Published in final edited form as:

Med Image Comput Comput Assist Interv. 2019 October ; 11764: 201–208. doi:
10.1007/978-3-030-32239-7_23.

Active Appearance Model Induced Generative Adversarial Network for Controlled Data Augmentation

Jianfei Liu, Christine Shen, Tao Liu, Nancy Aguilera, Johnny Tam

National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Abstract

Data augmentation is an important strategy for enlarging training datasets in deep learning-based medical image analysis. This is because large, annotated medical datasets are not only difficult and costly to generate, but also quickly become obsolete due to rapid advances in imaging technology. Image-to-image conditional generative adversarial networks (C-GAN) provide a potential solution for data augmentation. However, annotations used as inputs to C-GAN are typically based only on shape information, which can result in undesirable intensity distributions in the resulting artificially-created images. In this paper, we introduce an active cell appearance model (ACAM) that can measure statistical distributions of shape and intensity and use this ACAM model to guide C-GAN to generate more realistic images, which we call A-GAN. A-GAN provides an effective means for conveying anisotropic intensity information to C-GAN. A-GAN incorporates a statistical model (ACAM) to determine how transformations are applied for data augmentation. Traditional approaches for data augmentation that are based on arbitrary transformations might lead to unrealistic shape variations in an augmented dataset that are not representative of real data. A-GAN is designed to ameliorate this. To validate the effectiveness of using A-GAN for data augmentation, we assessed its performance on cell analysis in adaptive optics retinal imaging, which is a rapidly-changing medical imaging modality. Compared to C-GAN, A-GAN achieved stability in fewer iterations. The cell detection and segmentation accuracy when assisted by A-GAN augmentation was higher than that achieved with C-GAN. These findings demonstrate the potential for A-GAN to substantially improve existing data augmentation methods in medical image analysis.

Keywords

Generative adversarial network; Active appearance model; Data augmentation; Cell segmentation; Cell detection; Adaptive optics retinal imaging

1 Introduction

Deep learning has fundamentally changed the field of medical image analysis. Central to many deep learning-based applications is data augmentation [7,16], an important step because annotated medical imaging datasets are not only difficult and costly to generate, but also because they quickly become obsolete as imaging technology improves. Existing

methods for data augmentation are based on operations such as image rotation, shearing, translation, flipping, and Gaussian blurring [7]. Image retrieval has also been suggested for augmenting medical data through the selection of unlabeled medical images that are similar to labeled ones [16]. However, it is rare that these data augmentation strategies alone can lead to large improvements in results.

Recently, generative adversarial networks (GAN) [5], which combine generative and discriminative networks to generate data similar to real data, have shown great promise for data augmentation in medical imaging analysis [8], as in the case of spine segmentation [6] or liver lesion classification [4]. Among GAN techniques, image-to-image conditional GAN (C-GAN) [9] is of particular interest because it can generate artificial medical images from input annotations, which can substantially reduce the need for a large database of human annotations. For example, C-GAN has been adapted to transfer annotations between cardiovascular CT and MRI images and to align segmentations between two modalities [17], as well as to pair PET/CT and MRI/CT images [1]. Artificial lesions have also been created through C-GAN to balance training samples when there were insufficient numbers of pathological cases [10,15].

However, these existing data augmentation approaches are based primarily on shape information, which can result in unrealistic intensity distributions when there are anisotropic intensity distributions (e.g. adaptive optics retinal imaging [13,14], a rapidly-advancing technology for visualizing photoreceptor cells, shown in Fig. 1A). Images of photoreceptor cells exhibit low image contrast at boundaries (white arrows, Fig. 2), which is a common feature of many medical imaging modalities, including MRI and CT images. Whereas data augmentation often utilizes arbitrary data transformations that could easily give rise to datasets that are not representative of real data (e.g. different shape distributions that do not fall within the expected characteristics of real data), we propose that a more realistic approach would be to generate a population of images that captures the natural variations observed within real data. An additional challenge is that it is difficult for C-GAN to capture strong anisotropic intensity distributions because this information is not contained within binary cell contour input annotations. To overcome these challenges, we propose the use of an active cell appearance model (ACAM) [3] to statistically model the shape and intensity distributions of cells. We integrate ACAM with C-GAN, which we call A-GAN.

The contributions of A-GAN are as follows. A-GAN can utilize both shape information as well as intensity distributions for input annotations, a key step for generating more realistic images. This intensity information is useful not only for forcing the generative network to produce objects with similar intensities, but also for helping to stabilize the generative process. Augmented data from A-GAN is more representative of the original population of images because the transformation is applied locally to each cell in a manner that is consistent with the statistical distributions contained in ACAM, as opposed to arbitrary transformations that are applied globally. Evaluation of A-GAN for cell detection and segmentation from medical imaging data showed a clear improvement in performance using A-GAN-based data augmentation compared to existing methods.

2 Methods

An overview of A-GAN is shown in Fig. 1. The two main elements are (1) ACAM to compute the statistical distributions of cell contour shapes and intensities and (2) an improved C-GAN to generate realistic cell images by importing cell contours with their respective intensity information.

2.1 Active Cell Appearance Model

The goal of this step is to create ACAM in order to compute statistical distributions of cell shapes and intensities in medical imaging data. A set of cell contours and corresponding image patches are extracted from annotated medical images (Figs. 1A and B). Along each contour, N evenly-spaced landmarks (circles, bottom row of Fig. 1B) are defined as a one-dimensional vector $\mathbf{s} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$. Given M training contours, $\mathcal{S} = \{\mathbf{s}_1, \dots, \mathbf{s}_M\}$, we can derive a linear contour shape model $\{\bar{\mathbf{s}}, \mathbf{U}_s^k\}$, where $\bar{\mathbf{s}}$ is the mean shape vector of $\mathcal{S}$ and $\mathbf{U}_s^k$ is an orthonormal basis of the first k eigenvectors using principal component analysis (PCA). The corresponding k eigenvalues are in the matrix Λ_s^k . A new cell shape instance can be generated as $\mathbf{s}(\mathbf{p}) = \bar{\mathbf{s}} + \mathbf{U}_s^k \cdot \mathbf{p}$, where $\mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_k)$ is a shape parameter vector (Fig. 1C).

In addition to shape, anisotropic intensity information at cell contours needs to be embedded into ACAM. We follow a similar procedure as was done for shape. For a set of image patches from $\mathcal{S}$ defined as $\mathcal{I} = \{I_1, \dots, I_M\}$, we again sample a set of sub-image patches $\mathcal{L}_i = \{L_{i,1}, \dots, L_{i,N}\}$ of I_i around all contour points of $\mathbf{s}_i$ and align all $\mathcal{L}$ because we are concerned with the intensity distribution near cell contours. Vectoring $\mathcal{L}_i$ as $\mathbf{a}_i$ and applying the PCA results $\{\bar{\mathbf{a}}, \mathbf{U}_a^k\}$, where $\bar{\mathbf{a}}$ is the mean appearance vector and $\mathbf{U}_a^k$ is the orthonormal basis of the first k eigenvectors, we obtain a second eigenvalue matrix Λ_a^k . A new appearance instance can be created as $\mathbf{a}(\mathbf{t}) = \bar{\mathbf{a}} + \mathbf{U}_a^k \cdot \mathbf{t}$, where $\mathbf{t} = (t_1, \dots, t_k)$ is an appearance parameter.

Using these two components of ACAM, we can create a new cell with shape and appearance according to their statistical models, which when repeated over a large number of cells results in a more realistic variation in cell parameters. To create a contour image with a defined intensity distribution, we randomly assign a set of image points into the image plane with a separation distance that is close to the cell diameter. These cell points are used as the cell centroids. A cell instance is then created by sampling $\mathbf{p} \in [-\Lambda_s^k, \Lambda_s^k]$ and $\mathbf{t} \in [-\Lambda_a^k, \Lambda_a^k]$, and translating it to a cell centroid. Thus, we can create a set of contours with intensity values for C-GAN.

2.2 Conditional Generative Adversarial Networks

This step incorporates the image of contours generated from ACAM to guide a C-GAN resulting in A-GAN. Similar to C-GAN [9], A-GAN utilizes U-Net [12] as the generative network and PatchGAN [9] as the discriminative network. However, a key difference between C-GAN and A-GAN is the objective function to optimize networks which is

different because the input of A-GAN carries partial intensity information for the image to be generated.

During training, a contour image C (bottom image, left panel of Fig. 1D) is created by assigning intensity values from real images to contour positions while all other positions are assigned with the background value (0 in this case). The main purpose of A-GAN is thus to learn a mapping from C and a random noise vector $\mathbf{n}$ to a real image of cells, I , i.e., $G: \{C, \mathbf{n}\} \rightarrow I$. The generative network G attempts to produce realistic medical images I that are indistinguishable from the adversely discriminative network D , while D tries to improve its ability to flag fake medical images from G during the adversarial process. We define the first component of the objective of A-GAN as

$$F_{GAN}(G, D) = \mathbb{E}_{C, I}[\log D(C, I)] + \mathbb{E}_{C, \mathbf{n}}[\log(1 - D(C, G(C, \mathbf{n})))] \quad (1)$$

where G attempts to minimize Eq.1 while D tries to maximize it. Next, A-GAN also incorporates an L1 distance to encourage less blurring (similar to C-GAN),

$$F_{L1}(G) = \mathbb{E}_{C, I, \mathbf{n}}[\|I - G(C, \mathbf{n})\|_1] \quad (2)$$

Finally, since C contains the partial intensity information of I at cell boundaries, it is also used as an additional term for the objective function of A-GAN,

$$F_C(G) = \mathbb{E}_{C, \mathbf{n}}[\|C - G(C, \mathbf{n})\|_1], \text{ where } C \neq 0 \quad (3)$$

Here, $C \neq 0$ means Eq.3 only considers the image points that do not belong to the background. Combining Eqs.1–3 leads to the objective function for A-GAN:

$$F(G, D) = F_{GAN}(G, D) + \lambda_1 F_{L1}(G) + \lambda_2 F_C(G) \quad (4)$$

where λ_1 and λ_2 are two constants to balance the terms in Eq.4. The expected G is $G^* = \arg \min_G \max_D F(G, D)$.

During testing, the contour image created by ACAM is used as the input for A-GAN to produce realistic images (right panel, Fig. 1D). Importantly, the contour image from ACAM contains intensity information, which is a key feature of our approach.

2.3 Experiments and Validation Methods

Medical imaging data obtained using an adaptive optics ophthalmoscope [13,14] from 20 human subjects (age: 24.8 ± 5.1 years) was used to validate A-GAN. We created four groups of training datasets with 32, 64, 128, and 256 medical images (150×150 pixels), as well as one separate group of test data with 256 images. The purpose of using a small number of training images was to facilitate the observation of data augmentation effects. All images were randomly selected from these subjects and then cells were annotated by experienced graders.

To evaluate the robustness of A-GAN, we generated four different models of ACAM from the four training groups. ACAM models were used to generate 32,64,128, and 256 contour images with and without intensity values. Paired image sets with and without intensity values were used in A-GAN and C-GAN, respectively. In this manner, we obtained four pairs of results from C-GAN and A-GAN. We visually evaluated the robustness of these GANs using the generated images.

We assessed whether A-GAN-based data augmentation improved cell detection and cell segmentation results. Here, we used U-Net to perform cell detection by computing the centroid of each cell from the annotated contour images and creating centroid images as the labeled training data. Similarly, cell segmentation was performed using the contour image as the labeled training data. The test dataset was used to evaluate detection and segmentation accuracy. For each training dataset, we also compared three different types of U-Net trained with real medical images incorporating conventional data augmentation, doubled medical images from C-GAN, and doubled medical images from A-GAN. Recall and precision were used to evaluate detection accuracy and Dice coefficient and area difference for segmentation accuracy.

3 Experimental Results

3.1 Robustness Evaluation

Across the four training groups, A-GAN quickly achieved stability and generated a realistic medical image with only 20 iterations, while C-GAN resulted in many unrealistic cells even after 400 iterations (Fig. 2). These results clearly demonstrate that intensity information from ACAM is a key factor for stabilizing the generative network and improving the accuracy of generated images.

3.2 Cell Detection

Improvements in detection accuracy through data augmentation of medical images was most readily achieved through A-GAN when compared to both conventional and C-GAN-based data augmentation (Table 1). For conventional data augmentation we applied random rotations, translations, scaling, and shearing to the images. The highest recall and precision values were achieved through A-GAN-based data augmentation.

3.3 Cell Segmentation

A-GAN-based data augmentation also resulted in the highest segmentation accuracy (Table 2). Despite the extremely low number of images used for training (as few as 32), A-GAN-based data augmentation was able to outperform existing methods that utilized 8 times more training data (256 images), which demonstrates the effectiveness of our approach. Similar to above, conventional data augmentation was also based on applying random rotations, translations, scaling, and shearing.

4 Conclusion and Future Work

This work introduces A-GAN, an improved GAN method to augment medical images in a controlled manner that converges to realistic images faster than existing methods and that generates images with characteristics that are statistically representative of real data. This is achieved by using an active cell appearance model (ACAM) to constrain cell shapes and appearances in a statistical space. Unlike arbitrary global data transformation by conventional data augmentation, cell operations based on ACAM are performed locally with variances that are necessarily representative of real data. The intensity values of contours from ACAM are a key factor in defining the object of A-GAN as well as in stabilizing the image generation process. The improvement in cell detection and segmentation accuracy from data augmentation (A-GAN) is substantially greater than that from conventional data augmentation and C-GAN. The proposed methods may be useful for other applications, including fluorescent cell analysis [11], spine segmentation [6], or melanoma segmentation [2], in which anisotropic intensity distributions can be modeled using active appearance models [3]. These results demonstrate that A-GAN is an improved, efficient strategy for augmenting medical imaging data in situations where pixel intensity information is critical in addition to shape information.

Acknowledgments.

This work was supported by the Intramural Research Program of the National Institutes of Health, National Eye Institute.

References

1. Armanious K, Yang C, Fischer M, et al.: MedGAN: medical image translation using GANs. CoRR abs/1806.06397 (2018)
2. Chi Y, Bi L, Kim J, Feng D, Kumar A: Controlled synthesis of dermo-scopic images via a new color labeled generative style transfer network to enhance melanoma segmentation. In: Proceedings of the IEEE Conference on Engineering in Medicine and Biology Society, pp. 2591–2594 (2018)
3. Cootes T, Edwards G, Taylor C: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell 23(6), 681–685 (2001)
4. Frid-Adar M, Diamant J, Klang E, et al.: GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing 321, 321–331 (2018)
5. Goodfellow I, Pouget-Abadie J, Mirae M, et al.: Generative adversarial networks. In: NIPS, pp. 2672–2680 (2014)
6. Han Z, Wei B, Mercado A, et al.: Spine-GAN: semantic segmentation of multiple spinal structures. Med. Image Anal 50, 23–35 (2018) [PubMed: 30176546]
7. Hussain Z, Gimenez F, Yi D, Rubin D: Differential data augmentation techniques for medical imaging classification tasks. In: AMIA Annual Symposium Proceedings, pp. 979–984 (2017) [PubMed: 29854165]
8. Iqbal T, Ali H: Generative adversarial network for medical images (MI-GAN). J. Med. Syst 42(11), 1–11 (2018)
9. Isola P, Zhu J, Zhou T, Efros A: Image-to-image translation with conditional adversarial networks. In: IEEE CVPR (2017)
10. Liu S, Gibson E, Grbic S, et al.: Decompose to manipulate: manipulable object synthesis in 3D medical images with structured image decomposition. CoRR abs/1812.01737 (2018)

11. Ramesh N, Tasdizen T: Cell segmentation using a similarity interface with a multi-task convolutional neural network. *IEEE J. Biomed. Health Inform* 23(4), 1457–1468 (2019) [PubMed: 30530343]
12. Ronneberger O, Fischer P, Brox T: U-Net: convolutional networks for biomedical image segmentation In: Navab N, Hornegger J, Wells WM, Frangi AF (eds.) *MICCAI 2015. LNCS*, vol. 9351, pp. 234–241. Springer, Cham (2015). 10.1007/978-3-319-24574-4_28
13. Roorda A, Duncan J: Adaptive optics ophthalmoscopy. *Ann. Rev. Vis. Sci* 1, 19–50 (2015) [PubMed: 26973867]
14. Scoles D, Sulai Y, Langlo C, et al.: In vivo imaging of human cone photoreceptor inner segments. *Invest. Ophthalmol. Vis. Sci* 55(7), 4244–4251 (2014) [PubMed: 24906859]
15. Wu E, Wu K, Cox D, Lotter W: Conditional infilling GANs for data augmentation in mammogram classification In: Stoyanov D, et al. (eds.) *RAMBO/BIA/TIA-2018. LNCS*, vol. 11040, pp. 98–106. Springer, Cham (2018). 10.1007/978-3-030-00946-5_11
16. Zhang C, Tavanapong W, Wong J, de Groen PC, Oh JH: Real data augmentation for medical image classification In: Cardoso MJ, et al. (eds.) *LABELS/CVII/STENT-2017. LNCS*, vol. 10552, pp. 67–76. Springer, Cham (2017). 10.1007/978-3-319-67534-3_8
17. Zhang Z, Yang L, Zheng Y: Translating and segmenting multimodal medical volumes with cycle- and shape-consistency generative adversarial network. In: *IEEE CVPR* (2018)

Fig. 1.

Overview of A-GAN, which augments medical imaging data. (A) Cells (cone photoreceptors) imaged in a patient using an adaptive optics ophthalmoscope [13, 14]. (B) Cells and their contours are extracted. (C) Active cell appearance model (ACAM). (D) Sampling the intensity distributions from ACAM generates cell contours with representative intensity information, which are used to guide an image-to-image conditional GAN (C-GAN).

Fig. 2.

Comparison of generated medical images from C-GAN (top row) and A-GAN (bottom row) following different numbers of iterations using a set of contours with the same shapes, but with intensity information used by A-GAN (no intensity information used by C-GAN). A-GAN quickly achieved stability after only 20 iterations while C-GAN yielded unrealistic cells (white arrows) even after 400 iterations. The scale bar is 20 μm .

Table 1.

Comparison of cell detection accuracy across the four training datasets based on conventional, C-GAN, and A-GAN-based data augmentation (mean $\pm$ SD).

Training images	Conventional		C-GAN		A-GAN	
	Recall (%)	Precision (%)	Recall (%)	Precision (%)	Recall (%)	Precision (%)
32	39.8 $\pm$ 41.8	38.4 $\pm$ 40.7	71.7 $\pm$ 38.9	68.3 $\pm$ 35.9	71.9$\pm$31.9	77.4$\pm$30.6
64	41.6 $\pm$ 41.8	38.8 $\pm$ 39.6	71.7 $\pm$ 39.6	66.2 $\pm$ 38.0	73.6$\pm$34.1	78.0$\pm$33.7
128	41.8 $\pm$ 38.8	41.5 $\pm$ 39.2	72.5 $\pm$ 39.3	65.9 $\pm$ 37.6	75.0$\pm$34.3	77.8$\pm$33.5
256	65.6 $\pm$ 41.4	62.1 $\pm$ 39.5	72.3 $\pm$ 39.3	67.0 $\pm$ 38.0	79.5 $\pm$ 33.4	79.0 $\pm$ 32.1

Table 2.

Comparison of cell segmentation accuracy using Dice coefficient (DC) and area difference (AD) across the four training datasets based on conventional, C-GAN and A-GAN-based data augmentation (mean $\pm$ SD).

Training images	Conventional		C-GAN		A-GAN	
	DC (%)	AD (%)	DC (%)	AD (%)	DC (%)	AD (%)
32	52.1 $\pm$ 12.7	46.0 $\pm$ 22.7	64.3 $\pm$ 17.5	41.7 $\pm$ 48.9	75.5 $\pm$ 10.5	27.9 $\pm$ 38.5
64	64.4 $\pm$ 17.8	48.8 $\pm$ 45.5	65.6 $\pm$ 17.4	41.8 $\pm$ 34.8	80.3 $\pm$ 9.9	25.5 $\pm$ 23.5
128	63.8 $\pm$ 17.3	38.0 $\pm$ 48.3	66.2 $\pm$ 18.3	40.0 $\pm$ 37.8	85.3 $\pm$ 9.2	20.1 $\pm$ 34.7
256	75.2 $\pm$ 12.3	21.9 $\pm$ 27.9	65.7 $\pm$ 19.4	35.9 $\pm$ 40.7	86.4 $\pm$ 9.8	20.2 $\pm$ 29.7