A Study on Medical Image Data Augmentation Using Learning Techniques

Vanita D. Jadhav and **Lalit V. Patil**

Abstract One disadvantage of computer-assisted detection systems is the massive quantity of data needed to train them, which is costly in the medical industry. A big training dataset is critical in deep learning since it enhances training accuracy. Even with a big amount of data, a weak algorithm can be more accurate than a strong algorithm with a little amount of data. When verified on a different unobserved dataset, data augmentation generates new data which is used to train the model and enhances performance. We presented a thorough evaluation of the literature in which data augmentation was employed to train a learning model using lung CT images. Basic and deep learning data augmentation techniques were used to categorize the articles. The term "data augmentation" states a group of approaches for increasing the volume and quality of training datasets. Geometric transformations, kernel filters, color space augmentations, random erasing, mixing pictures, adversarial training, feature space dataset augmentation, meta-learning, and GAN-based networks are among the image augmentation processes explored in this paper. Students will learn how to employ data augmentation to improve model productivity and expand small datasets in order to take advantages of big data.

Keywords Training accuracy · Data augmentation · Big data · Deep learning · Image data · GANs

1   Introduction

In discriminative challenges, Deep Learning representations have achieved unbelievable growth. The development of deep network designs, commanding compute, and access to massive data have all contributed to this. Through the advancement of convolutional neural networks, deep neural nets have been effectively useful to image classification, picture segmentation and object identification. The spatial properties of images are reserved using parameterized, loosely linked kernels in these neural

V. D. Jadhav (✉) · L. V. Patil
SKNCOE, Savitribai Phule Pune University, Vadgaon(Bu), Pune, Maharashtra, India
e-mail: vdjadhav@coe.sveri.ac.in

nets. Convolutional layers reduce the 3-D resolution of pictures while increasing the deepness of their feature plots in a sequential manner. This sequence of convolutional changes can produce far lower-dimensional and further useful image illustrations than handmade ones. Convolutional Neural Networks' success has raised interest in using deep learning to solve computer vision issues. The validation error must decline with the training error in order to figure out effective deep learning prototypes. Data augmentation is a very effective way to accomplish this. In this study, we conducted a comprehensive evaluation of the literature in which data augmentation was used to train a deep learning model using lung CT images. Main objective of this study is to prepare live dataset representing CT scan images along with standard dataset and data augmentation methods.

2 Literature Survey

Geometric transformations, feature space augmentation, kernel filters, color space transformations, random erasing, mixing pictures, adversarial training, GAN-based augmentation, meta-learning methods, and neural style transfer are among the augmentations listed in this paper. This segment will describe how each augmentation algorithm works and analyses the method's downsides.

2.1 Geometric Transformations

This section explains how to use geometric transformations to create various augmentation techniques.

a. **Flipping**

A flip is a motion in geometry in which an object is turned over a straight line to form a mirror image. A flip is also called a reflection. Flipping the X axis is more common than flipping the Y axis. It is one of the easier augmentation techniques which has been recognized to work on datasets like ImageNet and CIFAR-10. It is not a label-protective transformation on datasets [1].

b. **Color space**

A tensor of the dimension is commonly used to encode digital image data. Another method which is quite feasible to implement is doing augmentations in the color channels space. Color augmentation can be as simple as separating a single-color channel like R, G, or B. A picture can be rapidly transformed into another picture in one color channel by dividing a matrix and addition of two zero matrices from the additional color stations. Using basic matrix operations, RGB values can be simply changed to enhance or lower the image's brightness.

Adjustments are made by altering the intensity levels in the histograms, similar to those found in photo editing software [2].

c. **Cropping**

Cropping photographs with varied height and width proportions are used as a valued processing step for image data. Furthermore, Random cropping can be used to create a similar effect to translations [1]. Random cropping varies from translations as it decreases the size of the input, whereas translations keep the image's 3-D dimensions.

d. **Rotation**

The image can be rotated right or left on an axis between 1° and 359° in rotation augmentations. The rotation degree parameter has a great influence on the safety of rotation augmentations [2].

e. **Translation**

Continuously shifting images right, left, down, and up can be a highly beneficial alteration for avoiding data positional bias. For instance, if all images in a dataset are perfectly positioned, the model must be validated on such pictures. The leftover space can be occupied with a constant value for example 0 s or 255 s, when the original picture is translated in a direction. This filling keeps image's 3-D dimensions after it is been augmented.

f. **Noise injection**

It is the process of inserting a matrix of arbitrary values, generally derived from Gaussian distribution. Addition of noise to photos can aid CNNs in recognizing further distinct features [3]. An excellent approach to cope with training data positional biases is Geometric transformation. There are a variety of biases that might cause training data distributions to deviate from testing data distributions. Geometric transformations are also advantageous because they are simple to apply. Increased memory, additional training time, and transformation computation costs are some of the drawbacks of geometric transformations [4].

2.2 Color Space Transformations

The picture data is divided in 3 matrices, which has a different size. These matrices signify the pixel values for each RGB color. Igniting biases are one of the most common problems that image recognition face. As a result, determining the efficiency of color space modifications (photometric transformations) is quite straightforward. Twisting over the photos and reducing or growing the pixel values by a fixed value is an easy solution for very bright or shady images. Another simple color space transformation is merging out individual RGB color matrices. Limiting pixel values to a defined least or extreme value is another adjustment. The inherent

representation of color in digital photographs allows for a wide range of augmentation approaches. This transformation can also be used in image-editing software [5]. Converting RGB matrices into a single grayscale image simplifies the representation of picture datasets. There are various ways to express digital color, such as HSV, besides RGB versus grayscale photos [6].

Color space conversions have several drawbacks, including bigger memory, cost of transformation, and time required for training. Because color modifications might potentially remove important color information, they are not necessarily a label-preserving alternative [7].

2.3 Kernel Filters

These are a type of image processing technique that can be used to improve and shape images. These filters use Gaussian blur filter to move a nxn matrix across an image, resulting in sharpy image along the edges. When pictures are blurred on the fly for data augmentation, they may be more resistant to gesture blur in testing. Furthermore, when images are refined for data augmentation, additional details about things of interest may be recorded. Sharpening and blurring are two examples of how kernel filters can be used on pictures. Kernel filters function better as a network layer rather than as a data augmentation dataset addition [8].

2.4 Mixing Images

A method of data augmentation that involves combining images by be around their pixel values is paradoxical. To a human viewer, the visuals created by this method will not appear to be a useful alteration. Another finding of the study is that when photos from the complete training set were mixed instead of instances from the same class, better results were obtained [9, 10].

This strategy has the obvious disadvantage of making little sense from a human standpoint. It is tough to comprehend or express the performance improvement that comes from combining images. One possibility is that as the dataset size grows, low-level properties like lines and edges become more robustly represented. The act of this strategy in comparison to pertaining methods and transfer learning is an intriguing field for further research. Other strategies for learning low-level properties in CNNs include transfer learning and pertaining [11, 12].

2.5 Random Erasing

It is a data augmentation technique invented by Zhong et al. Random erasing is related to dropout regularization, and it is based on dropout regularization mechanics. This method was developed to overcome occlusion-related image identification challenges. When some sections of an object are obscured, this is referred to as occlusion. Random erasure prevents this through encouraging the model to acquire additional graphic features of image, avoiding it from becoming fixated on a single graphic component [13]. Random erasing is a potential approach for guaranteeing that a network considers entire image instead of just a section of image, aside from the occlusion visual problem. Random erasing selects nxm pixels in a picture at random and masks them with 0 s, 255 s, random values or mean pixel values [14]. Additional augmentation methods, like color filters or horizontal flip, can be built on top of this augmentation method. Random erasing has the drawback of not necessarily being a label-preserving change. [15].

2.6 Feature Space Augmentation

All the above augmentation approaches are used on images in the input space. It is particularly impressive that neural networks can transform high-dimensional inputs into lower-dimensional pictures. In fattened layers, neural networks can map pictures to binary classes. Neural network's intermediate representations can be isolated from the network as a whole by altering the network's sequential processing. It is possible to separate and isolate the lower-dimensional pictures of visual input in fully-connected layers. The lower-dimensional pictures contained in the high-level layers of a CNN are stated to as the feature space. SMOTE is a well-liked augmentation for fixing concerns of class imbalance. By combining the K-nearest neighbors, this method is utilized to build new instances in the feature space. Feature space augmentation can also be accomplished by separating vector images from a CNN [16]. It is accomplished by slicing the network's output layer, resulting in a low-dimensional vector as the output instead of a class label. In the future, the effectiveness of this technique will be examined further [17]. A difficulty of feature space augmentation is that it is very difficult to interpret the vector data.

2.7 Adversarial Training

This is a method of using 2 or more than two networks with loss functions that have contrasting purposes set in them. Noise search or augmentation search which is application of adversarial training is still a new notion that has not been extensively explored. Though it has been demonstrated that employing adversarial search to insert

noise increases performance on adversarial instances, it is unknown whether this is also effective in reducing overfitting. The link between adversarial attack resistance and actual performance on test datasets will be the focus of future research [18].

2.8 *GAN-Based Data Augmentation*

Generative modeling is additional fascinating data augmentation approach. The exercise of constructing artificial instances from a dataset with parallel features to the unique set is known as generative modeling. The GAN-based data augmentation context can be expanded to advance the excellence of auto-encoder models. The outstanding performance of GANs has sparked a renewed interest in how they may be used for data augmentation. Neural networks are having the capability to provide more training data, resulting in more accurate classification models. GANs have disadvantage of requiring a big quantity of data to train [19].

2.9 *Neural Style Transfer*

It is one of the most fascinating demonstrations of deep learning skills. Fundamental impression is to use CNN-generated image representations. Although it is most known for its aesthetic uses, Neural Style Transfer can also be utilized for data augmentation. The approach manipulates the subsequent images in a CNN in such a way that the style of one image can be transferred to another while the original content is preserved [20].

A quick taxonomy of the data augmentations is shown below in Fig. 1.

3 Conclusion

This study classifies data augmentation hypotheses for the occurrence of overfitting in DL models due to a shortage of data. To avoid overfitting, DL models trust on large amounts of data. The benefits of large data in the restricted data realm can be achieved by precisely expanding datasets by means of the approaches mentioned in this survey. Data augmentation is a powerful tool for improving dataset quality. Deep neural networks' layered architecture opens up a lot of possibilities for data augmentation. The input layer is where the majority of the augmentations surveyed function. Some, however, are generated from hidden layer representations. The label space and the space of intermediate representations are two not yet explored areas of data augmentation with promising outcomes. Although many of these approaches and principles can be applied to other data domains, this study concentrates on applications for medical picture data. Data augmentation has a bright future ahead of
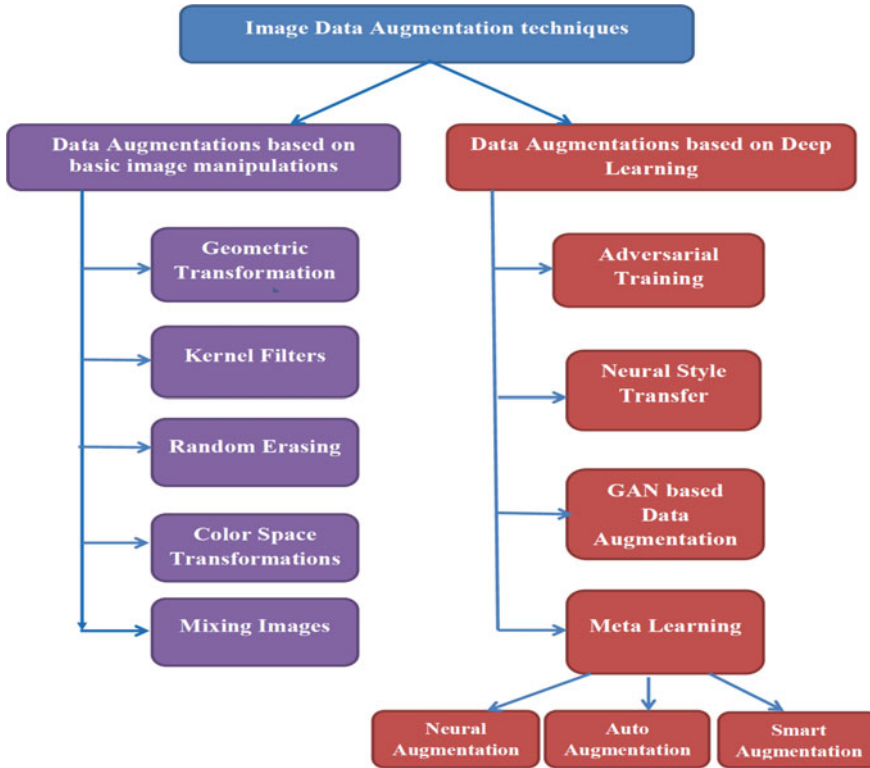
Fig. 1 Taxonomy of the data augmentations

it. The potential for using search algorithms that combine data warping and over-sampling methods is immense. Deep neural networks' layered architecture opens up a lot of possibilities for data augmentation. Main objective of this study is to prepare live dataset representing CT scan images along with standard dataset and data augmentation methods.

References

1. Olaf R, Philipp F, Thomas B (2015) U-Net: convolutional networks for biomedical image segmentation. In: MICCAI. Springer, p 234–41
2. Ekin DC, Barret Z, Dandelion M, Vijay V, Quoc VL (2018) Auto augment: learning augmentation policies from data. ArXiv preprint
3. Francisco JM-B, Fiammetta S, Jose MJ, Daniel U, Leonardo F (2018) Forward noise adjustment scheme for data augmentation. ArXiv preprints
4. Dua D, Karra TE (2017) UCI machine learning repository. University of California, School of Information and Computer Science, Irvine, CA. http://archive.ics.uci.edu/ml

5. Ken C, Karen S, Andrea V, Andrew Z (2014) Return of the devil in the details: delving deep into convolutional nets. In: Proceedings of BMVC
6. Mark E, Luc VG, Christopher KIW, John W, Andrew Z (2008) The pascal visual object classes (VOC) challenge. http://www.pascal-network.org/challenges/VOC/voc2008/workshop/
7. Aranzazu J, Miguel P, Mikel G, Carlos L-M, Daniel P (2010) A comparison study of different color spaces in clustering base image segmentation. IPMU
8. Guoliang K, Xuanyi D, Liang Z, Yi Y (2017) Patch Shuffle regularization. arXiv preprint
9. Hiroshi I (2018) Data augmentation by pairing samples for images classification. ArXiv eprints
10. Cecilia S, Michael JD (2018) Improved mixed-example data augmentation. ArXiv preprint
11. Daojun L, Feng Y, Tian Z, Peter Y (2018) Understanding mixup training methods. In: IEEE access
12. Ryo T, Takashi M (2018) Data augmentation using random image cropping and patches for deep CNNs. arXiv preprints
13. Zhun Z, Liang Z, Guoliang K, Shaozi L, Yi Y (2017) Random erasing data augmentation. ArXiv e-prints
14. Terrance V, Graham WT (2017) Improved regularization of convolutional neural networks with cutout. arXiv preprint
15. Agnieszka M, Michal G (2018) Data augmentation for improving deep learning in image classification problem. In: IEEE 2018 international interdisciplinary Ph.D. Workshop
16. Tomohiko K, Michiaki I (2018) Icing on the cake: an easy and quick post-learning method you can try after deep learning. arXiv preprints
17. Terrance V, Graham WT (2017) Dataset augmentation in feature space. In: Proceedings of the international conference on machine learning (ICML), workshop track
18. Seyed-Mohsen MD, Alhussein F, Pascal F (2016) DeepFool: a simple and accurate method to fool deep neural networks. arXiv preprint
19. Jiawei S, Danilo VV, Sakurai K (2018) One pixel attack for fooling deep neural networoks. arXiv preprints
20. Leon AG, Alexander SE, Matthias B (2015) A neural algorithm of artistic style. ArXiv