FACE AND EMOTION RECOGNITION UNDER COMPLEX ILLUMINATION CONDITIONS USING DEEP LEARNING WITH MORPHOLOGICAL PROCESSING

[1]. ATLURI BHAVYA SRI, [2].SK.ASHRAF ALI

[1]. M.Tech, Dept. of ECE, NRI Institute of Technology, Agiripalli, Vijayawada,

[2]. Associate Professor, Dept. of ECE, NRI Institute of Technology, Agiripalli, Vijayawada.

ABSTRACT: The main objective of this concept is to design an efficient and enhanced face detection and emotion recognition system. A novel method based on deep learning to solve the adverse impact imposed by illumination variation in the face recognition is proposed in this paper. In order to improve the adverse effects of intense illumination changes on face images, illumination preprocessing is applied. Log-Gabor filter is used to obtain the Log-Gabor feature images of different scales and directions, and LBP features of images are extracted. Then, these texture features are learned by DBN network to complete the classification and the recognition. Further this project is extended by introducing emotion recognition by using morphological operations.

KEYWORDS: Face recognition, Log- Gabor Filter, illumination, Local Binary Patterns, Neural Network, Deep Belief Network, Morphological Operation.

INTRODUCTION: Recent research has demonstrated that sparse coding (or sparse representation) is a powerful image representation model. The idea is to represent an input signal as a linear combination of a few items from an over-complete dictionary *D*. It achieves impressive performance on image classification . Dictionary quality is a critical factor for sparse representations. The sparse representationbased coding (SRC) algorithm [2] takes the entire training set as dictionary. However, sparse coding with a large dictionary is computationally expensive. Hence some approaches [1] focus on learning compact and discriminative dictionaries. The performance of algorithms like image classification is improved dramatically with a well-constructed dictionary and the encoding step is efficient with a compact dictionary. The performance of these methods deteriorates when the training data is contaminated(*i.e.*, occlusion, disguise, lighting variations, pixel corruption). Additionally, when the data to be analyzed is a set of images which are from the same class and sharing common (correlated) features (e.g. texture), sparse coding would still be performed for each input signal independently. This does not take advantage of any structural information in the set. Low-rank matrix recovery, which determines a low-rank data matrix from corrupted input data, has been successfully applied to applications including salient object detection [4], segmentation and grouping [13, 6], background subtraction [7], tracking [4], and 3D visual recovery [13]. However, there is limited work [5] using this technique for multi-class classification. [5] uses low lowrank matrix recovery to remove noise from the training data class by class. This process becomes tedious as the class number grows, as in face recognition. Traditional PCA and SRC are then employed for face recognition. They simply use the whole training set as the dictionary, which is inefficient and not necessary for good recognition performance [12] presents a discriminative low-rank dictionary learning for sparse representation (DLRD SR) to learn a low-rank dictionary for sparse representation-based face recognition. A sub-dictionary *Di* is learned for each class independently; these dictionaries are then combined to form a dictionary *D* = [*D*1,*D*2, ...*DN*] where *N* is the number of classes. Optimizing sub-dictionaries to be lowrank, however, might reduce diversity across items within each sub-dictionary. It results in a decrease of the dictionary's representation power. We present a discriminative, structured low-rank framework for image classification. Label information from training data is incorporated into the dictionary learning process by adding an ideal-code regularization term to the objective function of dictionary learning. Unlike [9], the dictionary learned by our approach has good reconstruction and discrimination capabilities. With this high-quality dictionary, we are able to learn a sparse and structural representation by adding a sparseness criteria into the low-rank objective function. Images within a class have a low-rank structure, and sparsity helps to identify an image's class label. Good recognition performance is achieved with only one simple multi-class classifier, rather than learning multiple classifiers for each pair of classes [2]. In contrast to the prior work [5, 1] on classification that performs low-rank recovery class by class during training, our method processes all training data simultaneously. Compared to other dictionary learning methods [15] that are very sensitive to noise in training images, our dictionary learning algorithm is robust. Contaminated images can be recovered during our dictionary learning process. Recent years have witnessed the rapid development of face recognition thanks to enormous theoretical breakthrough and increasingly strong computing capability. Nonetheless, most existing algorithms work fine under laboratory conditions but fail dramatically in practical, largely because of the variations they suffer in misalignment, illumination, expression and partial occlusion. Among all these problems, occlusion is viewed as the most common and challenging one. It is necessary for a practical face recognition system to handle occlusion, such as hats, scarvesor sunglasses. Sometimes, exaggerated facial expression is also treated as another kind of occlusion. These occlusions may destroy essential discriminative information, leading to misclassification, especially when the occlusion occupies large region.

LITERATURE SURVEY: In the literature of face recognition, traditional holistic feature based approaches [1, 2] are sensitive to outlier pixels, performing poorly against occlusion. Local feature based approaches are roughly divided into two categories in terms of patch-dependent [3-5] or data-dependent [6, 7]. Combined with partial matching methods [8, 9], the local features are more robust because they are not extracted from the entire image. Nonetheless, they are still inevitably affected by invalid pixels and far from being robust enough in practical classification tasks. An alternative solution addresses occlusion via a two-stage approach. It first identifies and discards the occluded pixels, and then performs the classification on the rest [10, 11]. As one can imagine, its classification performance is greatly determined by the occlusion identification accuracy. If too much discriminative information is abandoned, the following classification becomes difficult. To enhance the accuracy of occlusion identification, [11] adopts the prior that the occlusion is spatially continuous and consequently achieves excellent performance. However, such unsupervised approach might cause misestimate when occlusion is severe. For instance, a scarf larger than half of the testing face may be considered as a useful signal, and therefore face pixels may be discarded in each iteration. We call it a degenerate solution. Besides, the algorithm in [11] has to be carried out subject-by-subject and exhaustively search the class with the minimum normalized error, which is time-consuming and detrimental to real-time applications. Recently, several occlusion dictionary based approaches [13] for robust face recognition have been attached more and more importance. This kind of method is capable of efficiently handling various occlusions. They exploit characteristics of non-occluded and occluded region, assuming that both of them can be coded over the corresponding part of dictionary [12]. These methods act in the similar way with each other. Concretely, an occlusion dictionary is concatenated to the original dictionary to perform occlusion coding. The goal is to jointly represent the occluded image. illustrates how occlusion dictionary methods work. By seeking a sparse solution, the occluded image successfully decomposes into face and occlusion. The classification is carried out via the corresponding coefficients. Hence, they cast the recognition problem with occlusion as the one without occlusion, since occlusion is regarded as an extra and special class in training samples.

HISTOGRAM EQUALIZATION

Histogram equalization is used to enhance contrast. It is not necessary that contrast will always be increase in this. There may be some cases were histogram equalization can be worse. In that cases the contrast is decreased. First we have to calculate the PMF (probability mass function) of all the pixels in this image. If you donot know how to calculate PMF, please visit our tutorial of PMF calculation. Our next step involves calculation of CDF (cumulative distributive function). Again if you donot know how to calculate CDF , please visit our tutorial of CDF calculation. Calculate CDF according to gray levels.

GAUSSIAN FILTER:

A Gaussian filter is a good general-purpose filter and it is the current standardized approach for the separation of the roughness and waviness components from a primary surface. In electronics and signal processing, a Gaussian filter is a filter whose impulse response is a Gaussian function (or an approximation to it, since a true Gaussian response is physically unrealizable). Gaussian filters have the properties of having no overshoot to a step function input while minimizing the rise and fall time. This behavior is closely connected to the fact that the Gaussian filter has the minimum possible group delay.

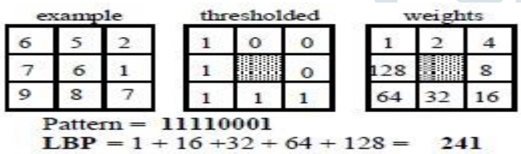
LOG GABOR FILTER:

In signal processing it is useful to simultaneously analyze the space and frequency characteristics of a signal. While the Fourier transform gives the frequency information of the signal, it is not localized. This means that we cannot determine which part of a (perhaps long) signal produced a particular frequency. It is possible to use a short time Fourier transform for this purpose, however the short time Fourier transform limits the basis functions to be sinusoidal. To provide a more flexible space-frequency signal decomposition several filters (including wavelets) have been proposed. The

Log-Gabor[1] filter is one such filter that is an improvement upon the original Gabor filter.[2] The advantage of this filter over the many alternatives is that it better fits the statistics of natural images compared with Gabor filters and other wavelet filters. The Log-Gabor filter is able to describe a signal in terms of the local frequency responses.

LOCAL BINARY PATTERNS:

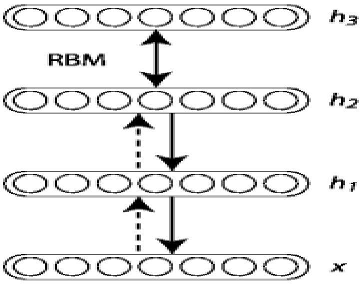
Local binary patterns (LBP) is a type of feature used for classification in computer vision. LBP is the particular case of the Texture Spectrum model proposed in 1990.LBP was first described in 1994. It has since been found to be a powerful feature for texture classification; it has further been determined that when LBP is combined with the Histogram of oriented gradients (HOG) classifier, it improves the detection performance considerably on some datasets. Local Binary Pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel with the value of the center pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. LBP is a binary code that describes the local texture pattern. It is built by thresholding a neighbourhood by the gray value of its center. The idea is illustrated in the figure below.



DEEP BELIEF NETWORKS: It Shows that RBMs can be stacked and trained in a greedy manner to form so-called Deep Belief Networks (DBN). DBNs are graphical models which learn to extract a deep hierarchical representation of the training data. They model the joint distribution between observed vector x and the ℓ hidden layers h^k as follows:

$$P(x,h^1,\ldots,h^\ell)=\left(\prod_{k=0}^{\ell-2}P(h^k|h^{k+1})\right)P(h^{\ell-1},h^\ell)$$
(1)

where $x=h^0$, $P(h^{k-1}|h^k)$ is a conditional distribution for the visible units conditioned on the hidden units of the RBM at level k, and $P(h^{\ell-1},h^\ell)$ is the visible-hidden joint distribution in the top-level RBM. This is illustrated in the figure below.



The principle of greedy layer-wise unsupervised training can be applied to DBNs with RBMs as the building blocks for each layer . The process is as follows:

1. Train the first layer as an RBM that models the raw input $x=h^{(0)}$ as its visible layer.
2. Use that first layer to obtain a representation of the input that will be used as data for the second layer. Two common solutions exist. This representation can be chosen as being the mean activations $p(h^{(1)}=1|h^{(0)})$ or samples of $p(h^{(1)}|h^{(0)})$.
3. Train the second layer as an RBM, taking the transformed data (samples or mean activations) as training examples (for the visible layer of that RBM).
4. Iterate (2 and 3) for the desired number of layers, each time propagating upward either samples or mean values.
5. Fine-tune all the parameters of this deep architecture with respect to a proxy for the DBN log- likelihood, or with respect to a supervised training criterion.

In this tutorial, we focus on fine-tuning via supervised gradient descent. Specifically, we use a logistic regression classifier to classify the input x based on the output of the last hidden layer $h^{(l)}$ of the DBN. Fine-tuning is then performed via supervised gradient descent of the negative log-likelihood cost function. Since the supervised gradient is only non-null for the weights and hidden layer biases of each layer (i.e. null for the visible biases of each RBM), this procedure is equivalent to initializing the parameters of a deep MLP with the weights and hidden layer biases obtained with the unsupervised training strategy.

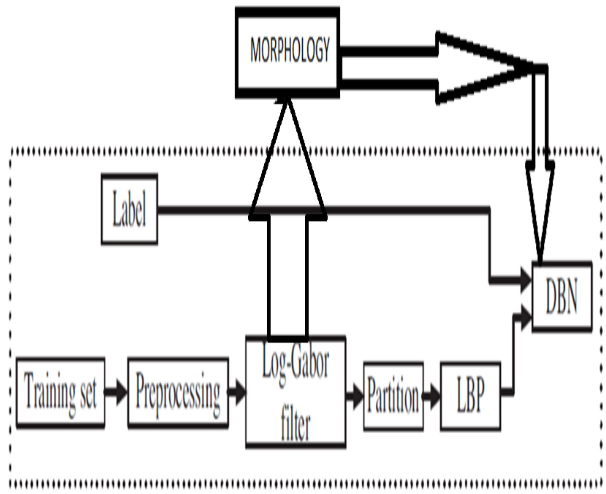


Fig: Proposed block diagram

Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. Morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. Morphological operations can also be applied to greyscale images such that their light transfer functions are unknown and therefore their absolute pixel values are of no or minor interest.

EROSION AND DILATION

The **erosion** of a binary image *f* by a structuring element *s* (denoted $f \ominus s$) produces a new binary image $g = f \ominus s$ with ones in all locations (*x,y*) of a structuring element's origin at which that structuring element *s* fits the input image *f*, i.e. $g(x,y) = 1$ is *s* fits *f* and 0 otherwise, repeating for all pixel coordinates (*x,y*).

Erosion with small (e.g. 2×2 - 5×5) square structuring elements shrinks an image by stripping away a layer of pixels from both the inner and outer boundaries of regions. The holes and gaps between different regions become larger, and small details are eliminated

RESULTS:

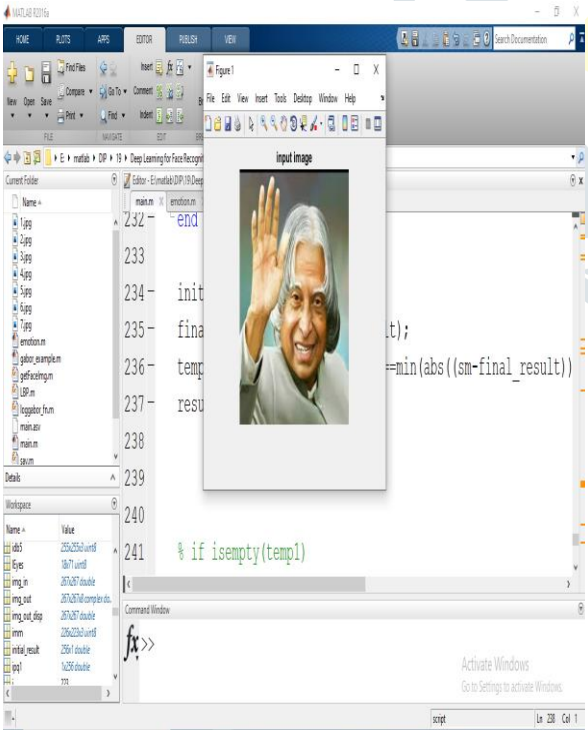


Fig1: Input image

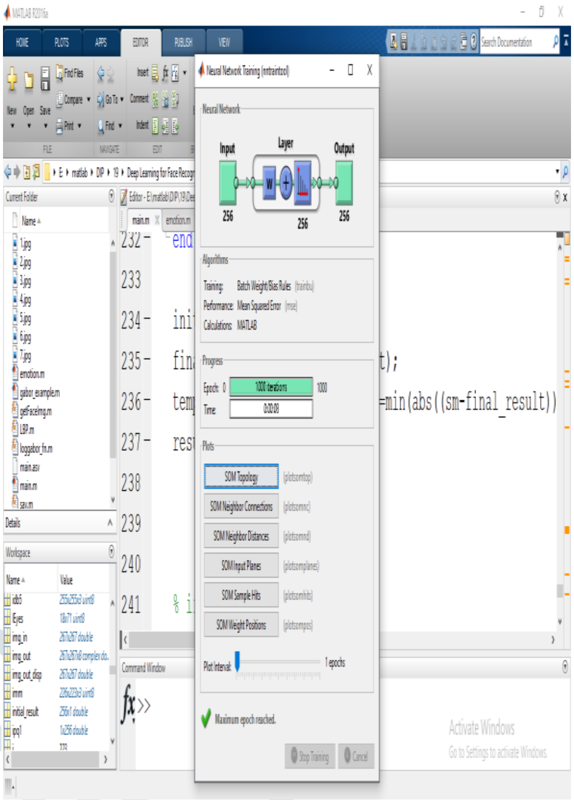


Fig2: Deep Belief Network training

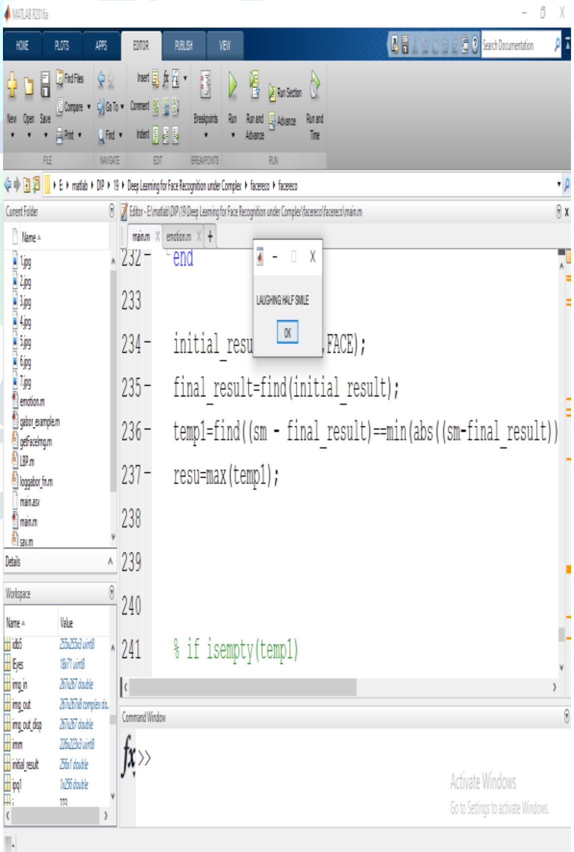


Fig3: Emotion Recognition

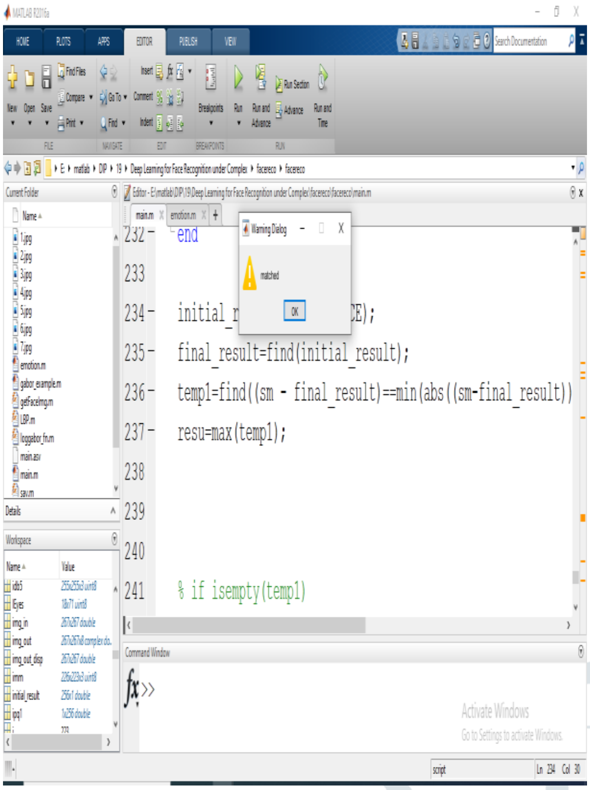


Fig4: Person recognition

CONCLUSION:

This paper implemented an automatic system to extract human face and facial features for facial expression recognition. **Proposed algorithm** uses local information, extracted from representative facial fiduciary points, with a set of Gabor filters, Local Binary patterns and Morphological operations. Results obtained from the experiments shows that morphological operations can be efficiently used for facial feature extraction.

REFERENCES:

[1] Adini. Y, Moses. Y, and Ullman. S, (Jul. 1997) "Face recognition: The problem of compensating for changes in illumination direction," IEEE Trans.Pattern Anal. Mach. Intel., vol. 19, no. 7, pp. 721–732.

[2] Ahonen. T, Hadid. A and Pietikainen. M, (Dec. 2006) "Face description with local binary patterns: Application to face recognition," IEEE Trans.pattern Anal. Mach. Intel., vol. 28, no. 12, pp. 2037– 2041.

[3] Dalal. N and Trigg's N, (2005) "Histograms of oriented gradients for human detection," in Proc. CVPR, Washington, DC, pp. 886–893.

[4] Lee. K, Ho. J, and Kriegman. D, (May 2005) "Acquiring linear subspaces for face recognition under variable lighting," IEEE Trans. Pattern Anal. www.iosrjournals.org 88 | Page

[5] Ojala. T, Pietikainen. M and Maenpaa. T, (Jul. 2002) "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,"IEEE Trans. Pattern Anal. Mach. Intel., vol. 24, no. 7, pp. 971–987.

[6] Rafael C. Gonzalez and Richard E Woods,(2005)"Digital Image Processing", Person Education Asia.

[7] Shan. S, Gao, Cao. B, and Zhao. D, (2003) "Illumination normalization for robust face recognition against varying lighting conditions," in Proc. AMFG, Washington, DC, p. 157.

[8] Tan. X and Triggs. B, (2007) "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in Proc. AMFG, pp. 168–182.

[9] Satyanadh Gundimada and Vijayan K Asari " A Novel Neighborhood Defined Feature Selection on Phase Congruency Images for Recognition of Faces with Extreme Variations" International Journal of Information Technology Volume 3 Number 1.

[10] D. Yi, Z. Lei, and S. Z. Li. Towards pose robust face recognition. In CVPR, 2013

[11] Y. Bengio, Learning deep architectures for AI. Foundations and Trends in Machine Learning, 2009

[12] Yi Sun, Xiaogang Wang, Xiaoou Tang, Deep Learning Face Representation by Joint Identification-Verification, June 2014

[13] J. van de Weijer and Th. Gevers, Tensor Based Feature Detection for Color Images, 2014

[14] Ahmadreza Baghaie and Zeyun Yu, Structure Tensor Based Image Interpolation Method, 2014

[15] H. Li, G. Hua, Z. Lin, J. Brandt, and J. Yang, Probabilistic elastic matching for pose variant face verification. In CVPR, 2013

[16] T. Hassner, Viewing real-world faces in 3D. In International Conference on Computer Vision (ICCV), Dec. 2013

[17] Y. Sun, X. Wang, and X. Tang, Deep convolutional network cascade for facial point detection. In CVPR, 2013

[18] Voruganti Ravi Kumar, Sk Subhan, Devireddy Venkatarami Reddy, A Novel Approach for Facial Expression Recognition Rate (FER) By Using Tensor Perceptual Color Framework, 2014

[19] T. Berg and P. N. Belhumeur. Tom-vs-pete classifiers and identity preserving alignment for face verification. In BMVC, 2012