Applied Soft Computing 28 (2015) 132–137

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Automatic identification of butterfly species based on local binary patterns and artificial neural network

Yılmaz Kaya[a,*], Lokman Kayci[b], Murat Uyar[c]

[a] Department of Computer Engineering, Siirt University, 56100 Siirt, Turkey
[b] Department of Biology, Siirt University, 56100 Siirt, Turkey
[c] Department of Electrical and Electronics Engineering, Siirt University, 56100 Siirt, Turkey

ARTICLE INFO

Article history:
Received 10 September 2012
Received in revised form 13 October 2014
Accepted 30 November 2014
Available online 8 December 2014

Keywords:
Butterfly identification
Local binary patterns
Texture features
Artificial neural network

ABSTRACT

Butterflies are classified firstly according to their outer morphological qualities. It is required to analyze genital characters of them when classification according to outer morphological qualities is not possible. Genital characteristics of a butterfly can be determined by using various chemical substances and methods. Currently, these processes are carried out manually by preparing genital slides of the collected butterfly through some certain processes. For some groups of butterflies molecular techniques should be applied for identification which is expensive to use. In this study, a computer vision method is proposed for automatically identifying butterfly species as an alternative to conventional identification methods. The method is based on local binary pattern (LBP) and artificial neural network (ANN). A total of 50 butterfly images of five species were used for evaluating the effectiveness of the proposed method. Experimental results demonstrated that the proposed method has achieved well recognition in terms of accuracy rates for butterfly species identification.

1. Introduction

Insects, the most crowded family in the animal kingdom are represented with one and a half million species. The *Lepidoptera* order that involves butterflies and moths, who are distinguished from their closest relatives *Trichoptera* with flakes rather than trichome on their wings and with absorbent mouth parts in adolescents, is one of the richest teams among insects with its more than 170,000 species. The wing shape, textures and colors vary in butterflies with a great range. The figures on the wings of butterflies mostly have important roles in distinction of species at first glance. While these types of features are used as taxonomic characters as long as they are used within the species, examination of genitals organs' outer structural features of especially the male individual is required when distinguishing species very similar to each other at first glance [1]. On the other hand, these techniques are difficult to apply and time-consuming. In recent years, molecular studies are added to these identification characteristics [2]. All of the studies carried out previously, although not completely decisive, have been on supporting characters for morphological characters can be used in butterfly identifications. The aim of this study is to design an automatic machine viewing (computer vision) system that correctly identifies butterfly species based on texture features of butterfly images. To our knowledge, there are no enough studies with machine vision and machine learning in the literature to identify the butterfly species.

In previous studies, we used the energy spatial Gabor filtered (GF) (different orientations and frequencies) method with various classification methods for identification of five butterfly species. The highest obtained accuracy is 97% by ELM for five butterfly species and the accuracy obtained by ANN is 92% [3]. In another study, we obtained 96.3% classification accuracy, while employing gray-level co-occurrence matrix (GLCM) with multinomial logistic regression (MLR) for classification of 10 butterfly species and the accuracy obtained by ANN is 93.2% [4]. Additionally, 92.85% accuracy was obtained by GLCM with ANN methods for 14 butterfly species methods [5]. Furthermore, we employed GLCM and local binary pattern (LBP), the highest accuracy of identification of 19 butterfly species was obtained for these feature extraction methods is observed while employing ELM with 98.25%, 96.45% accuracy, respectively, and the accuracy obtained for ANN are 93.16% and 89.47%, respectively [6].

For classification of butterflies, the texture features on surface of butterfly wings were used. Texture analysis plays an important role in many image analysis applications. Texture can be defined as the visual or tactile surface characteristics of objects [7]. It can

* Corresponding author. Tel.: +90 506 488 49 90.
E-mail address: yilmazkaya1977@gmail.com (Y. Kaya).

http://dx.doi.org/10.1016/j.asoc.2014.11.046
1568-4946/

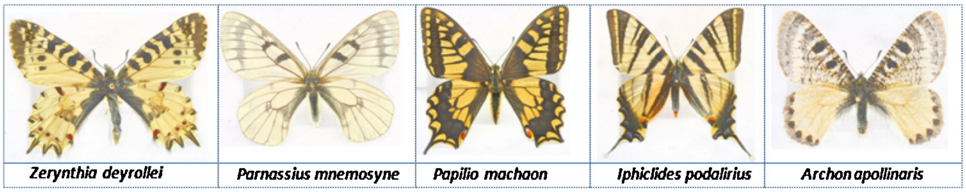

Fig. 1. The selected samples from five butterflies species.

also be formed by a single surface via variations in shape, illumination, shadows, absorption and reflectance [8,9]. Although in general there is no information on the cause of the variations, differences in image pixels provide a practical means of analyzing the textural properties of objects [10–14]. In this paper, we used the local binary pattern (LBP) operator to analyze patterns of the butterfly species and extract their texture properties. The LBP texture analysis operator was introduced as a robust descriptor of microstructures in images [15] and it has already been used in a large number of computer vision applications such as visual inspection, image retrieval, remote sensing, biomedical image analysis, face image analysis, motion analysis, environment modeling and outdoor scene analysis [16–21]. The main advantages of this operator are its tolerance to illumination changes, its computational simplicity, which makes it possible to analyze images in challenging real-time settings [17].

In present study, it is tried to prove that texture features of organisms are also decisive in outer morphological features used in identification. This study was formed with two stages; first, texture features, which were obtained from butterfly images and then classification was realized through ANN (artificial neural network). In the past years, ANNs have seen an increasingly interests in image processing. The advantage of using ANN is able to build the nonlinear and requires only input and output of data without knowing the processes in ANNs clearly. As a result of this study, identification and classification of butterfly species by using LBP texture features through ANN is showed a significant success.

The rest of this paper is organized as follows. Section 2 gives an overview of butterflies species used in study; Section 3 describes the LBP operator; Section 4 describes the feature extraction based on LBP. In Section 5, we describe the application of artificial neural network (ANN) for pattern recognition. In Section 6 the proposed method is given. The experimental results are presented in Section 7, and section 8 provides concluding remarks.

2. Data set

In this study, species belonging to family *Papilinidae* was collected from Mount Erek, Van between May 2002 and August 2003. Field studies were carried out between the attitudes of 1800–3200 m. In the field, butterflies were caught by using net trap. After being killed in jars containing ethyl acetate the butterflies were put into special envelopers prepared in advance together with labels including collection information. At the end of the field studies, samples in temporary storage boxes were put in softening containers. Samples of the butterflies softened in it for 2–3 days were pinned with standard insect pins of the appropriate numbers (0 and 1 pins). Then samples were stretched in stretching boards and dried for being prepared as standard museum materials. The drying process lasted one week in an incubator fixed to 50–55 °C. Locality labels were added to the samples rejected from stretching boards and these samples were then classified in collection drawers. External and internal morphological genital features were considered in the identification of samples. Together with morphological features of extremities and organs on head and chest, textures and colors on upper and lower sides of wings were considered in the identification made regarding external morphological features. Slides of male genital organs of some samples, which could not be identified with the external morphological features, were prepared. The identification was made by comparison of genital structures of related literature. Various handbooks, revision studies and comparison materials are used in identification, as follows in alphabetical order: Carbonell [22], Hesselbarth et al. [23], Skala [24] and Tolman [25]. Identification labels, on which the scientific names were written, were pinned on the samples at the end of the identification process. For an identification of texture features of the species used in the study 10 images for each of 5 species were used shot with a Nikon Professional camera. Butterfly species belonging to *Papilinidae* family that is spread in Van Lake basin were used in the study, as shown in Fig. 1.

3. Local binary pattern

The local binary pattern (LBP) operator was proposed to measure the local contrast in texture analysis [15]. This operator is defined as a gray scale invariant texture measure, derived from a general definition of texture in a local neighborhood [17]. The LBP can be seen as a unifying texture model that describes the structure of a texture with micro-textons and their statistical distribution rules. The original LBP operator, introduced by Ojala et al. [26], is a powerful means of texture description and it is defined as a grayscale invariant texture measure, derived from a general definition of texture in a local neighborhood. For each pixel in an image, a binary code is produced by thresholding its value with the value of the center pixel. The basic version of the LBP operator considers only the eight neighbors of a pixel, but the definition has been extended to include all circular neighborhoods with any number of pixels [15,26]. Different LBP operators can be defined according to the neighbors (see Fig. 2).

In general, $LBP_{P,R}$, where R is defined by a set of different multiscale models, P is the number of neighbors; R indicates the radius of the model. At a given pixel position (x_k, y_k), LBP operator labels the pixels of an image by using the value of the center pixel as a threshold value of the neighborhood of each pixel. If the neighboring pixel value is greater than or equal to the center pixel value this pixel takes the value 1 otherwise it takes 0. Then an LBP code for a neighborhood is formed (Fig. 3). Fig. 3 shows a basic LBP where P and R are 8 and 1, respectively. The decimal value of this binary code gives the local structural information around the given pixel.

The mathematical formulation of LBP for a pixel is as follows:

$$\mathrm{LBP}(x) = \sum_{i=0}^{P} S(G(x_i) - G(x))2^{i-1} \quad (1)$$

$$S(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (2)$$

where x is the location of the center pixel. x_i is the location of the ith neighboring pixel and $G(\cdot)$ is the pixel intensity value. Note that each bit of the LBP code has the same significance level and that two successive bit values may have a totally different meaning. Actually, the LBP code may be interpreted as a kernel structure index. By definition, the LBP operator is unaffected by any monotonic