Image Augmentation Using Hybrid RANSAC Algorithm

Aswathy K Cherian
SRM Institute of Science and Technology, Kattankulathur, India.

E. Poovammal
SRM Institute of Science and Technology, Kattankulathur, India.
E-mail: jrp557791@gmail.com

Received November 05, 2020; Accepted December 15, 2020
ISSN: 1735-188X
DOI: 10.14704/WEB/V18SI02/WEB18069

Abstract

The process of augmenting the number of images in a dataset is called Image Augmentation. Data volume is essential to process and generate digital outputs from a variety of features. This work focuses on the image augmentation using a hybrid RANSAC algorithm. The features extracted is used to join or merge the images by the blending of images. The proposed RANSAC algorithm is used to extract features from four images and produce the desired mosaiced image. A mosaiced picture is best suited for aerial photos and real-world objects. The blur metric of the proposed method is 185.2587 and which is 2.86% higher than the feathering blending algorithms. The total number of images in the dataset is 2100. The number of images after augmentation is 6300 with average accuracy of 95.6%. The reported remarkable results are beneficial to all the stakeholders on image augmentation.

Keywords

Image Augmentation, RANSAC Algorithm, Augmented Reality, Image Processing, Feature Extraction.

Introduction

Image augmentation is the method of augmenting or improving the number of images for study. In certain situations where we apply deep learning approach for specific tasks, the size of the training set required is huge. In such situation there is a need to improve the dataset volume in order to successfully produce the accurate output. For example, consider the application of medical data. The availability of data of healthy person is easily available, whereas the data for unhealthy or malicious person is unavailable to that extend. In such cases, for better accuracy of the classifier, we require almost equal number of training data for both healthy and unhealthy class. This situation thus explains a clear situation where augmentation of unhealthy data is important. Similarly, another application is the case for remote sensing images – like underwater images. The dataset of images for underwater classification is less. This is mainly due to loss of image due to lighting effect, blur or dust in water. In such scenario, augmentation is required to increase the size of the training data set. In such situation data augmentation play a major role. The increase in size of dataset is mainly done by introducing a minor variation to the existing dataset. Minor changes include changing the orientation of the images, size, contrast etc. [Mikołajczyk, A., & Grochowski, M. (2018)]. These processes increase the size of the dataset and hence enable in achieving improved accuracy while the training is done with the convolutional neural network. The major advantage of using convolutional neural network for classification is because of its robust feature. The robustness of CNN is its ability to treat the augmented images of a training data as a new training data irrespective of the orientation of the image. This property is known as invariance of CNN [Perez-Munuzuri, V., Perez-Villar, V., & Chua, L.O. (1993)]. CNN is said to be invariant to translation, size change, illumination of the training data set fed into the network. This property of CNN led to the introduction of image augmentation to the real-world scenarios. Thus, image augmentation is the synthetic process of creating data set from the existing datasets. It is also known as synthetically modified data set [Zhang, K., Zuo, W., Gu, S., & Zhang, L. (2017)].

Shorten [Shorten, C., & Khoshgoftaar, T.M. (2019)] discusses a survey on image augmentation using kernel filters, random erasing, mixing images adversarial training, generative adversarial networks, color space augmentations, feature space augmentation, neural style transfer, geometric transformations and meta-learning. Frid-Adar [Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018)] presented a method for generating synthetic medical images using Generative Adversarial Networks (GANs). The synthesized data improved the performance of CNN for medical data classification. Shima [Shima, Y. (2018)] proposed a method that combines the Support vector and convolutional neural network for image category classification. Okafor [Okafor, E., Schomaker, L., & Wiering, M.A. (2018)] proposed a novel image augmentation method in which the new images produced were rotated copies of the original image. This method creates a grid matrix in which every cell contained the rotated images randomly and applied a natural background to the newly created image. The output was tested using a deep learning classifier and was found to produce higher accuracy.

Mohammed [Abdallah, Y., & Yousef, R. (2015)] proposed method to produce the gray levels in both enhanced and unenhanced images and noise variance of medical x ray images. This method used the technique of Contrast enhancement filtering and Deblurring Images Using the Blind Deconvolution Algorithm. Tang [Tang, Y.B., Oh, S., Tang, Y.X., Xiao, J.,

& Summers, R.M. (2019)] suggested the augmentation of images using generative adversarial network (GAN) to produce a large number of realistic images of CT scan of lymph node mask. A pix2pix GAN model is preferred as the strength for image generated is huge with more details about the structural and contextual information of lymph nodes and their surrounding tissues from CT scans.

Dataset

The dataset used for this study is the UC Merced land-use dataset. It consists of 21land use scene class with 100 images in class, totaling to 2100 images. Each of these images in the data set measures 256×256 pixels. The spatial resolution of these images is 0.3 m per pixel. The images were downloaded from the United States Geological Survey (USGS) National Map of different areas of the US. The 21 land-use classes are agricultural, airplane, baseball diamond, beach, buildings, chaparral, dense residential, forest, freeway, golf course, harbor, intersection, medium density residential, mobile home park, overpass, parking lot, river, runway, sparse residential, storage tanks, and tennis courts [Mikołajczyk, A., & Grochowski, M. (2018)]. The images from the class dense residential, medium residential and sparse residential are highly overlapping which differs in density thus making it challenging and richer. This dataset finds wide application in the area of remote sensing image scene classification. The data set is available for download from http://weegee.vision.ucmerced.edu/datasets/landuse.html. Fig.1 shows the sample images of each classes of the data set.

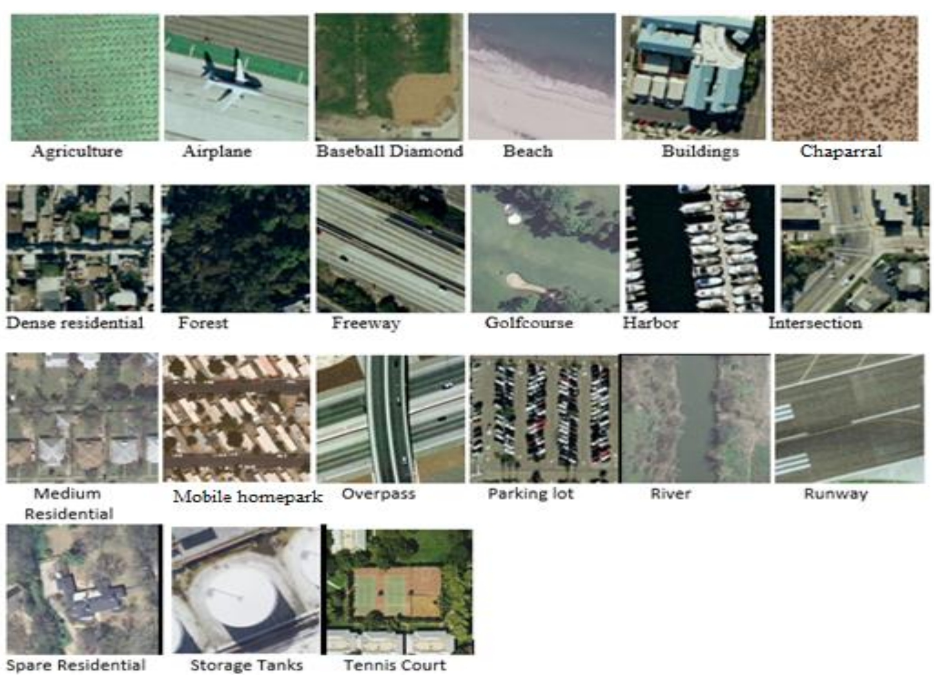


Fig. 1 Shows sample images of the above stated dataset

Methods of Data Augmentation

Traditional Methods

Traditional methods involve augmentation of the training dataset without changing the size of the image. The image size (pixel value) of the augmented images is the same as that of the existing data set. The number of images augmented using the traditional methods are twice the number of images in the original data set.

Some of the common image augmentation methods are explained in the below sections.

(A) Rotation

Rotation is the process of changing the orientation of the image. This does not alter the size of the image. Images are augmented by changing the image position by angles of 90, 45, 180 etc. Here for every image, using a single angle, another image can be produced. Thus, the data set size increases by 2 [Mikołajczyk, A., & Grochowski, M. (2018)]. Fig.2(b) shows the rotation of original image with 90 degrees.

(B) Reflection

Reflected image is produced when the reflection of the existing image is added to the data set. This is done when the image is rotated by 180 degrees. Fig.2(c) shows the reflection of original image. It is similar to rotation at 180 degrees.

(C) Flipping

Flipping of images can be done horizontally or vertically. A vertical flip is similar to rotating the image at 180 degrees and then performing horizontal flip. Flip can be a left or a right flip as well. Flipping does not include any change in size of the original image. The flipped image of ground image is shown in Fig.2 (d).

(D) Scaling

Scaling is a process that includes variation in the size of the image. The orientation of the image remains the same. This method is also called as zooming. This may be done inward or outward. An inward scaling results in image smaller in size and outward image results in image larger in size. Fig.2(e) shows the scaled image with scaling factor 0.2.

(E) Cropping

Cropping is the method of taking random sample of a section of the original image. The sample is then resized to the size of the original image [Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018)]. This is known as random cropping. Fig. 2(f) shows a sample of cropped image.

(F) Translation

Translation is the method of moving the image over the x or y direction. This method is the most used method as the object in the images can be placed anywhere in the image, thus forcing the Convolutional Neural Network to look everywhere. Fig.2 (g) is a translated image of 2.

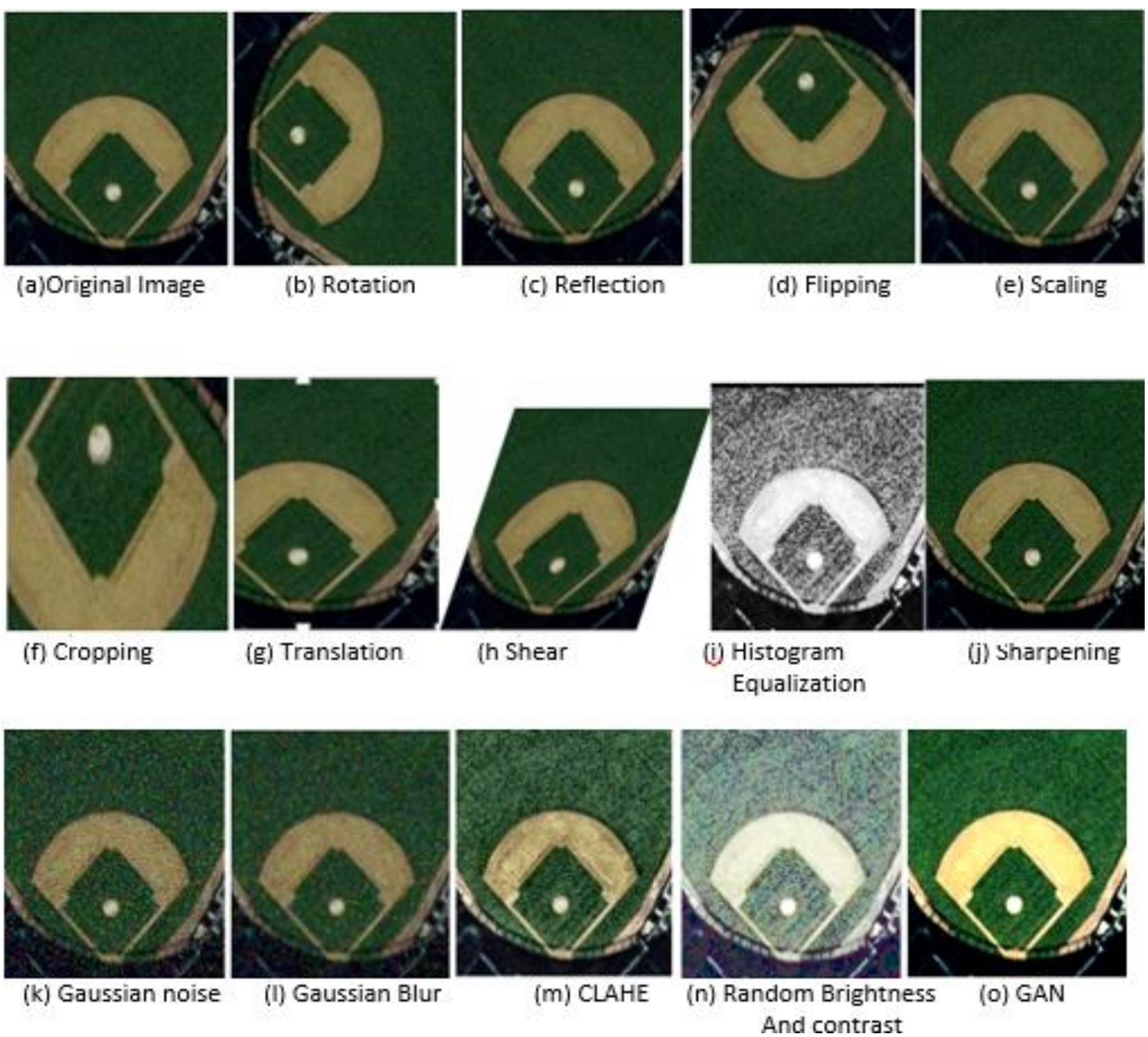


Fig. 2 Traditional and advanced Data Augmentation methods

(G) Shear

The method in which the image is augmented by keeping the base of image fixed and slanting the body of the image at different angles is called as shear. The size of the image remains the same. The sheared image of the original image is depicted in Fig.2(h).

(H) Histogram Equalization

Histogram Equalization produces a higher contrast image of the original image by taking the highest and lowest pixel value of the image to produce a subtle change in the shade. This method is mainly used for grayscale images. A histogram plot is created based on the distribution of pixel densities in an image and thus increasing the contrast of the image. Fig.2(i) is the histogram equalized image of the original image. The distribution of this histogram is then analyzed and if there are ranges of pixel brightness's that aren't currently being utilized, the histogram is then "stretched" to cover those ranges, and then is "back projected" onto the image to increase the overall contrast of the image [Shorten, C., & Khoshgoftaar, T.M. (2019].

(I) Sharpen

Sharpening is the process of increasing the contrast along edges of the images. This is done with help of kernels or filters. The filters are moved along the image with a Gaussian blur matrix to produce a blur image. Similarly, in order to produce a sharpened image, a high contrast filter is moved along vertical and horizontal edges. Also, sharpening of images results in attaining more details about the objects in the image for Data Augmentation. A typical example of sharpening of an image is shown in Fig.2(j).

B. Advanced Methods

(A) GAN

Generative Adversarial Network is a method that produces a high-resolution image using the min –max strategy [Mikołajczyk, A., & Grochowski, M. (2018)]. GANs are used for text-to-image synthesis [Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016)], super-resolution [Nair, V., & Hinton, G.E. (2010)], image-to-image translation [Isola, P., Zhu, J.Y., Zhou, T., & Efros, A.A. (2017)], image blending [Wu, H., Zheng, S., Zhang, J., & Huang, K. (2019)], and image in-painting [Yeh, R.A., Chen, C., Yian Lim, T., Schwing, A.G., Hasegawa-Johnson, M., & Do, M.N. (2017]. The GAN consists of two (G(z) and D(x)).

The Generator G(z) produces a realistic image in order to fool the other net. The discriminator D(x) is trained to distinguish fake images from the real ones Fig.3 Simplified architecture of Generative Adversarial Network. The disadvantages of GAN include poor maximum likelihood, problems with compliance with reality, lack of idea of three-dimensional perspective etc. [Goodfellow, I. (2016)]. Fig.2 (o) shows the output of an image from GAN, the augmented image has a high resolution compared to the original image.

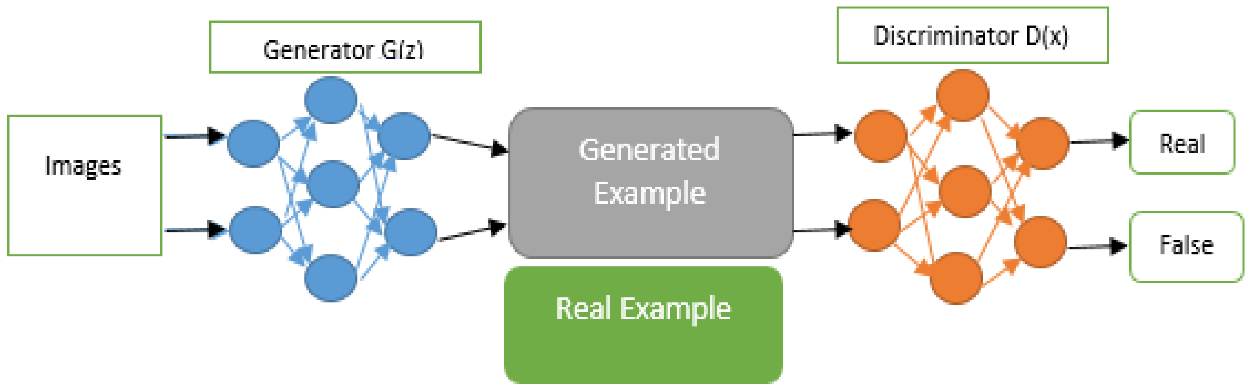


Fig. 3 Generative Adversarial Network

(B) Gaussian Noise

Gaussian Noise is a type of noise in which the probability Density Function (PDF) of the signal is equal to its Normal Distribution [Russo, F. (2003)]. This is also termed as Gaussian distribution.

For a random variable z, its probability density function (p) is given by equation 3.1.

$$p(z)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(z-\mu)^2}{2\sigma^2}} \qquad (3.1)$$

Here, standard deviation is represented as σ and mean is μ. As a part of data augmentation, the images applied with small amount of Gaussians noise can be included for classification of images. Fig.2(k) shows the noisy image with gaussian noise.

(C) CLAHE

CLAHE is a method that is applied on tiles, which are small region in the image rather than the whole image. The variance in luminance or color that makes an object different from other objects within the same field of view is defined as contrast [Setiawan, A.W., Mengko, T.R., Santoso, O.S., & Suksmono, A.B. (2013)]. The intensity of an image is adjusted by the histogram equalization which helps in enhancing the contrast of an image. In Adaptive

Histogram Equalization, every region of the image has their respective histograms thus causing every region of the image to be enhanced separately [Sahu, S., Singh, A.K., Ghrera, S.P., & Elhoseny, M. (2019)]. This also enables to redistribute the lightness values of the image. The major advantage of AHE is that it improvises the local contrast and the edge sharpness in every region of the image. Fig.2 (m) is the CLAHE applied image of the original image.

(D) Blur

Blurring is the process of smoothening the image by a function. This method is widely used to reduce noise in the image which eventually reduces the details of the image. The blur normally removes the outliers which are noise in the images [Popkin, T., Cavallaro, A., & Hands, D. (2010)]. This is done by applying low pass filter to the images. There are different types of blurring an image like the Gaussian Blur, Vector Blur, Radial Blur, Box Blur etc. The blurred image of the original data set also contribute to image augmentation. In Fig.2(l) gaussian blur is applied on the gaussian noise image.

(E) Random Brightness and Contrast

Random brightness is the process of applying random brightness/contrast to images. This type of augmentation includes changing of colors. Though the contrast is affected, the textual features of the image remains unchanged. Four aspects of an image can be altered: contrast, hue, brightness and saturation. Fig.2(n) shows the image augmented with brightness value of 10.

C. Existing Algorithm - SIFT Algorithm

Feature extraction refers the process of extracting features from the images that are unique to the images. Mosaicing of images is a technique to join images in order to produce a resultant image which is has all the features of the original images. In mosaicing the elements of the input images which could be matched together are used as features of the images. For this purpose, normally patches which are group of pixels are taken from the images. The key point descriptors contain the image features. In order to match another image for mosaicing, it checks for similar key-point descriptors using the Nearest Neighbor Search (NNS) for the blending. The key point descriptors are the potential interest points of the image which are detected by scaling over the input image using a difference of Gaussian function. The scale and location for all interest points is determined. A stable key point will withstand distortion of the images. The direction gradient of the key point is determined by the SIFT algorithm in order to assign orientation [Shi, G., Xu, X., & Dai, Y.

(2013)]. Based on the gradient direction of the image, the key points are assigned. The Best Bin First (BBF) algorithm is used for image blending by estimating the initial matching points between the input frames. SIFT algorithm being invariant in both rotation and scaling, is very suitable for object recognition in images with high resolution. It is also a robust algorithm for image comparison though the computation time for comparison is high.

Algorithm of SIFT Framework

1. Determine the SIFT key point which is the minima/maxima of Difference of Gaussian (DoG) taken at different scales of the images.

$$DoG(x,y,\sigma) = L(x,y,ki\sigma) - L(x,y,kj\sigma) \quad (3.2)$$

where L (x,y,kσ) is the convolution of the image at scale kσ.

2. Predict the nearby best fit of the keypoints for accurate location and scale. Thus, rejecting points with contrast mismatch and outliers.
3. The accurate position is calculated by interpolation.
4. Discard value of low contrasts.
5. Orientation Assignment for an image sample L (x, y) at scale σ, the gradient magnitude m (x, y) and orientation α (x, y)

$$m(x,y)=\sqrt{(L(x+1,y)-L(x-1,y))^2+(L(x,y+1)-L(x,y-1))^2} \quad (3.3)$$

$$\alpha(x,y)=\mathrm{atan2}(L(x,y+1) - L(x,y-1), L(x+1,y) - L(x-1,y)) \text{ Eq} \quad (3.4)$$

Proposed Method – Hybrid RANSAC

Image mosaicing [Vaghela, D., & Naina, P. (2014)] is the processes of fusing two or more pictures/images to form a single image. The resulting image is called as the mosaiced image. Fig.4 shows the flow chart of the proposed hybrid method of image mosaicing. The first step is the Feature Extraction which extracts features from the input images. The extracted features are then used to join or merge the images. This process is called the blending of images. In our proposed system we have used the hybrid RANSAC algorithm to extract features and blend the images. RANSAC algorithm has been used to blend 2 images to form the mosaiced image. In the proposed system, up to 4 images are merged to form the mosaiced image. This enables the classifier to learn more features from the resultant image to improve the accuracy of training. The existing method for image mosaicing is the SIFT algorithm [Pandey, A., & Pati, U.C. (2013)]. The major disadvantage of the SIFT algorithm is the complicated and high computational speed and also does not work well for blurred images. A mosaiced image is best suited for Aerial Images, Real World Objects and Low Object distribution. Mosaicing cannot be applied for Written Documents, Large and prominent objects and fixed location objects. A detailed explanation on each of these methods is made on the following sections.

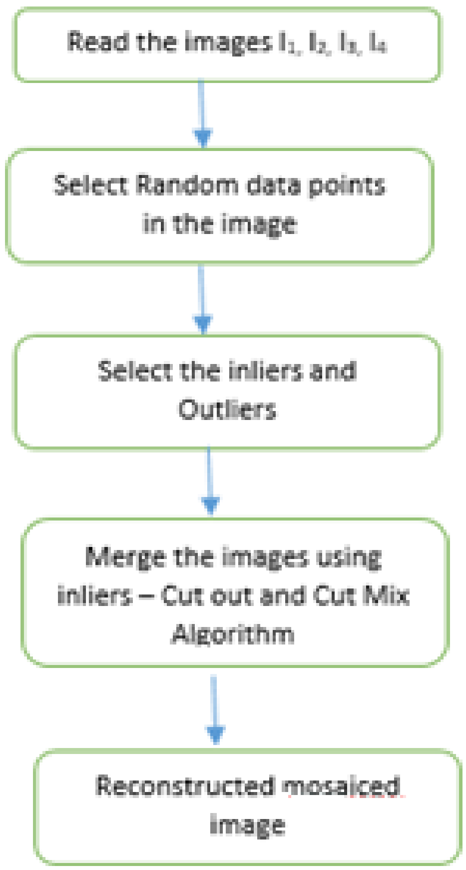
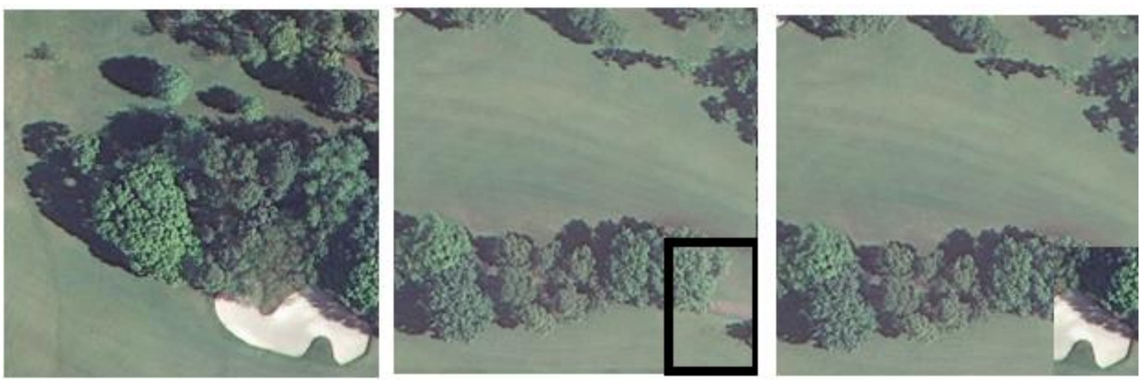


Fig. 4 Flowchart of the proposed image mosaicing approach

The RANSAC "RANdom Sample Consensus" is an algorithm for robust fitting of models in the presence of many data outliers. RANSAC is a hybrid algorithm cable of performing both homography followed by image blending [Bochkovskiy, A., Wang, C.Y., & Liao, H.Y.M. (2020)]. Homography is a transformation matrix which maps the points in one image to the corresponding points to the other image. The basic property of a RANSAC is to eliminate all outliers of the images to be merged. It follows an iterative method to predict the parameters. RANSAC is a robust algorithm which create a best model with all the data outliners of the image. The RANSAC algorithm is best suited for feature matching. The main strategy of RANSAC is to determine the best transformation that includes the greatest number of match features (inliers). The initial step of RANSAC is to compute all inliers of the image.

Image warping is the process of digitally processing a photo such that any shapes portrayed in the photo have been significantly distorted [Lati, A., Belhocine, M., & Achour, N. (2020)]. Wrapping is followed by blending which vanishes the seams/distortions seen in the image after wrapping. The simplest step is to use the weighted averaging color values to blend the overlapping pixels. In the proposed method, we have implemented a hybrid algorithm combining RANSAC with Cut-Out CutMix procedure (Fig.5). The image 5 (a) and 5 (b) are input images from the original dataset. 5 (c) is the augmented image with feature mapping technique of the proposed method. The rectangle in the image 5(b) is the cut part and mixed with image 5(a). The CutMix procedure which combines images by cutting parts from one image and pasting them onto the augmented image.



(a) (b) (c)

Fig. 5 Cut Out and Cut Mix Method of the proposed algorithm in a sample image

The cutout image enables the model to learn to predict with help of a robust number of features. In CutMix, the cutout is replaced with a part of another image along with the second image's ground truth labeling. For this the image is first resized to equal sizes. [Derpanis, K.G. (2010)] A mosaiced image is formed by joining 4 ground images. The ground images are cut to fit the same size as the original image. That is, size is 1/4th the original image. This method is capable of producing any number of mosaiced images. If the augmentation value is set to 2, then it would produce double the images in the data set. Similarly, if augmentation is set to 3, then a new data set with augmented images trice that of the dataset is produced. A sample mosaiced image on applying RANSAC CutMix is shown in the Fig.6.

Algorithm of RANSAC

RANSAC algorithm can be summarized as follows:

1. Choose the images to determine the transformation parameters. (For image mosaicing; select four images).

2. Predict the homography parameters.
3. Determine the number of key points from the set using a threshold value.
4. If the value of total count of inliers by the total number of points is more that the threshold value set in step 3, then consider the identified inliers and terminate.
5. Otherwise, repeat steps 1 through 4. Repeat for N iterations where n is number of key points.

The outstanding property of RANSAC is its robust estimation of parameters even if the number of outliers is high in the dataset [Shi, G., Xu, X., & Dai, Y. (2013)]. It is very simple and general and also applicable to different kinds of problems. The output of RANSAC is considerably higher accuracy compared to the existing algorithms. The major drawback of RANSAC is its high computational speed and also may not produce the best model if the number of iterations is limited. Another disadvantage of RANSAC is that it requires the setting of problem-specific thresholds.

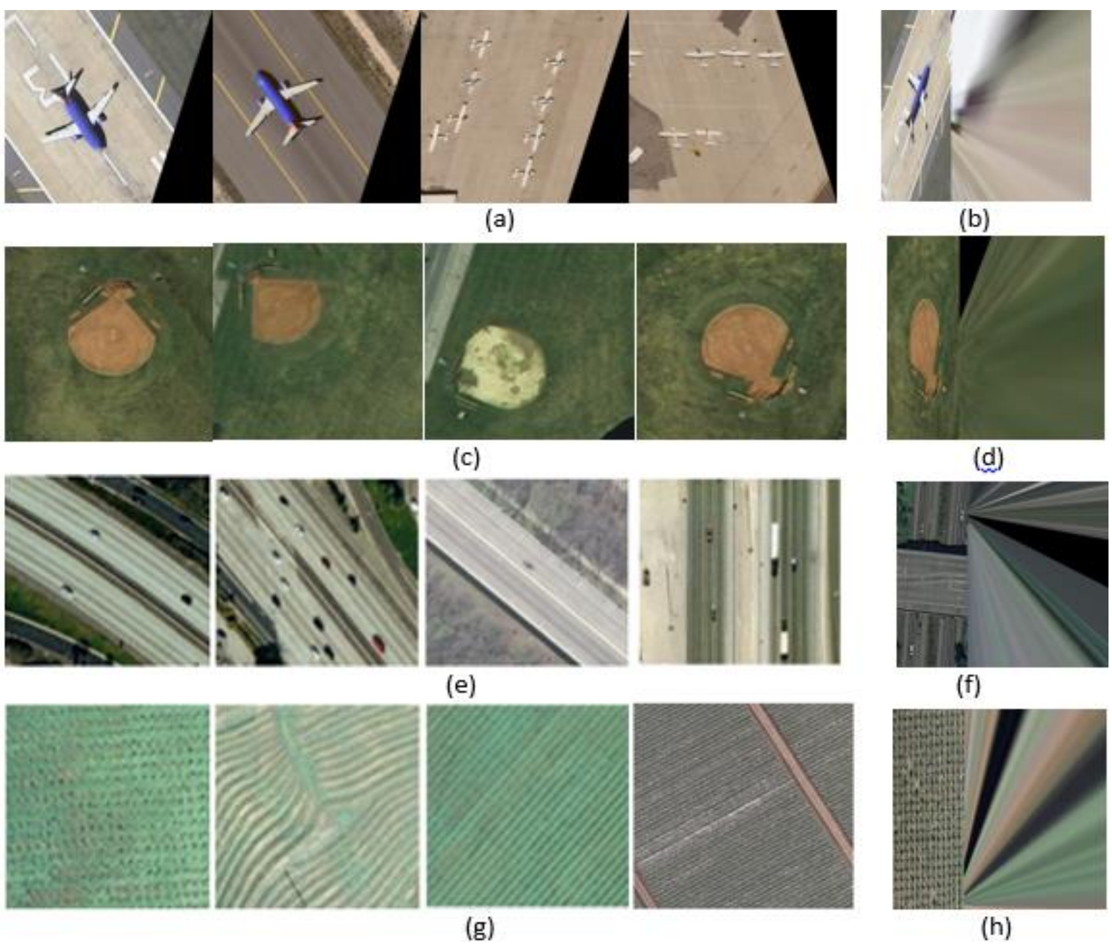


Fig. 6 Sample 4 input images from (a) Vegetation dataset (c) Airplane Dataset (e) Intersection Dataset (g) River Dataset and (b), (d), (f), (h) are the mosaiced images of the respective dataset using the proposed method

```
Input:
m : No of data points
k : Maximum number of iterations
t : threshold value to determine data points
that best fit the model
d : no of close

result:
para_model : model parameters which best
fit the data

Start:
    Initialize iterations = 0

do
   Inliers : n randomly selected values from
data.
   Model : model parameters fitted to Inliers.
   Set Inlier_set = 0


   for every point in data not in Inliers do
        if point fits Model with an error less
than threshold t
              add point to Inliers_set
   end for

     if the number of points in the Inlier_set >
  d, then
          //found good model
         betterModel = model parameters
  fitted to all points in Inliers and Inliers_set
         Err = a measure of how well
  betterModel fits these points
           if Err < bestErr then
             para_model:= betterModel
             bestErr :=Err
          end if
     end if
  increment iterations till iterations<k

  end while
  return para_model
```

Classification of Images Using CNN

Classification of images produces the at most accuracy using a Convolutional Neural Network. Convolution Neural Network is an artificial neural network with a number of layer of units called neurons. These neurons are similar to the neurons of the nervous system of human being. For a good classification, the training samples should be at least nine times the number of parameters with good resolution. A multi layered neural network is called as a convolutional neural network. For our experiment, we have used the AlexNet for classification of the images. The architecture of AlexNet consists of 8 layers: five convolutional layers and three fully-connected layers. The multiple convolutional kernels are used to extract features from the image. The first convolutional layer of AlexNet consists of 96 kernels of size 11x11x3. In an AlexNet the length and width are similar, whereas the depth refers to the number of channels [Krishna, M.M., Neelima, M., Mane, H., & Matcha, V.G.R. (2018)]. An overlapping max pool layer is attached after the second and fifth convolutional layers. The output of AlexNet goes through a series of fully connected layers.

Fig.7 shows sample AlexNet Used in our experiment. After every layer of the convolution, a ReLU activation function is applied. Max Pooling layers enable to reduce the width and height of the tensor, whereas the depth is unchanged. Overlapping Max Pool layers are done by overlapping the adjacent window on which the max is computed. Thus, the architecture uses windows of size 3×3 with a stride of 2 between the adjacent windows. The application of ReLU with Nonlinearity is an important feature of AlexNet. This makes the CNN to train faster than other activation functions like tanh or sigmoid.

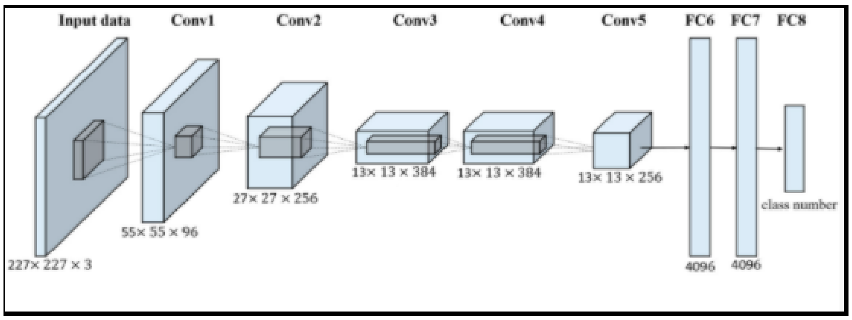


Fig. 7 Architecture of AlexNet

Result and Analysis

The significant results of the proposed hybrid RANSAC algorithm are presented in this section. The experiment is carried out in one of the images of the dataset. The quality of the image is assessed using the PSNR value, Entropy and Blur Metric on the mosaiced image shown in Fig.6. Table 1 shows the qualitative assessment of the blending algorithms. The blur metric of the proposed method is 185.2587 and which is 2.86% higher than the existing type feathering and SIFT blending algorithms.

Table 2 indicates the percentage increase in the data of the image augmented. The proposed method yielded a 300% increase in image augmentation data compared to other traditional and advanced techniques.

Table 1 Qualitative Assessment

Blending Algorithms	Quality Assessment		
	PSNR (dB)	Entropy	Blur Metric
Feathering	15.4029	5.9585	180.1148
SIFT Algorithm	15.6886	6.2056	184.9643
Proposed Method	16.0124	6.3257	185.2587

This enables the classifier to achieve more accuracy in training model compared with model of traditional dataset.

Table 2 Percentage of increase in images for Augmentation

Method	**Number of images in the dataset**	**Number of images after Augmentation**	**Percentage of Data increase**
Traditional Methods* (9 Methods)	2100	2100	100%
Advanced Method**(5 Methods)	2100	2100	100%
Proposed Method (augmentation = 3)	2100	6300	300%

**Rotation, flipping, Shear, Reflection, Scaling, Cropping, translation, Histogram Equalization, Sharpening.*

*** Gaussian noise, GAN, CLAHE, Gaussian Blur, Random brightness and Contrast.*

Table 3 shows the accuracy of classification of images augmented using different methods. The proposed approach by hybrid RANSAC produced an accuracy of 95.6% and the highest among the other methods.

Table 3 Accuracy of classification Algorithm

Method	**Number of images in the dataset**	**Number of images after Augmentation**	**Average Accuracy**
Traditional Methods	2100	2100	92.8%
Advanced Method	2100	2100	94.3%
Proposed Method (augmentation = 3)	2100	6300	95.6%

Another experimental result to claim is in terms of the quality of image and the computational speed. Though the proposed method is a robust algorithm, it claims to have a higher computational time and thus the speed of computation is low. This is one of the major drawbacks of the RANSAC algorithm. The results depict that the SNR value of the produced image is higher than the traditional existing system, but the computational time is high. This drawback is overcome when the number of images augmented by the proposed algorithm has a better resolution and quality, thus enabling the classifier to extract all features of the images during the training of data.

Conclusions

Image augmentation is the process of increasing the size of the dataset artificially from the existing dataset. In this paper, a hybrid RANSAC technique for augmenting images is proposed. The RANSAC algorithm estimates the parameters with a high degree of accuracy even when a substantial number of outliers existing in the data set. The major conclusions are provided below.

The feature extraction is used to join or merge the images by the blending of images. The blur metric of the proposed method is 185.2587 and which is 2.86% higher than the feathering blending algorithms.

A mosaiced picture is best suited for aerial photos and real-world objects. The proposed method yielded a 300% increase in image augmentation data compared to other traditional and advanced techniques.

The number of images in the dataset is 2100. The number of pictures after augmentation is 6300, with an average accuracy of 95.6%.

The reported results are further fine-tuning to deploy in real-world environments. The proposed image augmentation methodology's significant findings are most beneficial to all the community stakeholders who work on applications with limited data set (images).

References

Mikołajczyk, A., & Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. *In international interdisciplinary PhD workshop (IIPhDW)*, 117-122.

Perez-Munuzuri, V., Perez-Villar, V., & Chua, L.O. (1993). Autowaves for image processing on a two-dimensional CNN array of excitable nonlinear circuits: flat and wrinkled labyrinths. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, *40*(3), 174-181.

Zhang, K., Zuo, W., Gu, S., & Zhang, L. (2017). Learning deep CNN denoiser prior for image restoration. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 3929-3938.

Shorten, C., & Khoshgoftaar, T.M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), 1-48.

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, 321*, 321-331.

Shima, Y. (2018). Image augmentation for object image classification based on combination of pre-trained CNN and SVM. *Journal of Physics: Conference Series, Volume 1004, 2nd International Conference on Machine Vision and Information Technology (CMVIT 2018), 1004*, 1-8.

Okafor, E., Schomaker, L., & Wiering, M.A. (2018). An analysis of rotation matrix and colour constancy data augmentation in classifying images of animals. *Journal of Information and Telecommunication, 2*(4), 465-491.

Abdallah, Y., & Yousef, R. (2015). Augmentation of X-rays images using pixel intensity values adjustments. *International Journal of Science and Research (IJSR), 4*(2), 2425-2430.

Tang, Y.B., Oh, S., Tang, Y.X., Xiao, J., & Summers, R.M. (2019). CT-realistic data augmentation using generative adversarial network for robust lymph node segmentation. *In Medical Imaging 2019: Computer-Aided Diagnosis, 10950, 109503. International Society for Optics and Photonics.*

http://weegee.vision.ucmerced.edu/datasets/landuse.html

Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016). Generative adversarial text to image synthesis. In *International Conference on Machine Learning*, 1060-1069. PMLR.

Nair, V., & Hinton, G.E. (2010). Rectified linear units improve restricted boltzmann machines. *Proceedings of the 27th International Conference on Machine Learning*, Haifa, Israel, 2010.

Isola, P., Zhu, J.Y., Zhou, T., & Efros, A.A. (2017). Image-to-image translation with conditional adversarial networks. *In Proceedings of the IEEE conference on computer vision and pattern recognition,* 1125-1134.

Wu, H., Zheng, S., Zhang, J., & Huang, K. (2019). Gp-gan: Towards realistic high-resolution image blending. *In Proceedings of the 27th ACM international conference on multimedia,* 2487-2495.

Yeh, R.A., Chen, C., Yian Lim, T., Schwing, A.G., Hasegawa-Johnson, M., & Do, M.N. (2017). Semantic image in painting with deep generative models. *In Proceedings of the IEEE conference on computer vision and pattern recognition,* 5485-5493.

Goodfellow, I. (2016). *Nips 2016 tutorial: Generative adversarial networks.* arXiv preprint arXiv:1701.00160.

Russo, F. (2003). A method for estimation and filtering of Gaussian noise in images. *IEEE Transactions on Instrumentation and Measurement, 52*(4), 1148-1154.

Setiawan, A.W., Mengko, T.R., Santoso, O.S., & Suksmono, A.B. (2013). Color retinal image enhancement using CLAHE. *In International Conference on ICT for Smart Society*, 1-3.

Sahu, S., Singh, A.K., Ghrera, S.P., & Elhoseny, M. (2019). An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. *Optics & Laser Technology, 110,* 87-98.

Popkin, T., Cavallaro, A., & Hands, D. (2010). Accurate and efficient method for smoothly space-variant Gaussian blurring. *IEEE Transactions on image processing, 19*(5), 1362-1370.

Shi, G., Xu, X., & Dai, Y. (2013). SIFT feature point matching based on improved RANSAC algorithm. *In 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, 1*, 474-477.

Vaghela, D., & Naina, P. (2014). *A review of image mosaicing techniques.* arXiv preprint arXiv:1405.2539.

Pandey, A., & Pati, U.C. (2013). A novel technique for non-overlapping image mosaicing based on pyramid method. *In Annual IEEE India Conference (INDICON)*, 1-6.

Bochkovskiy, A., Wang, C.Y., & Liao, H.Y.M. (2020). *Yolov4: Optimal speed and accuracy of object detection.* arXiv preprint arXiv:2004.10934.

Lati, A., Belhocine, M., & Achour, N. (2020). Robust aerial image mosaicing algorithm based on fuzzy outliers rejection. *Evolving Systems*, *11*(4), 717-729.

Derpanis, K.G. (2010). Overview of the RANSAC Algorithm. *Image Rochester NY*, *4*(1), 2-3.

Shi, G., Xu, X., & Dai, Y. (2013). SIFT feature point matching based on improved RANSAC algorithm. *In 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, 1,* 474-477.

Krishna, M.M., Neelima, M., Mane, H., & Matcha, V.G.R. (2018). Image Classification using Deep Learning. *International Journal of Engineering & Technology, 7*(2.7) 614-617.