A review of medical image data augmentation techniques for deep learning applications

Phillip Chlap,[1,2,3] (iD) Hang Min,[1,2,4] (iD) Nym Vandenberg,[5] Jason Dowling,[1,4] (iD) Lois Holloway[1,2,3,5,6] (iD) and Annette Haworth[5] (iD)

1 South Western Sydney Clinical School, University of New South Wales, Sydney, New South Wales, Australia
2 Ingham Institute for Applied Medical Research, Sydney, New South Wales, Australia
3 Liverpool and Macarthur Cancer Therapy Centre, Liverpool Hospital, Sydney, New South Wales, Australia
4 The Australian e-Health and Research Centre, CSIRO Health and Biosecurity, Brisbane, Queensland, Australia
5 Institute of Medical Physics, University of Sydney, Sydney, New South Wales, Australia
6 Centre for Medical Radiation Physics, University of Wollongong, Wollongong, New South Wales, Australia

P Chlap MSc; **H Min** PhD; **N Vandenberg** MSc; **J Dowling** PhD; **L Holloway** PhD; **A Haworth** PhD.

Correspondence

Mr Phillip Chlap, Ingham Institute, 1 Campbell Street, Liverpool 2170, NSW, Australia.
Email: phillip.chlap@unsw.edu.au

Summary

Research in artificial intelligence for radiology and radiotherapy has recently become increasingly reliant on the use of deep learning-based algorithms. While the performance of the models which these algorithms produce can significantly outperform more traditional machine learning methods, they do rely on larger datasets being available for training. To address this issue, data augmentation has become a popular method for increasing the size of a training dataset, particularly in fields where large datasets aren't typically available, which is often the case when working with medical images. Data augmentation aims to generate additional data which is used to train the model and has been shown to improve performance when validated on a separate unseen dataset. This approach has become commonplace so to help understand the types of data augmentation techniques used in state-of-the-art deep learning models, we conducted a systematic review of the literature where data augmentation was utilised on medical images (limited to CT and MRI) to train a deep learning model. Articles were categorised into basic, deformable, deep learning or other data augmentation techniques. As artificial intelligence models trained using augmented data make their way into the clinic, this review aims to give an insight to these techniques and confidence in the validity of the models produced.

Key words: CT; data augmentation; deep learning; medical imaging; MRI.

Introduction

The use of deep learning (DL) within artificial intelligence research for processing and analysis of medical images has risen in popularity over recent years.[1] Convolutional Neural Networks (CNNs) are a type of deep learning approach which are able to automatically learn a set of feature detectors, usually over a number of layers (making the model 'deep'), from a labelled dataset.[2] CNNs have been utilised in medical imaging tasks including image classification (e.g. discriminating malignant from benign tumours in computer-aided diagnosis in radiology) and image segmentation (e.g. automatic delineation of anatomical structures in radiation therapy) and

have shown state-of-the-art performance.[3] Deep learning applications for these tasks can potentially assist clinicians in decision making, treatment planning and delivery, and improving operational efficiency.[4]

To train a deep learning model, the primary dataset should be split into a training and test set. The training set contains the data which is passed through the deep learning network during training over a number of iterations (called epochs) and the parameters of the network are optimised in an attempt to improve the desired result in each epoch. Once training is complete, the test dataset is used to assess the final model's performance. Deep learning methods generally require large amounts of data to train a model to prevent overfitting which is a
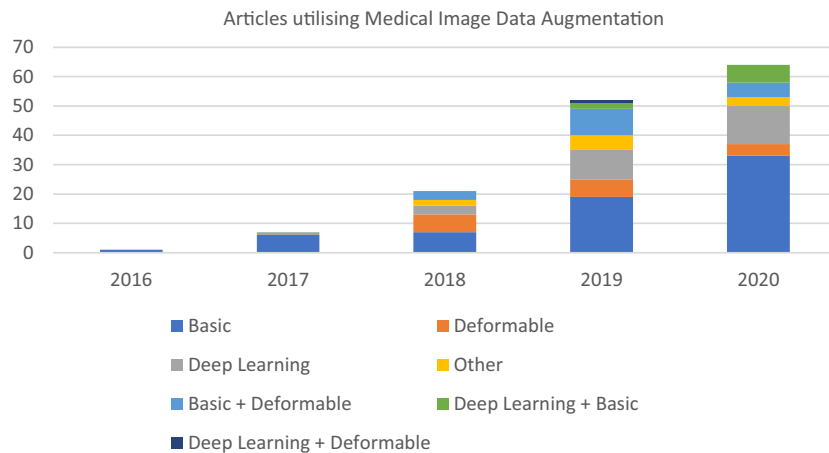
Fig. 1. Articles included in review by year and category (2016–2020).

common concern when the model is fitted to a limited training set which results in a model that does not generalise well to new testing data. However, applying deep learning in medical image analysis is often challenged by limited training data, since it can be expensive and time-consuming to acquire well-annotated medical data. To address this problem, data augmentation is often applied to increase the size and diversity of the training set in deep learning, which can be regarded as a form of regularisation technique to reduce the model's generalisation error.[5] Data augmentation uses images in the training set and applies modifications to these cases to generate further representative samples which simulate changes in acquisition and anatomical variation of patients. The additional data augmented should help the model avoid learning features which are too specific to the original training data and therefore make the model more generalisable and ultimately improve performance on the test set. Another common concern when training such a model, is class imbalance. This refers to one or more classes to be predicted being under-represented in the dataset, which can lead to the model having a bias towards the over-represented class. Data augmentation is commonly used to overcome this issue by augmenting additional data from the under-represented class.

There have been several review articles previously published on the topic of data augmentation.[6–10] However, these articles mainly focus on generative adversarial networks (GANs) based augmentation,[9,10] a specific task of either image classification or segmentation,[6,7] or general data augmentation for computer vision.[8] To the best of our knowledge, this work is the first to provide a comprehensive overview of contemporary data augmentation methods for common radiological tasks including both medical image classification and segmentation on CT and MRI which are two major imaging modalities in radiation oncology. This review presents augmentation methods based on basic image manipulation, deformable

techniques and DL-based approaches not limited to GAN networks, followed by a discussion of the characteristics and impact of these methods with considerations for future work.

Methods

To provide a thorough overview of data augmentation techniques used in medical imaging deep learning applications, a systematic review was conducted. The area of deep learning is quickly evolving, so to ensure relevance of the articles the concept of a rapid review was adopted.[11] Appropriate constraints were placed on the literature search and article inclusion criteria to provide a snapshot of data augmentation techniques that are currently used in radiology and radiation therapy applications of deep learning.

The constrained search terms used the combination of the keywords: 'data augmentation' and 'deep learning' as well as one or both of the keywords 'CT' or 'MRI'. The search was conducted on Scopus, Medline/Pubmed and ScienceDirect. Only articles published in 2016 or later were included. Figure 1 illustrates the growth in utilisation of data augmentation throughout this period. Next, the model training process described in each article was examined to determine what data augmentation was applied, if any. Articles which didn't ultimately utilise data augmentation or describe data augmentation techniques used in detail were excluded. Additionally, any work which was not applied to a deep learning application did not apply data augmentation to CT or MRI images or imaged non-human patients were also excluded. Each remaining article was categorised into one or more of the following categories: basic augmentation techniques, deformable augmentation techniques, deep learning augmentation techniques and other augmentation techniques. Each of these techniques is described in the following sections of this article. Due to the basic augmentation techniques being so commonly

used, and often applied in a very simplistic way, a further constraint was applied which excluded references describing only one type of basic augmentation. There were 395 different studies identified (after duplicates were removed) and this was reduced to 149 after all constraints were applied. Since articles were assigned to more than one category, this finally resulted in 93, 35, 38 and 10 studies in the basic, deformable, deep learning and other categories, respectively (Fig. 2).

Examples of data augmentation are provided where possible and include data from the Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI),[12] which consists of lung cancer diagnostic and screening CT scans along with annotations of nodules. This dataset is publicly available on The Cancer Imaging Archive. A second dataset is also used to demonstrate the performance of data augmentation and is derived from a clinical study for prostate cancer radiotherapy treatment,[14] which contains T2-weighted MR images from 39 patients. This dataset provides semantic ground truth labels for anatomical structures of the male pelvis including body, bone, urinary bladder, rectum and prostate.
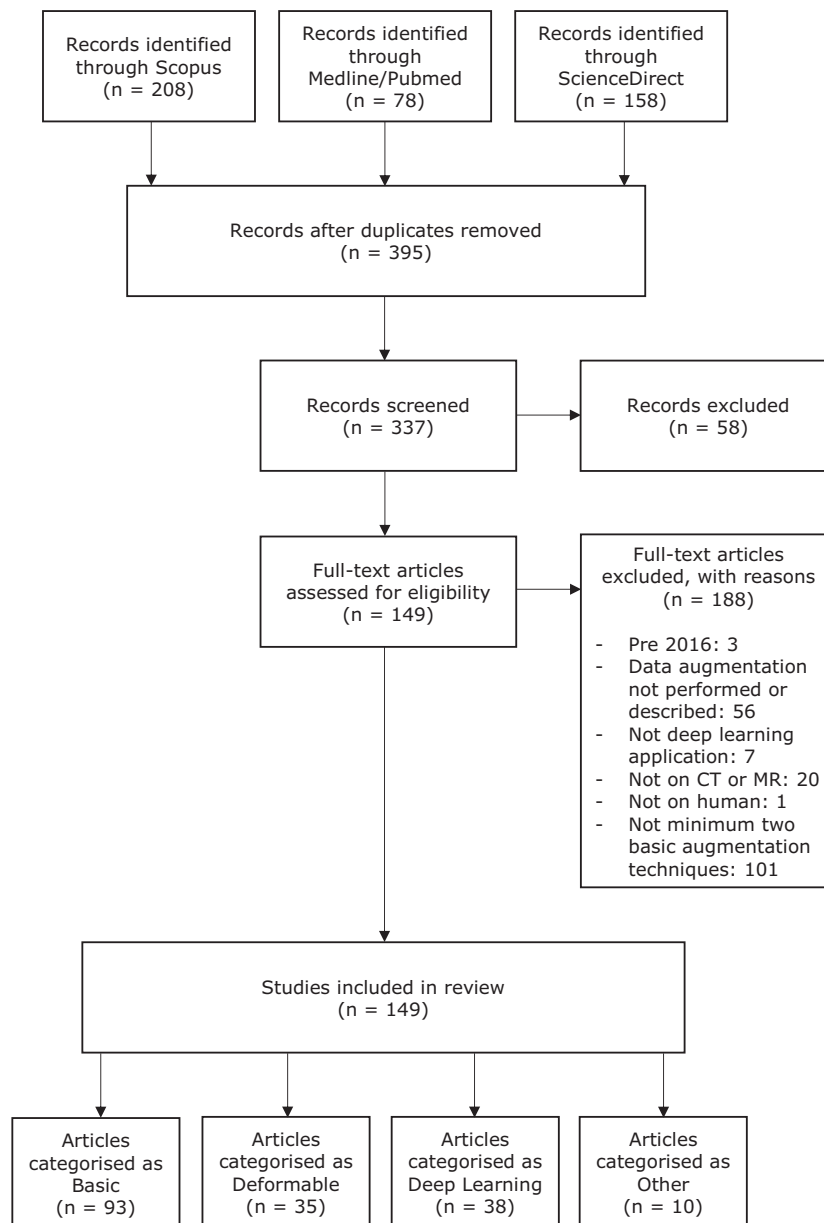
Fig. 2. Prisma chart describing article inclusion/exclusion process of the systematic review conducted.

Basic augmentation techniques

Basic augmentation techniques include those which apply a transformation to an image which maps the points of the image to a different position, or manipulates the image intensity values, to produce an augmented image. The operation will be performed on one image from the existing dataset and once modified it is added back into the data pool to increase the overall size of the dataset. These techniques are simple to apply but can be effective in improving the trained model's performance. Of the 149 articles included in this review, 93 of these applied a basic augmentation approach. This section provides an overview of these techniques and how they are typically applied. Figure 3 shows some examples of many of these techniques applied to one original image (Fig. 3a) and are denoted throughout this section.

Geometric transformations

Geometric transformations are the most common basic augmentation used. These transformations are most commonly affine such as scaling (3b), translating (3c), rotating (3d), reflecting (3e) and shearing (3f) or sometimes perspective transforms including skew (3g). The parameters of the transformation may be predefined or can be randomly sampled. If an image has one or more contours associated with it, the same transformation is applied to the contours. Geometric transformations are so common that they were utilised by 92 of the 93 basic augmentation studies.[15–106]

Cropping

Cropping is a technique where patches are randomly selected from an existing image (3h). These random patches are then added back into the dataset to increase the size. This technique is typically used when there is a class imbalance. Patches are generated from the underrepresented class to even the balance. 30 articles made use of cropping.[15–17,19,23,32,33,35,41,43,48,50,58,59,66,68,69,71,74,78,80,82,84,85,90,97,98,102,104,105]

Occlusion

Occlusion is a similar concept, except patches of an image are removed from an image to generate an augmented image (3i) and was used by Kompanek *et al*.[40]

Intensity operations

Intensity operations manipulate the values of the pixels/voxels within the image. This is often done by modifying the brightness or contrast of the image. Gamma correction (3j),[107] linear contrast (3k) and histogram equalisation (3l) are common methods to adjust the contrast of an image. Twenty-eight studies utilised intensity operations for data augmentation.[18,20,21,28,34,38,40–42,45,46,55,62,67,70,75–77,79,84,87,88,94,102–104,106,108]

Noise injection

Noise injection is another popular data augmentation technique which aims to simulate noisy images. Injection of Gaussian noise is most common where image intensities are modified by randomly sampling a Gaussian distribution (3m). Uniform noise, which modifies values by randomly sampling a uniform distribution (3n) or salt and pepper noise, where pixels are randomly set to black or white (3o) have also be used. 25 studies injected noise for data augmentation.[24,25,27,30,31,36,39,51,55,57,67,73,75,81,83,87,88,91,93,99,100,103–106]

Filtering

Filtering an image is done by using convolution. To achieve this a convolutional kernel is moved across the image to modify the intensities at each pixel based on the values of the surrounding pixels. Using this method, an image can be sharpened (3p), blurred (3q) or smoothed and thereby produce an augmented image. Some kernels are able to detect and intensify the edges of objects found within the image and have also been used for augmentation (3r). 16 studies reported use of filtering for data augmentation.[20,22,24,33,41,49,57,63,72,91,96,98,99,104,106,108]

Combination

Combination is a data augmentation method that generates a new image by combining 2 or more original images. Two studies made use of the 'mixup' technique, which works by randomly selecting two images from the dataset and blending the intensities of the corresponding voxels of the two images.[60,89,109] Nishio *et al*. also utilised Random Image Cropping And Patching (RICAP), where, rather than blending the images, four images are randomly selected, cropped and then patched together.[89,110]

Unlike other augmentation techniques, basic augmentation methods do not aim to produce 'realistic' images. Some geometric transformations of a medical image could be considered realistic such as scale or translation (simulating a differently sized or positioned patient), as well as some of the image intensity operations such as noise injection (simulating a noisy image). However, many other basic augmentation techniques are not aiming to generate a realistic image, but instead are purely intended to encourage the model to learn more general features. This is the case for image operations such as occlusion and mixup. When a contour is attached to an image, it has the same geometric transformation applied to it so the resulting contour will still delineate the same pixel data. For the other basic transformation
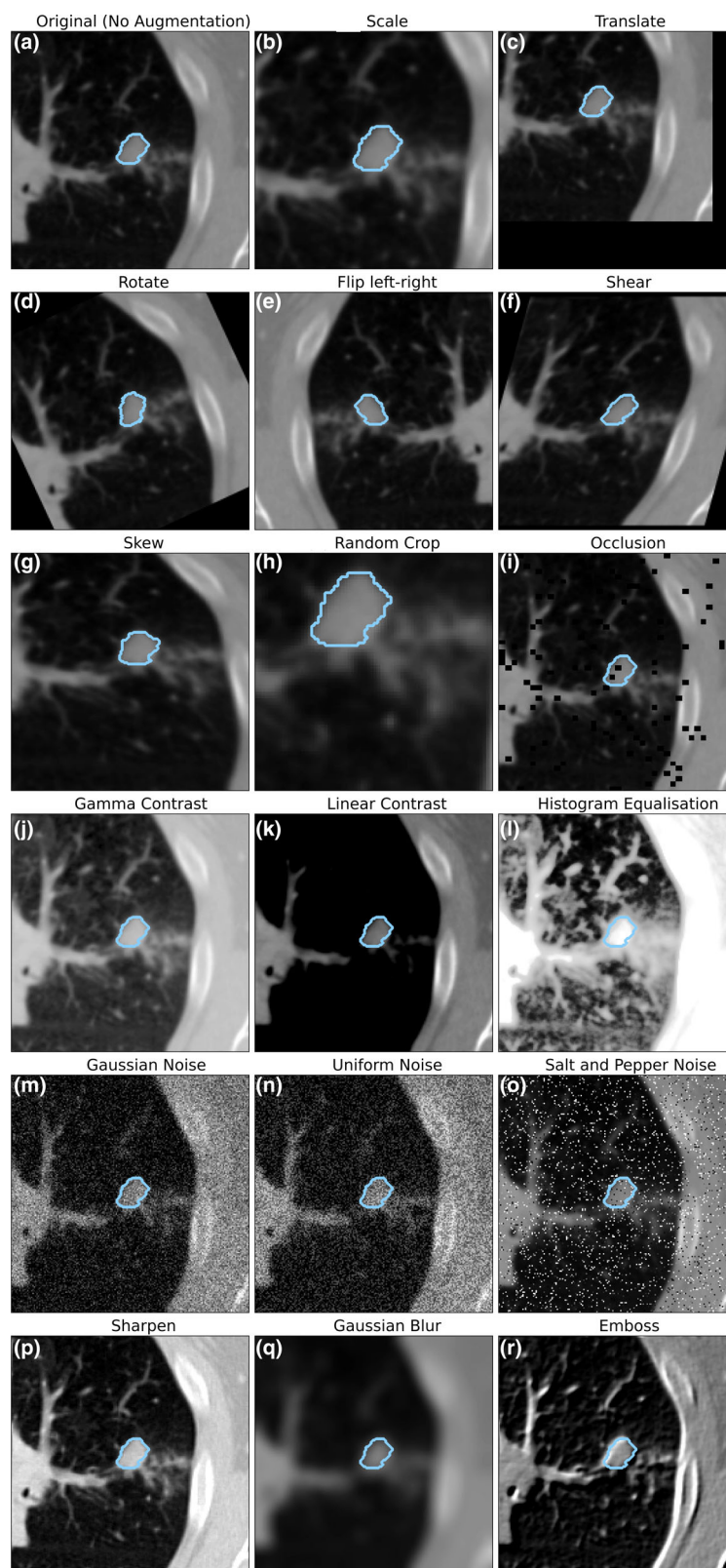
Fig. 3. Examples of commonly used basic data augmentation techniques with the corresponding contour overlayed (blue).

techniques, only the pixel data is modified, not the contours.

Deformable augmentation techniques

Deformable augmentation techniques may be used when basic augmentation techniques do not provide sufficient variability to make the subsequent model generalisable. The scale of deformation is typically constrained within user-defined limits to ensure the resulting augmentations are clinically plausible. Applying such operations can also assist in simulating a range of plausible variations that could be encountered in clinical scans, such as tissue deformations, or image artefacts created by patient movement. Of the 149 articles included in our review, 35 were found to include some form of image deformation for the purpose of augmentation. While some articles do not detail their method beyond stating that elastic deformations were used,[29,43,52,57,81,111] the majority of these publications can be categorised in one of four approaches, as outlined in this section.

Randomised displacement field

A landmark development in deformable augmentation techniques was the publication of Simard *et al*.'s[112] method, which describes the use of a randomly generated displacement field that may be applied to a 2D image to create variations of geometric shape. In this method, each pixel is shifted by a random value in the x and y directions within the range $(-1, 1)$, and the displacement fields Δx and Δy are convolved with a Gaussian. The standard deviation σ of the Gaussian influences how smoothly the deformation occurs, with intermediate σ values leading to a smooth elastic deformation (see Fig. 4b).

Of the articles reviewed, several have employed this method.[37,47,54,64,75,106,113–117]

Spline interpolation

Spline interpolation is a mathematical operation which uses a piecewise polynomial function in order to estimate new values between existing data points, often with superior results to, for example, linear interpolation. When applied to deformable image augmentation, this allows for a means of calculating a smooth deformed image to generate new image data. The two most commonly used splines in the literature reviewed were B-splines (demonstrated in Fig. 4c) and thin plate splines, with 9 articles employing such methods.[23,39,44,45,92,95,118–120]

Deformable image registration

A common use of deformation techniques is in the process of image registration, where one image (referred to as the moving image) is altered in order to more closely match another (referred to as the fixed image). This is typically used for comparing different imaging modalities for a single patient (e.g. CT and MRI); however, it has also been employed as a data augmentation technique (see Fig 4e/f).
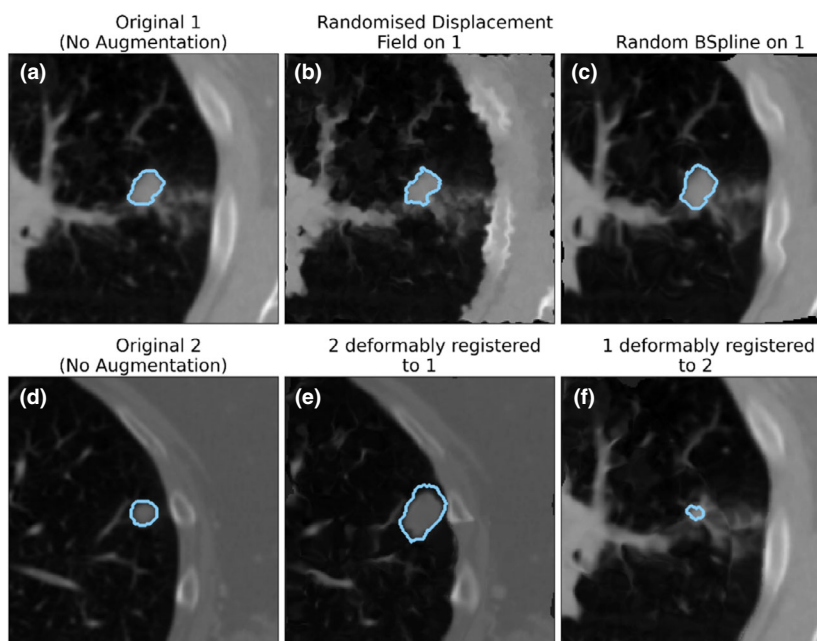
Fig. 4. Examples of augmented images using deformable techniques with the corresponding contour overlayed (blue).

Krivov et al.[26] developed a method wherein brain lesions could be mapped onto healthy patient scans, by performing deformable image registration between the original patient and a healthy one. Nalepa et al.[121] demonstrated an approach using diffeomorphic image registration, co-registering pairs of lesion images to create augmented data which improved the generalisation of their DL models when combined with an affine augmentation.

Tustison et al.[122] used a template (or atlas) based approach to generate augmented data. By using a deformable registration of each image to the template, the images were deformed to each other to produce the new augmented images.

Statistical shape models

Statistical shape models utilise data that is combined to generate a mathematical model that describes the shape variability across the dataset. This model can then be used to generate deformations within the range of plausible parameters. This approach has demonstrated an improvement in segmentation performance.[123–127]

Other deformable augmentation techniques

Javaid et al.[37] proposed a methodology for CT segmentation that aimed to simulate intra- and interobserver variability. In addition to basic and elastic deformations, they augmented the contours made on the training data, rather than the images themselves. This process generated randomised contour augmentations within a specified range and applied a smoothing function. Of the models tested (no augmentation, data augmentation only, contour augmentation only, data and contour augmentation combined), the combination of contour and data augmentation was the strongest performer.

Lindner et al.[128] presented an approach for inserting synthetic lesions into brain MRI data, by randomly displacing vertices of a polyhedron through several iterations. The brain MRI was then deformed around the synthetic lesion to simulate the presence of a tumour, including its effect on the surrounding tissue.

Deep learning augmentation techniques

The DL-based augmentation approaches can automatically learn the representations of images and generate realistic images to increase the model's generalisability and reduce overfitting during training. This process of generating synthetic images is often referred as image synthesis. Thirty-eight articles were identified that describe DL-based augmentation techniques as listed in Table 1. Thirty-three of these articles proposed methods to assist image classification and segmentation which are two of the most common tasks for medical image analysis. In this survey, we include both the articles

using DL augmentation as a training data enrichment step for these tasks and the ones that only proposed an image synthesis method with the potential of being used for image augmentation. The most commonly reported DL networks for data augmentation are the generative adversarial network (GAN) and its variants, which were adopted in 29 articles as shown in Table 1. The remaining articles proposed different adversarial learning and generative techniques from GANs, which will be described separately.

GAN-based augmentation methods

GAN network. The generative adversarial networks (GAN)[129] are a class of generative models that can generate plausible images via the adversarial learning of a generator G and a discriminator D. The structure of a basic GAN is shown in Figure 5(a). The generator G is typically a deep network which learns to map a fixed random distribution $p(z)$ to the distribution of the data of interest $p_r(x)$. The input z to G is a sample from the fixed prior distribution $p(z)$. The discriminator D is a network that learns to discriminate the output of the generator x_g from the real data sample x_r.[10] Adversarial learning/training refers to the competition between these two networks, which drives them to improve their performance during training. Ideally, an equilibrium state is eventually reached, where the generator can approximate data from the target data distribution and the discriminator predicts 'real' or 'generated' for its input data with 50% probability.[129,130] Realistic synthetic data can be generated by the generator via sampling the fixed distribution $p(z)$ for data augmentation. Figure 5(b) shows several examples of real and GAN-generated CT lung nodule images.

GAN-based augmentation for image classification. Nineteen articles were identified that used GAN networks to augment training data for image classification. The main challenges for GAN-based image synthesis typically are maintaining training stability and generating high quality (clear, high-resolution) images. To overcome these challenges, many previously proposed GAN variants, with improved network architecture and mathematical optimisation, have been adopted in Ref.53,87,91,131–140 as listed in Table 1. There is another subtype of GAN networks, often referred as conditional GAN (CGAN), which allows conditional image generation from the generator and outputs images with desired properties.[10,141] CycleGAN and Pix2Pix GAN are two CGAN networks proposed for image-to-image translation, which modify an input image into a new synthetic image satisfying a certain condition. These GAN networks can be used to create new data through domain adaptation which means the trained model can embed

Table 1. Publications that used DL-based methods for image synthesis only, data augmentation for classification (Classif.), segmentation (Segm.) and other downstream tasks, and the imaging modality they were applied on.

Ref.	Type	Network	Modality	Image Synth.	For Classif.	For Segm.	For others
177		GAN[129]	MR	Y	Y		
131	GAN-based	Deep Convolutional GAN (DCGAN)[179]	CT	Y			
132			CT	Y	Y		
133			MR	Y	Y		
87			Radiograph	Y	Y		
176			MR	Y	Y		
91			CT	Y	Y		
135		GAN with Laplacian Pyramid (LAPGAN)[180]	MR	Y			
136		WGAN/WGAN-GP[147,149]	MR	Y	Y		
137			CT	Y	Y		
138		a-GAN[181]	MR	Y			
53		Self-attention GAN (SA-GAN)[182]	CT	Y	Y		
139		Multi-scale gradient GAN (MSG-GAN)[183]	MR	Y	Y		
140		StyleGAN[184]	CT, MR	Y			
134		Conditional GAN (CGAN)[141]	MR	Y	Y		
83			CT	Y	Y		
185		Mask-guided CGAN[186]	CT	Y	Y		
86		CycleGAN[187]	CT	Y	Y		
119			CT	Y		Y	
143		Semi-supervised attention guided CycleGAN	MR	Y	Y		
152		Pix2Pix GAN[150]	MR	Y		Y	
101		Combination of GANs	CT	Y	Y		
144		A pairwise GAN	MR	Y	Y		
145		StitchLayer + GAN	MR	Y	Y		
146		Capsule network + GAN	MR	Y	Y		
153		Two branch GAN for multimodal image synthesis	CT	Y		Y	
154		Cross modality GAN	CT, MR	Y		Y	
155		Sketching-rendering unconditional GAN (SkrGAN)	CT, MR, x-ray, RCF	Y		Y	
156		GAN modelling deformation and intensity transformations	MR	Y		Y	
158	Others	Variational auto-encoder[159]	MR, ultrasound	Y	Y	Y	
157		Encoder-decoder unet	MR	Y	Y	Y	
100		Flow-based[160] generative model	CT	Y	Y		
161		Adversarial learning	CT	Y	Y		
162			MR	Y		Y	
163			CT, MR	Y	Y	Y	
56		Image registration	MR			Y	
164		Deep reinforcement learning[188]	CT			Y	
165		Variational dropout[166]	MR				Y

RCF, retinal colour fundus.

an adaptation/shift from a certain source of data onto a different source of data.[142] An example is Muramatsu et al.[86] which adopted CycleGAN to synthesise mammographic masses from lung nodules for a breast mass classification task.

Instead of adopting GAN networks proposed in previous studies, a few articles[101,143–146] designed their own novel augmentation methods based on GAN for classification tasks. Xu et al.[143] proposed a semi-supervised attention guided CycleGAN to augment MRI images for brain tumour classification, which can generate tumour images from normal images and the other way around. Wang et al.[101] demonstrated that using the combination of WGAN[147] and progressive growing[148] achieved

better pulmonary adenocarcinoma classification result compared with WGAN with gradient penalty (WGAN-GP)[149] and Pix2Pix GAN.[150] Ge et al.[144] proposed a pairwise GAN to generate multimodal synthetic MRI images, which enables the enlargement of a multimodal training set for brain tumour classification. Wang et al.[145] proposed the introduction of a stitch layer into the generator in GAN (StitchGAN) to stitch the low-dimensional images into a full size image to alleviate the complexity of full size image mapping, which synthesises apparent diffusion coefficient (ADC)-T2-weighted image pairs to enlarge the training set for clinically significant prostate cancer classification. Yu et al.[146] proposed the use of a Capsule network[151] to
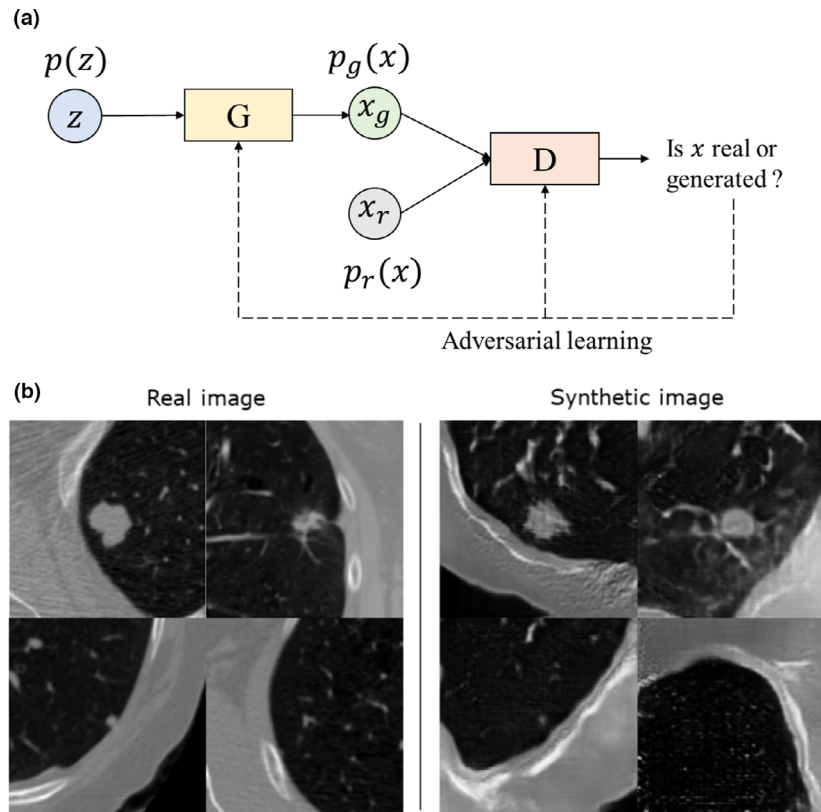
Fig. 5. (a) The diagram of a basic GAN. (b) Real CT images from the LIDC lung nodule dataset[12] and synthetic images generated by a GAN network.

replace the convolutional neural network as the discriminator in GAN to better model the hierarchical relationships of parts of the image.

GAN-based augmentation for image segmentation.

The goal of data augmentation for image segmentation tasks is to generate synthetic image-label pairs rather than synthetic images only as for classification tasks. Here, a label is normally an image with its pixels/voxels assigned with a category index representing a certain meaningful object. Six articles[119,152–156] were identified that proposed GAN-based augmentation for medical image segmentation tasks. These proposed methods are mostly based on the concept of domain adaptation mentioned in the last section. An example is to use semantic labels of anatomical structures to generate artificial medical images. Here, a semantic label is a label image with every pixel/voxel assigned with a category index. Figure 6 shows an example of using semantic labels to generate synthetic MRI images using a CGAN. The generated images can be paired with the labels and join the training set for a segmentation network. This type of label image translation technique was adopted by Cao el al.[153] and Shin et al.[152] to synthesise PET and CT images, and abnormal brain MRI images respectively for

tumour segmentation. Augmented labels where tumours' location, shape and size can be changed were used as the input to generate synthetic images with various anatomical appearance. This cross-domain translation can also happen between different imaging modalities. Sandfort et al.[119] used a CycleGAN to translate contrast-enhanced CT images into non-contrast-enhanced ones for organ segmentation, and Jiang et al.[154] performed pseudo MR image generation from CT images which is constrained by two GANs trained in parallel for lung tumour segmentation in MRI. In these articles, the structure correspondence can be maintained between the synthetic images and the labels drawn on the original source of images.

Zhang et al.[155] proposed a sketching-rendering unconditional GAN (SkrGAN), which can not only augment grayscale images, such as X-Ray, CT and MRI, but also colour images such as retinal colour fundus photography for various segmentation tasks. It can generate structural sketch from random noise and use intensity/colour render mapping to fill the sketch representations to resemble the real image. Chaitanya et al.[156] proposed a unique augmentation method which uses GAN to map deformation and intensity transformations for image-label pair augmentation.
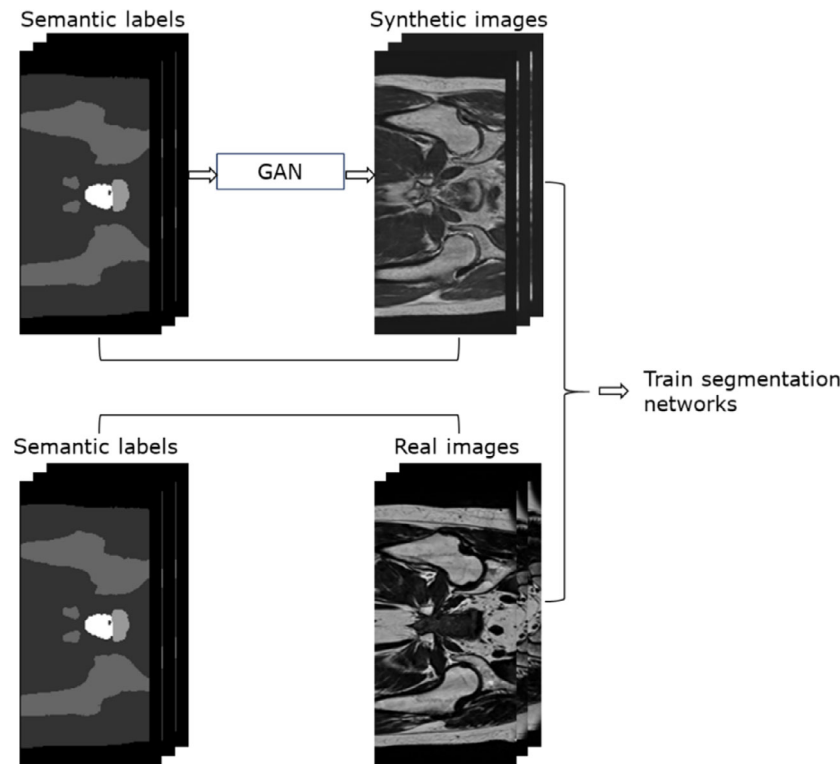
Fig. 6. An example of generating synthetic MR images from semantic labels of body, bone, bladder, rectum and prostate using a CGAN proposed in Ref.178 The image is from a prostate MRI dataset derived from Dowling *et al*.[14]

Other DL-based augmentation methods

To augment data for brain lesion detection and segmentation, Salem *et al*.[157] proposed to add brain lesions into normal brain MRI T1 weighted and FLAIR images via encoding the lesion information with a stack of discrete binary intensity masks. Pesteie *et al*.[158] proposed a generative model based on variational auto-encoder[159] for augmentation in both image classification and segmentation tasks. Uemura *et al*.[100] proposed using a different distribution mapping technique from GAN, the flow-based generative model,[160] for colorectal polyp synthesis in CT colonography, which outperformed GAN-based augmentation in polyp classification. Yang *et al*.,[161] Dorent *et al*.[162] and Anand *et al*.[163] proposed augmentation methods based on the concept of adversarial learning. Among these articles, Yang *et al*.[161] proposed an interesting class-aware adversarial network which can condition the types (malignant, benign) of image generated. This approach can synthesise both malignant and benign lung nodules to alleviate class imbalance and improve malignant nodule classification performance. Instead of using generative model to produce synthetic images, Qin *et al*.[164] and Xu *et al*.[56] proposed to augment data though embedding spatial transformations in a DL network which is combined with a segmentation network in an end-to-end fashion. Apart from the common image classification

and segmentation tasks, Tanno *et al*.[165] adopted variational dropout[166] as a form of data augmentation for uncertainty modelling in DL for image enhancement.

Other augmentation techniques

While the three categories of data augmentation techniques described above encompass the vast majority of techniques commonly used in the literature, there are some lesser utilised approaches which are worth noting. One group of these techniques make use of specific properties of an imaging modality. For MRI, augmenting images with motion artefacts is possible by manipulating the k-space. When the image is then reconstructed from this space motion artefacts will be visible, which is useful for training a model to detect or correct these artefacts or to train a model which is more robust under these conditions.[167–169] Liu *et al*.[170] trained a deep learning model for MRI reconstruction and were able to augment their data by undersampling the k-space. A similar approach can be applied to images captured using a spectral (dual energy) CT scanner since two separate energy spectra are captured, these can be combined in different ways to produce different augmentations.[171,172] Omigbodun *et al*.[173] also augmented images simulating different parameters available on a CT scanner, such as slice thickness or dose.

Lucena et al.[174] solved their problem of not having sufficient 'gold standard' manual annotations to train a deep learning model for skull stripping in brain MRI by generating so-called 'silver standard' annotations by combining the results of eight non-deep learning brain extraction methods. The authors claim that this produced a model which is more generalisable than those methods used to generate the annotations, and also has significant run-time advantages. Momeni et al.[175] generated synthetic cerebral microbleeds from MRI by simulating shape variation using Gaussian functions.

Discussion

The basic augmentation techniques described are generally simple to apply and relatively fast to compute. These techniques can be applied to 2D slices, or the entire 3D volume depending on the architecture of the deep learning model being trained. Typically, basic augmentations are added only to the training set and with the availability of fast computation speeds, are often applied using randomly sampled parameters on-the-fly (during training of the model). This can help generalise the model even further than simply increasing the size of the training set with augmented data before training. Many authors accept data augmentation as a necessity and apply basic augmentation without reporting in their work the effect this might have had on the training of the model. There are works such as that from Zhang et al.[106] which compare the results of their model with different augmentation techniques applied during training. For the task of segmenting the prostate from MRI, when using data augmentation their trained model performance improved by up to 12–47% when assessed on different test sets. Due to the simplicity and ease of use of these basic augmentation techniques, it is to be expected that researchers will continue to apply these to help generalise their deep learning models.

Deformable techniques allow for complex forms of data augmentation. They have an advantage over basic augmentation techniques for their ability to imitate types of variation that can be encountered in clinical environments, such as distortion from patient movement, or interobserver variability in organ and tumour contouring. They also contribute to the generalisability of deep learning models. It is important when implementing deformable augmentation techniques to set parameters for deformation that remain within physically plausible limits. Since deformable techniques can create greater variation in training data than basic techniques, but are computationally less expensive than deep learning techniques, it is expected that their use will gradually become more commonplace in model design.

Compared with basic geometric and deformable augmentation techniques which may generate highly correlated image data, the DL-based augmentation opens up more possibilities of producing new synthetic data without predefined augmentation rules. The DL-based methods can augment image data across different imaging modalities and sequences, and with anatomical variance (e.g. change of size, shape and location of tumours),[86,119,144,145,152–154,157] which conventional augmentation methods generally do not account for. Sometimes, geometric, deformable and intensity-based augmentation can also be applied to the data used to train the DL-based augmentation networks[56,91,101, 138,146,152,154] or used alongside the DL-based methods.[53,83,86,87,100,119] The majority of DL-based augmentation approaches are based on adversarial training (including GAN-based and other adversarial learning networks), where a discriminator network is usually used to review the generated images, and the training process iteratively bridges the gap between generated and real images. However, the uncertainty of whether the generated synthetic images can represent realistic radiological features in medical images is still a main concern for DL-based augmentation since these artificial variations are not from a real-world clinical environment. One way to validate the quality of the synthetic images is through downstream tasks such as classification and segmentation.[10] Many articles demonstrated the effectiveness of the DL-based augmentation by showing a higher performance improvement on real testing data for classification or segmentation tasks compared with no augmentation, basic geometric or deformable augmentation approaches.[56,86,91,100,101,119,132–134,136,143–146,152, 154–158,164,176,177] Another way of ensuring the realism of the synthetic data is to have the generated data reviewed by clinicians[86,101,119,133,135,136,144–146]; however, this approach could be resource intensive, time consuming and difficult to scale.[10] DL-based augmentation can also be challenged by its high computational cost. The majority (27 out of 38) of the proposed augmentation methods were developed on 2D images, even when the original data are in 3D imaging modalities. Currently, the research on DL-based augmentation is still at an early stage. However, it is rapidly evolving and has shown encouraging results in many studies in comparison to traditional augmentation methods.

Table 2. Common software libraries and frameworks for deep learning image data augmentation.

Ref.	Name	Geometric	Intensity	Deformable
189	ImgAug	Y	Y	Y
190	Augmentor	Y	N	Y
191	Keras ImageDataGenerator	Y	Y	N
192	PyTorch Transforms	Y	Y	N
193	TorchIO	Y	Y	Y
194	Albumentations	Y	Y	Y
195	CLoDSA	Y	Y	Y
196	MONAI	Y	Y	Y

Basic and deformable augmentation methods are generally easy to apply on medical data (2D, 3D, single or multimodal data) and the process is relatively transparent. There are many software libraries available, as listed in Table 2, providing basic and deformable data augmentation. While DL-based methods can cross domains to generate new data (e.g. synthesise mammographic masses from lung nodules,[86] cross-modality image synthesis[154]) and provide more variability, they can be challenging to train and difficult to interpret from the perspective of humans. Currently there is no out-of-the-box library for DL-based augmentation methods. It is difficult to say which type of methods is the best strategy for data augmentation, since the design of the augmentation stage is highly dependent on the given data and tasks. In fact, a hybrid of these types of methods is often used in practice.[53,83,86,87,100,119] For future work, we believe basic and deformable augmentation will continue to be regularly used for training DL networks. New developments on imitating a wider range of clinical variances can also be expected for these two types of methods. The DL-based augmentation is evolving rapidly and will likely gain even more popularity given its ability to tackle complex data synthesis problems such as cross-domain image synthesis. Reducing computational complexity to handle higher dimensional medical data could be one future aim for DL-based methods. Moreover, more sophisticated approaches for automatically validating the realism of synthetic data generated by DL-based augmentation may also be a research area for future work. We also expect to see more research on augmenting temporal medical data such as 4D-CT and dynamic contrast-enhanced MRI.

Conclusion

Data augmentation of medical images is becoming commonplace for deep learning applications. A suite of basic techniques (e.g. scaling, rotation and flipping of images) are widely used in model design, and as more advanced techniques, such as deformable augmentation (e.g. randomised displacement fields) and deep learning-based approaches (e.g. GAN) are refined, they are becoming more widely utilised in model design. While realism can be a goal for augmented data, it is not always a necessity as unrealistic augmentations may still produce a more generalisable model. Data augmentation is particularly advantageous in scenarios where there is insufficient training data available, to help correct overfitting of a model. The key factor in assessing the performance of a deep learning model is its results on a test set of real unseen data. If this test set represents a population sufficiently, then confidence can be had in the assessment of the model performance and the training process of the model, including any data augmentation that may have been performed.

References

1. Cai L, Gao J, Zhao D. A review of the application of deep learning in medical image classification and segmentation. *Ann Transl Med* 2020; **8**: 713.
2. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imag* 2018; **9**: 611–29.
3. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng* 2017; **19**: 221–48.
4. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Z Med Phys* 2019; **29**: 102–27.
5. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. The MIT Press, Cambridge, MA, 2016.
6. Mikołajczyk A, Grochowski M, editors. Data augmentation for improving deep learning in image classification problem. 2018 international interdisciplinary PhD workshop (IIPhDW), 2018: IEEE.
7. Nalepa J, Marcinkiewicz M, Kawulok M. Data augmentation for brain-tumor segmentation: a review. *Front Comput Neurosci* 2019; **13**. https://doi.org/10.3389/fncom.2019.00083
8. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data* 2019; **6**: 1–48.
9. Sorin V, Barash Y, Konen E, Klang EJAR. Creating artificial images for radiology applications using generative adversarial networks (GANs)–a systematic review. *Acad Radiol* 2020; **27**: 1175–85.
10. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med Image Anal* 2019; **58**: 101552.
11. Featherstone RM, Dryden DM, Foisy M *et al*. Advancing knowledge of rapid reviews: an analysis of results, conclusions and recommendations from published review articles examining rapid reviews. *Syst Rev* 2015; **4**: 50.
12. Armato SG, McLennan G, Bidaut L *et al*. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. *Med Phys* 2011; **38**: 915–31.
13. Armato SG III, McLennan G, Bidaut L *et al*. Data from LIDC-IDRI. *TCIA* 2015. https://doi.org/10.7937/K9/TCIA.2015.LO9QL9SX

14. Dowling JA, Sun J, Pichler P *et al*. Automatic substitute computed tomography generation and contouring for magnetic resonance imaging (MRI)-alone external beam radiation therapy from standard MRI sequences. *Int J Radiat Oncol Biol Phys* 2015; **93**: 1144–53.

15. Pereira S, Pinto A, Alves V, Silva CA. Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Trans Med Imaging* 2016; **35**: 1240–51.

16. Chen L, Bentley P, Rueckert D. Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks. *Neuroimage Clin* 2017; **15**: 633–43.

17. Huang X, Shan J, Vaidya V. Lung nodule detection in CT using 3D convolutional neural networks. In: *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017; 379–83.

18. Jamaludin A, Kadir T, Zisserman A. SpineNet: automated classification and evidence visualization in spinal MRIs. *Med Image Anal* 2017; **41**: 63–73.

19. Stefan LD, Dicente Cid Y, Jimenez-del-Toro O, Ionescu B, Müller H. Finding and classifying tuberculosis types for a targeted treatment: MedGIFT-UPB participation in the ImageCLEF 2017 tuberculosis task. 18th Working Notes of CLEF Conference and Labs of the Evaluation Forum, CLEF 2017, 2017.

20. Sun J, Chong P, Xiang Y, Tan M, Binder A. ImageCLEF 2017: ImageCLEF tuberculosis task – The SGEast submission. *18th Working Notes of CLEF Conference and Labs of the Evaluation Forum, CLEF 2017*, 2017.

21. Wang S, Shen Y, Chen W, Xiao T, Hu J. Automatic Recognition of Mild Cognitive Impairment from MRI Images Using Expedited Convolutional Neural Networks. *Artificial Neural Networks and Machine Learning – ICANN 2017. Lecture Notes in Computer Science*, 2017; 373–80.

22. Aderghal K, Khvostikov A, Krylov A, Benois-Pineau J, Afdel K, Catheline G. Classification of Alzheimer disease on imaging modalities with deep CNNs using cross-modal transfer learning. In: *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*, 2018; 345–50.

23. Drozdzal M, Chartrand G, Vorontsov E *et al*. Learning normalized inputs for iterative estimation in medical image segmentation. *Med Image Anal* 2018; **44**: 1–13.

24. Fang T. A Novel Computer-Aided Lung Cancer Detection Method Based on Transfer Learning from GoogLeNet and Median Intensity Projections. In: *2018 IEEE International Conference on Computer and Communication Engineering Technology (CCET)*, 2018; 286–90.

25. Gsaxner C, Pfarrkirchner B, Lindner L *et al*. Exploit 18F-FDG enhanced urinary bladder in PET data for deep learning ground truth generation in CT scans. In: Gimi B, Krol A (eds). *Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging*. SPIE, Houston, TX, 2018.

26. Krivov E, Pisov M, Belyaev M. MRI augmentation via elastic registration for brain lesions segmentation.

27. McCrackin L. Early detection of Alzheimer's disease using deep learning. Advances in Artificial Intelligence. Lecture Notes in Computer Science, 2018; 355–9.

28. Milde S, Liebgott A, Wu Z *et al*. Graphical user interface for medical deep learning – application to magnetic resonance imaging. In: *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2018; 838–47.

29. Näppi JJ, Hironaka T, Yoshida H. Detection of colorectal masses in CT colonography: application of deep residual networks for differentiating masses from normal colon anatomy. Medical Imaging 2018: Computer-Aided Diagnosis, 2018.

30. Perone CS, Cohen-Adad J. Deep semi-supervised segmentation with weight-averaged consistency targets. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Lecture Notes in Computer Science, 2018; 12–9.

31. Pfarrkirchner B, Gsaxner C, Lindner L *et al*. Lower jawbone data generation for deep learning tools under MeVisLab. Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging, 2018.

32. Afzal S, Maqsood M, Nazir F *et al*. A data augmentation-based framework to handle class imbalance problem for Alzheimer's stage detection. *IEEE Access* 2019; **7**: 115528–39.

33. Brahim I, Fourer D, Vigneron V, Maaref H. Deep Learning Methods for MRI Brain Tumor Segmentation: a comparative study. In: *2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2019; 1–6.

34. Chen A, Zhu L, Zang H, Ding Z, Zhan S. Computer-aided diagnosis and decision-making system for medical data analysis: a case study on prostate MR images. *J Manag Sci Eng* 2019; **4**: 266–78.

35. de Vos BD, Berendsen FF, Viergever MA, Sokooti H, Staring M, Isgum I. A deep learning framework for unsupervised affine and deformable image registration. *Med Image Anal* 2019; **52**: 128–43.

36. Gsaxner C, Roth PM, Wallner J, Egger J. Exploit fully automatic low-level segmented PET data for training high-level deep learning algorithms for the corresponding CT data. *PLoS One* 2019; **14**: e0212550.

37. Javaid U, Dasnoy D, Lee JA, Angelini ED, Landman BA. Semantic segmentation of computed tomography for radiotherapy with deep learning: compensating insufficient annotation quality using contour augmentation. Medical Imaging 2019: Image Processing, 2019.

38. Khalili N, Lessmann N, Turk E *et al*. Automatic brain tissue segmentation in fetal MRI using convolutional neural networks. *Magn Reson Imaging* 2019; **64**: 77–89.

39. Khened M, Kollerathu VA, Krishnamurthi G. Fully convolutional multi-scale residual DenseNets for

Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Lecture Notes in Computer Science, 2018; 369–80.

cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. *Med Image Anal* 2019; **51**: 21–45.

40. Kompanek M, Tamajka M, Benesova W. Volumetric data augmentation as an effective tool in MRI classification using 3D Convolutional Neural Network. In: *2019 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2019: 115–9.

41. Lee J, Oh JE, Kim MJ, Hur BY, Sohn DK. Reducing the model variance of a rectal cancer segmentation network. *IEEE Access* 2019; **7**: 182725–33.

42. Liu Y, Zhou J, Chen S, Liu L. Muscle segmentation of L3 slice in abdomen CT images based on fully convolutional networks. In: *2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2019; 1–5.

43. Nguyen N-Q, Lee S-W. Robust boundary segmentation in medical images using a consecutive deep encoder-decoder network. *IEEE Access* 2019; **7**: 33795–808.

44. Pan H, Feng Y, Chen Q, Meyer C, Feng X. Prostate segmentation from 3D MRI using a two-stage model and variable-input based uncertainty measure. In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019; 468–71.

45. Payer C, Stern D, Bischof H, Urschler M. Integrating spatial configuration into heatmap regression based CNNs for landmark localization. *Med Image Anal* 2019; **54**: 207–19.

46. Perez Malla CU, Valdes Hernandez MDC, Rachmadi MF, Komura T. Evaluation of enhanced learning techniques for segmenting ischaemic stroke lesions in brain magnetic resonance perfusion images using a convolutional neural network scheme. *Front Neuroinform* 2019; **13**: 33.

47. Ribalta Lorenzo P, Nalepa J, Bobek-Billewicz B *et al*. Segmenting brain tumors from FLAIR MRI using fully convolutional neural networks. *Comput Meth Prog Biomed* 2019; **176**: 135–48.

48. Rundo L, Han C, Nagano Y *et al*. USE-Net: incorporating squeeze-and-excitation blocks into U-Net for prostate zonal segmentation of multi-institutional MRI datasets. *Neurocomputing* 2019; **365**: 31–43.

49. Sajjad M, Khan S, Muhammad K, Wu W, Ullah A, Baik SW. Multi-grade brain tumor classification using deep CNN with extensive data augmentation. *J Comput Sci* 2019; **30**: 174–82.

50. Song T, Huang N. Integrated extractor, generator and segmentor for ischemic stroke lesion segmentation. In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Lecture Notes in Computer Science*, 2019; 310–8.

51. Sultan HH, Salem NM, Al-Atabany W. Multi-classification of brain tumor images using deep neural network. *IEEE Access* 2019; **7**: 69215–25.

52. Teng L, Li H, Karim S. DMCNN: a deep multiscale convolutional neural network model for medical image segmentation. *J Healthc Eng* 2019; **2019**: 8597606.

53. Uemura T, Näppi JJ, Lu H *et al*. Ensemble 3D residual network (E3D-ResNet) for reduction of false-positive polyp detections in CT colonography. Medical Imaging 2019: Computer-Aided Diagnosis, 2019.

54. Wang Y, Li C, Zhu T, Zhang J. Multimodal brain tumor image segmentation using WRN-PPNet. *Comput Med Imaging Graph* 2019; **75**: 56–65.

55. Winkels M, Cohen TS. Pulmonary nodule detection in CT scans with equivariant CNNs. *Med Image Anal* 2019; **55**: 15–26.

56. Xu Z, DeepAtlas NM. DeepAtlas: Joint Semi-supervised Learning of Image Registration and Segmentation. Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Lecture Notes in Computer Science, 2019; 420–9.

57. Yang F, Zhang Y, Lei P *et al*. A deep learning segmentation approach in free-breathing real-time cardiac magnetic resonance imaging. *Biomed Res Int* 2019; **2019**: 5636423.

58. Yang Q, Chao H, Nguyen D, Jiang S. A Novel Deep Learning Framework for Standardizing the Label of OARs in CT. Artificial Intelligence in Radiation Therapy. Lecture Notes in Computer Science, 2019; 52–60.

59. Zhang Y, Yap PT, Qu L, Cheng JZ, Shen D. Dual-domain convolutional neural networks for improving structural information in 3T MRI. *Magn Reson Imaging* 2019; **64**: 90–100.

60. Zhao W, Yang J, Ni B *et al*. Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning. *Cancer Med* 2019; **8**: 3532–43.

61. Aghnia Farda N, Lai JY, Wang JC, Lee PY, Liu JW, Hsieh IH. Sanders classification of calcaneal fractures in CT images with deep learning and differential data augmentation techniques. *Injury* 2021; **52**: 616–24.

62. Bamba U, Pandey D, Lakshminarayanan V, Azar FS, Intes X, Fang Q. Classification of brain lesions from MRI images using a novel neural network. Multimodal Biomedical Imaging XV, 2020.

63. Campello VM, Martín-Isla C, Izquierdo C, Petersen SE, Ballester MAG, Lekadir K. Combining Multi-Sequence and Synthetic Images for Improved Segmentation of Late Gadolinium Enhancement Cardiac MRI. Statistical Atlases and Computational Models of the Heart Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. Lecture Notes in Computer Science, 2020; 290–9.

64. Chandrashekar A, Handa A, Shivakumar N *et al*. A deep learning pipeline to automate high-resolution arterial segmentation with or without intravenous contrast. *Ann Surg* 2020. https://doi.org/10.1097/SLA.0000000000004595.

65. Chen Y, Ruan D, Xiao J *et al*. Fully automated multiorgan segmentation in abdominal magnetic resonance imaging with deep neural networks. *Med Phys* 2020; **47**: 4971–82.

66. Elazab A, Wang C, Gardezi SJS *et al*. GP-GAN: brain tumor growth prediction using stacked 3D generative adversarial networks from longitudinal MR Images. *Neural Netw* 2020; **132**: 321–32.

67. Gong J, Liu J, Hao W *et al*. A deep residual learning network for predicting lung adenocarcinoma

manifesting as ground-glass nodule on CT images. *Eur Radiol* 2020; **30**: 1847–55.

68. Hemke R, Buckless CG, Tsao A, Wang B, Torriani M. Deep learning for automated segmentation of pelvic muscles, fat, and bone from CT studies for body composition assessment. *Skeletal Radiol* 2020; **49**: 387–95.

69. Herzog L, Murina E, Durr O, Wegener S, Sick B. Integrating uncertainty in deep neural networks for MRI based stroke analysis. *Med Image Anal* 2020; **65**: 101790.

70. Hong J, Feng Z, Wang S-H et al. Brain age prediction of children using routine brain MR images via deep learning. *Front Neurol* 2020; **11**: 584682.

71. Hua W, Xiao T, Jiang X et al. Lymph-vascular space invasion prediction in cervical cancer: exploring radiomics and deep learning multilevel features of tumor and peritumor tissue on multiparametric MRI. *Biomed Signal Process Control* 2020; **58**: 101869.

72. Irsheidat S, Duwairi R. Brain Tumor Detection Using Artificial Convolutional Neural Networks. In: *2020 11th International Conference on Information and Communication Systems (ICICS)*, 2020; 197–203.

73. Jeong YU, Yoo S, Kim YH, Shim WH. De-identification of facial features in magnetic resonance images: software development using deep learning technology. *J Med Internet Res* 2020; **22**: e22739.

74. Jing B, Deng Y, Zhang T et al. Deep learning for risk prediction in patients with nasopharyngeal carcinoma using multi-parametric MRIs. *Comput Meth Prog Biomed* 2020; **197**: 105684.

75. Karani N, Erdil E, Chaitanya K, Konukoglu E. Test-time adaptable neural networks for robust medical image segmentation. *Med Image Anal* 2021; **68**: 101907.

76. Kermani S, Ghelich Oghli M, Mohammadzadeh A, Kafieh R. NF-RCNN: heart localization and right ventricle wall motion abnormality detection in cardiac MRI. *Phys Med* 2020; **70**: 65–74.

77. Khan HA, Jue W, Mushtaq M, Mushtaq MU. Brain tumor classification in MRI image using convolutional neural network. *Math Biosci Eng* 2020; **17**: 6203–16.

78. Klages P, Benslimane I, Riyahi S et al. Patch-based generative adversarial neural network models for head and neck MR-only planning. *Med Phys* 2020; **47**: 626–42.

79. Kofler A, Dewey M, Schaeffter T, Wald C, Kolbitsch C. Spatio-Temporal deep learning-based undersampling artefact reduction for 2D radial cine MRI with limited training data. *IEEE Trans Med Imaging* 2020; **39**: 703–17.

80. Kumar R, Wang WenYong, Kumar J et al. An Integration of blockchain and AI for secure data sharing and detection of CT images for the hospitals. *Comput Med Imaging Graph* 2020; **87**: 101812.

81. LaLonde R, Xu Z, Irmakci I, Jain S, Bagci U. Capsules for biomedical image segmentation. *Med Image Anal* 2020; **68**: 101889.

82. Lin S, Han ZE, Li D et al. Integrating model- and data-driven methods for synchronous adaptive multi-band image fusion. *Inf Fusion* 2020; **54**: 145–60.

83. Loey M, Manogaran G, Khalifa NEM. A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images. *Neural Comput Appl* 2020: 1–13. https://doi.org/10.1007/s00521-020-05437-x.

84. Lu Z, Bai Y, Chen YI et al. The classification of gliomas based on a Pyramid dilated convolution resnet model. *Pattern Recogn Lett* 2020; **133**: 173–9.

85. Medina G, Buckless CG, Thomasson E, Oh LS, Torriani M. Deep learning method for segmentation of rotator cuff muscles on MR images. *Skeletal Radiol* 2021; **50**: 683–92.

86. Muramatsu C, Nishio M, Goto T et al. Improving breast mass classification by shared data with domain transformation using a generative adversarial network. *Comput Biol Med* 2020; **119**: 103698.

87. Mutasa S, Varada S, Goel A, Wong TT, Rasiej MJ. Advanced deep learning techniques applied to automated femoral neck fracture detection and classification. *J Digit Imaging* 2020; **33**: 1209–17.

88. Nayak DR, Dash R, Majhi B, Pachori RB, Zhang Y. A deep stacked random vector functional link network autoencoder for diagnosis of brain abnormalities and breast cancer. *Biomed Signal Process Control* 2020; **58**: 101860.

89. Nishio M, Noguchi S, Fujimoto K. Automatic pancreas segmentation using coarse-scaled 2D model of deep learning: usefulness of data augmentation and deep U-net. *Applied Sciences* 2020; **10**: 3360.

90. Olin AB, Hansen AE, Rasmussen JH et al. Feasibility of multiparametric positron emission tomography/magnetic resonance imaging as a one-stop shop for radiation therapy planning for patients with head and neck cancer. *Int J Radiat Oncol Biol Phys* 2020; **108**: 1329–38.

91. Onishi Y, Teramoto A, Tsujimoto M et al. Investigation of pulmonary nodule classification using multi-scale residual network enhanced with 3DGAN-synthesized volumes. *Radiol Phys Technol* 2020; **13**: 160–9.

92. Rigaud B, Anderson BM, Yu ZH et al. Automatic segmentation using deep learning to enable online dose optimization during adaptive radiation therapy of cervical cancer. *Int J Radiat Oncol Biol Phys* 2020; **109**: 1096–110.

93. Roth H, Zhu W, Yang D, Xu Z, Xu D. Cardiac Segmentation of LGE MRI with Noisy Labels. Statistical Atlases and Computational Models of the Heart Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. Lecture Notes in Computer Science, 2020; 228–36.

94. Schramm G, Rigie D, Vahle T et al. Approximating anatomically-guided PET reconstruction in image space using a convolutional neural network. *NeuroImage* 2020; **224**: 117399.

95. Shen C, Wang C, Roth HR et al. Spatial information-embedded fully convolutional networks for multi-organ segmentation with improved data augmentation and instance normalization. Medical Imaging 2020: Image Processing, 2020.

96. Shinohara Y, Takahashi N, Lee Y, Ohmura T, Kinoshita T. Development of a deep learning model to identify hyperdense MCA sign in patients with acute ischemic stroke. *Jpn J Radiol* 2020; **38**: 112–7.

97. Song TA, Chowdhury SR, Yang F, Dutta J. PET image super-resolution using generative adversarial networks. *Neural Netw* 2020; **125**: 83–91.

98. Tam CM, Zhang D, Chen B, Peters T, Li S. Holistic multitask regression network for multiapplication shape regression segmentation. *Med Image Anal* 2020; **65**: 101783.

99. Tekchandani H, Verma S, Londhe ND. Mediastinal lymph node malignancy detection in computed tomography images using fully convolutional network. *Biocybern Biomed Eng* 2020; **40**: 187–99.

100. Uemura T, Nappi JJ, Ryu Y, Watari C, Kamiya T, Yoshida H. A generative flow-based model for volumetric data augmentation in 3D deep learning for computed tomographic colonography. *Int J Comput Assist Radiol Surg* 2021; **16**: 81–9.

101. Wang Y, Zhou L, Wang M *et al*. Combination of generative adversarial network and convolutional neural network for automatic subcentimeter pulmonary adenocarcinoma classification. *Quant Imaging Med Surg* 2020; **10**: 1249–64.

102. Wodzinski M, Banzato T, Atzori M, Andrearczyk V, Cid YD, Muller H. Training deep neural networks for small and highly heterogeneous MRI datasets for cancer grading. *Annu Int Conf IEEE Eng Med Biol Soc* 2020; **2020**: 1758–61.

103. Yasar H, Ceylan M. A novel comparative study for detection of Covid-19 on CT lung images using texture analysis, machine learning, and deep learning methods. *Multimed Tools Appl* 2021; **80**: 5423–47.

104. Zhao X, Xie P, Wang M *et al*. Deep learning-based fully automated detection and segmentation of lymph nodes on multiparametric-mri for rectal cancer: a multicentre study. *EBioMedicine* 2020; **56**: 102780.

105. Zhang L, Dong DI, Zhang W *et al*. A deep learning risk prediction model for overall survival in patients with gastric cancer: a multicenter study. *Radiother Oncol* 2020; **150**: 73–80.

106. Zhang L, Wang X, Yang D *et al*. Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation. *IEEE Trans Med Imaging* 2020; **39**: 2531–40.

107. Poynton C. *Digital Video and HD: Algorithms and Interfaces*. Elsevier, Waltham, MA, 2012.

108. Ly B, Cochet H, Sermesant M. Style Data Augmentation for Robust Segmentation of Multi-modality Cardiac MRI. Statistical Atlases and Computational Models of the Heart Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. Lecture Notes in Computer Science, 2020; 197–208.

109. Zhang H, Cisse M, Dauphin YN, Lopez-Paz D. mixup: Beyond Empirical Risk Minimization. 2017 October 01, 2017:[arXiv:1710.09412 p.]. Available from URL: https://ui.adsabs.harvard.edu/abs/2017arXiv171009412Z

110. Takahashi R, Matsubara T, Uehara K. Data augmentation using random image cropping and patching for deep CNNs. *IEEE Trans Circuits Syst Video Technol* 2020; **30**: 2917–31.

111. Chaudhary J, Rani R, Kamboj A. Deep learning-based approach for segmentation of glioma sub-regions in MRI. *Int J Intell Comput Cybern* 2020; **13**: 389–406.

112. Simard PY, Steinkraus D, Platt JC. Best practices for convolutional neural networks applied to visual document analysis. IEEE, 2003; 958–63.

113. Ganaye PA, Sdika M, Triggs B, Benoit-Cattin H. Removing segmentation inconsistencies with semi-supervised non-adjacency constraint. *Med Image Anal* 2019; **58**: 101551.

114. Novosad P, Fonov V, Collins DL. Alzheimer's disease neuroimaging I. Accurate and robust segmentation of neuroanatomy in T1-weighted MRI by combining spatial priors with deep convolutional neural networks. *Hum Brain Mapp* 2020; **41**: 309–27.

115. Peng Y, Chen S, Qin AN *et al*. Magnetic resonance-based synthetic computed tomography images generated using generative adversarial networks for nasopharyngeal carcinoma radiotherapy treatment planning. *Radiother Oncol* 2020; **150**: 217–24.

116. Thyreau B, Sato K, Fukuda H, Taki Y. Segmentation of the hippocampus by transferring algorithmic knowledge for large cohort processing. *Med Image Anal* 2018; **43**: 214–28.

117. Thyreau B, Taki Y. Learning a cortical parcellation of the brain robust to the MRI segmentation with convolutional neural networks. *Med Image Anal* 2020; **61**: 101639.

118. Kim YC, Kim KR, Choi K, Kim M, Chung Y, Choe YH. EVCMR: a tool for the quantitative evaluation and visualization of cardiac MRI data. *Comput Biol Med* 2019; **111**: 103334.

119. Sandfort V, Yan K, Pickhardt PJ, Summers RM. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. *Sci Rep* 2019; **9**: 16884.

120. Zeng G, Zheng G. Holistic decomposition convolution for effective semantic segmentation of medical volume images. *Med Image Anal* 2019; **57**: 149–64.

121. Nalepa J, Cwiek M, Dudzik W *et al*. Data Augmentation via Image Registration. In: *2019 IEEE International Conference on Image Processing (ICIP)*, 2019; 4250–4.

122. Tustison NJ, Avants BB, Lin Z *et al*. Convolutional neural networks with template-based data augmentation for functional lung image quantification. *Acad Radiol* 2019; **26**: 412–23.

123. Corral Acero J, Zacur E, Xu H *et al*. SMOD – Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI. Functional Imaging and Modeling of the Heart. Lecture Notes in Computer Science, 2019; 361–9.

124. Karimi D, Samei G, Kesch C, Nir G, Salcudean SE. Prostate segmentation in MRI using a convolutional

neural network architecture and training strategy based on statistical shape models. *Int J Comput Assist Radiol Surg* 2018; **13**: 1211–9.

125. Tang Z, Chen K, Pan M, Wang M, Song Z. An augmentation strategy for medical image processing based on statistical shape model and 3D thin plate spline for deep learning. *IEEE Access* 2019; **7**: 133111–21.

126. Bhalodia R, Elhabian SY, Kavan L, Whitaker RT. DeepSSM: a deep learning framework for statistical shape modeling from raw images. *Shape Med Imaging* 2018; **2018**: 244–57.

127. Bhalodia R, Goparaju A, Sodergren T, Whitaker R, Morris A, Kholmovski E et al. Deep Learning for End-to-End Atrial Fibrillation Recurrence Estimation. 2018 Computing in Cardiology Conference (CinC), 2018.

128. Lindner L, Egger J, Schmalsteig D, Gsaxner C, Pfarrkirchner B. TuMore: generation of synthetic brain tumor MRI data for deep learning based segmentation approaches. Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications, 2018.

129. Goodfellow I, Pouget-Abadie J, Mirza M et al. Generative adversarial nets. *Adv Neural Inf Process Syst* 2014; **27**: 2672–80.

130. Luc P, Couprie C, Chintala S, Verbeek J, editors. Semantic Segmentation using Adversarial Networks. NIPS Workshop on Adversarial Training, 2016.

131. Javaid U, Lee JA. Capturing variabilities from Computed Tomography images with Generative Adversarial Networks. European Symposium on Artificial Neural Networks – ESANN'18; Bruges, Belgium, 2018.

132. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 2018; **321**: 321–31.

133. Gao X, Wang X. Performance of deep learning for differentiating pancreatic diseases on contrast-enhanced magnetic resonance imaging: a preliminary study. *Diagn Interv Imaging* 2020; **101**: 91–100.

134. Kaur S, Aggarwal H, Rani R. Diagnosis of Parkinson's disease using deep CNN with transfer learning and data augmentation. *Multimed Tools Appl* 2021; **80**: 10113–39.

135. Calimeri F, Marzullo A, Stamile C, Terracina G. Biomedical Data Augmentation Using Generative Adversarial Neural Networks. Artificial Neural Networks and Machine Learning – ICANN 2017. Lecture Notes in Computer Science, 2017; 626–34.

136. Konidaris F, Tagaris T, Sdraka M, Stafylopatis A. Generative Adversarial Networks as an Advanced Data Augmentation Technique for MRI Data. In: *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2019; 48–59.

137. Wang Q, Zhou X, Wang C et al. WGAN-based synthetic minority over-sampling technique: improving semantic fine-grained classification for lung nodules in CT images. *IEEE Access* 2019; **7**: 18450–63.

138. Kwon G, Han C, Kim D-S. Generation of 3D Brain MRI Using Auto-Encoding Generative Adversarial Networks. Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Lecture Notes in Computer Science, 2019; 118–26.

139. Deepak S, Ameer PM. MSG-GAN Based Synthesis of Brain MRI with Meningioma for Data Augmentation. In: *Proceedings of CONECCT 2020 - 6th IEEE International Conference on Electronics, Computing and Communication Technologies*, 2020.

140. Fetty L, Bylund M, Kuess P et al. Latent space manipulation for high-resolution medical image synthesis via the StyleGAN. *Z Med Phys* 2020; **30**: 305–14.

141. Mirza M, Osindero SJ. Conditional generative adversarial nets, 2014.

142. Sankaranarayanan S, Balaji Y, Castillo CD, Chellappa R, editors. Generate to adapt: Aligning domains using generative adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

143. Xu Z, Qi C, Xu G. Semi-supervised attention-guided CycleGAN for data augmentation on medical images. In: *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2019; 563–8.

144. Ge C, Gu IY-H, Jakola AS, Yang J. Enlarged training dataset by pairwise GANs for molecular-based brain tumor classification. *IEEE Access* 2020; **8**: 22560–70.

145. Wang Z, Lin Y, Cheng KT, Yang X. Semi-supervised mp-MRI data synthesis with StitchLayer and auxiliary distance maximization. *Med Image Anal* 2020; **59**: 101565.

146. Yu H, Zhang X. Synthesis of prostate MR images for classification using capsule network-based GAN model. *Sensors* 2020; **20**: 5736.

147. Martin Arjovsky S, Bottou L, editors. Wasserstein generative adversarial networks. In: *Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia*, 2017.

148. Karras T, Aila T, Laine S, Lehtinen JJ. Progressive growing of gans for improved quality, stability, and variation; 2017.

149. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC, editors. Improved training of wasserstein gans. Advances in neural information processing systems, 2017.

150. Isola P, Zhu J-Y, Zhou T, Efros AA, editors. Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

151. Sabour S, Frosst N, Hinton GE, editors. Dynamic routing between capsules. Advances in neural information processing systems, 2017.

152. Shin H-C, Tenenholtz NA, Rogers JK et al. Medical Image Synthesis for Data Augmentation and Anonymization Using Generative Adversarial Networks. Simulation and Synthesis in Medical Imaging. Lecture Notes in Computer Science, 2018; 1–11.

153. Cao K, Bi L, Feng D, Kim J. Improving PET-CT Image Segmentation via Deep Multi-modality Data Augmentation. Machine Learning for Medical Image Reconstruction. Lecture Notes in Computer Science, 2020; 145–52.

154. Jiang J, Hu Y-C, Tyagi N *et al*. Cross-modality (CT-MRI) prior augmented deep learning for robust lung tumor segmentation from small MR datasets. *Med Phys* 2019; **46**: 4392–404.

155. Zhang T, Fu H, Zhao Y *et al*. SkrGAN: Sketching-Rendering Unconditional Generative Adversarial Networks for Medical Image Synthesis. Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Lecture Notes in Computer Science, 2019; 777–85.

156. Chaitanya K, Karani N, Baumgartner CF, Becker A, Donati O, Konukoglu E. Semi-supervised and Task-Driven Data Augmentation. Information Processing in Medical Imaging. Lecture Notes in Computer Science, 2019; 29–41.

157. Salem M, Valverde S, Cabezas M *et al*. Multiple sclerosis lesion synthesis in MRI using an encoder-decoder U-NET. *IEEE Access* 2019; **7**: 25171–84.

158. Pesteie M, Abolmaesumi P, Rohling RN. Adaptive augmentation of medical data using independently conditional variational auto-encoders. *IEEE Trans Med Imaging* 2019; **38**: 2807–20.

159. Kingma DP, Welling MJ. Auto-encoding variational bayes, 2013.

160. Kingma DP, Dhariwal P, editors. Glow: Generative flow with invertible 1×1 convolutions. Advances in neural information processing systems, 2018.

161. Yang J, Liu S, Grbic S *et al*. Class-aware adversarial lung nodule synthesis in CT images. In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019; 1348–52.

162. Dorent R, Booth T, Li W *et al*. Learning joint segmentation of tissues and brain lesions from task-specific hetero-modal domain-shifted datasets. *Med Image Anal* 2021; **67**: 101862.

163. Anand D, Tank D, Tibrewal H, Sethi A. Self-supervision vs. transfer learning: robust biomedical image analysis against adversarial attacks. In: *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020; 1159–63.

164. Qin T, Wang Z, He K, Shi Y, Gao Y, Shen D. Automatic data augmentation via deep reinforcement learning for effective kidney tumor segmentation. In: *ICASSP 2020 – 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020; 1419–23.

165. Tanno R, Worrall DE, Kaden E *et al*. Uncertainty modelling in deep learning for safer neuroimage enhancement: demonstration in diffusion MRI. *NeuroImage* 2020; **225**: 117366.

166. Kingma DP, Salimans T, Welling MJA. Variational dropout and the local reparameterization trick. *Proceedings of the 28th International Conference on Neural Information Processing Systems*, 2015; **2**: 2575–83.

167. Oksuz I, Ruijsink B, Puyol-Antón E *et al*. Deep Learning Using K-Space Based Data Augmentation for Automated Cardiac MR Motion Artefact Detection. Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Lecture Notes in Computer Science, 2018; 250–8.

168. Oksuz I, Ruijsink B, Puyol-Antón E *et al*. Automatic CNN-based detection of cardiac MR motion artefacts using k-space data augmentation and curriculum learning. *Med Image Anal* 2019; **55**: 136–47.

169. Shaw R, Sudre CH, Varsavsky T, Ourselin S, Cardoso MJ. A k-space model of movement artefacts: application to segmentation augmentation and artefact removal. *IEEE Trans Med Imaging* 2020; **39**: 2881–92.

170. Liu F, Samsonov A, Chen L, Kijowski R, Feng L. SANTIS: Sampling-Augmented Neural neTwork with Incoherent Structure for MR image reconstruction. *Magn Reson Med* 2019; **82**: 1890–904.

171. Bruns S, Wolterink JM, van Hamersvelt RW *et al*. Improving myocardium segmentation in cardiac CT angiography using spectral information. Medical Imaging 2019: Image Processing, 2019.

172. Lartaud P-J, Rouchaud A, Rouet J-M, Nempont O, Boussel L. Spectral CT Based Training Dataset Generation and Augmentation for Conventional CT Vascular Segmentation. Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Lecture Notes in Computer Science, 2019; 768–75.

173. Omigbodun AO, Noo F, McNitt-Gray M, Hsu W, Hsieh SS. The effects of physics-based data augmentation on the generalizability of deep neural networks: demonstration on nodule false-positive reduction. *Med Phys* 2019; **46**: 4563–74.

174. Lucena O, Souza R, Rittner L, Frayne R, Lotufo R. Convolutional neural networks for skull-stripping in brain MR imaging using silver standard masks. *Artif Intell Med* 2019; **98**: 48–58.

175. Momeni S, Fazllolahi A, Bourgeat P *et al*. Data Augmentation Using Synthetic Lesions Improves Machine Learning Detection of Microbleeds from MRI. Simulation and Synthesis in Medical Imaging. Lecture Notes in Computer Science, 2018; 12–9.

176. Kaur S, Aggarwal H, Rani R. MR image synthesis using generative adversarial networks for Parkinson's disease classification. In: *Proceedings of International Conference on Artificial Intelligence and Applications. Advances in Intelligent Systems and Computing*, 2021; 317–27.

177. Ma D, Lu D, Popuri K, Wang L, Beg MF. Alzheimer's disease neuroimaging I. Differential diagnosis of frontotemporal dementia, Alzheimer's disease, and normal aging using a multi-scale multi-type feature generative adversarial deep neural network on structural magnetic resonance images. *Front Neurosci* 2020;**14**:853.

178. Liu S, Dowling JA, Engstrom C, Greer PB, Crozier S, Chandra SSJ. Manipulating Medical Image Translation with Manifold Disentanglement, 2020.

179. Radford A, Metz L, Chintala SJ. Unsupervised representation learning with deep convolutional generative adversarial networks, 2015.

180. Denton EL, Chintala S, Fergus RJA. Deep generative image models using a Laplacian pyramid of adversarial networks. *Adv Neural Inf Process Syst* 2015; **28**: 1486–94.

181. Rosca M, Lakshminarayanan B, Warde-Farley D, Mohamed SJ. Variational approaches for auto-encoding generative adversarial networks, 2017.

182. Zhang H, Goodfellow I, Metaxas D, Odena A, editors. Self-attention generative adversarial networks. In: *International Conference on Machine Learning*, 2019: PMLR.

183. Karnewar A, Wang OJ. MSG-GAN: multi-scale gradient GAN for stable image synthesis, 2019.

184. Karras T, Laine S, Aila T, editors. A style-based generator architecture for generative adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019.

185. Wang Q, Zhang X, Chen W, Wang K, Zhang X. Class-Aware Multi-window Adversarial Lung Nodule Synthesis Conditioned on Semantic Features. Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Lecture Notes in Computer Science, 2020; 589–98.

186. Gu S, Bao J, Yang H, Chen D, Wen F, Yuan L, editors. Mask-guided portrait editing with conditional gans. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019.

187. Zhu J-Y, Park T, Isola P, Efros AA, editors. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision*, 2017.

188. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge, MA, 2018.

189. Jung AB, Wada K, Crall J *et al*. imgaug 2020. Available from URL: https://github.com/aleju/imgaug

190. Bloice MD, Roth PM, Holzinger A. Biomedical image augmentation using Augmentor. *Bioinformatics* 2019; **35**: 4522–4.

191. Chollet F. Keras 2015. Available from URL: https://keras.io

192. Paszke A, Gross S, Massa F *et al*. Pytorch: An imperative style, high-performance deep learning library. arXiv preprint arXiv:191201703, 2019.

193. Pérez-García F, Sparks R, Ourselin S. TorchIO: a Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. arXiv preprint arXiv:200304696, 2020.

194. Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M, Kalinin AA. Albumentations: fast and flexible image augmentations. *Information* 2020; **11**: 125.

195. Casado-García Á, Domínguez C, García-Domínguez M *et al*. CLoDSA: a tool for augmentation in classification, localization, detection, semantic segmentation and instance segmentation tasks. *BMC Bioinformatics* 2019; **20**: 1–14.

196. Consortium TM. Project MONAI. Zenodo, 2020.