Few-Shot Learning using Supervised Non-Associative Autoencoders and Correlation Techniques

Ehsan Sedgh-Gooya
L@bISEN, VISION Lab, Yncréa ouest
20 rue Cuirassé Bretagne CS 42807
29228 Brest Cedex 2, France
Email: ehsan.sedgh-gooya@isen-ouest.yncrea.fr

Ayman Alfalou
L@bISEN, VISION Lab, Yncréa ouest
20 rue Cuirassé Bretagne CS 42807
29228 Brest Cedex 2, France
Email: ayman.al-falou@isen-bretagne.fr

***Abstract*—Deep learning, while very effective today, traditionally requires very large amounts of labeled data to perform the classification task. In an attempt to solve this problem, the few-shot learning concept, which uses few labeled samples by class, becomes more and more useful. In this paper, we propose a new low-shot learning method, dubbed Supervised Non-Associative Auto-Encoder (SNAAE) to perform classification. Complementary to prior studies, SNAAE represents a shift of paradigm in comparison with the usual few-shot learning methods, as it does not use any prior knowledge neither unlabeled data. SNAAE is based on stacking layers of an autoencoder, which are trained in a supervised way to rebuild a single version representing their inputs. The reconstructed output is then classified outside of the neural network by correlation plane quantification metric. To perform the classification, the rebuilt output is compared with the initial versions used as target to train the SNAAE. We demonstrate empirically the efficiency of our proposed approach on the well known handwritten digits Modified National Institute of Standards and Technology database (MNIST) database.**

Keywords–Neural Networks; Few-shot learnong; Semi-Supervised learning; Autoencoders.

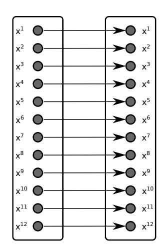
I. INTRODUCTION

At a time when unlabelled data is becoming increasingly common, manual labeling of all these data is expensive, time consuming and inefficient. Moreover, when we place ourselves on the side of the humans, we need few data to learn new concepts with very little supervision. Hence, the few-shot learning concept becomes increasingly important. The aim of these concepts is to improve the generalization capabilities of learning models so that they can achieve very good performance using a few labeled samples. For a maximal efficiency, the state-of-the-art few-shot learning algorithms [1] [2] typically make use of prior knowledge and large amounts of unlabeled data. In this paper, we address the above problem, and we propose a new few-shot learning classification method based on Supervised Non-Associative Auto-Encoder (SNAAE). Furthermore, SNAAE does not need at all any prior knowledge neither unlabeled data. We organize our article by first describing a general autoencoder framework. Then in the following section we define and explain our proposed method. Finaly, the efficiency of our proposed approach is tested on MNIST.

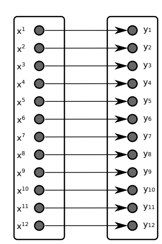
II. A GENERAL AUTOENCODER FRAMEWORK

Autoencoders [3]–[5] are simple learning circuits which aim to transform inputs into outputs with the least possible amount of distortion. While conceptually simple, they play an important role in machine learning. Autoencoders were first introduced in the 1980s by Hinton and the PDP group [6] to address the problem of backpropagation without a teacher, by using the input data as the teacher. Together with Hebbian learning rules [7], autoencoders provide one of the fundamental concepts for unsupervised learning and for beginning to address the mystery of how synaptic changes induced by local biochemical events can be coordinated in a self-organized manner to produce global learning and intelligent behavior. To derive a general framework an $n/p/n$ autoencoder [8] is defined by a tuple $n, p, m, \mathbb{F}, \mathbb{G}, \mathcal{A}, \mathcal{B}, \mathcal{X}, \Delta$ where:

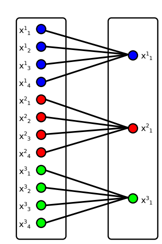
1) $\mathbb{F}, \mathbb{G}$ are sets;
2) n and p are positive integers. Here we consider primarily the case where $0 < p < n$.
3) $\mathcal{A}$ is a class of functions from $\mathbb{G}^p$ to $\mathbb{F}^n$.
4) $\mathcal{B}$ is a class of functions from $\mathbb{F}^n$ to $\mathbb{G}^p$.
5) $\mathcal{X} = \{x_1, \ldots, x_m\}$ is a set of m (training) vectors in $\mathbb{F}^n$. When external targets are present, we let $\mathcal{Y} = \{y_1, \ldots, y_m\}$ denote the corresponding set of target vectors in $\mathbb{F}^n$.
6) Δ is a dissimilarity or distortion function (e.g. L_p norm, Hamming distance) defined over $\mathbb{F}^n$.



(a) Associative Autoencoder



(b) Non-Associative Autoencoder



(c) SNAAE

Figure 1. Autoencoders

A. Associative Auto-Encoder

For any $A \in \mathcal{A}$ and $B \in \mathcal{B}$, the autoencoder transforms an input vector $x \in \mathbb{F}^n$ into an output vector $A \circ B(x) \in \mathbb{F}^n$ (Figure 1-a). The corresponding autoencoder problem is to find $A \in \mathcal{A}$ and $B \in \mathcal{B}$ that minimize the overall distortion function:

$$\min E(A,B) = \min_{A,B} \sum_{t=1}^{m} E(x_t) = \min_{A,B} \sum_{t=1}^{m} \Delta(A \circ B(x_t), x_t) \tag{1}$$

B. Non-Associative Auto-Encoder

In the non auto-associative case, when external targets y_t are provided, the minimization problem becomes:

$$correct : Figure2 \min E(A,B) = \min_{A,B} \sum_{t=1}^{m} E(x_t, y_t) = \min_{A,B} \sum_{t=1}^{m} \Delta \tag{2}$$

III. Supervised Non-Associative Autoencoders (SNAAE)

As any system based on neural network, two operations are performed by SNAAE: offline and online phase. In the following subsections, we describe these two phases.

A. Offline phase

Let us consider $X = \{x_1^1 \cdots, x_1^M, \cdots, x_p^1, \cdots, x_p^M, \cdots x_P^1, \cdots, x_P^M\}$ the set of training set where M is the number of labeled images for training phase and P is the cardinality of classes. Among $\{x_p^1, \cdots, x_p^M\}$ (samples images corresponding to class p), a reference image is chosen to be the target image (Figure 1-c). Offline phase is then performed according to 2.

B. Online phase

During online phase, when an input query image x_q is presented, the SNAAE first reconstruct the reference image $\hat{x}$. $\hat{x}$ class is then evaluated by corelation techniques [9] between all reference images and $\hat{x}$.

IV. Experiments and results

In this set of experiment, we have first evaluated SNAAE on the MNIST datasets [10]. We consider just 1, 3, 5, 10, 100, 500 randomly chosen labeled samples per class. We then evaluated our approach on the test dataset. Table I reports the results. We may observe that on MNIST datasets, our SNAAE proposed model achieve good accuracies, outperforming the ones obtained by the Convolutional Neural Network (CNN) models [11]. In Figure 2, the performance of our approach is compared to CNN more precisely.

TABLE I. FEW-SHOT LEARNING: CLASSIFICATION ACCURACY OF SNAAE AGAINST BASELINE CNN ON THE MNIST

	SNAAE	CNN
1-shot	**49.38** ±4.26	18.83 ±4.26
3-shot	**61.26** ±4.26	20.63 ±2.75
5-shot	**69.99** ±1.15	20.95 ±4.02
7-shot	**76.75** ±2.13	28.33 ±1.87
10-shot	**83.06** ±1.82	31.52 ±5.73
100-shot	**96.49** ±0.19	83.53 ±4.14
200-shot	**97.63** ±0.15	90.54 ±0.54
500-shot	**98.71** ±0.16	94.78 ±0.30
full-shot	**99.37** ±0.05	98.62 ±0.09

As shown in Figure 2, our method greatly exceeds the performance of a convolutional network when few data are available.

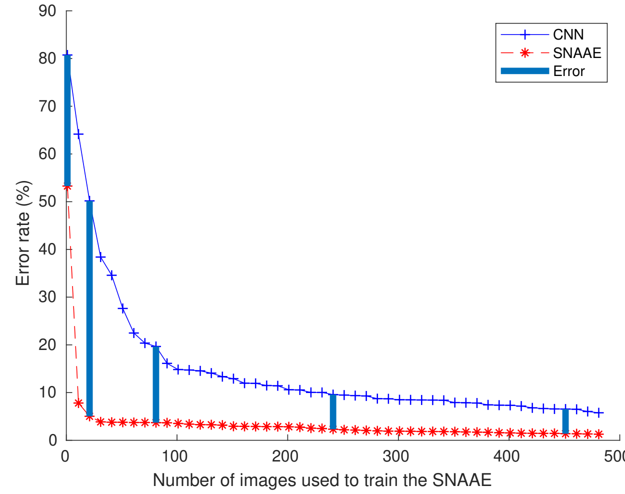


Figure 2. Error rate curve comparing SNAAE to CNN. As illustrated in this figure, for 1, 21, 81, 241, 451 images used to train the SNAAE, the difference between CNN's error rate and SNAAE one are respectively 27.46, 45.14, 15.96, 7.26, 5.16.

V. Conclusion

In this paper, we introduce SNAAE (Supervised Non-Associative Autoencoders), taking inspiration from the human world. SNAAE is capable to successfully perform the few-shot learning task, without the need of having prior knowledge neither unlabeled data. When unlabeled data is unavailable, SNAAE offers very good performance. In our proposed method, SNAAE, evaluated on MNIST, need only 500 images per class (less than 10% of the whole dataset) to surpass the cnn's performance.

References

[1] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," in ICML Deep Learning Workshop, vol. 2, 2015.

[2] O. Vinyals, C. Blundell, T. Lillicrap, D. Wierstra et al., "Matching networks for one shot learning," in Advances in neural information processing systems, 2016, pp. 3630–3638.

[3] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in Advances in neural information processing systems, 2007, pp. 153–160.

[4] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," arXiv preprint arXiv:1312.6114, 2013.

[5] Y. Pu, Z. Gan, R. Henao, X. Yuan, C. Li, A. Stevens, and L. Carin, "Variational autoencoder for deep learning of images, labels and captions," in Advances in neural information processing systems, 2016, pp. 2352–2360.

[6] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," California Univ San Diego La Jolla Inst for Cognitive Science, Tech. Rep., 1985.

[7] E. Oja, "Simplified neuron model as a principal component analyzer," Journal of mathematical biology, vol. 15, no. 3, 1982, pp. 267–273.

[8] P. Baldi, "Autoencoders, unsupervised learning, and deep architectures," in Proceedings of ICML workshop on unsupervised and transfer learning, 2012, pp. 37–49.

[9] A. Alfalou and C. Brosseau, "Understanding correlation techniques for face recognition: from basics to applications," in Face Recognition. InTech, 2010.

[10] Y. LeCun and C. Cortes, "MNIST handwritten digit database," 2010. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[11] F. Chollet et al., "Keras. h ps," github. com/fchollet/keras, 2015.