Title: Development of Vision Based Automatic Facial Expression Recognition Systems Using Machine Learning Algorithms

Author(s): Majumder, Anima

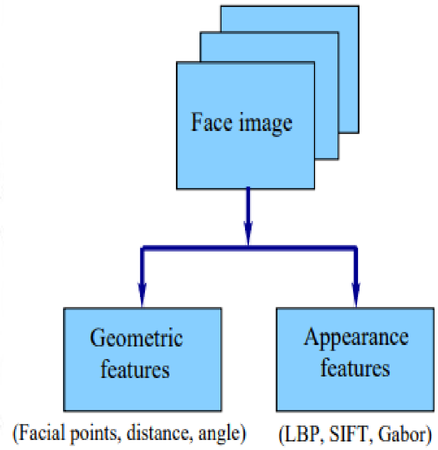
Supervisor(s): Venkatesh, K S
Behera, Laxmidhar

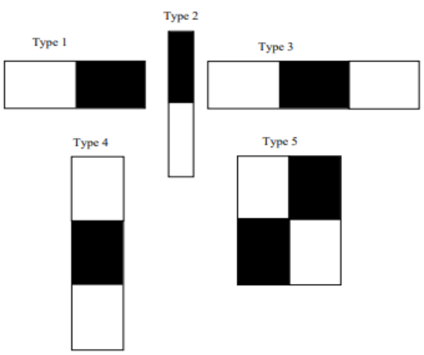
Keyword(s): Facial Expression
Self Organizing Map
Autoencoder
Imbalanced Data

Subject(s): Facial expressions recognition
Computer vision

Abstract: This thesis work is concerned with the development of an Automatic Facial Expressions Recognition System (AFERS) which can infer different emotions or intensities of emotion using facial expressions. Initial attention goes to recognition of six fundamental expressions such as happiness, sadness, disgust, anger, surprise and fear. The well established databases: MMI and extended Cohn-Kanade are used for experimental purposes. This thesis work also addresses the recognition of higher level emotions such as shoulder pain intensity, using the UNBC-McMaster shoulder pain expression archive database. The database has spontaneous, un-posed and unscripted, behavioral observations in individuals that have clinically relevant pain which make recognition task even more challenging, mainly due to lack of enough training and testing data. The major contributions of this thesis are as follows: 1. Emotion Recognition using Extended Self Organizing Map (ERESOM): An automatic facial expressions recognition system is proposed to recognize six fundamental emotions. ERESOM has two aspects: automatic facial features extraction and development of a multi-class classifier using extended self-organizing map (SOM). ERSOM uses 26 dimensional geometric facial features extracted from regions such as eye-brows, left and right eyes, nose and lips. A novel multi-class classification methodology is introduced using SOM. The learning approach takes advantage of both supervised and unsupervised techniques. The experimental results of ERESOM classifier using the geometric features proves its efficiency in recognizing six fundamental emotions while ensuring significant increase in average classification accuracy over RBFN, MLP and SVM. 2. A Emotion Recognition using Deep Network (ERDeN): The fusion of 26 geometric features and 236 LBP features (appearance features) using autoencoders is the hallmark of this approach - a novel idea to achieve a better representation of the facial attributes. An improved SOM based classifier is also introduced that further enhances classification accuracy. The fusion and classifier has been achieved in a deep network configuration consisting of autoencoders and extended SOM network. 3. Handling Multi-class Imbalanced Data (HaMID): A multi-class imbalanced data handling technique is proposed that includes sparse information extraction using directional sample space search and class prioritized synthetic data generation avoiding border violations. The technique is applied on Shoulder pain archive database to recognize 15 levels of shoulder pain intensity using facial expressions. An undersampling technique is proposed that operates on the majority classes to make them more compact by efficiently removing redundant from the class. The methodology also maintains the diversity of the data samples of the original class. It is achieved using the angular information among data samples, which is a novel contribution. An oversampling technique is proposed to generate synthetic data in the minority classes. The algorithm checks for the boundaries of other classes in the neighborhood of a data sample and generates desired number of samples in a minority class. A regression model is created using mixture of Gaussians to recognize 15 different intensity levels in

the final balanced dataset. Performance measures are analyzed for the proposed algorithms and compared with two existing techniques: SMOTE, ADASYN and random sampling.





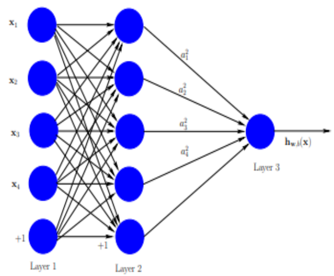


Figure 2.11: A basic architecture of a multilayer perceptron.

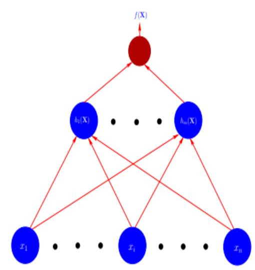


Figure 2.12: An example of a Radial Basis Function network.

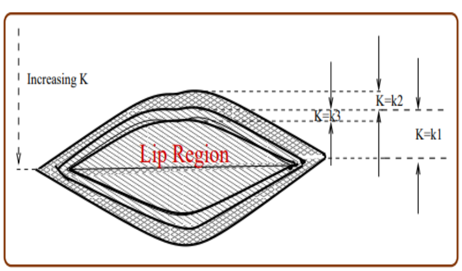
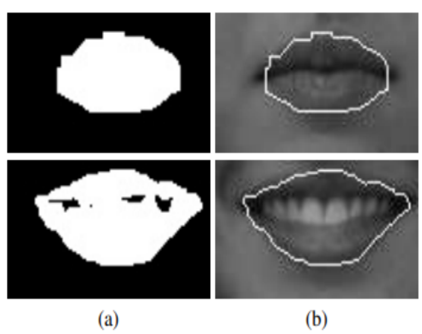
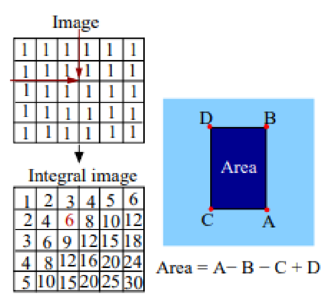
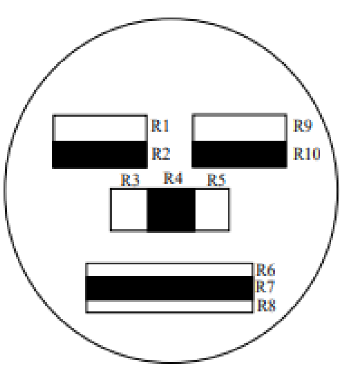


Figure 3.2: Lip region with varying constant k for ERESOM adaptive thresholding (modified Niblack's) algorithm.







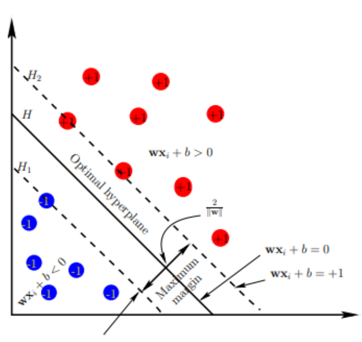


Figure 2.8: SVM based optimal hyperplane to separate data into two classes.

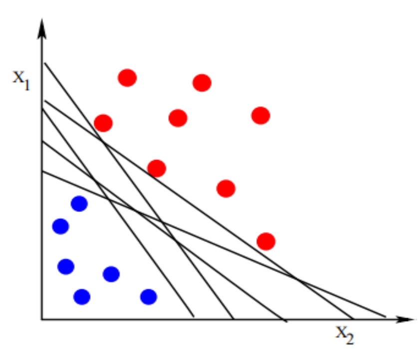


Figure 2.7: Various possible separating hyperplanes.