International Journal of Image Processing and Vision Science

Volume 1 | Issue 1 Article 11

July 2012

A Modified Neural Network system based on Morphological operations for detection of images with variation in Gray level intensity and facial expressions

Pratibha Rani
Department of Electronics & Comm Engg. Hi-Tech Institute of Technology, Ghaziabad (U.P), India, roy.pratibha85@gmail.com

Anshu Sirohi
Department of Electronics & Comm Engg., Hi-Tech Institute of Technology, Ghaziabad (U.P), India, anshu.sirohi.me@gmail.com

Manish Kumar Singh
K.I.E.T, Ghaziabad (U.P), India, kiet_manish@indiatimes.com

Recommended Citation
Rani, Pratibha; Sirohi, Anshu; and Singh, Manish Kumar (2012) "A Modified Neural Network system based on Morphological operations for detection of images with variation in Gray level intensity and facial expressions," *International Journal of Image Processing and Vision Science*: Vol. 1 : Iss. 1 , Article 11.
DOI: 10.47893/IJIPVS.2012.1010
Available at: https://www.interscience.in/ijipvs/vol1/iss1/11

A Modified Neural Network system based on Morphological operations for detection of images with variation in Gray level intensity and facial expressions

Pratibha Rani [1],
Department of Electronics & Comm Engg.
Hi-Tech Institute of Technology,
Ghaziabad (U.P), India
e-mail: [1]roy.pratibha85@gmail.com

Anshu Sirohi [2]
Department of Electronics & Comm Engg.
Hi-Tech Institute of Technology,
Ghaziabad (U.P), India
e-mail:[2]anshu.sirohi.me@gmail.coM

Manish Kumar Singh[3]
Department of Electronics & Comm Engg.
K.I.E.T,
Ghaziabad (U.P), India
e-mail: [3]kiet_manish@indiatimes.com

Abstract: **We introduce an algorithm based on the morphological shared-weight neural network. Which extract the features and then classify them. This type of network can work effectively, even if the gray level intensity and facial expression of the images are varied. The images are processed by a morphological shared weight neural network to detect and extract the features of face images. For the detection of the edges of the image we are using sobel operator. We are using back propagation algorithm for the purpose of learning and training of the neural network system. Being nonlinear and translation-invariant, the morphological operations can be used to create better generalization during face recognition. Feature extraction is performed on grayscale images using hit-miss transforms that are independent of gray-level shifts. The recognition efficiency of this modified network is about 98%.**

Keywords*: Morphological operations, Back propagation algorithm, Sobel operator, face detection and recognition.*

I. INTRODUCTION

In this paper, we present an MSNN-based face recognition system for recognizing front-view face images [1]. The standard backpropagation algorithm is re-designed to simultaneously perform hit-miss transformation, train, and classify features within the same iteration. The recognized image is determined by the corresponding output value that lies within a certain threshold. In its initial stage, the MSNN performs a combination of grayscale erosion and dilation known as the *hit-miss transform* [10]. Each input image is eroded by a hit structuring element and dilated by a miss structuring element separately. Both outputs are then subtracted to derive their overlapping difference. The result from this process forms the feature map, which becomes the direct input to a backpropagation network. The MSNN can have one or more layers, and each layer can also have one or more feature maps.

The output of the MSNN is interpreted through a clustering algorithm that associates groups of pixels and determines their centroids, resulting in the coordinates of possible faces in the input image. Each possible face is extracted from the input image, the average pixel value is adjusted to compensate for lighting conditions, and the resulting image is compared to the database images. Based on the comparisons with the database images, the face candidates are either labeled or discarded, and the results are superimposed on the original captured image. The following subsections describe each system component in more details. The paper is not just about being able to generalize facial expressions, orientations, and occlusions. Any healthy neural network has no problem with that .We aim for a more robust network. Criteria for evaluation mainly include the ability to generalize results to accommodate changes in gray-level intensity and noise.

Edge detection is a fundamental tool used in most image processing applications to obtain information from the frames as a precursor step to feature extraction and object segmentation. This process detects outlines of an object and boundaries between objects and the background in the image. Morphological edge detection algorithm selects appropriate structuring element of the processed image makes use of the basic theory of morphology including erosion, dilation, opening and closing. Operation and synthesization operations of them get clear image edge [5]. The effect of erosion and dilation operations is better for image edge by performing the difference between processed image and original image, but they are worse for noise filtering. As opposed to erosion and dilation, opening and closing operations are better for filtering. The classical operator such as Sobel and Prewitt which uses first derivative has very simple calculation to detect the edges and their orientations but has inaccurate detection sensitivity in case of noise.

The MSNN was trained using the "Olivitti" data set which has been used by many researchers to train their face detection systems. It contains face images taken at the Olivetti Research Laboratory in Cambridge, **UK**. There are 10 different images of 40 distinct subjects (a total of 400 face images). For some of the subjects, the images were taken at different times, varying lighting conditions, facial expressions (open/closed eyes, smiling/non-smiling) and facial details (glasses/no glasses).All the images are taken against a dark homogeneous background and the subjects are in up-right, frontal position. The background data set consists of 200images of natural scenes that do not include any human faces.

II. MORPHOLOGY SHARED WEIGHT NEUTRAL NETWORK

The word morphology commonly denotes a branch of biology that deals with the form and structure of animals and plants. We use the same word here in the context of mathematical morphology as a tool for extracting image component that are useful in the representation and description of region shape, such as boundaries, skeletons and the convex hull.

Initially developed for automatic target recognition, MSNN was created by modifying the *shared-weight neural network* (SSNN) [2]. SSNN's feature extraction layer performs linear convolutions on inputs instead of nonlinear convolutions. The MSNN implements mathematical morphology, which involves the erosion and dilation of an image by probing with neighborhood sets called structuring elements

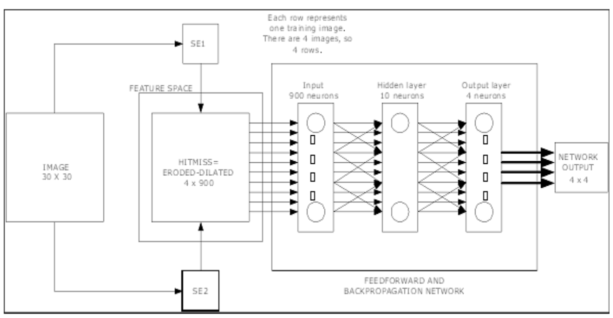


Fig.1 MSNN Architecture

The MSNN is composed of two stages: a feature extraction stage followed by a classification stage. The feature extraction stage can have one or more layers and each layer can have one or more feature maps. Each layer in this stage performs feature extraction using the morphological operation of gray-scale hit-miss transform with a pair of structuring elements. The output of the feature maps can be sub-sampled, which results in smaller size feature maps in the next higher layer. The output of the highest feature layer provides the input for the classification stage which is a standard feed-forward neural network (Figure 2). The structuring elements of the network are learned using a gradient descent method.

III. MORPHOLOGY OPERATIONS

Morphological operations [10] are nonlinear mask operations that can detect shapes. The word morphology refers to form and structure, but in digital image processing, it refers to the shape of a region. Mathematical morphology is mainly based on set theory. It is about adding or removing pixels from a binary image according to certain rules depending on neighborhood patterns. Dilation, erosion, closing, and opening are the more common morphological operations. As the names indicate, a dilation operation enlarges a region, while erosion makes it smaller.

A. Erosion

Let B_x denote the translation of B so that its origin is located at x. The *erosion* of X by B is defined as the set of all points x such that B_x is included in X [10]:

$X\Theta B = \{ x \mid B_x \quad X \}$

At each position where every 1-pixel of the structuring element (B) covers a 1-pixel of the binary image (X), the binary image pixel corresponding to the origin of the structuring element is *ORed* to the output image. An eroded sample image in given in Fig.2.

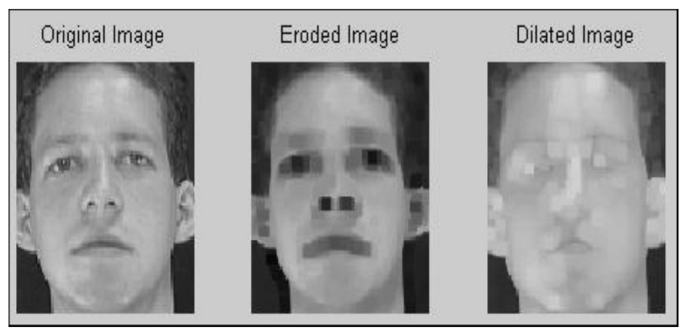


Fig. 2 output images of erosion and dilation (SE = 'Disk' , 3x3)

B. Dilation

Let B_x denote the translation of B so that its origin is located at x. The *dilation* of X by B is defined as the set of all points x such that B_x hits X – that is, they have a non-empty intersection

$X \quad B = \{ x \mid B_x \cap X \neq \ \}$

Each time the origin of the structuring element (B) touches a binary 1-pixel (X), the entire translated structuring element shape is *ORed* to the output image which has been initialized to all zeros. A dilated sample image is given above.

C. Image Resizing and Pixel Arrangement

The original ORL images are 112 by 92 pixels in dimension, which means each has a total pixel size of

10304 if we place them in a single array row. This size is not efficient for normal training, especially when we have large training sets. To avoid increased computation and long training times, it is necessary to resize the face images. Our images are therefore resized to 30×30 pixels immediately after they have passed through the *hit* and *miss* kernels - that is, before entering the classification stage. The procedure reads in images from data base, resizes them, and organizes every image into rows. Since each subject has an equal number of 4 face images in the training set and the test set, their input dimension to the MSNN is the same: 4 rows×900 columns pixel units. Each subject is trained individually.

IV. EDGE DETECTION

Edge detection is a critical element in image processing, since edges contain a major function of image information. The function of edge detection is to identify the boundaries of homogeneous regions in an image based on properties such as intensity and texture. Many edge detection algorithms have been developed based on computation of the intensity gradient vector, which, in general, is sensitive to noise in the image. The classical operator such as Sobel, and Prewitt [4].which uses first derivative has very simple calculation to detect the edges and their orientations.

Fig.3 output image of sobel edge detector

V. GRAY SCALE MORPHOLOGY

The concept of grayscale morphology is almost the same as that of binary morphology. Binary operations use logical 1 and 0, while grayscale operations use unsigned 8-bit integers from 0 to 255. This concept is thus carried forward to the grayscale hit-miss transform, which is a combination of both:

$$f(h,m) = (f \Theta h) - (f \quad m^*)$$

It measures how the structuring element *h* fits under the image *f* as well as how the structuring element *m* fits above *f*. Such a combination gives the hit-miss transform a shift-invariant property.

Hit-Miss Theorem: Let ε be the subtracted and added set of pixels for the eroded and dilated *f* respectively. Assuming that *h* and *m* have the same shape as *f*,

for $\varepsilon > 0$, $h(x) = f(x) - \varepsilon$, and $m^*(x) = f(x) + \varepsilon$

Since the structuring elements match the shape of *f* on the image, the hit-miss transform produces a peak. We should get a difference value that is very close to the width of the fitting boundary $\pm\,\varepsilon$

Shift Invariance Theorem: Let λ be the shift value by which *f* is shifted.

$$f(h,m,) = (f + \lambda),\ (h,m)$$

The hit-miss transform is invariant to shifts in grayscale because morphology is a set operation. The structuring element looks for sets that fit its own set.

VI. BACK PROPAGATION

Standard backpropagation uses the *gradient descent* algorithm (also known as the steepest descent) which is based on error correction. The algorithm is a nonlinear extension of the least mean square algorithm that is used to train multilayer networks with the same or different transfer functions. The difference between *least mean square* and *backpropagation* is the way in which these derivatives are evaluated. The backpropagation algorithm minimizes the performance index by recursively adjusting the complex weights of multilayer perceptrons. The gradient vector is calculated in the direction opposite to the flow of the output of each node. The backpropagation learning rule consists of two passes in every layer of the neural network – a feedforward and a backward pass.

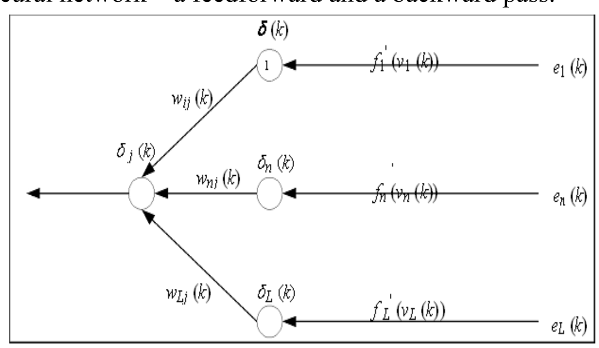


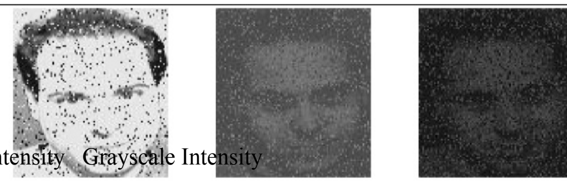
Fig: 4 Signal flow showing backpropagation of error signal

A. Inputs and Outputs

The training set must contain enough information to reveal the mapping structure. Its size depends on the user's decision on the number of inputs, hidden neurons, and output neurons. The number of inputs is usually determined by the dimension or the size of the data to be classified. The size of an input vector often corresponds to the number of features extracted from the previous stage. The number of output neurons is generally less than the input size. Training examples presented to the network should be randomized from one epoch to another for faster convergence.

B. Validation Set

We also keep a validation set of images to choose the network that trains and generalizes best. Training error is monitored with respect to both the training set and the validation set. The network that is most likely to generalize correctly to unseen data is the network with the lowest error over the validation set. It can also be used to determine the number of epochs that produces the lowest error over the validation, since this is the best indicator of network performance over unseen examples. Two copies of network weights should be kept: one for training, while the other stores the best-performing weights so far.



Grayscale Intensity	Grayscale Intensity	Grayscale Intensity	Grayscale Intensity	Grayscale Intensity
[0.1 0.4] [0.1 0.9]	[0.4 0.9] [0.3 0.5]	[0.3 0.9] [0.1 0.3]	[0.4 0.9] [0.8 0.9]	[0.6 0.9] [0.86 0.9]
Noise 0.1	Noise 0.2	Noise 0.3	Noise 0.3	Noise 0.2
MSNN 100%	MSNN 100%	MSNN 100%	MSNN 100%	MSNN 100%
NN 40%	NN 70%	NN 50%	NN 70%	NN 40%

Fig: 5 output with variation in gray level intensity

VII. OUTPUT IMAGES OF VARIATION IN GRAY LEVEL INTENSITY AND FACIAL EXPRESSION

Grayscale shifting is performed by altering pixel intensity. Normally, we use the MATLAB function imadjust to adjust the brightness or contrast of an image.To demonstrate invariance to gray-level shifts, all the images were gray-level shifted as shown in fig.5 and we trained the system using the MSNN and the standard backpropagation network. For the intensity value [0.1 0.4] [0.1 0.9], both networks can detect the images, and achieved perfect recognition. However, beyond this intensity, the backpropagation network behaved erratically, sometimes failing to converge. It never achieved 100% recognition. The backpropagation algorithm is also unable to handle images that are too bright. The brightest intensity it cannot handle perfectly is [0.6 0.9] [0.86 0.9].Whereas MSNN can achieved 100% recognition. In fig.6 when we train the MSNN with the variation in gray level intensity and facial expression of the 40 images where each image have 10 sets with variation in gray level intensity and facial expression. First we trained and then we test the MSNN for these images after that MSNN can able to detect them accordingly. Suppose we want to detect image no. 20 by MSNN then we put 20 in the command window as a input test image, then the MSNN search out the best matched image in all the trained set of images and provide the desired output.

VIII. CONCLUSION

This paper indicates that the MSNN is a robust classifier which can generalize better than the normal multilayer network. The results also prove that information or structures extracted by morphological operations are translation invariant. The MSNN is independent of shifts in gray level. It showed excellent performance under

Fig: 6 Output of detected image when facial expressions are changed

almost all the intensity levels we tried. As given in Fig.6, the MSNN perfectly recognizes all the barely-visible face images, even under the interference of noise. The outputs it generates are consistent.

ACKNOWLEDGEMENT

With a deep sense of gratitude, we wish to express my sincere thanks to my supervisor, Dr. Ranjit Singh for his immense help in planning and executing the paper work. The confidence and dynamism with which he guided the work requires no elaboration. His company and assurance at the time of crisis would be remembered lifelong. His

valuable suggestions as final words during the paper work are greatly acknowledged.
Our sincere thanks to Dr. Tarun kumar rawat for providing us constant encouragement.
We also want to thanks Prof. Kalyan Pal, our friends and parents who taught us the value of hard work by their own example.

IX. REFERENCES

[1] Y. Won et. al., Morphological Shared-Weight Networks with Applications to Automatic Target Recognition, Electronics and Telecommunications Research Institute, Daejon, South Korea, 1995. http://www.ieeexplore.ieee.org.
[2] Y. Won and P. Gader, Morphological Shared-Weight Neural Network for Pattern Classification and Automatic Target Detection, University of Missouri-Columbia, 1995. http://www.ieeexplore.ieee.org
[3] Y. Won, P. Gader, and P. Coffield, "Morphological Shared- Weight Neural Networks with Applications to Automatic Target Recognition," *IEEE Transactions Neural Networks*, Vol. 8, pp. 1195–1203, 1997.
[4] Raman Maini and J. S. Sobel, "Performance Evaluation of Prewitt Edge Detector for Noisy Images", GVIP Journal, Vol. 6, Issue 3, December 2006. www.icgst.com
[5] Davis, L. S., "Edge detection techniques", Computer Graphics Image Process. (4), 248-270, 1995.
[6] M. H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey," *IEEE Trans, Pattern. Anal. Machine. Intell*, vol. 24, no. 1, pp. 34-58, 2002.
[7] A. Tefas, C. Kotropoulos, and I. Pitas, "Face Verification based Morphological Shape Decomposition," *Proc. IEEE Int'l Conference Automatic Face and Gesture Recognition*, pp. 36–41, 1998
[8] Bovik, A. C., Huaung, T. S. and JR. D. C. M., "Non-parametric tests for edge detection noise", Pattern Recognition, 19:209-219, 1986.
[9] V. Starovoitov, D. Samal, and D. Briliuk, Three Approaches for Face Recognition, The 6-th International Conference on Pattern Recognition and Image Analysis, Velikiy Novgorod, Russia, 2002. http://www.ieeexplore.ieee.org
[10] E. R. Dougherty and R. P. Loce, "Optimal mean-absolute hit-or-miss filters," *Opt. Eng.*, vol. 32, no. 4, pp. 815–827, 1993.
[11] P. Gader, J. Keller, T. Jones, and J. Miramonti,

"MACE Prefiltering for neural-network-based automatic target recognition," in *Proc. IEEE Conf Neural Networks*, Orlando, FL, June 1994, pp. 4006–4011

[12] C. Daniell, D. Kemsley, W. Lincoln, W. Tackett, and G. Baraghi"Artificial neural networks for automatic target recognition," *Opt. Eng.*, vol. 31, no. 12, pp. 2521–2531, 1992.

[13] J. Wang, K. N. Plataniotis, J. Lu, and A.N. Venetsanopoulos, "On solving the face recognition problem with one training sample per subject," *Pattern Recognition*., vol. 39, no. 6, pp. 1746-1762, 2006.

[14] Kwok, N.M.; Ha, Q.P.; Dikai Liu; Gu Fang; Contrast Enhancement and Intensity Preservation for Gray-Level Images Using Multiobjective Particle Swarm Optimization, Vol.: 6 Issue: 1 , Jan. 2009,pp.: 145 - 155 .

[15] Download Site for Face Datasets, Olivetti Research Laboratory, and Surrey University.

http://www.ee.surrey.ac.uk/Personal/T.Windeatt/msc_projects/