Conference materials
UDC 528.854
DOI: https://doi.org/10.18721/JPM.173.243

Convolutional neural networks for image-free classification via single-pixel imaging

A.A. Reutov✉, D.V. Babukhin, D.V. Sych

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

✉ aleksey.reutov@phystech.edu

Abstract. The technology of single-pixel imaging extends visualization capabilities beyond pixel-matrix-based devices. One of possible applications for this technology is fast classification of objects without the need for reconstruction of an image. The single-pixel camera gathers light statistics and then a computational algorithm — such as a neural network — decides on what is the object been illuminated. We train a convolutional neural network on simulated data from single-pixel camera and demonstrate effectiveness of classification images of handwritten digits.

Keywords: Single-pixel imaging, convolutional neural networks, image-free classification

Funding: This study was funded by RSF grant no. 23-22-00381.

Citation: Reutov A.A., Babukhin D.V., Sych D.V., Convolutional neural networks for image-free classification via single-pixel imaging, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (3.2) (2024) 217–220. DOI: https://doi.org/10.18721/JPM.173.243

Материалы конференции
УДК 528.854
DOI: https://doi.org/10.18721/JPM.173.243

Сверточные нейронные сети для классификации без изображений с помощью однопиксельной визуализации

А.А.Реутов✉, Д.В.Бабухин, Д.В. Сыч

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

✉ aleksey.reutov@phystech.edu

Аннотация. Технология однопиксельной визуализации расширяет возможности визуализации за пределы устройств на основе пиксельной матрицы. Одним из возможных применений этой технологии является быстрая классификация объектов без необходимости реконструкции изображения. Однопиксельная камера собирает статистику освещенности, а затем вычислительный алгоритм, например нейронная сеть, принимает решение о том, какой объект был освещен. Здесь мы обучаем сверточную нейронную сеть на смоделированных данных с однопиксельной камеры и демонстрируем эффективность классификации изображений рукописных цифр.

Ключевые слова: однопиксельная визуализация, сверточные нейронные сети, классификация без полной визуализации

Финансирование: Исследование выполнено за счет гранта Российского научного фонда (проект № 23-22-00381).

Ссылка при цитировании: Реутов А.А., Бабухин Д.В., Сыч Д.В. Сверточные нейронные сети для классификации без изображений с помощью однопиксельной визуализации // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 3.2. С. 217–220. DOI: https://doi.org/10.18721/JPM.173.243

Introduction

Single-pixel imaging is a technique for image acquisition via illuminating an object with structured light and obtain an image via computational restoration [1]. This technique provides imaging capabilities beyond visible wavelengths and thus perspective for making specific devices. Another possible application of single-pixel imaging is image-free object classification [2]. Gathering structured light via single-pixel camera and processing the raw data without image acquisition allows fast detection and classification of objects in the range of single-pixel imaging applicability. In last years, neural networks were shown to be a convenient tool for computational processing of single-pixel-gathered data [3,4]. Here we investigate on capability of convolutional neural networks to classify handwritten digits, processed with a single-pixel camera.

Materials and Methods

The general idea of single-pixel can be explained as follows. The light is spatially modulated by set of patterns P, which can be represented as matrix with M rows (number of patterns) and N columns (number of elements of each pattern). A spatial light modulator can be realized in different ways, for example, it can be a digital micromirror device (DMD) consisting of an array of small controlled mirrors, each can take two positions with different deflection angles (providing reflection or transmitting a light). Patterns change over time and provide varying intensity spatial profile from the source. The modulated light is then reflected from the target object (with reflectance matrix I) and measured by the detector. Matrix I can be flattened and represented as one columns with length N. Each detector's record is a intensity value for the selected pattern. The measured signals S of single-pixel scheme can be expressed as $S = PI$. Another scheme of single-pixel imaging consider reflection from object as a first step (with intensities I) and spatial modulation by patterns P after that.

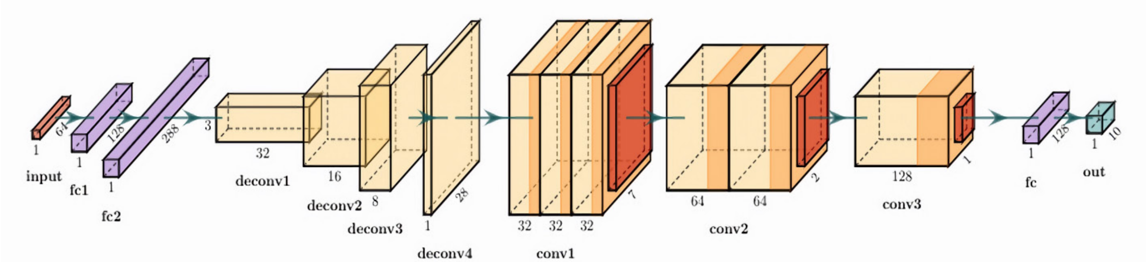

Fig. 1. Architecture of convolutional neural network for partial image reconstruction from synthetic detector signals

The second step in constructing single-pixel images is to reconstruct the object image I, using the detector signals S and known patterns P. There are several methods to reconstruct image: for example, matrix inversion of P, compressed sampling or neural network approach. The first technique is computationally complex task for high resolution images. The second and third approach provide promising results [1–4] as reconstruction algorithms. But instead of fully image reconstruction, we provide in our work approach for partial reconstruction, the recognition of the type of objects captured by the single-pixel camera. We use neural network methods suitable for classification task.

We trained a convolutional neural network on handwritten digits dataset MNIST. For every digit, we simulated a process of structured illumination with $M = 64$ spatial patterns, thus generating 64 intensity measurements S. For an image with $N = 784$ pixels, our setup corresponds to approximately 8% sampling rate. We used a dataset of 60000 images, where a neural network was trained on 48000 images, and 12000 images were used to calculate classification accuracy.

We designed and trained a convolutional neural network (CNN) on input data S and set of target output MNIST. Structurally, the CNN consists of several blocks, feedforward and convolutional (see Fig. 1). We choose ADAM optimizer for training process, its learning rate is 0,0003. The set of patterns P for synthetic signal data generation consists of binary matrices, with randomly generated and equiprobably values 0 or 1. The negative log-likelihood function was chosen as the loss function that determines the divergence between the expected and predicted value and used as optimization objective.

Results and Discussion

We provide the results of our simulation in Fig. 2. Here, we can see that after 30 epochs of training – full learning iterations through training dataset – our network classifies digits, based on intensity measurements only, with the final accuracy 94.1%. The accuracy is slightly varying during last trainings epochs, which can be handled by tuning of size of training minibatches and learning-rate scheduling. We also see that value of neural network loss – a characteristic of the learning process – has gradually decreased, which demonstrates a correct learning process. Both graphs show that machine learning model is almost trained after 15 epochs and both two metrics (accuracy and loss function) fluctuate within boundaries from 90% to 95% and from 0.2 to 0.3 respectively. Therefore, fine-tuning of the hyperparameters may improve and stabilize training results.

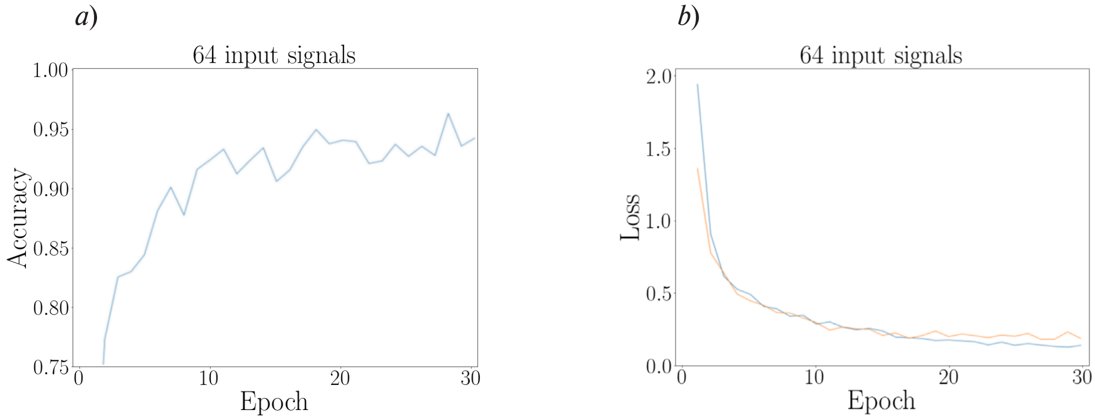

Fig. 2. Accuracy on test set (*a*) and dynamics of loss function (*b*) during training. The number of patterns and input signals to the CNN is 64, the accuracy by the end of training is 94.1%

Conclusion

In this work, we trained a convolutional neural network to classify objects, using intensity measurements from a single pixel camera and without image acquisition. We choose recognition of hand-written digits from MNIST dataset (as standard benchmark for classification tasks). Despite the low sampling rate (about 8%), our network demonstrated a high classification accuracy, thus serving as an efficient tool in single-pixel framework. Obtained metrics show the prospect of improving training results by fine-tuning. We further plan to investigate on image-free classification in the presence of device imperfections such as different noises of detector, overexposing or dropped patterns and signals.

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. **Sadao Ota et. al.,** Ghost cytometry, Science. (2018) 1246–1251.

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

4. **Zhang Z., Li X., Zheng S., Yao M., et al.,** Image-free classification of fast-moving objects using "learned" structured illumination and single-pixel detection. Optics Express. 28 (9) (2020) 13269.

5. **Diederik P.Kingma, Jimmy Ba**. Adam: A method for stochastic optimization. (2017).

THE AUTHORS

REUTOV Aleksei A.
aleksey.reutov@phystech.edu
ORCID: 0000-0001-5838-4770

SYCH Denis V.
denis.sych@gmail.com
ORCID: 0000-0002-4188-0951

BABUKHIN Danila V.
dv.babukhin@gmail.com
ORCID: 0000-0001-6511-7515

Received 30.07.2024. Approved after reviewing 02.08.2024. Accepted 02.08.2024.