Conference materials
UDC 53
DOI: https://doi.org/10.18721/JPM.173.221

Efficiency analysis of generative adversarial networks for single pixel imaging

D.V. Babukhin✉, D.V. Sych

P.N. Lebedev Physical Institute of the RAS, Moscow, Russia

✉ dv.babukhin@gmail.com

Abstract. The single-pixel camera provides a prospective tool for imaging beyond conventional pixel-matrix-based devices. In recent years, neural networks have become a part of single-pixel imaging as a method to restore an image from intensity measurements computationally. Generative adversarial networks (GANs) are particularly well suited for this task. In this paper, we investigate the performance of a generative adversarial least squares network in the task of image reconstruction from a single-pixel camera at low sampling rates. We demonstrate that stable successful image reconstruction is possible at sampling rates around 8%, and that the reconstructed images should match the structure of the images present in the training sample.

Keywords: Single pixel imaging, neural networks, image restoration

Funding: Grant no. 23-22-00381, https://rscf.ru/project/23-22-00381.

Citation: Babukhin D.V., Sych D.V., Efficiency analysis of generative adversarial networks for single pixel imaging, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (3.2) (2024) 112–115. DOI: https://doi.org/10.18721/JPM.173.221

Материалы конференции
УДК 53
DOI: https://doi.org/10.18721/JPM.173.221

Анализ эффективности генеративных состязательных сетей в задаче однопиксельной визуализации

Д.В. Бабухин✉, Д.В. Сыч

Физический институт имени П. Н. Лебедева РАН, Москва, Россия

✉ dv.babukhin@gmail.com

Аннотация. Однопиксельная камера представляет собой перспективный инструмент для получения изображений, выходящий за рамки традиционных устройств на основе пиксельных матриц. В последние годы нейронные сети стали частью однопиксельной визуализации как численный метод восстановления изображения по измерениям интенсивности. Генеративные состязательные сети (GAN) особенно хорошо подходят для решения этой задачи. В данной работе мы исследуем производительность генеративной состязательной сети наименьших квадратов в задаче восстановления изображения с однопиксельной камеры при низкой частоте дискретизации. Мы демонстрируем, что стабильная успешная реконструкция изображений возможна при частоте дискретизации около 8%, и что реконструированные изображения должны соответствовать структуре изображений, присутствующих в обучающей выборке.

Ключевые слова: Однопиксельная визуализация, нейронные сети, восстановление изображения

Финансирование: Грант № 23-22-00381, https://rscf.ru/project/23-22-00381.

Ссылка при цитировании: Бабухин Д.В., Сыч Д.В. Анализ эффективности генеративных состязательных сетей в задаче однопиксельной визуализации // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 3.2. С. 112–115. DOI: https://doi.org/10.18721/JPM.173.221

Introduction

Single-pixel imaging emerged as a mathematical development of approaches to image restoration [1]. The central problem of single-pixel imaging is the computational generation of an image from intensity measurements acquired via measuring a camera with a single light detector. These intensity data combined with specific light patterns, i.e., masks, responsible for spatial modulation of light, illuminating the object, allow the generation of an object image.

In the ideal circumstances (no noise in the device calibration and infinitely long exposition) we can simply obtain the image via inverse matrix multiplication. In practice, intensity measurements contain noise which makes a straightforward approach to image recovery inefficient and with bad quality of the produced image. One way to obtain an image with better quality is to use a compressed sensing approach [1] which uses an L_1-restricted optimization to produce the image.

In recent years, a conventional method for image generation was challenged via the use of neural networks [2] and in particular, generative adversarial networks [3]. Because neural networks can learn hard-to-find patterns from data and are highly flexible in terms of parameter choice (architecture and size), they provide a prospective tool in single pixel imaging. Here we analyze image restoration efficiency with low sampling rates for a least squares generative adversarial networks (GAN) – a neural network, consisting of two distinct networks that learn from each other outputs. We also demonstrate that a trained network can restore only images, matching image structure of the training set.

Materials and Methods

In this work we train a least squares GAN [4] on simulated intensity measurements of the MNIST dataset, processed with a single-pixel camera. A first part of the network, a generator, takes gathered intensity values as an input and produces a restored image of the target object (digits from 0 to 9 in our case). After initialization this network is untrained, and its restoration quality is unsatisfactory. The training procedure consists of producing the output of the generator and feeding it to the second part of the network, the discriminator, which evaluates the quality of the input image by labeling it with a number between 0 and 1, where 0 corresponds to false (generated image) and 1 corresponds to true (real image of a digit).

We use a set of digit images and simulated camera intensity values for every image to train a generator restoring digit images and a discriminator correct labeling of fake and real values via rounds of network training. Each round has two parts: generator training, when the parameters of the generator are optimized to produce a better image and the parameters of a discriminator are fixed, and discriminator training when the parameters of the generator are fixed and parameters of the discriminator are optimized to better labeling of images.

The difference between the two types of networks we use in this work is in the loss function we optimize. For least-squares GAN, the discriminator has a loss function

$$L_D^{LS}(D) = \frac{1}{2} E_x\left[\left(D(x)-1\right)^2\right] + \frac{1}{2} E_z\left[\left(D(G(x))\right)^2\right], \tag{1}$$

and generator has a loss function

$$L_G^{LS}(G)=\frac{1}{2}E_z\left[\left(D\left(G(z)\right)-1\right)^2\right]. \quad (2)$$

A set of images used for training, Z is a set of intensity vectors after camera process simulation, $E_{X(Z)}$ is averaging over set $X(Z)$ correspondingly. If only the above-mentioned loss functions are used for training a network, the generator may produce poor-quality images even after completion of training: images can be alike for different input vectors (mode collapse) or have spurious details, which look unrealistic to the human eye. To encourage the generator to produce plausible images, we use so-called content loss, an additional part of the generator loss function, which encourages the generator to produce images with realistic characteristics. The content loss is the following

$$L_G^{content}(G)=L_1(G)+L_{VGG}(G). \quad (3)$$

Here, the first component

$$L_1(G)=\frac{1}{WH}\sum\sum\left|x_{i,j}-G(z)_{i,j}\right|, \quad (4)$$

is a loss function, which encourages sparsity in a generated image. W and H denote width and height of the output image from a generator network, i and j are matrix indices. The second component

$$L_{VGG}(G)=\frac{1}{W_kH_k}\sum\sum\left(VGG_k\left(x_{i,j}\right)-VGG_k\left(G(z)_{i,j}\right)\right)^2, \quad (5)$$

is a loss function, which returns the squared difference between feature representations of a real image x and a generated image $G(z)$, calculated with a kth layer output of a pre-trained *VGG* network [6]. W_k and H_k denote width and height of the output image from the kth layer of *VGG* network. The latter loss function encourages the generator to produce images, that are similar to realistic images from another convolutional network point of view, which was trained for the classification task.

Results and Discussion

Here we provide result for image restoration via a trained least-square GAN. We trained several examples of this GAN for sampling rates 1%, 2%, 4% and 8%, where the sampling rate is defined as a ratio between the number of patterns, used to aquire intensity data, and the number of pixels in the target image. In Fig. 1 we provide several restored digits for every sampling rate along with ground truth images. We also provide restoration results for an image, which has different structure than images from the training data set (a pictorial smile).

From Fig. 1 one can see that for 8% single pixel camera sampling rate all restored images of digits have good quality. The situation changes as sampling rate decreases: images of digits with simple structure (“1”) can be restored up to lowest sampling rates, and images of digits with more complicated structure (“2”) cannot be restored with low sampling rate with our trained network. Nonetheless, for practical applications, 8% sampling rate is already a significant speed-up comparing to conventional methods (e.g., compressed sensing requires more than 30% sampling rate). It is a question of further research if we can find a neural network architecture, which requires lower sampling rate and does not require prohibitively large amount of training data and time.

Another point illustrated in Fig. 1 is that a least-squares GAN, trained to restore digits from intensity measurement via single pixel camera, cannot restore image of a pictorial smile, because there were no alike pictures in the training set. For practice, this result means that in order to successfully apply neural networks for single-pixel imaging we need to have a training data set, which includes images of target objects. As specific practical applications (for example, in medicine) can have only a limited amount of freely-available data, it is interesting to explore possibilities of overcoming the need of such data for training neural networks for imaging purposes.

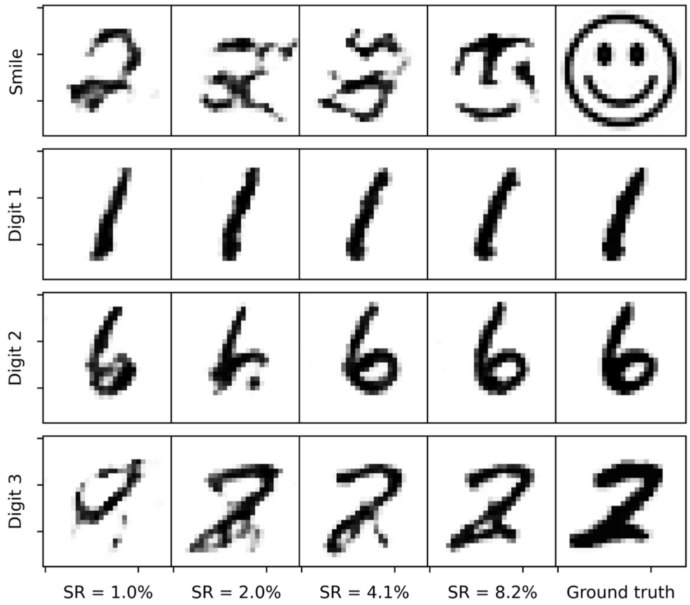

Fig. 1. Images from single-pixel camera simulation, restored via the least-squares GAN. Here we provide results for 1%, 2%, 4% and 8% sampling rates along with a ground truth image

Conclusion

In this work, we analyzed the efficiency of a least-squares generative adversarial network in the problem of image restoration in single-pixel imaging. We demonstrated how quality of the restored image depends on the sampling rate. We also demonstrated the importance of having target object images in the training set. Our work is aimed at advancing the development of computational tools for single-pixel imaging.

REFERENCES

1. **Duarte M.F., Davenport M.A., Takhar D., et al.,** Single-Pixel Imaging via Compressive Sampling, IEEE Signal Processing Magazine. 25 (83) (2008).

2. **Higham C.F., Murray-Smith R., Padgett M.J., Edgar M.P.,** Deep learning for real-time single-pixel video, Scientific Reports. 8 (2018).

3. **Zhao M., Li F., Huo F., Tian Z.,** Generative adversarial network-based single-pixel imaging, Soc for Inf Disp. 30 (8) (2022).

4. **Gui J., Sun Z., Wen Y., et al.,** A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications, IEEE Transactions on Knowledge and Data Engineering 35 (4) (2023).

5. **Mao X., et al.,** Least Squares Generative Adversarial Networks, IEEE International Conference on Computer Vision (ICCV), Venice, Italy. (2017) 2813–2821.

6. **Simonyan K., Zisserman A.,** Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv:1409.1556 (2014).

THE AUTHORS

BABUKHIN Danila V.
dv.babukhin@gmail.com

SYCH Denis V.
denis.sych@gmail.com

Received 31.07.2024. Approved after reviewing 12.08.2024. Accepted 12.08.2024.