Color space

The dimensions (height, breadth, and color channels) of digital picture data are often recorded as a tensor. Color channel augmentations are another approach that can be easily implemented. Isolating a single color channel such as R, G, or B is a simple way to enhance a picture's color. Isolating one color channel's matrix and adding two zero matrices from other channels can swiftly transform a picture into its representation in that color channel [9-10].

Cropping

Using a centre patch of each picture to crop an image can be a feasible processing step for photos with mixed height and width dimensions. As an alternative to translations, you may instead utilize random cropping.

Rotation

It is possible to add rotational augmentations to a picture by rotating it on an axis between 0° and 359°. The rotation degree parameter has a significant impact on the safety of rotational augmentations.

Translation

Images can be shifted to the left, right, up, or down in order to remove any positional bias from the data. Because face recognition datasets often contain perfectly centred photos, testing the model on those would need using only images that were also perfectly centred.

Noise injection

In order to introduce noise, a random matrix derived from a Gaussian distribution is used. Nine datasets from the UCI collection were used by Moreno-Barea et al. to examine the effects of noise introduction. CNNs can learn more robust features by introducing noise to pictures.

Color space transformations

Three stacked matrices, each with a height width of size height, encode the image data. An individual RGB color value is represented in these matrices. One of the most common obstacles to picture identification is lighting bias. Color space

transformations, also known as photometric transformations, are easy to understand because of their intuitive nature [10].

Kernel filters

Sharpening and blurring pictures with kernel filters is a common practice in the field of image processing. To apply one of these filters, you just move a $n \times n$ matrix over a picture, and the outcome is either a blurrier image with a Gaussian blur filter or a sharper image with a high contrast vertical or horizontal edge filter.

Mixing images

Data Augmentation should not be done by averaging the pixel values of many photos. To a human viewer, the pictures created by this will not appear to have undergone any meaningful modification. Despite this, I have shown how the pairing of samples could be turned into an effective method of augmenting. Two photos are randomly cropped from 256×256 to 224×224 and flipped horizontally in this experiment.

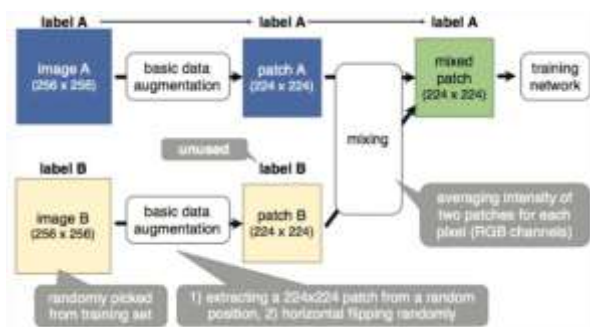


Fig. 1: Sample pairing augmentation strategy

Random erasing

A novel Data Augmentation technology called random erasing. For example, random erasing might be understood as a dropout regularization technique in the input data space rather than contained in the network. Specifically, this method was created to address picture identification issues caused by occlusion. An item is occluded when certain sections are obscured [11].

DATA AUGMENTATIONS BASED ON DEEP LEARNING

Feature space augmentation: A picture in the input space receives all of the augmentations described above. High-dimensional data may be reduced to simpler representations via neural networks. Images can be mapped to binary classes or $n \times 1$ vector in flattened layers using these networks. Sophisticated manipulations of neural network sequential processing can be used to decouple the network as a whole and its intermediate representations from each other.

Adversarial training: Adversarial training is a method for exploring the possibilities of augmentations. Training two or more networks with opposing goals contained in their loss functions is known as adversarial training. Adversarial training and the phenomena of adversarial attacks will be discussed in this section. Using a competitor network, an adversarial assault consists of learning to tamper with pictures in order to misclassify its rival classification network. The adversarial network has been surprised by the success of these assaults, which are limited to noise injections. This is startling, as our preconceived notions of how these models represent pictures are flatly contradicted.

GAN-BASED DATA AUGMENTATION

Generative modeling is another fascinating approach to Data Augmentation. Using a dataset as a starting point, generative modeling creates new objects with comparable properties to those in the original. GANs, a generative modeling framework based on the adversarial training concepts previously outlined, is a fascinating and very popular result of this research. A dataset may be "unlocked" using GANs, according to Bowles et al. Despite the fact that GANs aren't the only generative modeling approach out there, they're leading the way in both computation speed and quality of outcomes.

Variation auto-encoders are another effective generative modeling approach. It is possible to increase the quality of samples generated by variation auto-encoders using the GAN framework. Variation auto-encoders learn to represent data points in a low-dimensional way. A vector of size $n \times 1$ is the same as an image tensor of height x breadth and color channels in the image domain, as mentioned in the feature space augmentation section [11-12].

NEURAL AUGMENTATION

Using the Neural Style Transfer technique, there are two weights for the style and content loss. Neural Augmentation was developed by Perez and Wang as an algorithm for meta-learning a Neural Style Transfer method. The Neural Augmentation method takes two random pictures from the same class and augments them with additional information. A CNN with five layers, each with 16 channels, 3×3 filters, and ReLU activation functions maps them into a new picture. Neural Style Transfer is used to apply a different random picture to the image that was generated as a result of the augmentation. With the help of Cycle GAN, this style transition may be accomplished. Neural Augmentation nets are then updated based on the categorization model's errors, which are fed back into the model.

SMART AUGMENTATION

Similarly to the Neural Augmentation technique, Smart Augmentation employs a similar strategy. However, rather than employing the Neural Style Transfer technique, the mixture of pictures is produced only from the learnt parameters of a supposed CNN.

Meta-learning augmentations can also take the form of Smart Augmentation. Having two networks, Network-A and Network-B, is how this is accomplished. In order to train Network-B, Network-A maps two or more input images into a new image or images [12]. Network-error B's rate is subsequently transmitted to Network-A, which is updated as a result. An additional metric has been included in Network-A so that it may be compared to others in its class. The augmented image is created using a succession of convolution layers in Network-A.

CONCLUSION

The over fitting in Deep Learning models may be avoided by using Data Augmentation. Improved accuracy of deep CNNs may also be achieved without the need for additional model training expenditures by using a full-stage data-augmentation method proposed in this study. Convergence of the network and its capacity to generalize to unknown test samples may be ensured via simultaneous data augmentation during training and testing. To avoid over fitting, Deep Learning models use a lot of data. Using the approaches outlined in this study to artificially inflate datasets delivers the advantages of big data in a confined data area. The process of enhancing existing datasets with new data is known as data augmentation. An oversampling or a data warping methodology has been proposed for a variety of augmentation methods. A tiny dataset's biases can't be addressed by data augmentation.

REFERENCES

1. Chen S, Abhinav S, Saurabh S, Abhinav G (2017). Revisiting unreasonable effectiveness of data in deep learning era. In: ICCV; pp. 843–52.
2. Jaswal, D, Vishvanathan, S & Soman, KP (2014). 'Image Classification using convolutional neural Networks', International Journal of Scientific and Engineering Research, vol. 5, no.6, pp. 1661-1668.
3. Krizhevsky A, Sutskever I, Hinton GE (2012). Image Net classification with deep convolutional neural networks. Adv Neural Inf. Process Syst.; 25: pp. 1106–14.

4. Lemley J, Barzafkan S, Corcoran P. (2017). Smart augmentation learning an optimal data augmentation strategy. In: IEEE Access.
5. Lorris Nanni, Alessandra Luminiz & Sheryl Brahnam (2012). 'Survey on LBP based texture descriptors for image classification', Expert Systems with Applications, Vol. 39, pp. 3634-3641.
6. Luis P, Jason W (2017). The effectiveness of data augmentation in image classification using deep learning. In: Stanford University research report.
7. Luke T, Geoff N. (2017). Improving deep learning using generic data augmentation. arXiv preprint.
8. Maayan F-A, Eyal K, Jacob G, Hayit G. (2018). GAN-based data augmentation for improved liver lesion classification. arXiv preprint.
9. Nitish S, Geoffrey H, Alex K, Ilya S, Ruslan S. (2014). Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res.; 15(1): pp. 1929–58.
10. Onathan L, Evan S, Trevor D. (2014). Fully convolutional networks for semantic segmentation. CoRR, abs/1411.4038.
11. Park Soo Beom, Jae Won Lee & Sang Kyoon Kim (2004). 'Content-based image classification using a neural network', Pattern Recognition Letters, vol. 25, no. 3, pp. 287-300
12. Zhang, B. (2010). 'Detection of Micro aneurysms using Multiscale Correlation Coefficients', Pattern Recognition, vol. 43, pp. 2237-2248

Corresponding Author

Amit Kumar Pandey*

PhD Scholar, Department of Computer Science and Engineering, Shri Venkateshwara University Gajraula, Amroha, Uttar Pradesh