CHAPTER

Data augmentation for medical image analysis

14

He Zhao[a], **Huiqi Li**[a], **and Li Cheng**[b]
[a]*Beijing Institute of Technology, Beijing, China*
[b]*ECE, University of Alberta, Edmonton, AB, Canada*

14.1 Introduction

Medical image analysis has gained a lot of attention in the clinic. With the help of modern algorithms, computer-aided diagnosis assists doctors to deal with the large variations in pathology and intra- and inter-observers. Deep learning together with the high computational ability converts the human-designed features to a learning-based feature extraction process. Even without prior knowledge about the domain, the model can learn the meaningful features directly from data [1]. Therefore, a large amount of data are required to build the deep learning algorithms for learning a hierarchical feature representation. High accuracy by using a large dataset makes convolutional neural networks (CNNs) popular in medical image analysis tasks, such as liver lesion classification [2], brain analysis [3], and retinal image analysis [4]. For example, the method proposed by Google uses 1.28 million retinal images to train their systems to diagnose diabetic retinopathy [5], while deep CNNs also achieve a desirable result on skin lesion classification [6]. Unfortunately, such a large amount of data with labels is not always accessible in practical medical applications. There are two reasons which limit labeled data access in the medical image area. Firstly, it is a time-consuming and tedious task that requires experienced experts to spend a long time annotating. Secondly, it is also hard to get a desirable amount of disease images in the applications as the diseases are not common. Based on the type of variance, the features of images can be divided into two categories which are pertinent and non-pertinent features, respectively [7]. In medical images, pertinent features contain the most useful information to determine the organs or lesions, while non-pertinent features are those which vary between images such as the intensity difference. For most medical image applications, it is hard to build a large dataset, especially due to rare diseases and patient privacy. Training small datasets is harmful to the model to capture generalized pertinent features and may cause a performance drop due to the overfitting problem. In order to avoid the problem of overfitting and improve the performance of deep learning algorithms [8], researchers try to utilize data augmentation techniques [9–12]. Data augmentation is a way that can remove non-pertinent variance by feeding the model with different sources of data and it has been studied

Biomedical Image Synthesis and Simulation. https://doi.org/10.1016/B978-0-12-824349-7.00021-9
Copyright © 2022 Elsevier Inc. All rights reserved.

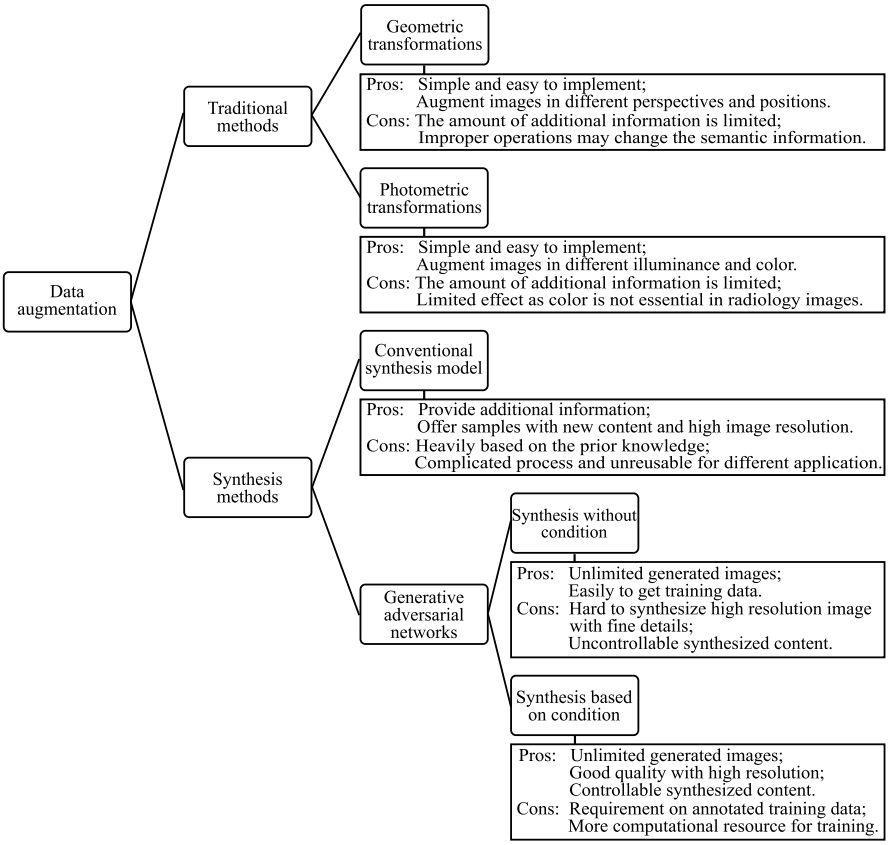


FIGURE 14.1

Summary of different data augmentation methods.

by the community to create more data for model training. In what follows, the works on data augmentation will be discussed, including the efforts on computer vision and medical image analysis. We will start from the traditional methods and end with the generative adversarial networks. The summary of methods is displayed in Fig. 14.1 which contains the pros and cons of each method.

14.2 Traditional methods for augmentation

The early works augment data by applying simple transformations on original images, where the techniques include geometric transformations and color space transformations.

14.2.1 Geometric transformations

Geometric transformation is one of the conventional methods to augment data, which aims to make the model invariant to the changes in position and orientation. Some typical operations include rotation with a certain angle, horizontal or vertical flipping, random patch cropping, scaling, shifting, and so on [13]. It provides a chance for the model to learn invariance without additional labeled data [14].

The rotation is done by rotating the images following a specific direction and angle between 0° and 360°. This operation changes the angles of an object in the images and helps the model to recognize the object with different angle locations. For most of the cases, rotation can improve the model performance, while sometimes it may not be a good choice when the direction of the object will influence the result, e.g., the digit numbers six and nine. Horizontal flipping is much more commonly used in retinal image analysis because usually there is a difference between right and left eyes. By using this operation, the model can learn the appearance from either left or right eyes without new data and labels. Random cropping is a technique to create a subset of images from the original image. This operation does not change the semantic information in the image but can provide a different situation of the object. It helps the model generalize better because the objects can be located in different positions and the contents are not the same in each cropping patch. Image shifting is another useful transformation to alleviate the position bias in the training data. Let us take optic disc segmentation in retinal fundus images as an example. When all the training images are centered on the optic disc, the model should recall the positions of optic discs in the center of the image. It will require the model to be tested on this kind of image to get perfect results. Unsatisfactory results may be generated for retinal images centered on the macular.

The geometric transformations are widely used in the augmentation of various kinds of images. They are easily implemented and helpful to overcome positional and angular biases. On the other hand, it is necessary when using geometric transformations to make sure the transformation has not changed the label of original images. Due to the complexity of medical data, the positional variances created by geometric transformations are not enough in some applications, which is a limitation of this method.

14.2.2 Photometric transformations

Another kind of data augmentation is carried on color space, which is called color space transformation or photometric transformation. This technique aims to make the model invariant to change of lighting and color. The color images have three channels, where each channel is with different pixel density. Lighting biases are the most frequent challenges for image classification or segmentation tasks, which may cause unsatisfactory results. To avoid this issue, one simple way is to edit the pixel with a constant value to increase or decrease the density. In [15], the authors not only apply geometric transformation but also change the values of RGB channels with a random

scale to enlarge the variance. The work of [16] has evaluated pixel value shuffling, which can help to improve the robustness of deep learning models. Krizhevsky et al. [17] apply principal component analysis (PCA) on the training images and alter the intensities of the RGB channels. They add the principal components with the magnitudes of the corresponding eigenvalues and a random variable from the Gaussian distribution. In [18], the authors randomly manipulate the brightness, color, and contrast with an arbitrary value of [0.5, 1.5]. They add an additional lighting noise as that has been done in [17]. This strategy has also been employed in [19] to augment skin images for melanoma analysis.

Compared with geometric transformations, the photometric transformations introduce other information which is illuminance and color diversity rather than spatial information of the object. In most applications, color may not be the essential characteristic and the improvement with color augmentation will be limited. Furthermore, this kind of augmentation can be hardly useful in most medical imaging data, as the radiology images are not sensitive to color.

The work of [20] has discussed the performance of several geometric and photometric transformation methods on natural images with a CNN model structure. There are also evaluations on medical image research [21,22] to discuss which augmentation technique is better. Interested readers may refer to these papers for more details. Some researchers also use label propagation for data augmentation [23,24]. The idea is based on Gaussian random fields, which allows the labels to be propagated to unlabeled data. A weighted undirected graph is constructed by the label propagation. The weights of graph edges reflect the similarity between the two data samples. By this operation, the unlabeled data can also be used to train models, which can also be regarded as a data augmentation method. Some researchers attempt to build a data augmentation method automatically, such as [25–27], where a policy is learned to choose which operation is suitable for the target dataset.

14.2.3 Augmentation on medical images

In this section, we will focus on the augmentation strategies on medical images. Hauberg et al. [28] propose a learned data augmentation approach on a class by class basis. They learn a probabilistic transformation model in the Riemanian sub-manifold of the Lie group of diffeomorphisms. In [29], the authors utilize two non-linear transformations called Simard transformation [30] and Ronneberger transformation [31] to generate new training images and corresponding vessel ground-truths in retinal fundus images. The new structures and scenes are created, and the generated images and vessel annotations can be used as new training data by applying these operations. In [32], a patch-based augmentation method is employed on the retinal vessel segmentation task, where random patches are selected for rotation, flipping, and nosing on both image and ground-truth. They state that this kind of strategy achieves a better performance than the usual operations. To augment the magnetic resonance imaging (MRI) volumes for prostate image segmentation, Milletari et al. employ non-

linear transformations by B-spline interpolation and histogram matching to vary the intensity distribution of data [33]. Roth et al. [34] enrich 3D computed tomography (CT) data by applying spatial deformations which include random translation, rotations, and non-rigid deformations. The non-rigid deformation is computed by fitting a thin-plate-spline to a regular grid of 2D control points and a deformed image can be generated using a radial basis function.

14.3 Synthesis-based methods

Although the conventional augmentation methods increase the number of training examples and are easily implemented, these strategies are highly sensitive to the choice of parameters [35] and have the limitation of emulating real variations [36]. In medical image research, careful consideration of the application will lead to which types of transformations are appropriate. For example, only sagittal reflection and intensity augmentation are used in [37] for brain lesion segmentation. In [31], the elastic deformations lead to the largest improvement for microscopy images, while this operation may not be useful in brain images. In addition, some patient-specific variations may not be removed by geometric transformations.

Besides the conventional methods which rely on existing images and apply transforms to create new training data, some researches are carried out to explore generating new data with different content and appearance-looking combinations. We call this kind of strategy a synthesis-based method in this chapter and it can be divided into methods based on domain knowledge and data-driven mechanisms.

14.3.1 Conventional synthesis model

The conventional synthesis model is heavily based on domain knowledge. To synthesize retinal fundus images, Fiorini et al. [38] focus on reconstructing the textural background from scratch. The background is generated by a patch-based tiling algorithm which is derived from the Image Quilting technique [39], where small examples of existing images are stitched together to obtain the phantom. Then a model learning the distributions of key morphometric quantities from real images is employed to reproduce optic disc on the retinal fundus. Finally, a model-based approach [40,41] is utilized to generate the vascular structures, where the features of vessel trees are included. The work of Menti et al. [42] aims to derive vessels and textures from real data utilizing active shape contours and Kalman filter techniques. For the neuronal images, GENESIS [43], NEURON [44], and L-Neuron [45] are the most well-known efforts. GENESIS is a simulation method for constructing realistic models of neurobiological systems. It was one of the first simulation systems specifically designed for modeling nervous systems. NEURON is developed similarly for modeling individual neurons and neuron networks. L-Neuron is based on a set of recursive rules that concisely describe dendritic geometry and topology through locally interrelated morphological parameters.

14.3.2 Generative adversarial networks

Leveraging the advances in generative adversarial networks (GANs) [46], some researchers try to synthesize realistic image and label pairs from a random noise vector that is generated from a simple distribution [7,47]. A generative adversarial network provides a new potential way to augment data making it possible to create augmented data without making decisions of which type to choose. Although the variational autoencoders (VAEs) [48] allow generating images with controllable latent variables, the synthesized samples tend to be blurry compared with generative adversarial networks. Nowadays, more efforts are put on GANs to synthesize realistic-looking images. It has been suggested that GANs can have a significant benefit when used for data augmentation in some classification tasks [2,49]. It can also augment more challenging variance data such as lesion shape or size. By providing a sufficient number of training samples with different shapes and sizes, the model will gain the ability to create samples with more variance. Performing the same augmentation results by the conventional methods will lead to a very complex model considering realistic shape, size, and the surroundings around the lesion. On the other hand, the generated image quality may be one disadvantage of GAN compared with conventional augmentation methods. But this limitation has less effect on improving the performance of other analysis tasks, as it has been proven that fully realistic images are not compulsory [50,51]. Balancing the pros and cons, the strategy using GANs is still an advanced method for data augmentation.

There are many efforts to utilize GANs to create artificial instances with similar characteristics to the original dataset. In the work of [52], Antoniou et al. discuss how the generative manifold can learn better classifiers and propose a class-based generative model to learn a representation for data augmentation. A conditional GAN is trained based on class-provided images with encoder–decoder model structure. The model learns a meaningful representation of training data and encapsulates the feature with a random vector engaging the variance to create extra augmented data. A synthetic refiner network is proposed in [53] to improve the quality of simulated images by GAN, where synthetic images created by the simulator are refined by the generator which is optimized by adversarial loss and self-regularization loss. The refined images are then used to train eye gaze estimation model.

In medical image research, some researchers explore image synthesis directly from a noise vector, while others decide to generate augmented images based on conditional information. The condition-based generative model has the potential to synthesize rare pathological cases, where the conditional information could be provided by medical experts through text descriptions or the masks of anatomical structures. All the attempts have shown the potentiality of GANs working on medical image analysis as a data augmentation method. The GAN models can offer an effective way to explore the manifold of training data and increase the data variance, which is hard to augment by other methods. However, this kind of augmentation cannot extend the distribution beyond the extremes of the training data.

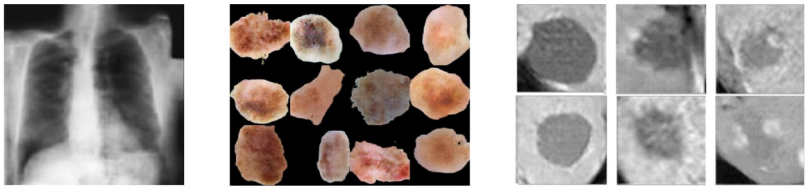
14.3.2.1 Synthesis without condition

Images generated without condition can also be known as unconditional synthesis, where no extra information is provided to synthesize the data and the images are generated directly from a random vector. The structure of deep convolutional GAN (DCGAN) [54] is widely employed by researchers to generate augmented data from random noise. In [2], the authors try to synthesize labeled lesions on CT scans for each class of lesion separately, including cysts, metastases, and hemangiomas. Three generators are employed for individual classes in their work. They find the generated lesions are helpful to improve the sensitivity and specificity when the augmented data are combined with real data. While this technique is also utilized for many applications such as chest X-ray [49], lung nodule CT images [55], retinal fundus images [56], and brain MR images [57,58]. In [56], the authors not only use the discriminator to distinguish the fake from real but also take it as a segmentor. The generated data together with the unlabeled data can be used to train the discriminator (segmentor) to achieve a better segmentation performance. It has been stated that the data generated from a random vector achieves a comparable quality to real ones on MR images although there is still a discrepancy in anatomic accuracy [57]. A visual Turing test has been carried out in [55] with two radiologists to evaluate the quality of generated nodules as augmented data.

Due to the limitation of DCGAN to synthesize large size images, other generative model structures have also been used for data augmentation, such as the Laplacian pyramid GAN [59], Wasserstein GAN [60], and progressive growing of GAN [61]. Bowles et al. [7] investigate the application of GAN in different modalities of medical images for the segmentation tasks as a way of data augmentation and compare the performance with rotation augmentation. Their model is based on a progressive growing of GAN, where the network can generate images with large sizes and stable training progress by a progressive growing training strategy. The same idea is also employed to create new training data for retinal fundus images with retinopathy of prematurity by [62] and dermoscopic images in [63]. In [63], a single source of noise is used instead of multiple sources for each scale and their model is trained in an end-to-end manner. As an improved version of GAN [46], the Wasserstein GAN has been used to improve the training stability and generate more plausible images. Several works have taken advantage of this variant of GAN to augment data on different image modalities, such as MR images [64], CT scans [65], histopathology images [66], or dermoscopic images [67]. Fig. 14.2 displays the examples of augmentation applications using unconditional GAN.

14.3.2.2 Synthesis based on condition

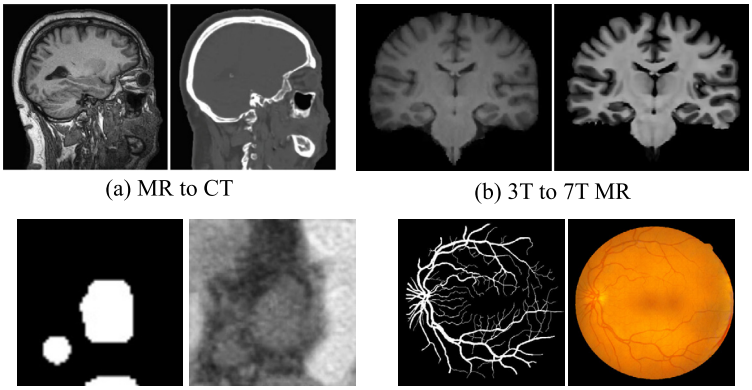
Although synthesis from a random vector needs less labeled data to train the generative model, this type of image generation is not controllable and the image quality is not satisfactory sometimes. To augment data with specific structure and position of lesions or vessels, researchers tend to introduce additional information to the generator where the conditional generative model [72] is utilized in their augmentation



(a) Chest x-ray (b) Dermoscopy image (c) Liver lesion CT images

FIGURE 14.2

Examples of applications using unconditional generative adversarial networks: (a) synthesized chest X-ray image [49]; (b) dermoscopy images generated by [67]; and (c) liver lesion images from [2].



(a) MR to CT (b) 3T to 7T MR

(c) Lymph mask to CT (d) Vessel to fundus image

FIGURE 14.3

Examples of applications using conditional generative adversarial networks: (a) CT image generated from an MR image [68]; (b) 7T MR image and its 3T MR input [69]; (c) augmented CT image and its input lymph mask [70]; and (d) synthesized retinal fundus image and its corresponding vessel map [71].

models. In this kind of method, pix2pix [73] and CycleGAN [74] are two model structures usually used to augment data. The pix2pix-based model is used for the cases where paired or aligned data are available, while the CycleGAN-based model is used to handle the tasks where no paired data are provided or the registration is challenging. One of the most common applications for this augmentation strategy is to augment data across modality such as from MRI to CT, or transformation from 3T to 7T MR images. Another kind of application is to create new data from existing annotations, e.g., from labeled masks to images. In what follows, we will introduce the work on these two kinds of applications with different data, while Fig. 14.3 illustrates the examples of images generated by state-of-the-art methods.

It is well known that CT is a critical imaging modality for many medical applications, but it exposes patients to radiation during image acquisition. As MRI does not involve radiation and is much safer than CT, the community is motivated to augment CT data from MR images to help training segmentation or classification methods on CT images. Within MR images, images acquired at 7T have a higher signal-to-noise ratio and better tissue contrast compared to 3T MRI, which leads to a more accurate disease diagnosis. However, 7T images are less available and may take a longer time to acquire. This raised the interest of researchers to generate 7T images from 3T MRI data. Nie et al. [69] propose a fully convolutional network based generative model to augment data from different sources with an additional constraint on the gradient similarity between real and synthesized images to generate high quality results. To engage the global information into their generator, the auto-context model is applied to achieve a context-aware GAN. To overcome the slice discontinuity and blurry problem on the boundary, Yu et al. [75] propose an edge-aware GAN to capture the image structure information and voxel-wise intensity to improve the quality of generated images. Besides the generator and discriminator, they investigate an edge detector to offer more information to the generator. In [68], the authors find that training with unpaired data with CycleGAN achieves better generated results than using the aligned images, where the reason is likely because the registration could not handle local alignment well. Several works further improve the CycleGAN model with extra loss and components. Hiasa et al. [76] extend the CycleGAN approach by adding the gradient consistency loss to improve the accuracy at the boundaries. The shape consistency loss is employed in [77] to avoid the geometric distortion, where two segmentors are utilized to extract semantic labels and the shape constraints on the anatomy during translation for both modalities. Other works [78,79] follow a similar idea but apply the segmentor to only one modality. Besides the works above, many papers have published cross-modality image synthesis approaches aiming to generate CT, positron emission tomography and MR images [80–85], but also different sequences in MR images [86–90]. Interested readers may refer to them for further reading.

Synthesizing images from labeled masks is another way to avoid the problem of lack of annotated data. The synthesized images together with their corresponding labeled masks can be regarded as new annotated training data. The authors in [91] propose a method to generate synthetic image–label pairs by learning the generative models of deformation fields and intensity transformations. Two generators are used for the deformation field and intensity field respectively. The deformation field generator creates a dense pixel-wise deformation field for both image and ground-truth. The intensity field generator generates an intensity mask to change the pixel intensity of an image. Tang et al. [70] utilize pix2pix GAN [73] to synthesize a large number of CT-realistic images from customized lymph node masks, where the model can learn the structural and contextual information of lymph nodes and surroundings. They state that by using the generative model, the augmented data achieves more diverse results than the ones generated by affine transformations. The simple pix2pix model is utilized in [47] to synthesize abnormal brain MR images from tumor labels. They

illustrate that the tumor segmentation model gains an improvement when leveraging the synthesized images as augmented data and achieves comparable results when trained only on the synthesized images. Mok et al. [92] use conditional GAN to augment training images for brain tumor segmentation. The generator is conditioned on a segmentation map and generates brain MR images in a coarse-to-fine manner, which also outputs tumor boundaries in the generation process to ensure the tumor is well delineated with a clear boundary in the generated image. Zhao et al. [93] generate labeled examples by using learning-based registration methods leveraging unlabeled images on brain MRI data. They employed two generators to learn spatial transformation and appearance transformation separately, where the anatomical and imaging diversities are captured in the unlabeled images. The new training examples are synthesized by sampling transformations and applying them to the existing labeled examples. A stylized GAN is proposed by [71], where the retinal fundus images and neuronal images are generated from the vascular or neuronal structures. A perceptual feature descriptor is introduced to extract content features and style features for the generator to synthesize stylized retinal images based on the input of style reference. As an improved version of [71], the authors in [94] propose a recurrent generative model to synthesize different style images within a single model, where the gated recurrent unit is engaged to control the information flow of different styles. The synthesized images have been used to train a supervised vessel segmentation model without annotated data. Instead of augmenting healthy retinal images, Zhou et al. [95] try to synthesize images with diabetic retinopathy. The proposed model is conditioned on vessel and lesion masks with adaptive grading vectors sampled from the latent grading space, which can be adopted to control the synthesized grading severity. To increase the quality of generated images, a multi-scale discriminator is designed to operate from large to small receptive fields. Finally, their model has achieved a charming result of synthesizing diabetic retinopathy images with rather high resolution and quality.

14.3.3 Pros and cons of synthesis methods

Compared with traditional augmentation methods, synthesis-based augmentation introduces a way to explore the features of existing data beyond the superficial information. It offers samples that do not appear in the training set but belong to the distribution. New contents are provided in the augmented data making the images more meaningful and powerful to improve model performance than the simple geometric and photometric transformations. However, the synthesis-based methods have their limitations. The conventional synthesis model is heavily based on prior knowledge and needs to be well designed for each component. Take the retinal fundus synthesis as an example. The conventional method needs to consider the orientation and branching of blood vessels to build a skeleton tree. The width of vessels on different locations should also be determined by a separate model with the texture appearance designed following other rules. All the processes make the conventional methods complicated and only usable for certain applications. On the other hand, the

GAN-based synthesis models are more flexible. They are constructed by data mining and learn to generate new images from either a random vector or a map. However, such models are not easy to train and require more data than the conventional synthesis method, which restrains their usage on a rather small amount of data. Compared with the conditional synthesis model, the unconditional GAN model can be easily used to synthesize additional data with only images required. It is much more convenient when it is hard to obtain images with mask annotations due to the laborious and time-consuming work in the medical imaging area. In general, the unconditional model cannot synthesize large size with fine details and it is harder to train than the conditional one. Although the BigGAN [96] provides the ability to generate high resolution images with a random vector as input, it requires a large amount of computational resources to achieve the goal. In addition, the content of generated images by the unconditional model cannot be controlled. Given a random vector, the GAN is not able to predict what the synthesized image will be like. On the other hand, the conditional model achieves higher resolution and an easy training process with the help of additional information that further guides the direction of the synthesized content. These advantages make the conditional GAN-based methods become a more and more popular data augmentation way in most applications.

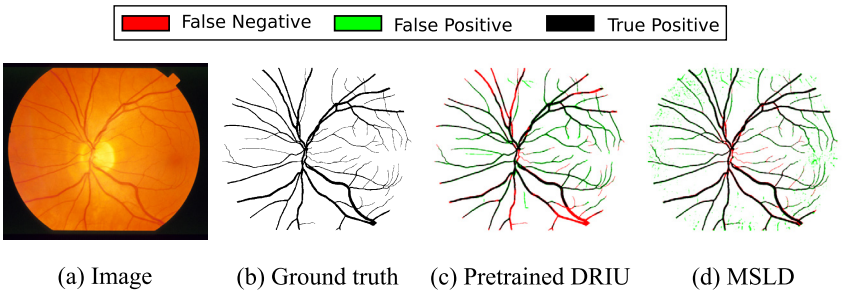
14.4 Case study: data augmentation for retinal vessel segmentation

In this section, a concrete example is given on retinal vessel segmentation to show how the augmentation method works on medical imaging analysis tasks.

Retinal vessel segmentation is a fundamental step in retinal image analysis. There is only a small set of annotated retinal image datasets available, e.g., DRIVE [97] contains 40 pixel-level annotated images and STARE [98] consists of 20 images with segmentation maps. Moreover, lots of the retinal fundus image datasets do not have any vessel segmentation. This situation is also faced by clinical practice. Lacking annotated training data on retinal images limits the usage of vessel segmentation methods. In the case of missing labeled training data, the most common solution is training the model with the existing annotated dataset and testing directly on the images. However, the performance usually drops significantly due to the discrepancy between different datasets. Table 14.1 and Fig. 14.4 show the performance of a supervised segmentation method (i.e., Deep Retinal Image Understanding (DRIU) [12]) pretrained on DRIVE and tested on STARE. The quantitative results show that the model trained on the existing labeled dataset, e.g., DRIVE, seems not to perform well on a new test set, e.g., STARE. Moreover, the performance of pretrained DRIU model is even worse than an unsupervised model, i.e., Multi-Scale Line Detector (MSLD) [99]. Fig. 14.4 displays the visual segmentation results of the two methods. The pretrained DRIU method has more false negative pixels compared with MSLD, which illustrates that the pretrained model with images from other datasets may not achieve a better performance than the unsupervised method. The experimental re-

Table 14.1 The performance of a supervised (DRIU) and unsupervised (MSLD) method on the STARE test dataset with F1-score (%), sensitivity (%), and specificity (%).

	F1	Sensitivity	Specificity
DRIU trained on DRIVE	68.32	67.12	98.03
MSLD	77.74	74.15	98.63



(a) Image (b) Ground truth (c) Pretrained DRIU (d) MSLD

FIGURE 14.4

The visual segmentation results of the pretrained supervised segmentation method (DRIU) and the unsupervised method (MSLD).

sults have shown that the discrepancy between different datasets does not allow a pretrained supervised model to achieve a satisfactory performance on another dataset.

To solve this problem, we select a conditional GAN-based method as the data augmentation way to generate images with the appearance of given data. The synthesized images together with the existing labels are used to train the supervised model to obtain a higher performance on the test set. The process can be divided into two steps as shown in Fig. 14.5. Step 1 focuses on the construction of a synthesized dataset with the desirable style appearance, while step 2 proceeds to learn a supervised model based on the generated images. In step 1, the synthesis method is built based on recurrent generative adversarial networks (R-sGAN) [94]. It takes random vessel trees as conditional information input and outputs the generated retinal fundus images. One of the biggest advantages of this method is that it is able to generate images with different style appearances. In what follows, we will give an introduction on how to use R-sGAN to generate new data for vessel segmentation. Interested readers may also refer to [94] for detailed information.

The core component of R-sGAN is the GRU module which consists of generator gate, reset gate, and update gate. The generator G of GAN is incorporated as the generator gate and the discriminator D is utilized as a part of losses. Each time point τ is treated as a cell of GRU with the input of current vessel trees $\boldsymbol{y}_\tau$ and cell state $\boldsymbol{h}_{\tau-1}$ while generating different style retinal images. The style information is stored in the cell state when training the model. As shown in Fig. 14.6(b), GRU consists of three parts: generator gate, reset gate, and update gate. The reset gate and update

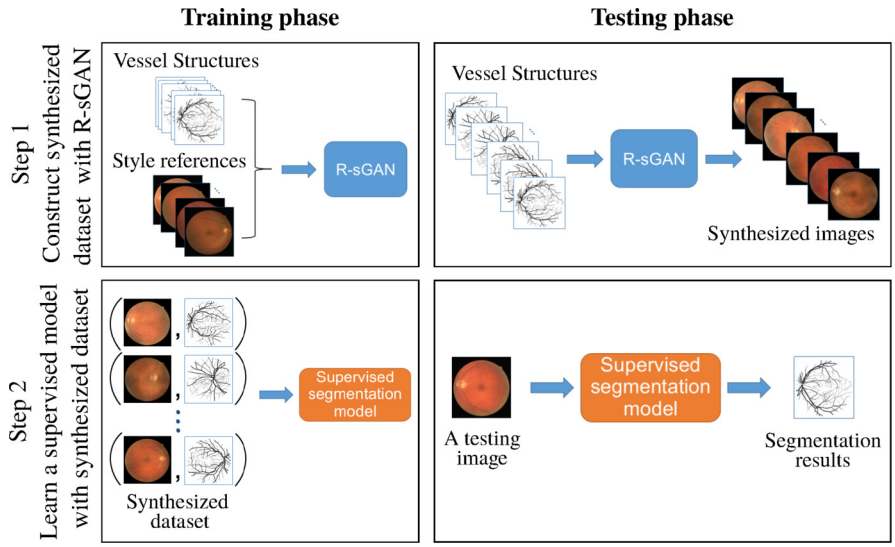


FIGURE 14.5

The flowchart of data augmentation on retinal vessel segmentation application.

gate control the information flow and the generator gate combines the current vessel structure and cell state to generate new images. By introducing different style reference images in the training stages, the model obtains the ability to synthesize multiple styles in different time states. Concretely, the reset gate and the update gate share the same fully convolutional network (FCN) structure but for different purposes, which are calculated by

$$\boldsymbol{r}_\tau = \sigma\left(\boldsymbol{f}_\gamma\left(\boldsymbol{y}_\tau, \boldsymbol{h}_{\tau-1}\right)\right), \quad \boldsymbol{u}_\tau = \sigma\left(\boldsymbol{f}_\mu\left(\boldsymbol{y}_\tau, \boldsymbol{h}_{\tau-1}\right)\right), \tag{14.1}$$

where $\boldsymbol{f}_\gamma$ and $\boldsymbol{f}_\mu$ are the FCN modules and σ is the sigmoid activation function. The cell state after the reset gate is changed to

$$\widetilde{\boldsymbol{h}}_{\tau-1} = \boldsymbol{r}_\tau \otimes \boldsymbol{h}_{\tau-1}, \tag{14.2}$$

where $\otimes$ indicates the element-wise multiplication. The generator gate takes the new cell state $\widetilde{\boldsymbol{h}}_{\tau-1}$, vessel structure $\boldsymbol{y}_\tau$, and a noise vector $\boldsymbol{z}_\tau$ as inputs and learns a mapping by a U-net-structured convolutional neural network G,

$$\widetilde{\boldsymbol{x}}_\tau = G\left(\boldsymbol{y}_\tau, \widetilde{\boldsymbol{h}}_{\tau-1}, \boldsymbol{z}_\tau\right). \tag{14.3}$$

By merging the signals from these three gates, the output becomes

$$\boldsymbol{h}_\tau = (1 - \boldsymbol{u}_\tau) \otimes \boldsymbol{h}_{\tau-1} \oplus \boldsymbol{u}_\tau \otimes \widetilde{\boldsymbol{x}}_\tau, \tag{14.4}$$

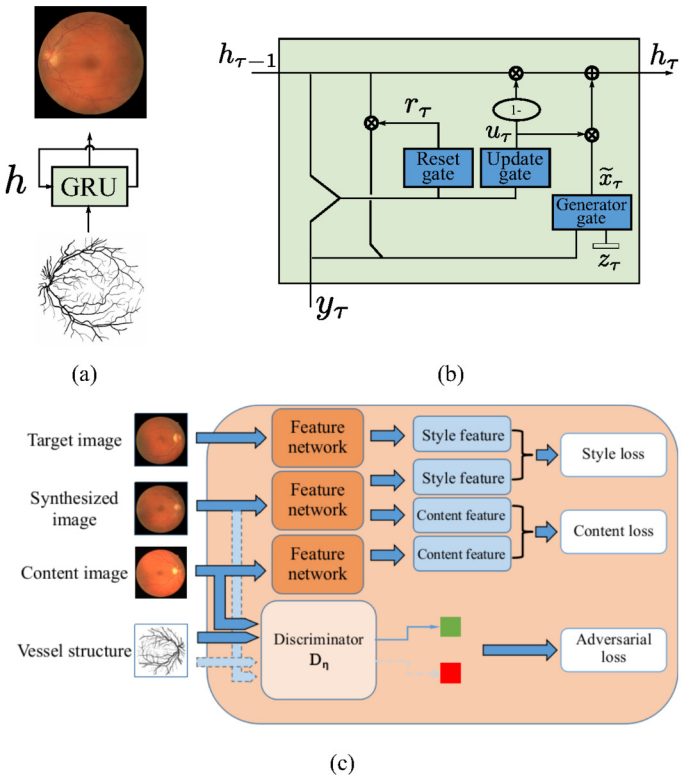


FIGURE 14.6

Model structure of R-sGAN: (a) the overall synthesis process; (b) detailed information of GRU network; and (c) the loss functions used to train the model.

where $\oplus$ refers to element-wise addition. With the help of reset gate and update gate, the GRU cell is able to capture the different styles and maintain the vessel structures over the sequence of time states.

Fig. 14.6(c) displays the main loss functions to train R-sGAN, which contain the adversarial loss, style loss, and content loss. The adversarial loss is used to guarantee the images are with a realistic appearance, which is trained by minimizing the following function:

$$\mathcal{L}_{\text{adv}} = -\sum_{\tau} \log D\left(\boldsymbol{h}_\tau, \boldsymbol{y}_\tau\right). \tag{14.5}$$

The style loss is engaged to evaluate how faithful the synthesized images is with respect to the style reference, while the content loss is used to enforce the generated images maintaining the vessel structure as the conditional information. These two

losses are calculated based on the feature representation generated by feature network (e.g., VGG [100]). For the style loss, the Gram matrix $G^l = (g^l_{mn})$ is utilized to capture the textural representation, which is defined as the inner product between the mth and nth vectorized feature maps in the lth layer,

$$g_{mn} = \sum_k \phi'_{mk} \phi'_{nk}, \tag{14.6}$$

where ϕ' is the vectorized version of feature map ϕ in feature network. With the style representation G^l the style loss can be defined as

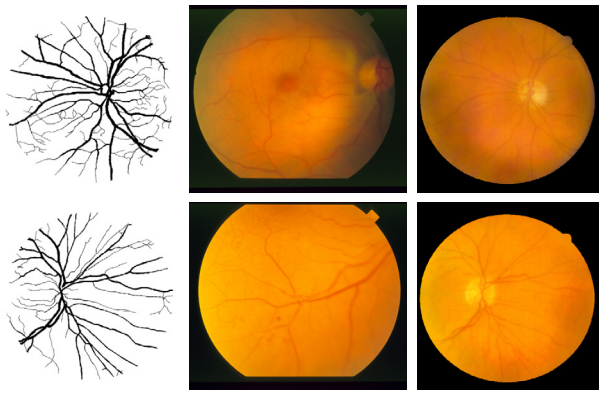
$$\mathcal{L}_{\text{sty}} = \sum_\tau \sum_l \left\| G^l(\boldsymbol{x}_\tau) - G^l(\boldsymbol{h}_\tau) \right\|^2, \tag{14.7}$$

where x_τ is the style reference image. Similarly, the content loss is calculated based on the difference of feature maps between the generated image and the vessel structure content, which is defined as

$$\mathcal{L}_{\text{cont}} = \sum_\tau \sum_l \left\| \phi^l\left(\boldsymbol{x}^c_\tau\right) - \phi^l(\boldsymbol{h}_\tau) \right\|^2, \tag{14.8}$$

where x^c_τ stands for the image providing the vessel structures. With all the components above, R-sGAN can generate realistic-looking retinal images with specific style appearance. This characteristic facilitates the segmentation method on the applications with less data or even no annotated data for training.

In Fig. 14.7, the augmented images generated by R-sGAN are displayed. The generated images maintain the vessel structures while obtaining the appearance-looking of style reference images from the test set. The segmentation performance improves with a significant margin by using the generated images as augmented inputs to train the supervised model. The quantitative results are shown in Table 14.2, while the visual results are shown in Fig. 14.8. Compared with the pretrained model, DRIU trained with augmented data generated by R-sGAN achieves superior performance with less false alarm and false positive points. To further reveal how large data augmentation benefits will be achieved in the segmentation tasks, visual comparisons on more challenging datasets will be presented. Fig. 14.9 displays the segmentation results of pretrained DRIU, DRIU trained with augmented data, and MSLD on Kaggle dataset [101] and images captured by mobile device [102]. In the Kaggle dataset, there is noticeable variability in contrast and luminance, mostly due to the presence of diabetic retinopathy. For the images captured by mobile devices, the imaging quality may differ a lot due to the less experienced user and lead to low quality and poorly illuminated fundus images. On these two datasets, the pretrained DRIU model and MSLD method cannot obtain satisfactory results, either by capturing the lesion or



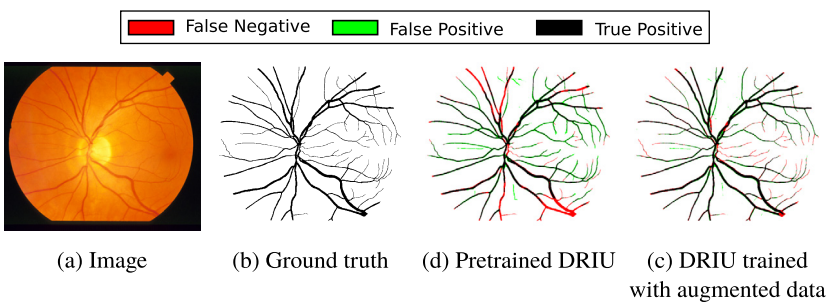
(a) Vessel structure (b) Style reference (c) Generated image

FIGURE 14.7

Two exemplar images generated by R-sGAN with vessel structures from the DRIVE dataset and style references from the STARE dataset.

Table 14.2 The performance of DRIU models with different training strategies on STARE dataset with F1-score (%), sensitivity (%), and specificity (%).

	F1	Sensitivity	Specificity
DRIU trained on DRIVE	68.32	67.12	98.03
DRIU trained on augmented data	79.60	79.49	98.36



(a) Image (b) Ground truth (d) Pretrained DRIU (c) DRIU trained with augmented data

FIGURE 14.8

Visual segmentation comparison of the supervised segmentation models (DRIU) with different training strategies.

missing the vessel structures. On the other hand, the model trained with augmented data achieves a better result with main vessel trunks and some detailed branches will less false alarm even on the blurry and less visible images.

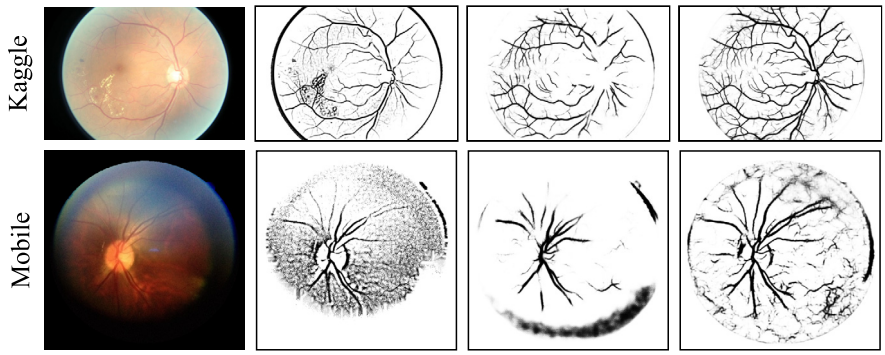


FIGURE 14.9

Visual segmentation comparison of the supervised models (DRIU) with different training strategies.

14.5 Research challenges and future work

Benefiting from the generative adversarial network, the augmentation methods have moved to a new stage. However, there are still challenges that need to be solved in medical imaging. Image quality is the most important thing in GAN-based methods to generate augmented data. Although most of the generated data have the desired appearance, the anatomic structures may not match with the clinical situation, such as the optic disc shape in the retinal fundus images or the appearance of tumor in brain MR images. The structure may be good enough from the image point of view, but the augmented image may be a failure case in the clinical sense. How to augment data with clinically meaningful structure is a future work to be explored. Most of the generative models gain good quality of augmented images by training with paired data where extensive labeling work needs to be done in order to create the training dataset. Some works utilize CycleGAN to loose the restriction where an unpaired translation is performed, while most of the recent works focus on synthesizing images from existing ground-truth. There is little work trying to augment ground-truth data. For some applications, such as retinal images or neuronal images, vascular or neuronal structures contain the most useful information. So augmenting the vascular or neuronal ground-truth is very helpful to create new data and it provides more potential for the generative models to learn the essential structures in this type of image. Another new direction is to augment data not only in the image space but also in the feature space, which works like domain adaptation. Finally, there is no effective way to measure the quality of generated images. Some of the traditional evaluation metrics such as structural similarity index measure and peak signal-to-noise ratio can be used for paired data evaluation, but the results do not always correspond to the real visual quality or match visual conception of the human. Interested readers may refer

to Chapter 25 for more information. The way to validate the augmented data by clinical experts is another choice but expensive and time-consuming. Thus the validation metrics of the generated images remain to be explored.

14.6 Summary

In this chapter, we have reviewed the augmentation methods used in medical image analysis. Deep learning methods in medical image analysis achieve a significant performance in many applications. However, they rely on a large amount of data to train the model to avoid overfitting problems and improve the model performance. Data augmentation is a very useful strategy to create a large dataset with variance, which agrees with the situation that labeled data in medical imaging is very limited. Several kinds of augmentation methods have been discussed, including conventional methods such as image rotation or flipping, and methods based on synthesis. Benefiting from the development of generative adversarial networks, the quality and variance of augmented data are substantially improved. Models trained with the augmented data obtain a higher performance in medical image analysis. The augmentation in medical image analysis has a bright future but comes with challenges that need to be explored.

References

[1] I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, Deep Learning, vol. 1, MIT Press, Cambridge, 2016.

[2] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, H. Greenspan, Synthetic data augmentation using GAN for improved liver lesion classification, in: IEEE International Symposium on Biomedical Imaging, IEEE, 2018, pp. 289–293.

[3] J.H. Cole, R.P. Poudel, D. Tsagkrasoulis, M.W. Caan, C. Steves, T.D. Spector, G. Montana, Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker, NeuroImage 163 (2017) 115–124.

[4] H. Zhao, H. Li, L. Cheng, Improving retinal vessel segmentation with joint local loss by matting, Pattern Recognition 98 (2020) 107068.

[5] V. Gulshan, L. Peng, M. Coram, M.C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, et al., Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs, JAMA 316 (22) (2016) 2402–2410.

[6] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118.

[7] C. Bowles, L. Chen, R. Guerrero, P. Bentley, R. Gunn, A. Hammers, D.A. Dickie, M.V. Hernández, J. Wardlaw, D. Rueckert, GAN augmentation: augmenting training data using generative adversarial networks, arXiv preprint, arXiv:1810.10863.

[8] L. Perez, J. Wang, The effectiveness of data augmentation in image classification using deep learning, arXiv preprint, arXiv:1712.04621.

[9] A.A.A. Setio, F. Ciompi, G. Litjens, P. Gerke, C. Jacobs, S.J. van Riel, M.M.W. Wille, M. Naqibullah, C.I. Sánchez, B. van Ginneken, Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks, IEEE Transactions on Medical Imaging 35 (5) (2016) 1160–1169.
[10] H.R. Roth, L. Lu, J. Liu, J. Yao, A. Seff, K. Cherry, L. Kim, R.M. Summers, Improving computer-aided detection using convolutional neural networks and random view aggregation, IEEE Transactions on Medical Imaging 35 (5) (2016) 1170–1181.
[11] J.-Z. Cheng, D. Ni, Y.-H. Chou, J. Qin, C.-M. Tiu, Y.-C. Chang, C.-S. Huang, D. Shen, C.-M. Chen, Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans, Scientific Reports 6 (1) (2016) 1–13.
[12] K.-K. Maninis, J. Pont-Tuset, P. Arbeláez, L. Van Gool, Deep retinal image understanding, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 140–148.
[13] M.D. Bloice, C. Stocker, A. Holzinger, Augmentor: an image augmentation library for machine learning, arXiv preprint, arXiv:1708.04680.
[14] H. Zhang, M. Cisse, Y.N. Dauphin, D. Lopez-Paz, Mixup: beyond empirical risk minimization, arXiv preprint, arXiv:1710.09412.
[15] D. Eigen, C. Puhrsch, R. Fergus, Depth map prediction from a single image using a multi-scale deep network, in: Advances in Neural Information Processing Systems, 2014, pp. 2366–2374.
[16] F.J. Moreno-Barea, F. Strazzera, J.M. Jerez, D. Urda, L. Franco, Forward noise adjustment scheme for data augmentation, in: IEEE Symposium Series on Computational Intelligence, IEEE, 2018, pp. 728–734.
[17] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[18] A.G. Howard, Some improvements on deep convolutional neural network based image classification, arXiv preprint, arXiv:1312.5402.
[19] C.N. Vasconcelos, B.N. Vasconcelos, Convolutional neural network committees for melanoma classification with classical and expert knowledge based image transforms data augmentation, arXiv preprint, arXiv:1702.07025.
[20] L. Taylor, G. Nitschke, Improving deep learning using generic data augmentation, arXiv preprint, arXiv:1708.06020.
[21] F. Perez, C. Vasconcelos, S. Avila, E. Valle, Data augmentation for skin lesion analysis, in: OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis, Springer, 2018, pp. 303–311.
[22] B. Abdollahi, N. Tomita, S. Hassanpour, Data augmentation in training deep learning models for medical image analysis, in: Deep Learners and Deep Learner Descriptors for Medical Applications, Springer, 2020, pp. 167–180.
[23] X. Zhu, Z. Ghahramani, J.D. Lafferty, Semi-supervised learning using Gaussian fields and harmonic functions, in: Proceedings of the International Conference on Machine Learning (ICML-03), 2003, pp. 912–919.
[24] X. Lu, B. Zheng, A. Velivelli, C. Zhai, Enhancing text categorization with semantic-enriched representation and training data augmentation, Journal of the American Medical Informatics Association 13 (5) (2006) 526–535.
[25] J. Lemley, S. Bazrafkan, P. Corcoran, Smart augmentation learning an optimal data augmentation strategy, IEEE Access 5 (2017) 5858–5869.

[26] T. Tran, T. Pham, G. Carneiro, L. Palmer, I. Reid, A Bayesian data augmentation approach for learning deep models, in: Advances in Neural Information Processing Systems, 2017, pp. 2797–2806.
[27] E.D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, Q.V. Le, AutoAugment: learning augmentation strategies from data, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 113–123.
[28] S. Hauberg, O. Freifeld, A.B.L. Larsen, J. Fisher, L. Hansen, Dreaming more data: class-dependent distributions over diffeomorphisms for learned data augmentation, in: Artificial Intelligence and Statistics, 2016, pp. 342–350.
[29] A. Oliveira, S. Pereira, C.A. Silva, Augmenting data when training a CNN for retinal vessel segmentation: how to warp?, in: IEEE Portuguese Meeting on Bioengineering, IEEE, 2017, pp. 1–4.
[30] P.Y. Simard, D. Steinkraus, J.C. Platt, et al., Best practices for convolutional neural networks applied to visual document analysis, in: ICDAR, vol. 3, 2003.
[31] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.
[32] Y. Jiang, H. Zhang, N. Tan, L. Chen, Automatic retinal blood vessel segmentation based on fully convolutional neural networks, Symmetry 11 (9) (2019) 1112.
[33] F. Milletari, N. Navab, S.-A. Ahmadi, V-Net: fully convolutional neural networks for volumetric medical image segmentation, in: International Conference on 3D Vision, IEEE, 2016, pp. 565–571.
[34] H.R. Roth, C.T. Lee, H.-C. Shin, A. Seff, L. Kim, J. Yao, L. Lu, R.M. Summers, Anatomy-specific classification of medical images using deep convolutional nets, in: IEEE International Symposium on Biomedical Imaging, IEEE, 2015, pp. 101–104.
[35] A. Dosovitskiy, P. Fischer, J.T. Springenberg, M. Riedmiller, T. Brox, Discriminative unsupervised feature learning with exemplar convolutional neural networks, IEEE Transactions on Pattern Analysis and Machine Intelligence 38 (9) (2015) 1734–1747.
[36] Z. Eaton-Rosen, F. Bragman, S. Ourselin, M.J. Cardoso, Improving data augmentation for medical image segmentation, in: International Conference on Medical Imaging with Deep Learning, 2018.
[37] K. Kamnitsas, C. Ledig, V.F. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Medical Image Analysis 36 (2017) 61–78.
[38] S. Fiorini, L. Ballerini, E. Trucco, A. Ruggeri, Automatic generation of synthetic retinal fundus images, in: STAG, 2014, pp. 41–44.
[39] A.A. Efros, W.T. Freeman, Image quilting for texture synthesis and transfer, in: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, 2001, pp. 341–346.
[40] J.A. Adam, Blood vessel branching: beyond the standard calculus problem, Mathematics Magazine 84 (3) (2011) 196–207.
[41] F. Oloumi, R.M. Rangayyan, A.L. Ells, Parabolic modeling of the major temporal arcade in retinal fundus images, IEEE Transactions on Instrumentation and Measurement 61 (7) (2012) 1825–1838.
[42] E. Menti, L. Bonaldi, L. Ballerini, A. Ruggeri, E. Trucco, Automatic generation of synthetic retinal fundus images: vascular network, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer, 2016, pp. 167–176.
[43] J.M. Bower, H. Cornelis, D. Beeman, Genesis, the general neural simulation system, 2014.

[44] N.T. Carnevale, M.L. Hines, The NEURON Book, Cambridge University Press, 2006.
[45] G.A. Ascoli, J.L. Krichmar, L-neuron: a modeling tool for the efficient generation and parsimonious description of dendritic morphology, Neurocomputing 32 (2000) 1003–1011.
[46] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: Advances in Neural Information Processing Systems, 2014, pp. 2672–2680.
[47] H.-C. Shin, N.A. Tenenholtz, J.K. Rogers, C.G. Schwarz, M.L. Senjem, J.L. Gunter, K.P. Andriole, M. Michalski, Medical image synthesis for data augmentation and anonymization using generative adversarial networks, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer, 2018, pp. 1–11.
[48] D.P. Kingma, M. Welling, Auto-encoding variational Bayes, in: International Conference on Learning Representations, 2014.
[49] A. Madani, M. Moradi, A. Karargyris, T. Syeda-Mahmood, Chest X-ray generation and data augmentation for cardiovascular abnormality classification, in: Medical Imaging 2018: Image Processing, vol. 10574, International Society for Optics and Photonics, 2018, p. 105741M.
[50] A. Dosovitskiy, P. Fischer, E. Ilg, P. Hausser, C. Hazirbas, V. Golkov, P. van der Smagt, D. Cremers, T. Brox, FlowNet: learning optical flow with convolutional networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 2758–2766.
[51] S.R. Richter, V. Vineet, S. Roth, V. Koltun, Playing for data: ground truth from computer games, in: European Conference on Computer Vision, Springer, 2016, pp. 102–118.
[52] A. Antoniou, A. Storkey, H. Edwards, Data augmentation generative adversarial networks, arXiv preprint, arXiv:1711.04340.
[53] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, R. Webb, Learning from simulated and unsupervised images through adversarial training, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2107–2116.
[54] A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, arXiv preprint, arXiv:1511.06434.
[55] M.J. Chuquicusma, S. Hussein, J. Burt, U. Bagci, How to fool radiologists with generative adversarial networks? A visual turing test for lung cancer diagnosis, in: IEEE International Symposium on Biomedical Imaging, IEEE, 2018, pp. 240–244.
[56] A. Lahiri, V. Jain, A. Mondal, P.K. Biswas, Retinal vessel segmentation under extreme low annotation: a GAN based semi-supervised approach, in: IEEE International Conference on Image Processing, IEEE, 2020, pp. 418–422.
[57] C. Bermudez, A.J. Plassard, L.T. Davis, A.T. Newton, S.M. Resnick, B.A. Landman, Learning implicit brain MRI manifolds with deep learning, in: Medical Imaging 2018: Image Processing, vol. 10574, International Society for Optics and Photonics, 2018, p. 105741L.
[58] A.K. Mondal, J. Dolz, C. Desrosiers, Few-shot 3D multi-modal medical image segmentation using generative adversarial learning, arXiv preprint, arXiv:1810.12241.
[59] E.L. Denton, S. Chintala, R. Fergus, et al., Deep generative image models using a Laplacian pyramid of adversarial networks, in: Advances in Neural Information Processing Systems, 2015, pp. 1486–1494.
[60] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein generative adversarial networks, in: Proceedings of the International Conference on Machine Learning, vol. 70, 2017, pp. 214–223.

[61] T. Karras, T. Aila, S. Laine, J. Lehtinen, Progressive growing of GANs for improved quality, stability, and variation, in: International Conference on Learning Representations, 2018.
[62] A. Beers, J. Brown, K. Chang, J.P. Campbell, S. Ostmo, M.F. Chiang, J. Kalpathy-Cramer, High-resolution medical image synthesis using progressively grown generative adversarial networks, arXiv preprint, arXiv:1805.03144.
[63] C. Baur, S. Albarqouni, N. Navab, MelanoGANs: high resolution skin lesion synthesis with GANs, arXiv preprint, arXiv:1804.04338.
[64] C. Han, H. Hayashi, L. Rundo, R. Araki, W. Shimoda, S. Muramatsu, Y. Furukawa, G. Mauri, H. Nakayama, GAN-based synthetic brain MR image generation, in: IEEE International Symposium on Biomedical Imaging, IEEE, 2018, pp. 734–738.
[65] Q. Wang, X. Zhou, C. Wang, Z. Liu, J. Huang, Y. Zhou, C. Li, H. Zhuang, J.-Z. Cheng, WGAN-based synthetic minority over-sampling technique: improving semantic fine-grained classification for lung nodules in CT images, IEEE Access 7 (2019) 18450–18463.
[66] B. Hu, Y. Tang, I. Eric, C. Chang, Y. Fan, M. Lai, Y. Xu, Unsupervised learning for cell-level visual representation in histopathology images with generative adversarial networks, IEEE Journal of Biomedical and Health Informatics 23 (3) (2018) 1316–1328.
[67] X. Yi, E. Walia, P. Babyn, Unsupervised and semi-supervised learning with categorical generative adversarial networks assisted by Wasserstein distance for dermoscopy image classification, arXiv preprint, arXiv:1804.03700.
[68] J.M. Wolterink, A.M. Dinkla, M.H. Savenije, P.R. Seevinck, C.A. van den Berg, I. Išgum, Deep MR to CT synthesis using unpaired data, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer, 2017, pp. 14–23.
[69] D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, D. Shen, Medical image synthesis with deep convolutional adversarial networks, IEEE Transactions on Biomedical Engineering 65 (12) (2018) 2720–2730.
[70] Y.-B. Tang, S. Oh, Y.-X. Tang, J. Xiao, R.M. Summers, CT-realistic data augmentation using generative adversarial network for robust lymph node segmentation, in: Medical Imaging 2019: Computer-Aided Diagnosis, vol. 10950, International Society for Optics and Photonics, 2019, p. 109503V.
[71] H. Zhao, H. Li, S. Maurer-Stroh, L. Cheng, Synthesizing retinal and neuronal images with generative adversarial nets, Medical Image Analysis 49 (2018) 14–26.
[72] M. Mirza, S. Osindero, Conditional generative adversarial nets, arXiv preprint, arXiv: 1411.1784.
[73] P. Isola, J.-Y. Zhu, T. Zhou, A.A. Efros, Image-to-image translation with conditional adversarial networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1125–1134.
[74] J.-Y. Zhu, T. Park, P. Isola, A.A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2223–2232.
[75] B. Yu, L. Zhou, L. Wang, Y. Shi, J. Fripp, P. Bourgeat, EA-GANs: edge-aware generative adversarial networks for cross-modality MR image synthesis, IEEE Transactions on Medical Imaging 38 (7) (2019) 1750–1762.
[76] Y. Hiasa, Y. Otake, M. Takao, T. Matsuoka, K. Takashima, A. Carass, J.L. Prince, N. Sugano, Y. Sato, Cross-modality image synthesis from unpaired data using CycleGAN, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer, 2018, pp. 31–41.

[77] Z. Zhang, L. Yang, Y. Zheng, Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 9242–9251.
[78] C. Chen, Q. Dou, H. Chen, P.-A. Heng, Semantic-aware generative adversarial nets for unsupervised domain adaptation in chest X-ray segmentation, in: International Workshop on Machine Learning in Medical Imaging, Springer, 2018, pp. 143–151.
[79] Y. Zhang, S. Miao, T. Mansi, R. Liao, Task driven generative modeling for unsupervised domain adaptation: application to X-ray image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 599–607.
[80] A. Chartsias, T. Joyce, R. Dharmakumar, S.A. Tsaftaris, Adversarial image synthesis for unpaired multi-modal cardiac data, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer, 2017, pp. 3–13.
[81] Y. Huo, Z. Xu, S. Bao, A. Assad, R.G. Abramson, B.A. Landman, Adversarial synthesis learning enables segmentation without target modality ground truth, in: IEEE International Symposium on Biomedical Imaging, IEEE, 2018, pp. 1217–1220.
[82] H. Yang, J. Sun, A. Carass, C. Zhao, J. Lee, Z. Xu, J. Prince, Unpaired brain MR-to-CT synthesis using a structure-constrained CycleGAN, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer, 2018, pp. 174–182.
[83] W. Wei, E. Poirion, B. Bodini, S. Durrleman, N. Ayache, B. Stankoff, O. Colliot, Learning myelin content in multiple sclerosis from multimodal MRI through adversarial training, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 514–522.
[84] Y. Pan, M. Liu, C. Lian, T. Zhou, Y. Xia, D. Shen, Synthesizing missing PET from MRI with cycle-consistent generative adversarial networks for Alzheimer's disease diagnosis, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 455–463.
[85] V. Sandfort, K. Yan, P.J. Pickhardt, R.M. Summers, Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks, Scientific Reports 9 (1) (2019) 1–9.
[86] B. Yu, L. Zhou, L. Wang, J. Fripp, P. Bourgeat, 3D CGAN based cross-modality MR image synthesis for brain tumor segmentation, in: IEEE International Symposium on Biomedical Imaging, IEEE, 2018, pp. 626–630.
[87] S. Olut, Y.H. Sahin, U. Demir, G. Unal, Generative adversarial training for MRA image synthesis using multi-contrast MRI, in: International Workshop on Predictive Intelligence in Medicine, Springer, 2018, pp. 147–154.
[88] Q. Yang, N. Li, Z. Zhao, X. Fan, E.-C. Chang, Y. Xu, et al., MRI image-to-image translation for cross-modality image registration and segmentation, arXiv preprint, arXiv:1801.06940.
[89] F. Liu, SUSAN: segment unannotated image structure using adversarial network, Magnetic Resonance in Medicine 81 (5) (2019) 3330–3345.
[90] S.U. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, T. Çukur, Image synthesis in multi-contrast MRI with conditional generative adversarial networks, IEEE Transactions on Medical Imaging 38 (10) (2019) 2375–2388.
[91] K. Chaitanya, N. Karani, C.F. Baumgartner, A. Becker, O. Donati, E. Konukoglu, Semi-supervised and task-driven data augmentation, in: International Conference on Information Processing in Medical Imaging, Springer, 2019, pp. 29–41.

[92] T.C. Mok, A.C. Chung, Learning data augmentation for brain tumor segmentation with coarse-to-fine generative adversarial networks, in: International MICCAI Brainlesion Workshop, Springer, 2018, pp. 70–80.
[93] A. Zhao, G. Balakrishnan, F. Durand, J.V. Guttag, A.V. Dalca, Data augmentation using learned transformations for one-shot medical image segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 8543–8553.
[94] H. Zhao, H. Li, S. Maurer-Stroh, Y. Guo, Q. Deng, L. Cheng, Supervised segmentation of un-annotated retinal fundus images by synthesis, IEEE Transactions on Medical Imaging 38 (1) (2018) 46–56.
[95] Y. Zhou, X. He, S. Cui, F. Zhu, L. Liu, L. Shao, High-resolution diabetic retinopathy image synthesis manipulated by grading and lesions, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 505–513.
[96] A. Brock, J. Donahue, K. Simonyan, Large scale GAN training for high fidelity natural image synthesis, in: International Conference on Learning Representations, 2018.
[97] J. Staal, M.D. Abràmoff, M. Niemeijer, M.A. Viergever, B. van Ginneken, Ridge-based vessel segmentation in color images of the retina, IEEE Transactions on Medical Imaging 23 (4) (2004) 501–509.
[98] A. Hoover, V. Kouznetsova, M. Goldbaum, Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response, IEEE Transactions on Medical Imaging 19 (3) (2000) 203–210.
[99] U.T. Nguyen, A. Bhuiyan, L.A. Park, K. Ramamohanarao, An effective retinal blood vessel segmentation method using multi-scale line detection, Pattern Recognition 46 (3) (2013) 703–715.
[100] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv:1409.1556.
[101] Diabetic retinopathy detection, www.kaggle.com/c/diabetic-retinopathy-detection, 2015.
[102] X. Xu, W. Ding, X. Wang, R. Cao, M. Zhang, P. Lv, F. Xu, Smartphone-based accurate analysis of retinal vasculature towards point-of-care diagnostics, Scientific Reports 6 (1) (2016) 1–9.