A comprehensive survey of recent trends in deep learning for digital images augmentation

Nour Eldeen Khalifa[1] · Mohamed Loey[2] · Seyedali Mirjalili[3,4]

Abstract

Deep learning proved its efficiency in many fields of computer science such as computer vision, image classifications, object detection, image segmentation, and more. Deep learning models primarily depend on the availability of huge datasets. Without the existence of many images in datasets, different deep learning models will not be able to learn and produce accurate models. Unfortunately, several fields don't have access to large amounts of evidence, such as medical image processing. For example. The world is suffering from the lack of COVID-19 virus datasets, and there is no benchmark dataset from the beginning of 2020. This pandemic was the main motivation of this survey to deliver and discuss the current image data augmentation techniques which can be used to increase the number of images. In this paper, a survey of data augmentation for digital images in deep learning will be presented. The study begins and with the introduction section, which reflects the importance of data augmentation in general. The classical image data augmentation taxonomy and photometric transformation will be presented in the second section. The third section will illustrate the deep learning image data augmentation. Finally, the fourth section will survey the state of the art of using image data augmentation techniques in the different deep learning research and application.

Keywords Data augmentation · Image augmentation · Deep learning · Machine Learning · GAN · Artificial Intelligence

✉ Mohamed Loey
  mloey@fci.bu.edu.eg

✉ Seyedali Mirjalili
  ali.mirjalili@gmail.com

  Nour Eldeen Khalifa
  nourmahmoud@cu.edu.eg

[1] Faculty of Computers and Artificial Intelligence, Cairo University, Cairo 12613, Egypt

[2] Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University, Benha 13518, Egypt

[3] Center for Artificial Intelligence Research and Optimization, Torrens University Australia, Fortitude Valley, Brisbane, QLD 4006, Australia

[4] Yonsei Frontier Lab, Yonsei University, Seoul, Korea

🐧 Springer

1 Introduction

Data augmentation (Dyk and Meng 2001) is greatly important to overcome the limitation of data samples and particularly image data-sets. Data is the raw material for every machine learning algorithm, such as the means used to feed the algorithm as illustrated in Fig. 1, any shortage in data and it is labeled may reflect on the accuracy of any proposed model in machine learning (Baştanlar and Özuysal 2014). Image augmentation is one effective training strategy to grow the collection of images for neural network models that also do not include additional images. The image data augmentation is extremely needed for the following reasons:

1. It is an inexpensive methodology if it is compared with regular data collection with its label annotation.
2. It is extremely accurate, as it is originally generated from ground-truth data by nature.
3. It is controllable, which helps in generating balanced data.
4. It contributes to overcoming the overfitting problem (Subramanian and Simon 2013).
5. It helps in achieving better testing accuracies.

Deep learning (DL) is a sub-category of machine learning (Baştanlar and Özuysal 2014) which consequently a subset of artificial intelligence (Nilsson 1981). DL instructs algorithms to learn by imitation. DL aims to simulate the functioning of the human brain, particularly when interacting with data and trends, to help decision-making. DL is the secret to computer vision (Ponce and Forsyth 2012), image classifications, object recognition (Papageorgiou et al. 1998), image segmentation (Pal and Pal 1993), and more. In deep learning, the data is the main source for learning, without sufficient data especially images, the DL model will not learn and produce an accurate model. Models that use deeper learning strategies are less likely to exhibit overfitting but often lack valid training data. Data augmentation, the subject of this survey, is just one approach used to minimize overfitting. Other methods are described below to help prevent overfitting in deep learning models. The paragraphs that follow provide alternative methods to prevent overfitting in DL models. The outcomes of our survey would demonstrate how oversampling of classes in picture data can be performed using Data Augmentation. The main contributions of the presented survey are (1) highlighting the importance of the data augmentation in general, (2) demonstrates the state-of-the-art methods and techniques of data augmentations for images which will help researchers to design more
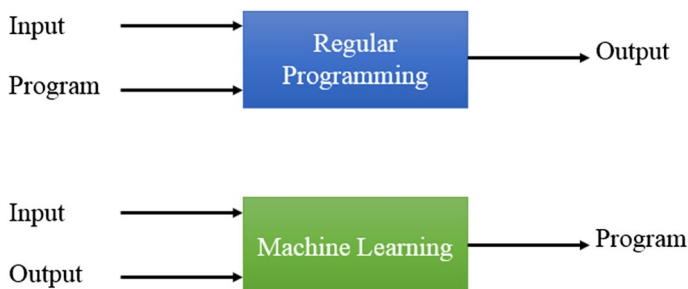

Fig. 1 Difference between regular programming and machine learning

robust and accurate deep learning models, and (3) listing the state of the art research's that successfully use image augmentation in their work.

This survey is organized as follows, Sect. 2 presents a survey over the classical image data augmentation techniques. Section 3 introduces image data augmentation techniques based on DL models. Section 4 illustrates the state of the art of using image data augmentation techniques in deep learning while Sect. 5 summarizes the paper.

2 Classical image data augmentation.

Classical image data augmentation may also be noted as "basic data augmentation" in other scientific researches. The classical image data augmentation consists of geometric transformation and photometric shifting. It includes primitive data manipulation techniques. Those techniques include flipping, rotation, shearing, cropping, and translation in the geometric transformation. It includes primitive color-shifting techniques in photometric shifting such as color space shifting, applying different image filters, and adding noise. Figure 2 represents the classical image data augmentation taxonomy.

2.1 Flipping

Flipping (Vyas et al. 2018) reflects an image around its vertical axis or horizontal axis or both vertical and horizontal axis. It helps users to maximize the number of images in a dataset without needing any artificial processing.. Figure 3 presents different flipping techniques.

Fig. 2 Classical image data augmentation taxonomy

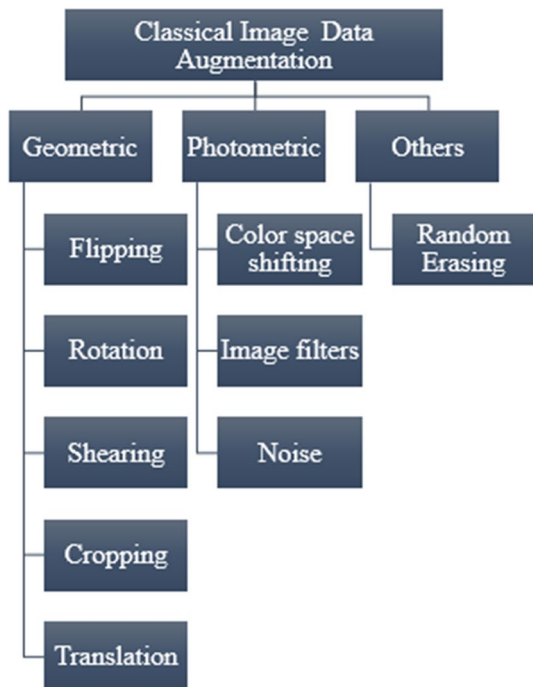**(a)**   **(b)**   **(c)**   **(d)**

Fig. 3 Flipping technique where **a** original image, **b** vertical flipping, **c** horizontal flipping, and **d** vertical and horizontal flipping

Vertical flipping The picture is rotated upside down so that the y-axis is on top and the x-axis is on the bottom. The f_x value and the f_y the value indicates the current coordinates of each pixel after flipping along the vertical axis as illustrated in Eq. (1).

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{1}$$

Horizontal Flipping The picture must be rotated horizontally wherein it's left and right sides. The f_x and f_y components are the pixel's current location after reflection along the horizontal y-axis, as shown in Eq. (2).

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{2}$$

Vertical and horizontal flipping The picture is rotated horizontally and vertically, where both horizontal and vertical columns are preserved. The f_y coordinate and the f_x coordinate is the current coordinates of each pixel after reflection along the vertical and horizontal axes as illustrated in Eq. (3).

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{3}$$

2.2 Rotation

Rotation (Sifre and Mallat 2013) is another type of classical geometric image data augmentation; the rotation process is done by rotating the image around an axis whether in the right direction or the left direction by angels between 1 and 359. Rotation may be applied to images by a certain angle degree in an additive way. For example, rotate the image at about 30-degree angles. It will produce 11 images by rotation with angles 30,60,90,120,150,180,210,240,270, 300,330 angles. The rotation equation is presented in Eq. (4). The f_x and f_y is the new position of each pixel after the rotation process and the x and y pair of coordinates is the raw image. Figure 4 illustrates a sample of the image with different rotation angles (φ).

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{4}$$

Fig. 4 Samples of rotated images

2.3 Shearing

Shearing (Vyas et al. 2018) is the change of the original image along the x direction as well as the y direction. It is an ideal technique to change the shape of an existing object in an image. Shearing contains two types. The first type of component is within the x-axis and the second type is within the y-axis. Equation (5) presents the shearing in the direction of the x-axis while Eq. (6) presents the shearing in the direction of the y-axis. The f_x and f_y is the new position of each pixel after shearing and the x and y of the picture coordinates. Figure 5 illustrates an example of shearing types.

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} 1 & shX \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ shY & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{6}$$

(a)   **(b)**   **(c)**

Fig. 5 Shearing operation for an image where **a** original image, **b** the sheared image in the direction of the X-axis, and **c** the sheared image in the direction of the Y-axis

2.4 Cropping

Cropping (Sifre and Mallat 2013) may also be noted as "zooming" or "scaling" in other scientific researches. Cropping is a process of magnifying the original image. This type of classical geometric image data augmentation consists of two different methods. The first operation is cutting an image from a start X, Y location to another X, Y location. For example, if the image size is 200 * 200 pixel, the image may be cut from (0, 0) to (150, 150) location or from (50, 50) to (200, 200) location. The second operation is the image scaling to its original size. In the above example, the image should be rescaled to 200* 200 pixels after the cutting operation. Equation (7) presents the scaling equation. The f_x and f_y is the new coordinates of each pixel after scaling operation and the x and y represent the coordinates of the original location on the image. Figure 6 illustrates examples of cropping.

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} \text{Xscale} & 0 \\ 0 & \text{Yscale} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{7}$$

2.5 2.5 Translation

Translation (Vyas et al. 2018) is a process of moving an object from one position to another in the image. In translation, geometric image data augmentation is preferred to leave a part of the image white or black after translation to preserve the image data or it is randomized, or it includes Gaussian noise. The translation can be operated in the X direction or the Y direction or X and Y direction at the same time. The left, right, up, and down direction of picture translation may be very useful for avoiding positional bias in the data. Equation (8)
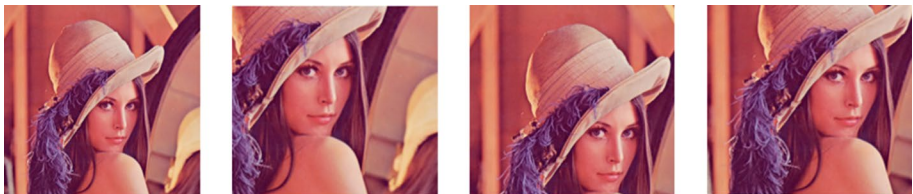

Fig. 6 Different cropping results for an example image

present the translation equation. The f_x and f_y is the new coordinates of each pixel after translation and the x and y represent the coordinates of the original location on the picture. Figure 7 illustrates examples of different translations.

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} \tag{8}$$

Using classical image data augmentation and especially geometric transformation may include some demerits, such as additional memory consumption, additional computing processing power, and more training time. Moreover, classical geometric image data augmentation such as cropping, and translation may remove important features from images, so it must be operated manually not in an automatic random cropping and translation process. In certain applications such as medical image processing, the training data isn't as similar to the test data as classical image data augmentation algorithms make it out to be. So, the scope of where and when to use classical image data augmentation can be practical is quite inadequate.

2.6 Color space shifting

Color space (Winkler 2013) shifting belongs to the family of classical photometric data augmentation. A color plane is a mathematical instrument used to construct and paint using colors. Humans distinguish shades by their color properties such as brightness and hue. Colors may be represented using the quantities of red, green, and blue light produced by the phosphor panel (Winkler 2013). In classical photometric data augmentation, color space-shifting is considered one of the important techniques for increasing the number of images and may reveal some important features in images that were hidden under a specific color space. There most famous color spaces are (Winkler 2013):

- CMY(K) {Cyan—Magenta -Yellow – Black}.
- YIQ, YUV, YCbCr, YCC {Luminance / Chrominance}.
- HSL {Hue—Saturation—Lightness}.
- RGB {Red—Green – Blue}.

The conversion between those color spaces is done by using color space shifting equations. The most used common color space is RGB, Eq. (9) presents the conversion from RGB to CMY, while Eq. (10) presents the conversion from RGB to HSL, finally, Eq. (11)

| (a) | (b) | (c) | (d) |

Fig. 7 Translation examples were **a** original image, **b** translation on the X-axis direction, **c** translation on the Y-axis direction, and **d** translation on the X and Y-axis direction

presents the conversion from RGB to YIQ. Figure 8 illustrates the different color space shifting for a sample image.

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{9}$$

$$h = \begin{cases} 0° & if \ \max = \min \\ 60° \times \frac{g-b}{\max - \min} + 0° & if \ \max = randg \geq b \\ 60° \times \frac{g-b}{\max - \min} + 360° & if \ \max = randg < b \\ 60° \times \frac{b-r}{\max - \min} + 120° & if \ \max = g \\ 60° \times \frac{r-g}{\max - \min} + 240° & if \ \max = b \end{cases} \tag{10}$$

$$s = \begin{cases} 0, & if \ max = 0 \\ \frac{max-min}{max} = 1 - \frac{min}{max}, & otherwise \end{cases}$$

$$l = max$$

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} . \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{11}$$

Color space-shifting is random of intellectualization. Bright or dark pictures are the problem to see by raising the pixel values by a constant value. Another neat function of color space manipulation is to independently process ordinary RGB color matrices. Another approach involves restricting the values for each pixel to a minimum or maximum. The advancement of the color of optical photographs without the need for advanced tools.

2.7 Image filters

Many common image processing techniques such as histogram equalization, brightness increases, sharpening, blurring, and filters are very widespread techniques. (Galdran, et al. 2017). These techniques and filters work by sliding an $n \times m$ matrix across the whole image. Histogram equalization (Premaladha and Ravichandran 2016) is a
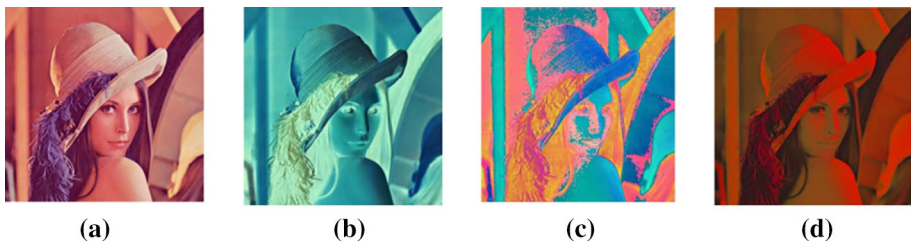

(a)　　　　　**(b)**　　　　　**(c)**　　　　　**(d)**

Fig. 8 Color space for **a** RGB, **b** CMY, **c** HSL, and **d** YIQ

technique for adjusting image intensities to enhance contrast, while the white balancing (Lam and Fung 2008) operates by changing the picture such that it is illuminated by a neutral light source. Special operations are typically conducted separately in the different spectral realms of the signal. Sharpening (Reichenbach et al. 1990) spatial filters are used to highlight fine details in an image or to enhance details that have been blurred, while blurring is averaging process for pixels by its neighbors as a process of integration. Using the sharpening and blurring filter (Reichenbach et al. 1990) can result in a distorted picture or high contrast horizontal or vertical edges which will aid in recognizing the details of an image (Shorten and Khoshgoftaar 2019).

The above-mentioned filters are operated using matrix multiplication of the original image with the filter matrix. Figure 9 presents the image filter output using different filters and they are histogram equalization (Premaladha and Ravichandran 2016), white balancing (Lam and Fung 2008), enhancing brightness sharpening (Reichenbach et al. 1990) accordingly.

2.8 Noise

Adding noise to a picture requires inserting a noise matrix that is created from a regular distribution. Four famous types of noise can be used as image data augmentation and they are Gaussian, Poisson, salt & pepper, and speckle noise. Many other noises exist, but in this research, the mentioned four types of noises were selected and investigated.

The first form of noise to be studied is additive noise since it's called additive. Gaussian noise continues to shift the color values that makeup pictures. The probability density function is provided in this article, based on this Eq. (12) (Boyat and Joshi Apr. 2015).

$$P(g_c) = \left(\sqrt{\frac{e}{2\pi\sigma^2}} \right)^{\frac{-(g_c - \mu)^2}{2\sigma^2}} \tag{12}$$

where g_c is the grey color meaning, σ and μ are the standard deviation and mean respectively. The mean value is zero, the spectrum is between zero and one, and the standard deviation is 0.1, and 256, as is seen in Fig. 10.

The second form of noise is the Poisson noise which is normally present in the electromagnetic frequencies encountered by humans. This x-ray and gamma-ray machine released a variety of photons continuously. The passion distribution is shown by the graph seen below (13) (Boyat and Joshi Apr. 2015).
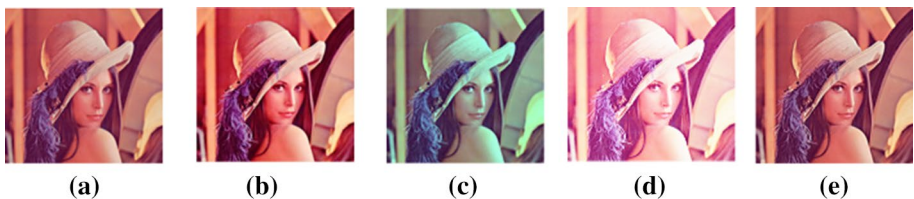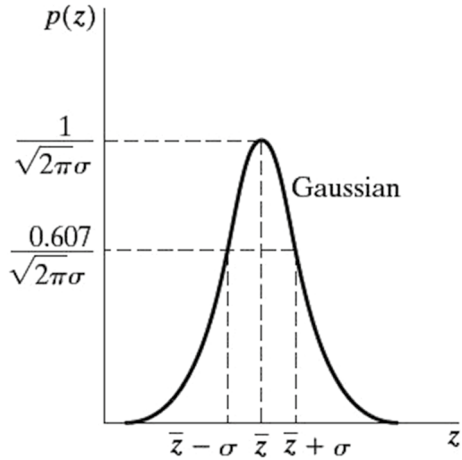

Fig. 9 Output images from applying different filters, **a** original image, **b** histogram equalization, **c** white balancing, **d** enhancing brightness, and **e** sharpening

Fig. 10 The probability density function for Gaussian noise

$$P(f_X = x) = e^{-\lambda} \frac{\lambda}{x!} where x = 0, 1, 2, 3, \ldots \qquad (13)$$

where e$=2.718282$, λ is the mean number of events per interval, and $f_X =$ The number of events in a given interval.

The third form of noise is called a salt & pepper noise in which the values of certain pixels in the picture are modified. In a noisy picture, some neighbor pixels do not shift, as shown by Eq. (14), which illustrates the salt & pepper noise (Boyat and Joshi Apr. 2015).

$$P(g) = \begin{cases} Pa & for\, g = a \\ Pb & for\, g = b \\ 0 & otherwise \end{cases} \qquad (14)$$

The last form of noise is speckle noise, a multiplicative/additive noise that is often written. Their appearance is related to optical devices such as lasers, radar, and sonar, etc. Speckle noise may occur in the same way as Gaussian noise. The PDF has a tail a gamma distribution that reflects speckle noise by Eq. (15) (Boyat and Joshi Apr. 2015).

$$g(n, m) = f(n, m)*u(n, m) + \delta(n, m) \qquad (15)$$

In the observed image, $g(n, m)$ is the speckle noise additive component, and u(n, m) is the multiplicative function. It is seen in Fig. 11 that numerous noises were applied to the original picture. the noises are Gaussian, Poisson, salt & pepper, and speckle noise in the order presented.

Using classical photometric image data augmentation may include some disadvantages, such as additional memory consumption, additional computing processing power, and more training time. Moreover, photometric image data augmentation may cause the elimination of features in an image, and those features are important and especially when it is a color feature that may be used to distinguish between different classes in the dataset. The recommendation hereby is to use photometric image data augmentation with care and after studying the features of the original dataset first.

A lot of researches such as (Khalifa et al. 2019a, 2019b, 2019c, 2018a; Khalifa et al. xxxx) uses a mixture of classical image data augmentation together. The mixture might

Fig. 11 An example of different noises **a** Gaussian noise, **b Poisson** noise, **c** salt & pepper noise, and **d** speckle noise

include two or three types of geometric transformation or include a mixture of one or two from geometric transformation along with one type of photometric transformation. Those mixtures are tested along with testing accuracy to prove their efficiency. There is no obvious rule state that classical image data augmentation is the most appropriate one as it depends on the characteristics of the original dataset.

2.9 Random erasing

The random erasing technique is one of the image data augmentation techniques introduced in Zhong et al. (2017). This is not part of geometry transformation. The basic principle of random erasing is to randomly erase one square in the square region of the picture. which proved to be effective as illustrated in Zhong et al. (2017). Figure 12 presents a sample of removing random rectangles from the original image.

3 Deep learning data augmentation

Deep learning (LeCun et al. 2015) has achieved remarkable breakthrough research's during the last decade. This was a result of the continuous contributions of researchers around the world through their deep learning architectures. Deep learning proved its efficiency in computer vision, image classifications, object detection, image segmentation, and more. It is expected that deep learning for image data augmentation will prove its efficiency again in this field. Deep learning image data augmentation consists of three main categories, the first category is Generative Adversarial Networks (GAN)(Goodfellow, et al. 2014), the second category is Neural Style Transfer (Jing et al. 2019), While the third category is meta metric learning (Frans et al. 2018). The third category consists of three models. The
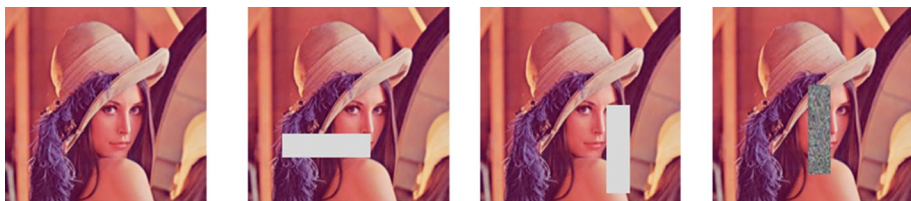

Fig. 12 Random erasing image data augmentation different samples

models are Neural Augmentation, Auto Augment, and Smart Augmentation. Figure 13 presents the structure of deep learning data augmentation.

## 3.1 Generative adversarial networks

One of the deep learning artificial intelligence picture data enhancement technologies is generative modeling. Generative modeling includes generating artificial images from the initial dataset and then utilizing them to predict features of the image. An example of a generative network is a generative adversarial network (GAN) (Yi et al. 2019). GANs are made of two distinct kinds of networks. The networks are educated concurrently. The network is trained to forecast indoor scenes while the network is trained to differentiate amongst them. GANs are called a specific form of the Deep Learning model.

GANs may learn representations from data not needing labeled datasets. It is extracted from competitive learning mechanisms involving a pair of neural networks (Yi et al. 2019). Academic and business fields have accepted adversarial preparation as a data-driven manipulation strategy due to its simplicity and usefulness in producing new pictures. GANs have made considerable strides and have brought in major changes in many applications. These applications include picture synthesis, style conversion, semantic image editing, image super-resolution, image classification.

3.2 GAN architecture

The key problem discussed in the paper is the two-player zero-sum scenario. The one who wins at the game gets the same sum of money as the other team. The networks lead to classes of GANs labeled discriminator and generator networks. The discriminator was developed to decide whether or not a sample was a true sample or a synthetic one. (Alqahtani et al. 2019). Alternatively, the generator will create a false sample of images to confuse the discriminator.

Fig. 13 Deep learning image data augmentation structure
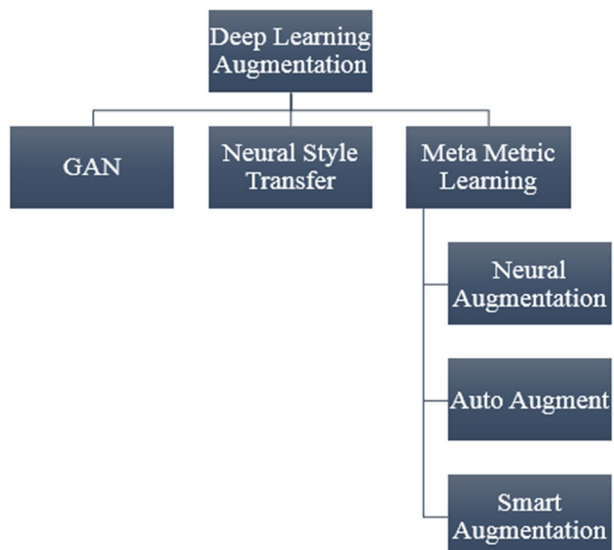
The Discriminator produces the likelihood that a given sample originated from a collection of real samples. A real sample has a strong chance of being true. Possibly false samples are suggested by their low likelihood. The generator may provide an optimum solution where the discriminator has almost no potential to distinguish true from false samples where the discriminator error rate is close to 0.5 (Alqahtani et al. 2019). The GAN structure is presented in Fig. 14. As data, a random sample is collected and this is used by the Generator to produce the output (Alqahtani et al. 2019).

The generator is a neural network that learns from noise in photos to generate pictures. The noise generated from the Generator is documented through the G (z). Gaussian noise is an input to the device in latent space. During the training phase, each neuron's G and D values are changed iteratively.

The discriminator The Discriminator is a neural network that is capable of identifying whether or not the picture it has memorized is indicative of real-world evidence. X is input into D and is the output (x) (Goodfellow, et al. 2014). The goal function of a two-person minimax game was described in equation form. (16).

$$\min_G \max_D V(D, G) = E_{x \sim P_{data}(x)} \big[\log(d(x)) \big] + E_{z \sim p_g(z)} \big[log(1 - D(G(z))) \big] \tag{16}$$

The popularity of GANs has generated a new curiosity in how these models could be applied to Data Augmentation. These networks enable the generation of novel training data which results in improved classifiers. Figure 15 provides examples of outputs of a GAN for an original picture.

3.3 Neural style transfer

Neural Style Transfer (NST) (Gatys et al. 2016) is another technique for generating images in deep learning data augmentation. It is an artificial model which was developed by using a deep neural network particularly a deep convolutional network. The model uses neural representations of material and style to isolate and recombine pictures, demonstrating a way to construct creative images computationally (Gatys et al. 2016). it is probably best known for its artistic domain applications. The NST was a base work for creating artistic works in applications such as real-time style transfer for super-resolution images (Johnson et al. 2016) (Hayat 2018). A set of equations can be found in Gatys et al. (2016) which is considered the mathematical foundation of the neural style transfer.
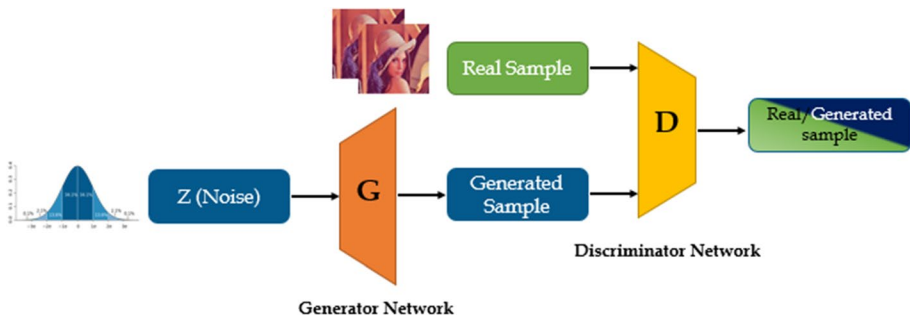

Fig. 14 Graphical representation of the generative adversarial network

Fig. 15 Samples of output image during the training of GAN

There are many artistic styles for neural style transfer such as the starry night by Vincent Van Gogh in 1889, the muse by Pablo Picasso in 1935, composition vii by Wassily Kandinsky in 1913, the great wave off by Kanagawa Hokusai from 1829–1832, the sketch style, and the Simpsons style (Johnson et al. 2016). Figure 16 presents a different NST sample image.

Selecting which trends may be incredibly difficult, particularly for specialists. For instances like self-driving vehicles, it's normal to think about data from day to night, summer to winter, or sunny to rainy days. Nevertheless, in other types of applications, the conventional styles to transfer into are not so clear (Shorten and Khoshgoftaar 2019). The choice of the neural style transfer again depends on the characteristics of the original dataset. The method of image data augmentation includes the careful collection of type images to be transmitted to various image datasets. If the sample range is so tiny, depending on the outcome, the results will be in danger of being skewed. (Shorten and Khoshgoftaar 2019).
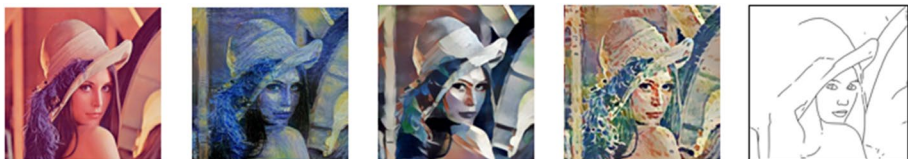

Fig. 16 Samples of neural style transfer images

3.4 Meta metric learning

The meta metric learning (Zoph and Le 2019) concept is the use of a neural network to optimize neural networks. This concept was applied firstly by Barretzoph and et al. in Zoph and Le (2019). They have a novel model for meta-learning architectures which uses a recurrent network to achieve the highest possible accuracy (Zoph and Le 2019). There are many research trials in this area. Three research trials were selected to be studied in this work. The three models are neural augmentation, auto augment, and smart augmentation. These models used a mixture of different techniques such as photometric transformation, geometric transformations, NST, and GAN.

3.4.1 Neural augmentation

Neural Augmentation (NA) is originally presented by Wang and Perez (Perez and Wang 2017). They suggest that a neural network should be trained with the best-fit augmentation strategy such that it can reach full accuracy. The authors proposed two different approaches for data augmentation. The author first manipulated the data to maximize identification before training the classifier. The researchers implemented GANs and simple transformations to construct a broader dataset. The second solution required learning from a neural net that was prepended to the input. At training time as presented in Fig. 17, this neural network takes in two random images and produces a single image that is in style or in context with a specified image it has been trained on. The gradient for training a convolutional network is then transferred onto the following layers of the network. During training, photos from the validation collection would be used to train the classifier (Perez and Wang 2017). They achieved remarkable accuracy and their result was very promising.

3.4.2 Auto augment

Auto Augment (AA) is originally presented by Cubuk and et al. (Cubuk et al. 2019). They created a software package AA that automatically searches for the right augmentation policies for a specified picture. In their implementation, they built a search space where a policy
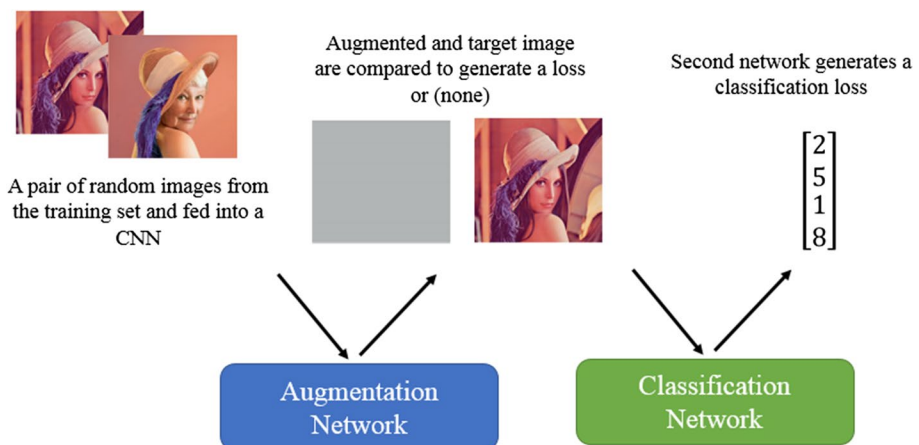

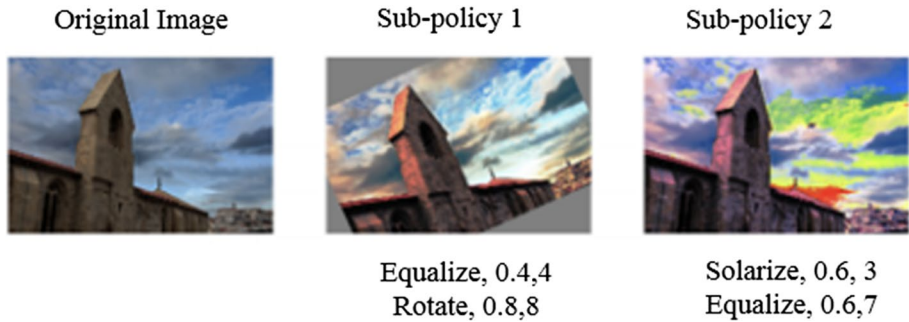**Fig. 17** Training model for Neural Augmentation

Original Image   Sub-policy 1   Sub-policy 2

Equalize, 0.4,4
Rotate, 0.8,8

Solarize, 0.6, 3
Equalize, 0.6,7

Fig. 18 Example of sub-policies for Auto Augment (Cubuk et al. 2019)

Table 1 Simple representation for Network A, Network B structure for smart augmentation

Network A	Input (multi channel)	Network B	Input (one channel)
	16 channel 5×5		16 channel 3×3
			Batch Normalization 1
	16 channel 7×7		Max pool (3,3)
			Batch Normalization 2
	32 channel 3×3		8 channel 3×3
			Batch Normalization 3
	32 channel 5×5		Max pool (3,3)
			Dropout
	Output (single channel)		Output

is comprised of several sub-policies. Each sub-policy is selected at random for each picture (Cubuk et al. 2019). Under the sub-policy, two image processing tasks, such as translating, rotating, or shearing, are done, and how much and how powerful these functions are implemented (Cubuk et al. 2019) as illustrated in Fig. 18. They built a series of rules centered on the effects of the neural network that produces the best validity accuracy. (Cubuk et al. 2019). They compared their proposed AA model with other related works such as GAN, their achieved result was very promising. This Auto Enhanced model achieved a classification rate of 98.52 percent for CIFAR-10. (Krizhevsky et al. 2009). Additionally, it attained a 16.5% top-1 error rate on the ImageNet dataset. (Krizhevsky et al. 2012).

3.4.3 Smart augmentation

Smart Augmentation (SA) is another type of meta metric learning under deep learning data augmentation. It was originally presented by Lemley and et al. (Lemley et al. 2017). They used a modern regression approach to boost training precision and minimize overfitting. SA succeeded by building a network that learned how to produce enhanced data through the training phase of a target network in a manner that minimized the sum of the target network's failure (Lemley et al. 2017) as presented in Table 1. The used ANN improved the accuracy of the presented model by minimizing the error of that network. SA showed strong and substantial improvements in precision on all datasets. As well as the opportunity

to reach comparable or better efficiency standards with smaller networks, it has been effective on many networks. (Lemley et al. 2017). The SA approach as illustrated is like the NA technique presented in the earlier sub-section. Their experiment shows that the augmentation method can be automated, especially in situations when two or more samples of a certain type can be mixed non-homogenously outcomes in a stronger generalization of a target network. They demonstrated that a deep learning algorithm might learn the enhanced task at the same time the task was being taught to the network (Lemley et al. 2017). The SA strategy was evaluated on learning tasks in conducting gender identification on a dataset (Phillips et al. 1998), the accuracy rose to 88.46%. The audience dataset responded with an improvement of 76.06%. By utilizing another face dataset, the findings changed to 95.66% (Lemley et al. 2017).

One of the drawbacks of meta metric learning for data augmentation, it is a relatively new concept and needs to be tested properly and extensively by researchers to prove its efficiency. Moreover, the implementation of meta metric learning for image data augmentation is relatively hard and consumes a lot of time in development.

4 Image data augmentation state of art researches

Image data augmentation with its two branches (classical, and deep learning) has attracted the attention of many researchers throughout the previous years. The section conducted its results based on the Scopus database in the field of computer science with keyword terms "data augmentation, image augmentation, and deep learning" from the year 2015 to 2020. Figure 19 presents the number of researches through the last 6 years in the image data augmentation within the computer science field using the Scopus database. It is clearly shown throughout the figure that the researches in image data augmentation are exponentially increasing. In the year 2020, the number of researches was 1269 which is 24 times larger than researches carried out in 2015 which only include 52 research papers, a large number of researches due to the effectiveness of data augmentation in producing accurate results.
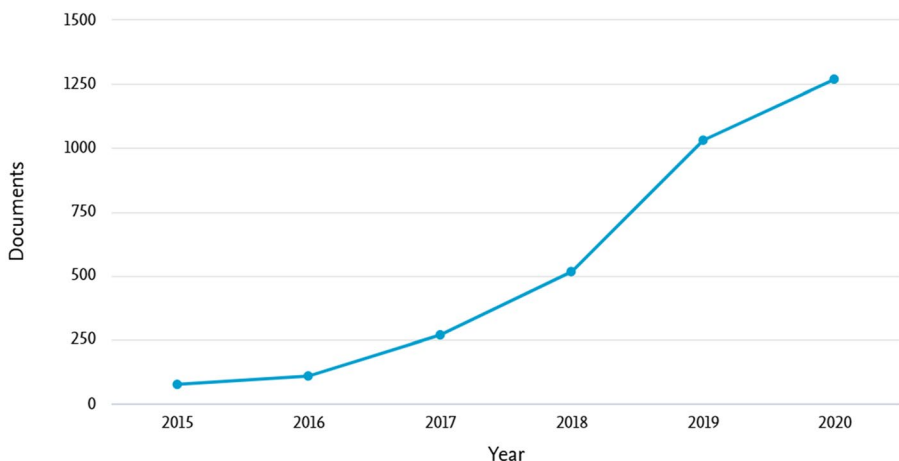

Fig. 19 Image data augmentation research number in the computer science field from 2015–2020

Figure 20 also shows that the image data augmentation attracted the institutions to support the researchers in the domain of image data augmentation related to computer science within the last 6 years. The figure also shows according to the Scopus database that the National Natural Science Foundation of China sponsored more than 430 research papers in the domain of image data augmentation which is related to the computer science field. The list of the institutions ordered by the number of researches is " National Natural Science Foundation of China, National Science Foundation, Ministry of Science and Technology of the People's Republic of China, Fundamental Research Funds for the Central Universities, Nvidia, Ministry of Education of the People's Republic of China, National Key Research and Development Program of China, Ministry of Finance, and National Institutes of Health."

4.1 Medical domain

Image data augmentation techniques in the medical domain have accomplished a distinguished contribution and a breakthrough. As the existence of a large medical dataset is a difficult task as it needs continuous efforts in the long term. Image data augmentation techniques help in generating medical images for diagnoses in an inexpensive way and accomplished the highest testing accuracy possible without the need for the existence of large medical datasets.

One of the breakthrough research in the medical domain which used image data augmentation techniques is the work presented in Ronneberger et al. (2015), the authors presented U-net: convolutional networks for biomedical image segmentation. They claim to have built a neural network and training algorithm which relies on the extensive use of image data augmentation. The image data augmentation with their proposed neural networks achieved 92% testing accuracy (Ronneberger et al. 2015).

In (Pereira et al. 2016), The paper suggested utilizing convolutional neural networks to segment brain tumors. The authors stated that they presented an automatic segmentation method based on convolutional neural networks and the use of image data
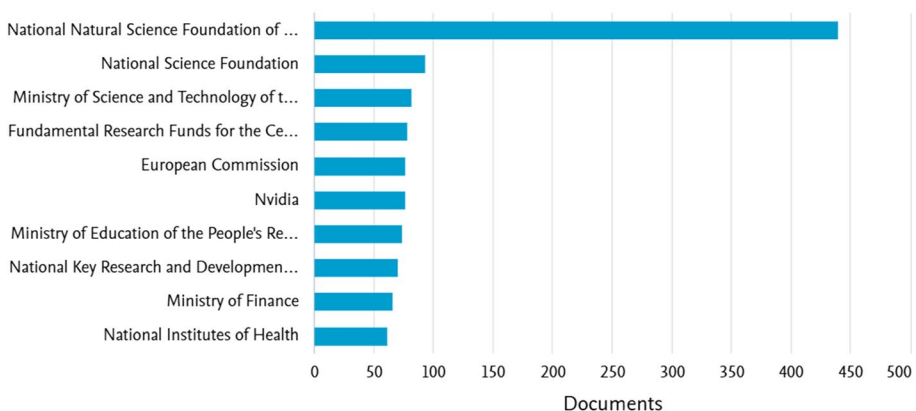

Fig. 20 Image data augmentation research number funded by an institution from 2015–2020

augmentation which is considered to be very effective for brain tumor segmentation in MRI images. They used classical image data augmentation (rotations with multiple 90 degrees). The achieved results relevant dice similarity coefficient metric (Thada and Jaglan 2013) were 0.78, 0.65, and 0.75.

According to the World Health Organization, the coronavirus (COVID-19) pandemic is placing healthcare services worldwide under unparalleled and growing strain. The scarcity of COVID-19 datasets, especially in chest X-ray and CT photos, is the primary motivation for some scientific researches such as (Loey et al. Apr. 2020a; Loey et al. 2020a). The primary objective is to capture all available x-ray and CT images for COVID-19 and to use classical data augmentation techniques in conjunction with GAN (Loey et al. Apr. 2020a) and CGAN (Loey et al. 2020a) to produce additional pictures to aid in the identification of the COVID-19. The combination of classical data augmentations and GAN significantly improves classification accuracy in all chosen models.

A lot of work in the medical field has been used for image data augmentation whether classical image data augmentation or deep learning augmentation. Table 2 summarizes selected research works that used the image data augmentation techniques in the medical domain.

4.2 Agriculture domain

Agriculture is an important domain which secures the human with the necessary foods for their living. Image data augmentation helps many researchers around the globe to enhance their models in the agriculture domain. In the presented work in Khalifa et al. 2020b, the authors presented different deep transfer models to classify 101 class of insect pests which are harmful to agriculture crops. They used classical augmentation techniques to raise the number of images to be 3 times larger than the original dataset. They adopted reflection as an augmentation technique for their dataset. Using the image data augmentation technique in their work raised their accuracy from 41.8% in testing accuracy to 89.33%.

In (Mehdipour Ghazi et al. 2017), The authors proposes a novel approach for plant identification utilizing deep neural networks adjusted using optimization techniques. To boost image precision, they have used methods such as rotation, conversion, reflection, and scaling. Their algorithm scored an accuracy of 80 percent on the validation collection and a rank ranking of 75.2% on the official test set. Table 3 summarizes selected research works that used the image data augmentation techniques in the agriculture domain (Loey et al. Apr. 2020c).

4.3 Remote sensing domain

Remote sensing is a critical field that involves detecting and tracking the physical properties of an environment (typically from satellite or aircraft) by the calculation of reflected and emitted radiation from a radius. Many researchers around the world use image data augmentation to improve their models in the remote sensing domain. Table 4 summarizes selected research works that used the image data augmentation techniques in the different remote sensing domains.

Table 2 Selected works in the medical domain used image data augmentation

	Year	Short description	Application	Augmentation techniques
Ronneberger et al. (2015)	2015	U-net: convolutional network model for biomedical images	Segmentation	Shifting and rotation
Pereira et al. (2016)	2016	Convolutional neural network model for MRI	Segmentation	Rotation
Khalifa et al. (2019c)	2019	Convolutional neural networks for limited bacterial colony dataset	Classification	Reflection, cropping, and noise
Khalifa et al. (2019a)	2019	Medical Diabetic Retinopathy detection using DL	Detection	Reflection
Loey et al. Apr. (2020b)	2020	Leukemia blood cells image recognizing based on DL	Classification	Shifting and rotation
Frid-Adar et al. (2018)	2018	A convolutional neural network model for liver lesion classification	Classification	GAN
Khalifa et al. (2020a)	2020	Deep transfer models for COVID-19 associated pneumonia from x rays chest dataset	Classification	GAN
Loey et al. Apr. (2020a)	2020	GAN with transfer learning models for Chest COVID-19 X-ray classification	Classification	GAN
Loey et al. (2020a)	2020	CGAN with data augmentation for Chest COVID-19 CT images classification	Classification	CGAN

Table 3 Selected research works that used image data augmentation in the agriculture domain

	Year	Short description	Application	Augmentation techniques
Bargoti and Underwood (2017)	2016	A regional convolutional network model for the detection, and counting of fruits in orchards	Detection, counting	Reflection, scaling, and color space shifting
Mehdipour Ghazi et al. (2017)	2017	Deep transfer learning for plant classification (1000 class)	Classification	Rotation, translation, reflection, and scaling
Khalifa et al. (2020b)	2019	Deep transfer learning for the classification of insect pests (101 class)	Classification	Reflection
Giuffrida et al. (2017)	2017	Arabidopsis Rosette Image Generator using GAN	Image Generator	GAN
Zhang et al. (2019)	2019	Convolutional neural networks for fruit classification (18 class)	Classification	Rotation, noise addition

Table 4 Selected research works that used image data augmentation in the remote sensing domain

	Year	Short description	Application	Augmentation techniques
Yan et al. (2019)	2019	Rotation Region Detection Network for detecting aircraft	Detection	Novel data augmentation method
Guirado et al. (2019)	2019	Convolutional neural networks for whale counting in remote sensing	Detection	Rotation, translation, reflection, and scaling
Shawky et al. Nov. (2020)	2020	Convolutional neural networks for remote sensing classification (21 class)	Classification	Rotation, translation, flipping
Yamashkin et al. (2020)	2021	Deep transfer learning for plant classification for remote sensing	Classification	different-scale images

4.4 Miscellaneous domains

The image data augmentation techniques not only help in the medical and agriculture domain. It contributed greatly to other domains. Those domains vary from human identification, art, and music to space technology. Table 5 summarizes selected research works that used the image data augmentation techniques in the different miscellaneous domains.

5 Summary

This survey started with the importance of the image data augmentation for the limited dataset and especially the image dataset. A collection of structured data augmentation approaches is suggested for dealing with the depth overfitting question in DL models. Data-intensive models depend on deep learning Applying approaches in this survey yield the same or superior performance in small datasets. Information augmentation is very useful for producing improved datasets. The survey was structured into three main sections. The first section was the classical image data augmentation, while the second section was deep learning data augmentation, and the third section was the image data augmentation state art researches. The classical image data augmentation taxonomy consisted of geometric transformation which included flipping, rotation, shearing, cropping, and translation. The photometric transformation included color space shifting, image filters, and adding noise. The deep learning image data augmentation included three types, the first type was the image data augmentation using GAN, while the second type was the Neural Style Transfer, and the third type was the meta metric learning. The meta metric included Neural Augmentation, Auto Augment, and smart Augmentation. Finally, the third main section illustrated the state-of-the-art researches in image data augmentation within different domains such as the medical domain, agriculture domain, and other miscellaneous domains. The prospect of data augmentation is extremely promising search algorithms that use data warping and oversampling approaches have enormous potential. The deep neural network's layered design provides multiple possibilities for Data Augmentation. Future study would aim to create a taxonomy of augmentation techniques, build up new quality standards for GAN samples, discover relationships between data augmentation, and further generalize the concepts of data augmentation.

Table 5 Selected research works that used image data augmentation in different domains

	Year	Short description	Application	Augmentation techniques
Eitel et al. (2015)	2015	Two deep CNN for RGB-D object recognition	Recognition	Cropping, and adding noise
Farfade et al. (2015)	2015	Deep CNN for face detection	Detection	Rotation, Flipping, adding noise
Yu et al. (2015)	2015	Deep transfer models sketch recognition	Recognition	Flipping, Translation, and Shifting
Boominathan et al. (2016)	2016	Combination of deep and shallow CNN for dense crowd counting	Counting	Scaling
Uhlich, et al. (2017)	2017	Deep neural networks for music source separation	Speech Segmentation	Flipping, and Scaling
Khalifa et al. (2018b)	2018	Deep CNN for galaxy morphology classifications	Classification	Reflecting, Rotation, and cropping
Lim et al. (2018)	2018	Unsupervised anomaly detection using GAN and CNN	Detection	GAN
Zhu et al. (2018)	2018	Emotion classification using generative adversarial networks and CNN	Classification	GAN
Loey et al. (2020b)	2020	Medical face mask detection based on DL	Detection	Flipping, Reflecting, Rotation
Loey et al. (2021)	2021	Deep learning for face mask classification	Classification	Reflecting, Rotation

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Alqahtani H, Kavakli-Thorne M, Kumar G (2019) Applications of generative adversarial networks (GANs): an updated review. Arch Comput Methods Eng. https://doi.org/10.1007/s11831-019-09388-y

Bargoti S, Underwood J (2017) "Deep fruit detection in orchards," in 2017 IEEE International Conference on Robotics and Automation (ICRA), May 2017, pp 3626–3633 https://doi.org/10.1109/ICRA.2017.7989417

Baştanlar Y, Özuysal M (2014) Introduction to machine learning. Methods Mol Biol. https://doi.org/10.1007/978-1-62703-748-8_7

Boominathan L, Kruthiventi SSS, and Venkatesh Babu R (2016) CrowdNet: a deep convolutional network for dense crowd counting 2016 https://doi.org/10.1145/2964284.2967300

Boyat AK, Joshi BK (2015) A review paper : noise models in digital image processing. Signal Image Process Int J 6(2):63–75. https://doi.org/10.5121/sipij.2015.6206

Cubuk ED, Zoph B, Mane D, Vasudevan V, Le QV (2019) Autoaugment: learning augmentation strategies from data https://doi.org/10.1109/CVPR.2019.00020

Dyk DAV, Meng XL (2001) The art of data augmentation. J Comput Graph Stat. https://doi.org/10.1198/10618600152418584

Eitel A, Springenberg JT, Spinello L, Riedmiller M, Burgard W (2015) Multimodal deep learning for robust RGB-D object recognition https://doi.org/10.1109/IROS.2015.7353446

Farfade SS, Saberian M, Li LJ (2015) Multi-view face detection using Deep convolutional neural networks https://doi.org/10.1145/2671188.2749408

Frans K, Ho J, Chen X, Abbeel P, Schulman J (2018) Meta learning shared hierarchies

Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H (2018) GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing. https://doi.org/10.1016/j.neucom.2018.09.013

Galdran A et al. (2017) Data-driven color augmentation techniques for deep skin image analysis pp 1–4

Gatys L, Ecker A, Bethge M (2016) A neural algorithm of artistic style. J vis. https://doi.org/10.1167/16.12.326

Giuffrida MV, Scharr H, Tsaftaris SA (2017) "ARIGAN: synthetic arabidopsis plants using generative adversarial network," *Proceedings - 2017 IEEE International Conference on Computer Vision Workshops, ICCVW 2017*, vol 2018-Janua, no. i, pp 2064–2071, 2017, doi: https://doi.org/10.1109/ICCVW.2017.242

Goodfellow IJ et al. (2014) Generative adversarial nets," in Proceedings of the 27th International Conference on Neural Information Processing Systems - Vol 2, 2014, pp 2672–2680

Guirado E, Tabik S, Rivas ML, Alcaraz-Segura D, Herrera F (2019) Whale counting in satellite and aerial images with deep learning. Sci Rep. https://doi.org/10.1038/s41598-019-50795-9

Hayat K (2018) Multimedia super-resolution via deep learning: a survey. Digital Signal Process Rev J. https://doi.org/10.1016/j.dsp.2018.07.005

Jing Y, Yang Y, Feng Z, Ye J, Yu Y, Song M (2019) Neural style transfer: a review. IEEE Trans Visual Comput Graphics. https://doi.org/10.1109/tvcg.2019.2921336

Johnson J, Alahi A, Fei-Fei L (2016) Perceptual losses for real-time style transfer and super-resolution https://doi.org/10.1007/978-3-319-46475-6_43.

Khalifa N, Loey M, Taha M, Mohamed H (2019a) Deep transfer learning models for medical diabetic retinopathy detection. Acta Informatica Medica 27(5):327. https://doi.org/10.5455/aim.2019.27.327-332

Khalifa N, Taha M, Hassanien A, Mohamed H (2019b) Deep Iris: deep learning for gender classification through iris patterns. Acta Informatica Medica 27(2):96. https://doi.org/10.5455/aim.2019.27.96-102

Khalifa NEM, Taha MHN, Hassanien AE, Hemedan AA (2019c) Deep bacteria: robust deep learning data augmentation design for limited bacterial colony dataset. Int J Reason-Based Intel Syst. https://doi.org/10.1504/ijris.2019.102610

Khalifa NEM, Loey M, Taha MHN (2020b) Insect pests recognition based on deep transfer learning models. J Theor Appl Inf Technol 98(1):60–68

Khalifa NEM, Taha MHN, Hassanien AE (2018) Aquarium family fish species identification system using deep neural networks, in Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2018 pp. 347–356 https://doi.org/10.1007/978-3-319-99010-1_32

Khalifa NE, Hamed Taha M, Hassanien AE, Selim I (2018) Deep galaxy V2: robust deep convolutional neural networks for galaxy morphology classifications," in 2018 International Conference on Computing Sciences and Engineering, ICCSE 2018 - Proceedings, Mar 2018, pp 1–6 doi: https://doi.org/10.1109/ICCSE1.2018.8374210

Khalifa NEM, Taha MHN, Hassanien AE, Elghamrawy S (2020) Detection of Coronavirus (COVID-19) associated pneumonia based on generative adversarial networks and a fine-tuned deep transfer learning model using chest X-ray dataset, *arXiv*, pp 1–15

Khalifa NEM, Taha MHN, Hassanien AE, Selim IM (2017) Deep galaxy: classification of galaxies based on deep convolutional neural networks. arXiv preprint. arXiv:1709.02245

Krizhevsky A, Nair V, Hinton G (2009) CIFAR-10 and CIFAR-100 datasets," https://www.cs.toronto.edu/~kriz/cifar.html

Krizhevsky A, Sutskever I, Geoffrey HE (2012) Imagenet. *Adv Neural Information Process Syst 25 (NIPS2012)* https://doi.org/10.1109/5.726791.

Lam EY, Fung GSK (2008) Automatic white balancing in digital photography. Single-Sens Imag Methods Appl Digital Cameras. https://doi.org/10.1201/9781420054538.ch10

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436. https://doi.org/10.1038/nature14539

Lemley J, Bazrafkan S, Corcoran P (2017) Smart augmentation learning an optimal data augmentation strategy. IEEE Access. https://doi.org/10.1109/ACCESS.2017.2696121

Lim SK, Loo Y, Tran NT, Cheung NM, Roig G, and Elovici Y (2018) DOPING: generative data augmentation for unsupervised anomaly detection with GAN https://doi.org/10.1109/ICDM.2018.00146

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

Loey M, Smarandache F, Khalifa NEM (2020a) Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning. Symmetry 12(4):651. https://doi.org/10.3390/sym12040651

Loey M, Manogaran G, Taha MHN, Khalifa NEM (2020) Fighting against COVID-19: a novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustain Cities Soc. https://doi.org/10.1016/j.scs.2020.102600

Loey M, Naman M, Zayed H (2020b) Deep transfer learning in diagnosing leukemia in blood cells. Computers 9(2):29. https://doi.org/10.3390/computers9020029

Loey M, ElSawy A, Afify M (2020c) Deep learning in plant diseases detection for agricultural crops: a survey. Int J Serv Sci Manag Eng Technol 11(2):41–58. https://doi.org/10.4018/IJSSMET.2020040103

Loey M, Manogaran G, Taha MHN, Khalifa NEM (2021) A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Measurement 167:108288. https://doi.org/10.1016/j.measurement.2020.108288

Mehdipour Ghazi M, Yanikoglu B, Aptoula E (2017) Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing. https://doi.org/10.1016/j.neucom.2017.01.018

Nilsson NJ (1981) Principles of artificial intelligence. IEEE Trans Pattern Anal Mach Intell. https://doi.org/10.1109/TPAMI.1981.4767059

Pal NR, Pal SK (1993) A review on image segmentation techniques. Pattern Recogn. https://doi.org/10.1016/0031-3203(93)90135-J

Papageorgiou CP, Oren M, Poggio T (1998) General framework for object detection https://doi.org/10.1109/iccv.1998.710772

Pereira S, Pinto A, Alves V, Silva CA (2016) Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans Med Imag. https://doi.org/10.1109/TMI.2016.2538465

Perez L, Wang J (2017) The effectiveness of data augmentation in image classification using deep learning. arXiv preprint. arXiv:1712.04621

Phillips PJ, Wechsler H, Huang J, Rauss PJ (1998) The FERET database and evaluation procedure for face-recognition algorithms. Image Vis Comput. https://doi.org/10.1016/s0262-8856(97)00070-x

Ponce J, Forsyth D (2012) Computer vision: a modern approach https://doi.org/10.1016/j.cbi.2010.05.017

Premaladha J, Ravichandran KS (2016) Novel approaches for diagnosing melanoma skin lesions through supervised and deep learning algorithms. J Med Syst 40(4):96. https://doi.org/10.1007/s10916-016-0460-2

Reichenbach SE, Park SK, Alter-Gartenberg R (1990) Optimal small kernels for edge detection https://doi.org/10.1109/icpr.1990.119330

Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation https://doi.org/10.1007/978-3-319-24574-4_28

Shawky OA, Hagag A, El-Dahshan E-SA, Ismail MA (2020) Remote sensing image scene classification using CNN-MLP with data augmentation. Optik 221:165356. https://doi.org/10.1016/j.ijleo.2020.165356

Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data. https://doi.org/10.1186/s40537-019-0197-0

Sifre L, Mallat S (2013) Rotation, scaling and deformation invariant scattering for texture discrimination https://doi.org/10.1109/CVPR.2013.163

Subramanian J, Simon R (2013) Overfitting in prediction models - is it a problem only in high dimensions? Contemp Clin Trials. https://doi.org/10.1016/j.cct.2013.06.011

Thada V, Jaglan V (2013) Comparison of jaccard, dice, cosine similarity coefficient to find best fitness value for web retrieved documents using genetic algorithm Int J Innovations Eng Technol

Uhlich S et al. (2017) Improving music source separation based on deep neural networks through data augmentation and network blending https://doi.org/10.1109/ICASSP.2017.7952158

Vyas A, Yu S, Paik J (2018) Fundamentals of digital image processing. Signals Commun Technol. https://doi.org/10.1007/978-981-10-7272-7_1

Winkler S (2013) Color space conversions. Digital Video Qual. https://doi.org/10.1002/9780470024065.app1

Yamashkin SA, Yamashkin AA, Zanozin VV, Radovanovic MM, Barmin AN (2020) Improving the efficiency of deep learning methods in remote sensing data analysis: geosystem approach. IEEE Access 8:179516–179529. https://doi.org/10.1109/ACCESS.2020.3028030

Yan Y, Zhang Y, Su N (2019) A novel data augmentation method for detection of specific aircraft in remote sensing RGB images. IEEE Access 7:56051–56061. https://doi.org/10.1109/ACCESS.2019.2913191

Yi X, Walia E, Babyn P (2019) Generative adversarial network in medical imaging: a review. Med Image Anal. https://doi.org/10.1016/j.media.2019.101552

Yu Q, Yang Y, Song YZ, Xiang T, Hospedales T (2015) Sketch-a-net that beats humans https://doi.org/10.5244/c.29.7

Zhang YD et al (2019) Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. Multimedia Tools Appl. https://doi.org/10.1007/s11042-017-5243-3

Zhong Z, Zheng L, Kang G, Li S, Yang Y (2017) Random erasing data augmentation

Zhu X, Liu Y, Li J, Wan T, and Qin Z (2018) Emotion classification with data augmentation using generative adversarial networks https://doi.org/10.1007/978-3-319-93040-4_28

Zoph B, Le QV (2019) Neural architecture search with reinforcement learning. arXiv preprint. arXiv:1611.01578