ITM Web of Conferences **64**, 01008 (2024) https://doi.org/10.1051/itmconf/20246401008
ICACS24

Enhancing Face Recognition for Security Systems: A n A pproach Usi ng G abor W avelet, t-SNE, and SVM

Mustafa Zuhaer Nayef Al-Dabagh[1,5*], *Hussein* Ibrahim Hussein[2,3], *Salar* Ameen Raheem[4], *Muhammed* Imran Ahmed[5], and *Nashwan* Adnan Othman[6]

[1]Department of Computer Science, College of Science, Knowledge University, Erbil 44001, Iraq
[2]Department of Computer Techniques Engineering, Alsafwa University College, Almamalie str, Karbala, Iraq
[3]Department of Information Security, College of Information Technology, University of Babylon, Hillah, Iraq
[4]Department of Information and Communication Technology Engineering, Erbil Polytechnic University, Erbil, Iraq
[5]Faculty of Electronic Engineering Technology, Universiti Malaysia Perlis, Perlis, Malaysia
[6]Department of Computer Engineering, College of Engineering, Knowledge University, Erbil 44001, Iraq

Abstract. Facial recognition is crucial for safety and security, especially for identifying people. This paper applies facial recognition to a database of facial images by analyzing the images and subsequently assigning a set of unique features to each one. The process of extracting features from the input image is accomplished using the gabor wavelet transform. t-SNE (t-distributed Stochastic Neighbor Embedding) select and reduce the dimension of features, thus specifying various aspects within the input image. These features are then used in a classification step, where a multiclass Support Vector Machine (SVM) is employed to categorize the face. Three popular databases (Yale, ORL and JAFFE) were the sources of the images used to evaluate the effectiveness of the proposed technique. The results show the system's high accuracy in identifying facial images. Specifically, our method achieved a 97.78% accuracy rate on the Yale, 97.50 % in the ORL databases and 100 % in the JAFFE databases, outperforming traditional methods by 2%. These results approved the system's accuracy in recognizing facial images.

1 Introduction

Many research studies found the broad-based implementation of the facial recognition technology in diversified areas from academic institutions and research centers to various facilities around the transport and other public service hubs. [1]. The purpose of these technologies is to detect individuals barred from access and to block their entry upon recognition. Numerous scholars have highlighted this issue, aiming to contribute to finding

* Corresponding author: mustafa.zuhaer@knu.edu.iq

definitive, effective methods for accurately identifying individuals deemed unwelcome or problematic.

The majority of studies concentrate on exploring innovative methods that primarily emphasize human traits. Consequently, a facial recognition technique has emerged, characterized by its rapid execution and ease of integration across multiple platforms. The Gabor wavelet transform is an important tool in many applications, such as image segmentation, pattern recognition, and computer vision. It combines the advantages of spatial and frequency domain analysis and extracts local frequency information, which is important for analyzing textures, and edges within an image. These properties make it a crucial component [2,3].

Classification is the next stage of the automatic recognition system that has been presented. The input image can be classified using a variety of classifiers in accordance with the matrix of data variances produced by the PCA or t-SNE. SVM, a supervised learning algorithm that is a binary classifier [4], is one of these techniques. Its goal is to locate the "optimal hyperplane" in the n-dimensional classification space, which provides the maximum margin between classes [5]. In order to accomplish this, SVM transforms the training set obtained in the event of complex nonlinear data into a higher-dimensional space in which a linear hyperplane can be used to divide the data. SVM is often reported to achieve better results than other classifiers.

The same field has seen many prior works using comparable methodologies. Some works have explored the robustness of Gabor wavelet in addressing the challenges posed by adverse conditions such as low resolution, difficult illumination, blur, and noise in face recognition. The central issue these works aim to address is the non-Euclidean nature of the covariance matrix derived from Gabor Wavelet subbands, which complicates the application of traditional Euclidean-based measures for face recognition. In the research [7], the authors propose two methods, LCMoG-CNN and LCMoG-LWPZ, utilizing Convolutional Neural Networks and Whitening Principal Component Analysis, respectively, to project covariance matrices into Euclidean space for improved feature extraction. Demonstrated across various challenging scenarios, these methods significantly enhance face recognition accuracy, integrating effectively with deep learning models to offer robust solutions for complex identification and verification needs.

Various works have endeavoured to enhance the features extracted from facial images through the development of methods. One of this interesting method combines the convolutional kernel extreme learning machine (CKELM) and the histogram of oriented gradients (HOG) to improve feature extraction. The method was evaluated based on accuracy and computational efficiency, and it achieved a recognition accuracy rate of 100% [6].

Among these innovative approaches are Dimension Reduction (DR) techniques like t-SNE, UMAP, TriMap, and the newly introduced PaCMAP. The challenge of striking a balance between maintaining local and global structures within datasets stands at the core of DR techniques. The efficacy of this algorithm is primarily assessed based on its ability to preserve these structures. According to the authors, PaCMAP has demonstrated outstanding performance, outperforming traditional methods by maintaining the integrity of local and global structures with enhanced efficiency. This will make a significant improvement in DR algorithm development [8].

Which caught the attention is the great progress that has been made in the field of pattern recognition, especially in the creation of a new two-stage classifier that improves face recognition accuracy by dealing with common problems like changes in pose, facial expressions, and lighting. Some studies suggest using K-NN and SVM classifiers together to increase the success rate of recognizing face images. The author in [9] first preprocessed face images before applying the K-NN classifier and performing feature extraction using the

HOG. Subsequently, images not recognized by the K-NN classifier are further processed using the SVM classifier, resulting in a notable increase in overall recognition accuracy.

One of the primary objectives of face recognition technology is safety and security. It can be used in personal settings like homes, where a robotic platform-based system is presented to patrol the covered area and issue a warning in the event that an unidentified person is spotted. A method known as Gabor wavelet transform was initially utilized in order to extract features of the face image. The contribution of this work is to the ongoing efforts in the field by offering an efficient, robust method for face recognition, potentially beneficial for various applications in biometric authentication, surveillance, and computer vision systems.

This research paper is structured into four main sections for clarity and depth. Section I offers a concise introduction to facial recognition systems and surveys related research, setting the stage for our contributions. Section II details the system's design, covering how we've built and implemented our facial recognition approach. In Section III, we examine the outcomes of our system, analyzing its performance and effectiveness. Finally, Section IV presents our conclusions, summarizing the key findings and implications of our work.

2 System Design

The developed system operates through four key stages to ensure effective facial recognition. Firstly, in the preprocessing step, we focus on enhancing the image quality by adjusting the image's brightness and improving its contrast until it's suitable for further processing. The second stage encompasses the core of our approach, which includes extracting features from the image using Gabor wavelet transform; the third stage is comprised of selecting and reducing the dimension of features through t-SNE and then classifying these features using SVM. Fig.1 outlines the sequence of these steps, illustrating the systematic flow from image preprocessing to the final classification.

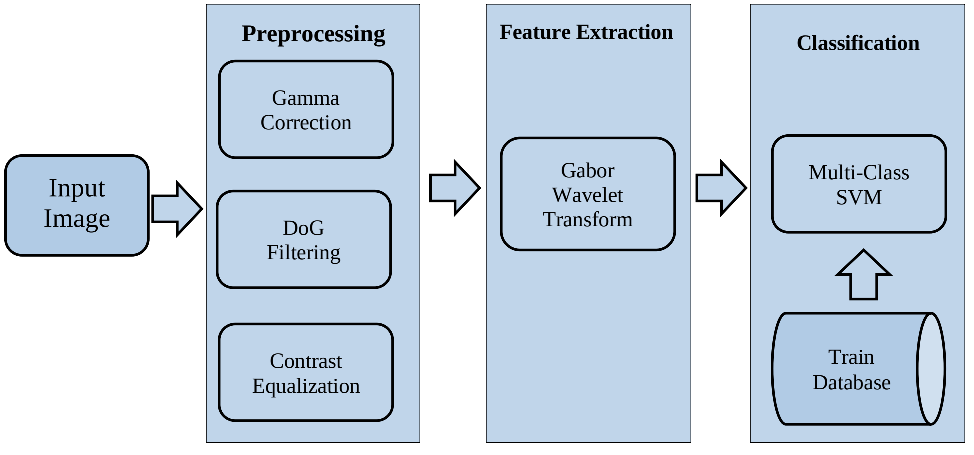

Fig. 1. The Design of System.

2.1 Preprocessing

This stage aims to enhance image quality and improve the system's performance. The process starts with the gamma correction, adjusting the brightness levels of images using a power-law function. This step is important because it gives the images the correct brightness. Then, the Difference of Gaussian filtering is used to sharpen image edges, highlight structural details at different scales, and make the features of the face more pronounced and easier to analyze. Finally, the contrast equalization technique is employed to enhance the image's overall contrast. This involves redistributing the pixel intensity values to make the features more distinguishable and pave the way for easy feature extraction and classification [10].

2.2 Feature Extraction

After the preprocessing stage, it moves to the extraction feature stage at which it aimed to extract features in the preprocessed images. In this stage, the work is focused on extracting distinct features from the preprocessed images. The information is processed within the system as the most distinctive embedded in the images, through the use of the Gabor wavelet transform method. This is a very essential process in bringing out the uniqueness lying in the face of every person for the sake of recognizing an individual through the analysis of small differences between the characteristics of the face.

2.2.1 Feature Extraction Using Gabor Wavelet Transform

The Gabor filter is a very helpful tool in image processing, particularly for identifying images. It works by using a Gaussian kernel function over a two-dimensional space, combined with a complex sinusoidal plane wave, as described below:

$$G(x,y) = \frac{f^2}{\pi\gamma\eta}\exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\exp(j2\pi f x' + \phi) \quad (1)$$

In this context, f stands for the frequency of the sinusoid, θ represents the direction of the normal to the stripes of the Gabor function, ϕ is the phase offset, σ is the standard deviation of the Gaussian envelope, and γ is the spatial aspect ratio that shapes the elliptical nature of the Gabor function's support. The coordinates x' and y' are determined by specific equations.

$$x' = \mathrm{x}\cos\theta + y\sin\theta \quad (2)$$
$$y' = -\mathrm{x}\sin\theta + y\cos\theta \quad (3)$$

Using Gabor wavelets for recognizing faces offers multiple benefits, including a certain level of consistency despite changes in position, angle, and size. They also have the capability to generalize and extract essential features from the data used for training, ensuring that the maximum amount of information about the object is encoded, given the network's size [11]. Other key advantages are the preservation of the spatial relationships among pixels, resistance to variations in lighting (provided the face has been properly normalized), resilience to noise, and straightforward updates. Additionally, recognition processes are quick, and they require a minimal amount of computing resources [12,13].

2.3 Feature Reduction

Feature reduction is one of the major preprocessing methods in machine learning and data analysis. This stage aims to change high-dimensional data into a lower dimensional space in such a way that the least possible important information is lost. Feature reduction, therefore, enhances computational efficiency and reduces overfitting. It generally aims at reducing performance variance of the model through the elimination of features that are redundant or irrelevant.

In this system, t-SNE [14,15] was employed as a method for feature reduction in the dataset. This method enabled the transformation of high-dimensional data into a lower-dimensional space, facilitating visualization and interpretation while preserving local similarities among data points.

2.3.1 Feature Reduction using t-SNE

This method has emerged as one of the premier methods for data reduction and data visualization, yielding an intuitive two- or three-dimensional representation of the data, while preserving the intricate local structure of the data set. Such a property would afford t-SNE a competitive edge versus other methods for its use in exploratory analysis of complex data sets, where the aim actually is to discover potential, inherent clusters, patterns, or trends.

At the core of the t-SNE algorithm lies the probabilistic pairwise similarity of all data points with respect to both the original high-dimensional space and the projected low-dimensional space. For each data point x_i in the high-dimensional dataset, t-SNE calculates a conditional probability $p_{j|i}$, which reflects the likelihood that x_i would choose x_j as its neighbor. This probability is derived using a Gaussian distribution centered at x_i, with the variance σ tuned to the density of the surrounding data points:

$$p_{j|i} = -\frac{\exp(-\left\|x_i - x_j\right\|^2 / 2\sigma_i^2)}{\sum_{k \neq i} exp\ (-||x_i - x_k\ ||^2 / 2\sigma\ _i^2)} \tag{4}$$

To ensure symmetry in the model, the joint probabilities p_{ij} are then defined as the mean of the conditional probabilities:

$$p_{ij} = -\frac{\mathrm{p}_{j|i} + \mathrm{p}_{i|j}}{2N} \tag{5}$$

Where N is the total number of data points.

In the reduced-dimensional space, t-SNE seeks to learn the coordinates y_i and y_j of each point such that the similarities p_{ij} are preserved. Unlike the high-dimensional similarities, the low-dimensional counterparts q_{ij} are modeled using a t-distribution with one degree of freedom to better accommodate the reduced space:

$$q_{ij} = -\frac{(1 + \left\|y_i - xy_j\right\|^2)^{-1}}{\sum_{k \neq i}(1 + ||y_k - y_i\ ||^2)^{-1}} \tag{6}$$

The essence of t-SNE's optimization process lies in minimizing the Kullback-Leibler (KL) divergence between the high-dimensional and low-dimensional probability distributions, P and Q respectively. The KL divergence serves as a non-symmetric measure of the difference between these distributions:

$$C = \mathrm{KL}(\mathrm{P}||\mathrm{Q}) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{7}$$

Minimization of this divergence through gradient descent allows t-SNE to iteratively adjust the positions of points in the low-dimensional space, effectively mirroring the original dataset's structure as closely as possible.

2.4 Classification

In the stage of classification, the obtained features are used as training for a machine-learning model, which will classify the images according to their characteristics. This classification task will be handed to several different classification approaches, such as KNN and SVM.

SVM is a supervised learning algorithm that can be used for both classification and regression tasks. SVMs have an effect on high-dimensional data and nonlinear relationships; they handle nonlinearity of relationships using kernel functions.

2.4.1 Classification using SVM

SVMs are powerful supervised learning models with very good usage in classification, but can also be used in regression. It is extended in the multi-class classification to support more than two classes in the training data set. The purpose of the multi-class SVM [16,17] is to learn a decision function, which correctly assigns input vectors to their class labels. This is actualized through optimization of a primal problem where the objective is set to minimize a combination of the norm of weight vectors with a regularization term, subject to constraints enforcing correct classification of training samples.

The optimization problem of a multi-class SVM is formulated as follows:

$$\min_{w_1, w_2, \ldots, w_K} \frac{1}{2}\sum_{k=1}^{K} ||w_k||^2 + C \sum_{i=1}^{N} \xi_i \tag{8}$$

Subject to the constraints:

$$y_i\left(w_{y_i}^T x_i + b_{y_i}\right) \geq 1 - \xi_i, \ \ \xi_i \geq 0, \ \forall_i = 1,2, \ldots, N \tag{9}$$

Where K is the number of classes, w and b are the weight vector and bias term for class k, C is the regularization parameter, ξ_i are slack variables that allow for misclassifications and y_i is the class label of sample i.

In this formulation, each class is associated with its own weight vector and bias term. The parameter C, by regularization, is responsible for balance between margin-maximizing and classification error-minimizing trade-off. Slack variables ξ_i are introduced to allow the SVM to perform a soft-margin classification, in case misclassifications are possible in the training data. The formulation of the optimization problem ensures that the decision function effectively separates the classes while minimizing the classification error.

The decision function of the multiclass SVM computes the score as a dot product of the input vector with the corresponding weight vector, adjusted by a bias. Mathematically, it can be represented as the following decision function:

$$\hat{y}(x) = \underset{k}{argmax} \left(w_k^T x + b_k \right) \tag{10}$$

Where $\hat{y}(x)$ represents the predicted class label for a new sample x.

It's important to note that multi-class SVMs can utilize various strategies to extend binary classification to multiple classes, such as one-vs-one or one-vs-all approaches. Additionally, optimization techniques like quadratic programming or gradient descent are employed to efficiently solve the primal problem. Besides, kernel methods find applications in handling non-linearly separable data via the implicit mapping of input space into a higher-dimensional feature space. Therefore, the kernel methods help to support improvements to the SVM in capturing the complex patterns lying within data.

3 Experimental Results and Discussion

This section presents an in-depth analysis and discussion of the results obtained from implementing a face recognition system integration of various processing steps: Preprocessing using the Tan-Triggs method, Feature extraction using the Gabor wavelet, Feature reduction using t-SNE, and Multiclass Classification using SVM. Experiments have been performed on two classical and well-recognized face databases: Yale, ORL and JAFFE Face Databases. These databases were selected due to their wide variety of facial expressions, lighting conditions, and poses.

3.1 Databases Description

- **Yale Face Database** [18]: This database contains 165 grayscale images of 15 individuals, with each person depicted under 11 different expressions or lighting conditions, including center-light, with glasses, happy, left-light, with no glasses, normal, right-light, sad, sleepy, surprised, and wink expressions. The diversity in expressions and accessories, such as glasses, provides a challenging set for facial recognition tasks.
- **ORL Face Database** [19]: Also known as the AT&T Face Database, this collection comprises 400 images of 40 subjects, with each subject represented by 10 images. The images vary in terms of time, lighting, facial expressions (e.g., open/closed eyes, smiling/not smiling), and facial details (glasses/no glasses). The images were taken at different times, providing variability in facial features and posing a comprehensive test for recognition algorithms.
- **JAFFE Face Database [**20-21**]:** The database includes 213 images of 10 Japanese female models showing seven different facial expressions: happiness, sadness, surprise, anger, disgust, fear, and neutral. This dataset is widely used in facial recognition research due to its diversity and accurate expression labels. In our study, we employed the JAFFE database to evaluate our method's performance in recognizing various facial expressions, demonstrating its effectiveness in real-world scenarios.

3.2 Evaluation Metrics

The performance of the facial recognition system was assessed using three metrics: accuracy, average Precision, Average and F1 Score. A straightforward measure of the overall system performance was provided by three metrics, indicating the proportion of correct predictions. Major key metrics that would show the classification performance of the model, giving an idea about the accuracy, precision, and general effectiveness of the predictions to make.

3.3 Results Analysis and Discussion

Discussion of the results, explaining their analysis, and discussion of the results are imperative parts of finding out the performance, strength, and weakness of this method. Further, in the section, it is assigned for the thorough discussion of performance metrics extracted from the application. The comparison is made of three benchmark databases, varying in complexity but aiming to provide an evaluation of the adaptability and success of the FRS under different conditions. The results are summarized in the table below:

Table 1. Experiments Results.

Database	Method	Accuracy %	Average Precision %	Average F1 Score %
Yale	Proposed Method	97.78	98.33	97.71
	LBP [22]	93.33	95	92.48
	HOG [23]	95.56	96.67	95.43
	BPPC[24]	82.22	86.67	81.02
ORL	Proposed Method	97.50	98.12	97.43
	LBP [22]	95.83	97.12	95.55
	HOG [23]	92.50	93.69	92.23
	BPPC[24]	93.33	95.50	93.07
JAFFE	Proposed Method	100	100	100
	LBP [22]	96.67	97.14	96.64
	HOG [23]	96.67	97.50	96.57
	BPPC[24]	90.00	92.74	89.56

The Proposed Method feature extraction method was better and more consistent than HOG, LBP and BPPC on all databases. Specifically, above or at 97% in any database, the Proposed Method has the highest accuracy, average precision, and F1 score. This superiority may most likely be because the Gabor Wavelet has the ability to capture spatial and frequency information with detailed and robust facial feature representation paramount for recognition.

On a comparative note, the HOG method showed moderate performance, with obtained results of over 95% in terms of accuracy, precision, and F1 score for both the databases. On the other hand, It can be said that HOG results to be moderate since it is quite simple and also computationally efficient. However, the performance showed a slightly lower performance in comparison to the Gabor Wavelet's, which might imply that it could capture less discriminative information towards the complexity of facial recognition tasks.

LBP, while effective for texture representation, may not be robust enough to ensure that the high-variance face recognition tasks are performed; this explains why it gets poor results at the evaluation in both accuracy and F1 scores, falling below the 95% mark in the ORL database and even deeper in the Yale database.

BPPC, while more modern than LBP and it is more effective, was not good enough to give high face recognition performance; this explains why it gets poor results.

From database to database, differences in performance stress the need for feature extraction techniques that can handle various degrees of complexity and variability in facial images. The Yale database, known for its challenging conditions, still saw Gabor Wavelet outperforming other methods, underlining its effectiveness in complex scenarios.

Summarily, the comparative analysis underlines a better performance of the Gabor Wavelet in face recognition tasks due to its comprehensive feature representation abilities with the wavelet transform. It also opens avenues for future research, particularly in optimizing Gabor Wavelet parameters and exploring hybrid approaches that could combine the strengths of different feature extraction methods for even better recognition accuracy and robustness against diverse and challenging conditions.

4 Conclusion

In this work, a new approach to facial recognition is described, which, through Tan-Triggs preprocessing, Gabor wavelet for feature extraction, t-SNE for feature dimension reduction, and multi-class SVM for classification, attains high performance. Testing on Yale, ORL, and JAFFE facial image databases revealed an approximate more than 97% accuracy rate, 98% average precision, and 97% average F1 score in all databases, which highlights that the method can certainly ascertain people under different conditions effectively and efficiently.

However, challenges arise in differentiating between individuals with similar features or those wearing accessories. This research significantly enriches the field of facial recognition, opening up new pathways for further development. It concludes that a holistic approach, integrating multiple processing stages, provides a strong framework for tackling facial recognition problems, offering scope for continued innovation towards more dependable recognition systems.

References

1. P. Hu, H. Ning, T. Qiu, Y. Xu, X. Luo, A.K. Sangaiah, Future Gen. Comput. Syst. **81**, 582-592 (2018)
2. A.S. Tolba, A.H. El-Baz, A.A. El-Harby, Int. J. Signal Process. **2**, 88-103 (2006)
3. L.C. Paul, A. Al Sumam, Int. J. Adv. Res. Comput. Eng. & Tech. (IJARCET) **1**, 135-139 (2012).
4. C. Campbell, Y. Ying, *Learning with Support Vector Machines* (Springer Nature, 2022).
5. E.Y. Boateng, J. Otoo, D.A. Abaye, J. Data Anal. Inform. Process. **8**, 341-357 (2020).
6. J. Anil, Meas. Sensors **30**, 100907 (2023).
7. C. Li, Y. Huang, W. Huang, F. Qin, Pattern Recogn. **119**, 108085 (2021).
8. Y. Wang, et al., J. Mach. Learn. Res. **22**, 1-73 (2021)
9. M. Chandrakala, P. Durga Devi, Mater. Today: Proc. **47**, 5771-5775 (2021)
10. C. Fernández, M. Vicente, R. Ñeco, R. Puerto, Image Preprocessing for Appearance-Based Face Recognition, 2011.
11. S. Ahmed, M. Frikha, T.D.H. Hussein, et al., BioMed Res. Int. **2021**, 6621540 (2021)
12. A.W. Muzaffar, et al., IEEE Access **11**, 60324-60334 (2023)
13. Y. Li, Y. Bi, W. Zhang, J. Ren, J. Chen, Appl. Sci. **13**, 9409 (2023)
14. L. van der Maaten, G. Hinton, J. Mach. Learn. Res. **9**, 2579-2605 (2008).
15. M. Skrodzki, H. van Geffen, N.F. Chaves-de-Plaza, T. Hollt, E. Eisemann, K. Hildebrandt, IEEE Trans. Vis. Comput. Graph. (2024).
16. E.J. Bredensteiner, K.P. Bennett, Comput. Optim. Appl. **12**(1), 53-79 (1999).
17. V. Vapnik, *Statistical Learning Theory* (Wiley, 1998)
18. Yale University, http://cvc.yale.edu/projects/yalefaces/yalefaces.html.
19. AT&T Laboratories Cambridge, ORL Face Database, https://cam-orl.co.uk/facedatabase.html
20. M.J. Lyons, M. Kamachi, J. Gyoba, Coding Facial Expressions with Gabor Wavelets (IVC Special Issue), Zenodo, 2020
21. M.J. Lyons, "Excavating AI" Re-excavated: Debunking a Fallacious Account of the JAFFE Dataset, arXiv preprint arXiv:2107.13998, 202
22. T. Ojala, M. Pietikainen, T. Maenpaa, IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 971-987 (2002).
23. N. Dalal, B. Triggs, in 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR'05), **1**, 886-893 vol. 1 (2005).
24. S. Shojaeilangari, W.-Y. Yau, J. Li, E.-K. Teoh, in Proc. 12th Int. Conf. Control Autom. Robotics & Vision (ICARCV), IEEE, 2012, pp. 166-170.