Deep Learning for Edge Detection

[1]Cihan Tiken and *[2]Tolga Ensari
[1]Department of Computer Engineering, Arel University, 34537, Istanbul, Turkey
*[2]Department of Computer Engineering, Istanbul University, 34320, Istanbul, Turkey

Abstract:

Deep learning and deep belief networks (DBNs) are one of the most used topics in machine learning and pattern recognition area in recent years. DBNs consist of stacked Restricted Boltzmann Machine (RBM) structure. RBM is energy based stochastic neural networks. DBNs have many hidden layers and it is optimized after fine-tuning process with Autoencoder (AE) architecture. AE transforms input space to new space. Edge detection is also one of the important issues in machine vision. It is generally done with gradient or Laplacian methods. Some of classical techniques, used in the literature, are canny, differential, sobel, prewitt, roberts or fuzzy logic methods. In this paper, we propose deep learning based edge detection method. Some hidden features is discovered with suitable DBNs architecture. In order to evaluate the performance of presented method we use handwritten character images from MNIST data set. Experimental results show that our approach improves the performance of edge detection process.

Key words: Deep learning, deep belief networks, edge detection, computer vision.

1. Introduction

We give brief information about deep learning (deep belief networks) and edge detection in this section. Deep learning is one of the most popular method in machine learning research [1-8]. Y. Bengio, G. Hinton, M. Ranzato, R. Salakhutdinov and their colleagues started to present next generation of neural networks as a deep learning concept [1, 2, and 5]. The disadvantages of error back propagation has been achieved via defining the deep belief networks.

1.1. Deep Learning and Deep Belief Networks

Deep learning is an emerging member of machine learning family, it has became popular recently. It is used in learning and training of large and complex probabilistic models very effectively. Especially in deep neural networks it gives impressive results according to conventional approaches [1-8]. Deep neural networks consist of a large number of nonlinear layers of processing units and it uses layer-wise unsupervised pre-training during learning of probabilistic model for the data. The layers has a hierarchical structure from low-level to high-level features. Deep neural networks, convolutional deep neural networks, and deep belief networks are some various architecture of deep learning, all these have been applied different areas such as automatic speech recognition , natural language processing, audio recognition and computer vision. A deep neural network generally includes stacked multiple Restricted

*Corresponding author: Address: Department of Computer Engineering, Istanbul University, 34320, Istanbul, TURKEY. E-mail address: ensari@istanbul.edu.tr, Phone: +902124737000 Fax: +902124737180

Boltzmann Machines (RBM). RBMs are connected to each other. A hidden layer of a RBM is a visible layer of an another RBM. Layer-wise pre-training of a RBM makes easier to find more accurate model for the data. Many studies proved that multi-stage learning structures give better results versus conventional learning algorithms. It is very important to have a useful method to train RBM, but training a RBM is really difficult. Learning parameters must be suited specific data sets and RBMs carefully if not learning algorithms can easily fail to model data distribution correctly. Decreasing likelihood during learning indicates this problem. This failures decrease using rate of RBMs in more sophisticated machine learning tasks. Fig. 1 shows input and hidden layers of network.

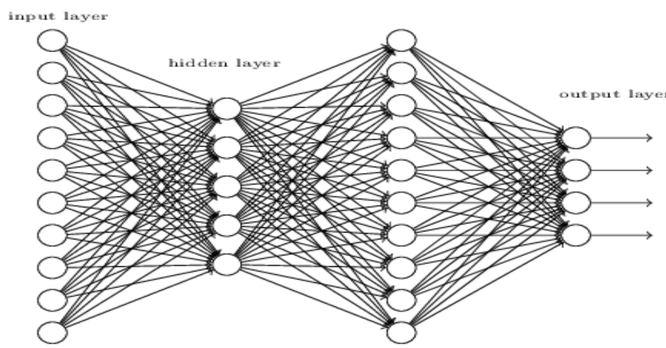


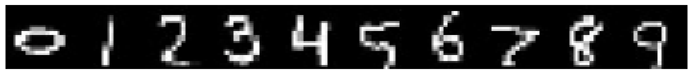
Figure 1. Layers of the deep network structure

1.2. Edge Detection

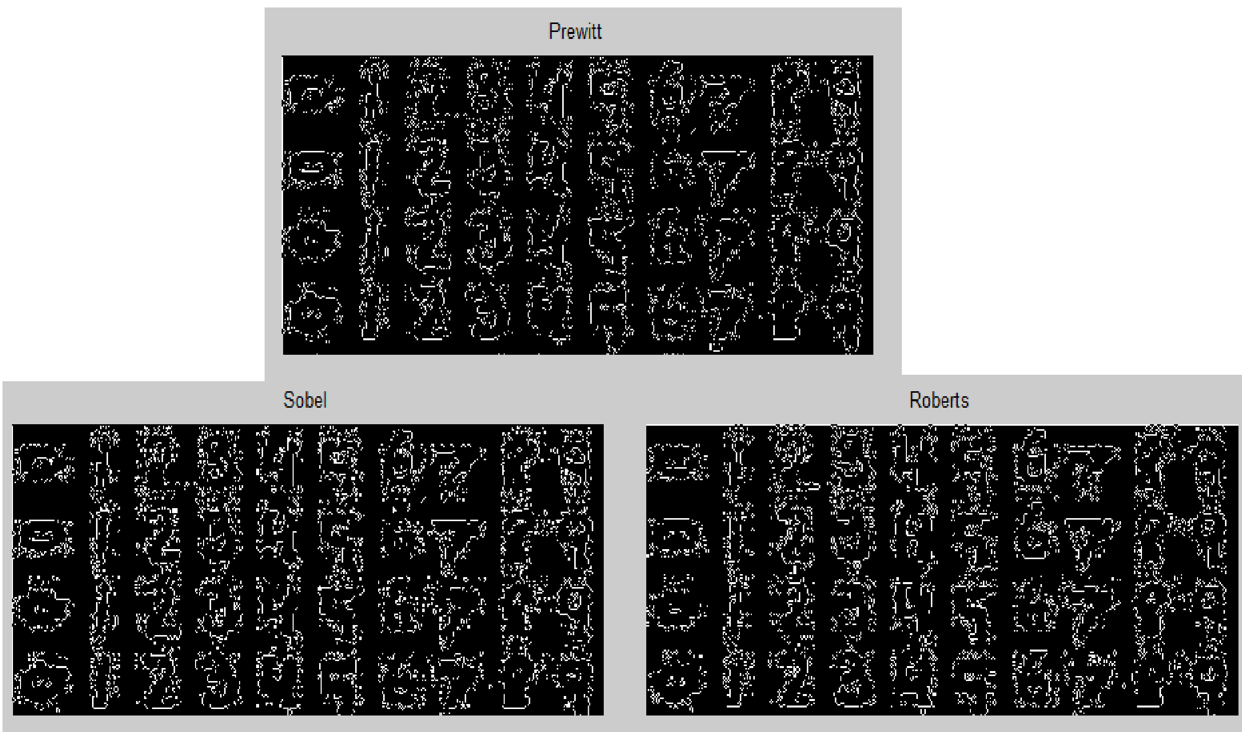
A set of processes which aim at identifying points in an image at which the image brightness changes more formally or sharply is called edge detection. Edges are the points at which brightness changes and they usually provide enough information for a human to understand the content of an image. It is a vector variable (magnitude of the gradient, direction of an edge). Edge information in an image is found by looking at the relationship a pixel has with its neighborhoods. The goal of detecting sharp changes in a digital image brightness is to capture important events. Some operators return orientation information. Others return only information about the an edge at each point whether it is exist at that point or not. Edge detection methods can be classified as this order: *Sobel operator* look for edges vertically and horizontally directions, the operator uses two 3×3 kernels which are convolved with the original image to calculate approximations of the derivatives for horizontal and vertical changes. *Canny edge operator* use Gaussian filter and finalize the detection of edges by suppressing all the other edges that are weak and not connected to strong edges. *Prewitt edge operator*, calculates an approximate value of the gradient of function. Vertical changes and horizontal changes kernels are used for this operator.*Roberts edge operator* uses discrete differentiation to approximate the gradient of an image. It is done by calculating sum of the squares of the differences between cross (diagonal) pixels.

2. Experimental Studies

Deep belief network structure is constructed in order to get the reconstructed images. Autoencoder is used for this process and we choose MNIST character data set for the experiments. In [8], reconstruction of handwritten digit images obtained with deep belief networks and we use these digit in our experiments. The size of these images are 14*x*14. Each model have 7 hidden layers and 6000 images has been used for training and 1000 images for testing part. The total number of epochs is 230 for the training phase [8].



a) Original images from MNIST character data set



(b) In each image first row shows original images, second row shows stacked autoencoder with RBM, third row shows traditional autoencoder with RBM and fourth row shows stacked autoencoder without RBM.

Figure 2. Prewitt, Sobel and Roberts edge detection of MNIST character data set.

The edge detection performance of the deep belief networks (output) can be seen from Fig. 2. Additionally, different types of deep structures tested in this process. These types are stacked autoencoder with RBM, traditional autoencoder with RBM and stacked autoencoder without RBM. It can be easily seen from the figure, the sobel edge operator has better performance than the others for deep belief network. Definition and calculation of performance metrics will be published in the future studies. Differences between network types also can be investigated.

3. Conclusions

We use deep belief networks output for edge detection process. This task has been never studied in the literature for this aim. Firstly, we get the reconstructed images output. Details about network are given in the experimental studies section. After that these images are used with prewitt, sobel and roberts edge detection algorithms. Original images are also used with these edge detectors in order to make comparison. So, the edge detection performance of original images and deep belief network output images is studied in this article.

References

[1] R. Salakhutdinov R., G. Hinton, An efficient learning procedure for deep Boltzmann machines, Neural Computation, vol. 24, issue 8, 2012, pp. 1967-2006.
[2] G. Hinton, S. Osindero, Y. Teh, A fast learning algorithm for deep belief nets, Neural Computation, vol. 18, issue 7, 2006, pp. 1527-1554.
[3] L. Deng, D. Yu, J. Platt, Scalable stacking and learning for building deep architectures, Int. Conf. on Acoustics, Speech and Signal Processing, 2012, pp. 2133-2136.
[4] J. Weston, F. Ratle, H. Mobahi, R. Collobert, Deep Learning via semi-supervised embedding, Lecture notes in Computer Science, vol. 7700, 2012, pp. 639-655.
[5] D. Erhan, Y. Bengio, A. Courville, P. A. Manzagol, P. Vincent, S. Bengio, Why does unsupervised pre-training help deep learning?, The Journal of Machine Learning Research, vol. 11, 2010, pp. 625-660.
[6] M. Spencer, J. Eickholt, J. Cheng, A deep learning network approach to ab initio protein secondary structure prediction, IEEE Transactions on Computational Biology and Bioinformatics, vol. 12, issue 1, 2015, pp. 103-112.
[7] K. Charalampous, A. Gasteratos, A tensor-based deep learning framework, Image and Vision Computing, vol. 32, issue 11, 2014, pp. 916-929.
[8] C. Tan, C. Eswaran, Reconstruction of handwritten digit images using autoencoder neural networks, Canadian Conference on Electrical and Computer Engineering, 2008, pp. 465-470.