Deep Neural Network for Fuzzy Automatic Melanoma Diagnosis

Wiem Abbes and Dorra Sellami
CEM Laboratory, National Engineering School of Sfax, Sfax University, Soukra Street, Sfax 3038, Tunisia

Keywords: Melanoma, Bag of Words, CAD System, Feature Extraction, Fuzzy C-Means, Deep Neural Network Classifier.

Abstract: Melanoma is the most serious type of skin cancer. We consider in this paper diagnosing melanoma based on skin lesion images obtained by common optical cameras. Given the lower quality of such images, we should cope with the imprecision of image data. This paper proposes a CAD system for decision making about the skin lesion severity. We first define the fuzzy modeling of the Bag-of-Words (BoW) of the lesion. Indeed, features are extracted from the skin lesion image related to four criteria inspired by the ABCD rule (Asymmetry, Border, Color, and Differential structures). Based on Fuzzy C-Means (FCM), membership degrees are determined for each BoW. Then, a deep neural network classifier is used for decision making. Based on a public database of 206 lesion images, experimental results demonstrate that the fuzzification of feature modeling presents good results in term of sensitivity (90.1%) and of accuracy (87.5%). A comparative study illustrates that our approach offers the best accuracy and sensitivity.

1 INTRODUCTION

Melanoma presents the most serious form of skin cancer. It is the most speedily developing cancer in the world (Bickers et al., 2006). The melanoma incidence continues to increase significantly, at a rate rapidly than that of any of the seven most common cancers (Ries et al., 2006). Nearly 178,560 melanoma cases will be diagnosed in 2018, with approximately 9,320 cases of death (One person dies of melanoma every hour) (Facts and Figures, 2018). There are three main types of skin cancer: Basal cell carcinoma, Squamous cell carcinoma, and Melanoma which represents only 4% of skin cancers but it is responsible for 75% of deaths from skin cancer (Haralick et al., 1973; Ferlay et al., 2010).

A diagnosis of skin lesions is performed by a two-step procedure. The first step is to differentiate between melanocytic and nonmelanocytic skin lesions and the second step is to identify benign melanocytic lesions from melanoma based on dermoscopy rules. Many dermoscopy rules are adopted by dermatologists : ABCD rule (Asymmetry, Border, Colors, and Differential structures) (Stolz, 1994), seven-point checklist (Argenziano et al., 1998), Menzies method (Menzies et al., 1996) and CASH algorithm (Henning et al., 2007). These dermoscopy rules allow making the diagnosis more reproducible and reliable. Despite using these rules, clinical analysis of malignant melanoma is still challenging.

In the last decade, various Computer-Aided Diagnosis (CAD) systems have been developed to solve this problem such as (Celebi and Zornberg, 2014; Codella et al., 2017; Lopez et al., 2017; Yu et al., 2017; Quang et al., 2017). These existing CAD systems achieve good results in term of specificity and of accuracy, but they are not satisfactory in term of sensitivity. Therefore, the automatic diagnostic process is still an open problem.

Several previous works are based on dermoscopy images. Indeed, a dermatoscope is an optical device that manipulates light features to elucidate subsurface information. This device is used by dermatologists to obtain an excellent quality of an image. Therefore, it allows identifying benign and malignant lesions. Unfortunately, the use of dermatoscopes is limited in Taiwan, USA, and India. According to recent survey reporting, only nearly half of dermatologists use dermatoscope in these countries (Engasser and Warshaw, 2010; Kuo et al., 2015; Kaliyadan et al., 2018). In fact, a dermatoscope presents an expensive solution in comparison with optical standard cameras. Accordingly, we consider, in this study, diagnosing melanoma based only on skin lesion images obtained by optical cameras. Therefore, these images have poor contrast and less saturated colors, adding a vagueness of image data knowledge.

In the field of medical image analysis, fuzzy modeling is adopted to manage the imprecision of analyzed data. In fact, the fuzzy logic theory is able to

Abbes, W. and Sellami, D.
Deep Neural Network for Fuzzy Automatic Melanoma Diagnosis.
DOI: 10.5220/0007697900470056
In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019)*, pages 47-56
ISBN: 978-989-758-354-4
Copyright © 2019 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

describe the qualitative concepts (e.g. poor, medium, and highly) and then to cope with the vagueness of experts' rules. Thus, it is well-adapted for the modeling step formulated by experts. To cope with the imprecision of extracted data, we adopt the fuzzification to model the Bag-of-Words (BoW) of the lesions.

In recent decades, CAD systems using Deep Neural Networks (DNN) has been reported to be an efficient tool for the diagnosis of skin lesions (Binder et al., 1994; Piccolo et al., 2002; Blum et al., 2004). In our work, a deep neural network classifier is used to identify the classes of benign common nevi and malignant melanoma.

In our work, we develop a CAD system for melanoma detection based on fuzzy BoW using DNN classifier. Firstly, we apply the fuzzification of features extracted from preprocessed lesion images using Fuzzy C-Means (FCM) in order to model the BoW. The result of this feature modeling step is the membership degrees of the lesion to each word of the BoWs. Then, these membership degrees are applied as input to a DNN classifier for decision making about the skin lesion severity.

The remainder of this paper is organized as follows: In section 2, we define related works on skin lesion classification. Then, in section 3, we describe a general overview of the proposed CAD system where a fuzzy modeling is proposed of BoW. Later in section 4, we give details on feature modeling step. Then, in section 5, we describe deep classification. Thus, in section 6, experimental results, based on public databases, are discussed. Finally, conclusions are drawn in section 7.

2 RELATED WORKS

Many studies on automatic diagnosis of melanoma have been developed in the last few decades where several methods and techniques have been proposed. Most of these previous researches consider only dermoscopy images (Celebi and Zornberg, 2014; Codella et al., 2017; Lopez et al., 2017; Yu et al., 2017; Quang et al., 2017). In (Celebi and Zornberg, 2014), authors present an approach for automatic estimation of clinically significant colors in dermoscopy images. The database of this study is composed of 914 images (272 melanomas and 642 non-melanoma) obtained from the EDRA Interactive Atlas of Dermoscopy. Based on color features, the rates of classification obtained are nearly 71.7% for accuracy, 75.8% for specificity and 61.6% for sensitivity.

In (Codella et al., 2017), authors adopted a CNN model in order to segment and classify the skin lesions for melanoma detection from dermoscopy skin lesion images. They yield an accuracy of %94 on segmentation and an accuracy of 85% on classification. In another study, Lopez et. al. use a CNN model for melanoma detection. They are based on transfer learning method and they acquired an accuracy of 81% on dermoscopy images (Lopez et al., 2017). In (Yu et al., 2017), authors use the fully convolutional residual network (FCRN) and deep residual network (DRN) for segmentation and classification of melanoma. They get accuracy rates of 94% and 85% for segmentation and classification, respectively. In (Quang et al., 2017), authors adopt a fully convolutional model to segment skin lesion and use a simple CNN and VGG-16 model based on transfer learning for skin lesion classification. This work achieves a segmentation accuracy of 87% and a classification accuracy of 79%. These previous works are working on ISIC Archive dataset (Marchetti et al., 2018).

According to the state of art, all the existing methods suffer from sensitivity low values. Such values result from inaccuracy modeling of melanoma so that they are confused to benign lesions. In this paper, we consider improving such modeling by taking into account the ambiguity of data in a fuzzy framework modeling. Besides, we will be based in our dataset on more samples of melanoma than those used in the state of the art, let it be half of the samples. Our dataset is then balanced. A further improvement of the accuracy of diagnosis can be obtained by applying a machine learning decision-making system, which is able to cope with the complexity of such a diagnosis of melanoma. Accordingly, a deep classifier is applied in decision making.

3 GENERAL OVERVIEW

In Figure.1, we present a general block-diagram of the proposed CAD system. Firstly, a skin lesion image is manually segmented. Secondly, the proposed system is composed of three main blocks: preprocessing, feature modeling, and classification. In the preprocessing step, a median filter (Chang et al., 2008) is applied. Briefly, the median filter is a nonlinear digital filter which is commonly adopted for noise removal. It is generally considered in digital image processing because under certain conditions, it preserves relevant image information while removing noise. Such noise decreasing is a typical preprocessing step to improve the results of subsequent steps. The second step consisting on a feature modeling aims at filling the gap between the low-level description and the diagnosis concept described by the BoWs. To cope with the in-

formation ambiguity aspect in low-level descriptors, fuzzy modeling fits better our context. Firstly, image is analyzed for feature extraction in order to define our BoW related to four specific criteria which are: asymmetry, border, color, and differential structures. Then, given these features, FCM is applied to define the membership degrees of lesion image to each word (concept) of the BoW. For classification, DNNs are adopted since they have been prowed to be the most efficient in modeling all complex decision making systems. These main steps are discussed in detail in the following sections.

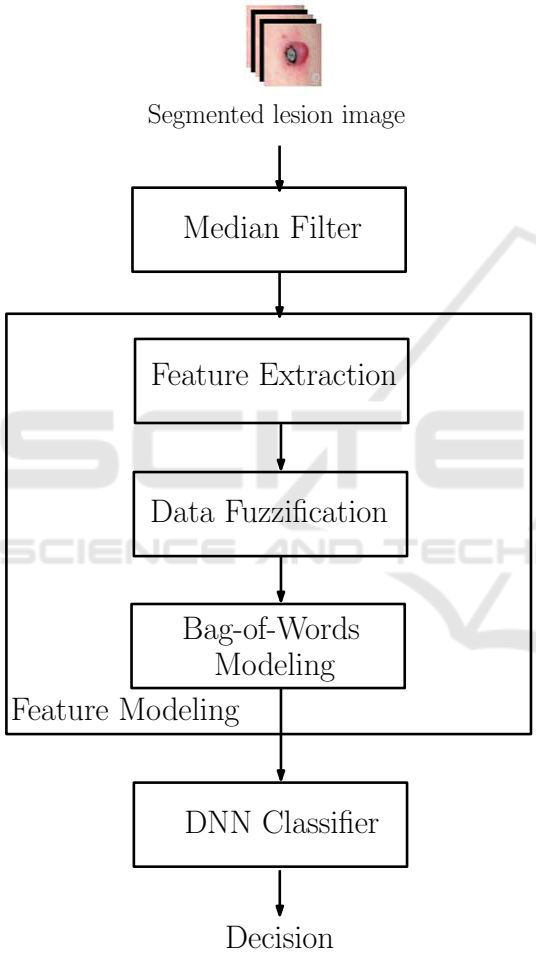


Figure 1: Graph representation of main steps of our CAD system for decision making about severity of the skin lesion.

4 FEATURE MODELING

4.1 Bag-of-Words Modeling

Dermatologists adopt some concepts to describe imaging signs related to melanoma. In our work, the BoW modeling is inspired by expert interpretation and diagnosis of melanoma. These concepts help dermatologists to identify the features from skin lesion images which are not visible to the naked eye. Besides, the feature modeling step is performed by FCM method in order to model the fuzzified BoW from extracted features (Abbes and Sellami, 2016; Abbes and Sellami, 2017).

For asymmetry, dermatologists use different modeled words to describe this concept. We can define these words by: "Poor Asymmetry", "Medium Asymmetry", and "Highly Asymmetry". Relating to the border, experts model the irregularity of the lesion border using different labels. We can describe these labels by: "Regular Border", "Fine Irregularities", and "Coarse Irregularities". For the color, the description of this concept is based on the complexity of the color distribution. In our work, two modeled words are defined for the color concept which are: "Simple Color" and "Complex Color". Based on the modeled words adopted by the dermatologists, the differential structure concept is described by two modeled words which are: "Uniform Structure" and "Composite Structure". Table.1 defines the description of each BoW.

4.2 High-level Descriptors

Inspired by the ABCD rule (Stolz, 1994), a feature set of the skin lesion image is extracted for each criterion (Asymmetry, Border, Color, and Differential structures) in order to determine the lesion severity.

4.2.1 Asymmetry

Dermatologists consider the lesion asymmetry, according to the principal axes, as an important descriptor. Indeed, they try to characterize the shape and color asymmetry of a skin lesion, since melanoma cases tend to be asymmetrically pigmented. While, the benign lesion cases have homogeneous color distributions. Moreover, melanoma cases have an asymmetrical shape, whereas, the benign lesions are elliptically shaped. Therefore, the aim of asymmetry features is to define a quantitative description of the shape and the color of the skin lesion. The asymmetry features are:

- **Central Shape Asymmetry Feature:** F_{CSA}: This feature describes the degree of lesion asymmetry relative to the center of gravity. It is computed in this work by the following method: The segmented lesion is divided by n lines passed by its center of gravity in order to achieve $2*n$ equal angles. For each sample, the average length of rays

Table 1: Description of the bag-of-words.

Concept	Bag-of-Words	Description
Asymmetry	Poor Asymmetry	It describes the poor asymmetry of the lesion
	Medium Asymmetry	It describes the medium asymmetry of the lesion
	Highly Asymmetry	It describes the high asymmetry of the lesion
Border	Regular Border	It describes the elliptical border of the lesion
	Fine Irregularities	It describes the border of the lesion that contains fine irregularities
	Coarse Irregularities	It describes the border of the lesion that contains coarse irregularities
Color	Simple Color	It describes the lesion that contains at most two colors
	Complex Color	It describes the lesion that contains at least three colors
Differential Structures	Uniform Structure	It describes the uniform texture of the lesion
	Composite Structure	It describes the composite texture of the lesion

of each angle is calculated as it is shown in Figure. 2.

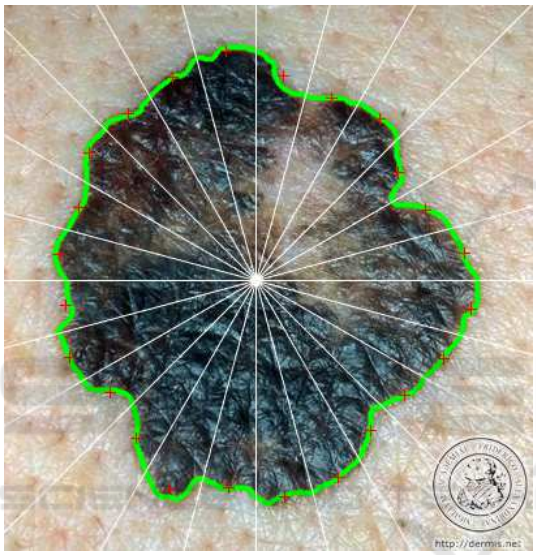


Figure 2: Lesion crossed by n lines(White), lesion border (Green), centre of gravity (White square), average rays of each angle of lesion (Red).

Then, the central shape asymmetry feature is determined by chi-square distance of two opposite average rays as it is expressed in the following equation:

$$F_{CSA} = \sum_{i=1}^{n} \frac{(\bar{r}_{(i+n)} - \bar{r}_i)^2}{\bar{r}_{(i+n)} + \bar{r}_i} \quad (1)$$

where $\bar{r}_i$ represents the average length of the rays of angle i. High values of this feature describe an asymmetric lesion, whereas, lower values define an elliptical shape.

- **Axial Shape Asymmetry:** F_{ASA}: This feature determines the lowest asymmetry of the skin lesion with respect to an axis passing by the center of gravity. The chi-square distance of the averages of two symmetric rays is calculated for each axis. The axial shape asymmetry feature is the minimum value that verifies the lowest asymmetry as it is shown in the following equation.

$$F_{ASA} = \min_{0 \le j \le (n-1)} \left(\sum_{i=(1+j)}^{(n+j)} \frac{(\bar{r}_{(2(n+j)-i)} - \bar{r}_i)^2}{\bar{r}_{(2(n+j)-i)} + \bar{r}_i} \right) \quad (2)$$

High values of this feature are achieved with an asymmetric lesion.

- **Color Asymmetry:** F_{CA}: This feature evaluates the symmetrical distribution of colors in a skin lesion. This color feature is based on the HSV color space (Smith, 1978). The skin lesion is divided into two sides where an axis of separation is the major axis of the lesion. Each side has a set of vectors specific to each space V_{ij}^c with c$\in$ $\{Hue, Saturation, Value\}$: $S1$ contains V_{11}^c,...,V_{1j}^c ,...,V_{1N}^c and $S2$ contains V_{21}^c,...,V_{2j}^c ,...,V_{2N}^c. Using the segmented lesion, the color distributions on each side are compared using chi-square distance. Therefore, the color asymmetry feature is determined by the following equation:

$$F_{CA} = \sum_{c \in \{H,S,V\}} \sum_{k=1}^{N} \frac{(V_{1j}^c(k) - V_{2j}^c(k))^2}{V_{1j}^c(k) + V_{2j}^c(k)} \quad (3)$$

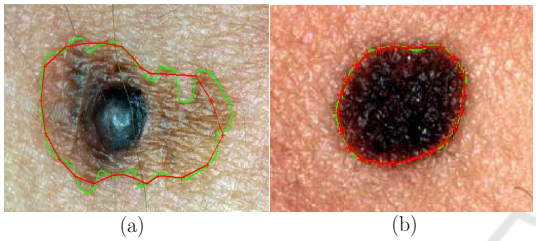
4.2.2 Border

The lesion border can be determined by the gray level transition between the inner lesion and the region surrounding it. Experts try to identify the irregular borders of the lesion, since melanoma cases generally tend to have highly irregular borders, while a benign skin lesion has a smooth border. To identify this irregularity, we characterize the following features:

- **Border Irregularity based on Area:** F_{BI}^1: Melanoma cases frequently have irregular localized patterns. To evaluate these irregularities, a smoothed method is applied to the border of the lesion. In fact, the average rays, computed in the previous section, are marked in the median of each angle. Therefore, the smoothed border is determined by the link between all extreme points of the average rays as it is illustrated in Figure.3.

Then, border irregularity feature based on the area can evaluate the fine irregularities based on the normalized area between the original shape and the smoothed shape. A regular border presents very little difference, whereas, a highly irregular border has large area differentials.

- **Border Irregularity based on the Perimeter:** F_{BI}^2: This feature evaluates the coarse irregularities of the skin lesion. It is defined by the difference of the smoothed shape perimeter and the original shape perimeter. A high value of this feature presents a highly irregular border.



(a) (b)

Figure 3: Smoothed lesion, original shape (Green), smoothed shape (Red). (a) melanoma lesion where F_{BI}^1=13.64 and F_{BI}^2=0.22 (b) benign lesion where F_{BI}^1=5.21 and F_{BI}^2=0.04.

- **Shape Signature:** F_{BS}: This feature is defined by the variance of the distance of boundary points from the center of gravity of the skin lesion. Then, it is normalized to make it scale-invariant, by dividing it by the maximum ray which is the maximum distance between the center of gravity and the border points. Finally, this feature is calculated by the sum of the difference between the normalized signature and its average.
- **Compactness:** F_{BC}: This feature is defined as the ratio of the area to the lesion perimeter as it is shown in the Equation.4. It is used to evaluate the connectedness of the skin lesion. Melanoma cases have low compactness values, whereas, benign lesions present highly compactness values.

$$F_{BC} = \frac{4\pi.A}{P^2} \tag{4}$$

where A is the lesion area and P presents the perimeter of the lesion shape.

4.2.3 Color

The scientific definition of melanoma is usually associated with the presence of different colors, which can be defined as non-uniform distributions of colors compared to the benign skin lesion. Dermatologists deal with a color number that is between one and six. Therefore, this color feature evaluates the number of colors in a skin lesion. A way to identify this feature in our study is to apply k-means method where k is the number of colors. The number of classes "k" is varied from one to six. The result of the k-means method is a reconstruction of the lesion image accordingly. For each reconstruction, the intra-class and inter-class variance are calculated. The best reconstruction is achieved with the lowest ratio of intra-class to inter-class variance. Therefore, it is associated with the optimal color class number.

4.2.4 Differential Structure

Texture analysis identifies the region characterization in an image by their texture data. It presents one of the useful methods for the discrimination of skin cancer with high precision. It consists on computing the statistics of pairs of neighboring pixels, using the co-occurrence matrix (Haralick et al., 1973). The Gray Level Co-occurrence Matrix (GLCM) is a relevant approach for texture description by mapping the grey level co-occurrence probabilities based on spatial relations of pixels in different orientations (Sheha et al., 2012). The differential structure features are extracted based on texture analysis using normalized GLCM with an angle value of 0° and a distance value of 1. To normalize the GLCM, we compute the sum of all the values in each GLCM in the array and divide each element by its sum.
The features based on GLCM in our work are as follows: Correlation, Autocorrelation, Cluster Prominence, Contrast, Cluster Shade, Difference variance, Dissimilarity, Difference entropy, Entropy, Energy, Maximum probability, Sum average, Sum entropy, Sum of squares Variance, Sum variance, Homogeneity, Inverse difference homogenous (INV), Inverse difference moment normalized, Inverse difference normalized (INN), Information measure of correlation 1, and Information measure of correlation 2.

4.3 Modeling using Fuzzy C-Means

4.3.1 Fuzzy C-Means

FCM was determined by Dunn in 1973 (Dunn, 1973) and developed by Bezdek in 1981 (Peizhuang, 1983). In our work, FCM is adopted for lesion annotation in order to cluster the lesions with respect to concepts of each criterion (Asymmetry, border, color, differential structures), given the features extracted from the skin lesion image. Data partitioning into clusters is performed by minimizing an objective function. The latter allows minimizing intra-cluster variance. Minimizing objective function means increasing similarity among all the components within a cluster. Therefore, the

adopted objective function is expressed by the following equation:

$$J = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^2 \|x_i - c_j\|^2 \quad (5)$$

where:

- x_i is the i^{th} sample of a d-dimensional measured data;
- c_j is the d-dimension center of the cluster j;
- u_{ij} is the degree of membership of x_i to cluster j;
- $\| * \|$ is a norm function expressing the Euclidean distance;
- N is the number of samples;
- C is the number of clusters.

FCM is carried out by an iterative optimization of the adopted objective function J, applying the update of the centers of clusters c_j and the membership degrees u_{ij}. Firstly, the matrix u_{ij} is randomly initialized and then it is modified in order to achieve $\sum_{j=1}^{C} u_{ij} = 1$. After that, the update of the cluster centers c_j and the membership u_{ij} is performed through the following expressions deduced from equation .5:

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^2 \times x_i}{\sum_{i=1}^{N} u_{ij}^2} \quad (6)$$

$$u_{ij} = \frac{1}{\|x_i - c_j\|^2 \sum_{k=1}^{C} \left(\frac{1}{\|x_i - c_k\|}\right)^2} \quad (7)$$

By iteratively updating the cluster centers and the membership degrees for each sample, FCM modifies the centers of clusters to the "perfect" location within a data set. This update iteration will stop when $\max_{i,j}\{|u_{ij}^{(k+1)} - u_{ij}^{(k)}|\} < \varepsilon$ where k presents the iteration steps, and ε, defining a termination criterion, must be very low.

4.3.2 Modeling Step

Once the features are extracted from a lesion image, the different words are modeled. Firstly, as a preprocessing step, we normalize the extracted data. Data normalization is useful because our data has varying scales and we use the algorithm that does not make assumptions about the distribution of extracted data. Normalization is the process of rescaling one or more features to the interval [0,1]. Accordingly, for each feature dimension, we calculate the mean of the feature and subtract it from the dataset. After that, we compute the standard deviation of each feature and divide each feature by its standard deviation as it is shown in the following equation:

$$fn = \frac{|f - \overline{f}|}{\sigma_f} \quad (8)$$

where:

- fn presents the normalized feature,
- f is the extracted feature,
- $\overline{f}$ defines the mean of the feature f,
- σ_f presents the standard deviations of the feature f.

For asymmetry, FCM was carried out to classify the skin lesions into three classes: "Poor asymmetry", "Medium asymmetry" and "Highly asymmetry", using the normalized features (central shape asymmetry, axial shape asymmetry, and color asymmetry). Accordingly, the membership degrees of the lesion are computed for each cluster.

Using the normalized feature of the border, FCM was applied to model the lesions by the three bag-of- words: "Regular border", "Fine irregularities" or "Coarse irregularities". Thus, the membership degrees of the lesion are calculated for each word.

Based on the color feature, FCM was performed to annotate the lesion with two bag-of-words: "Simple color" and "Complex color". Therefore, the membership degrees for each modeled feature are computed.

Based on the normalized GLCM features, FCM was performed to cluster the lesions into "Uniform structure" and "Composite structure". Then, membership degrees are deduced.

Figure. 4 illustrates the modeling step of each word of the BoW.

where:

- NF_{CSA}, NF_{ASA}, and NF_{CA} present the normalized asymmetry features of F_{CSA}, F_{ASA}, and F_{CA} , respectively.
- NF_{BI}^1, NF_{BI}^2, NF_{BS}, and NF_{BC} are the normalized border features of F_{BI}^1, F_{BI}^2, F_{BS}, and F_{BC}, respectively.
- NCF presents the normalized color feature.
- $NGLCM$ defines the normalized GLCM features.

5 CLASSIFICATION

In medical diagnosis areas, Artificial Intelligence (AI) techniques are the most able to cope with such com-

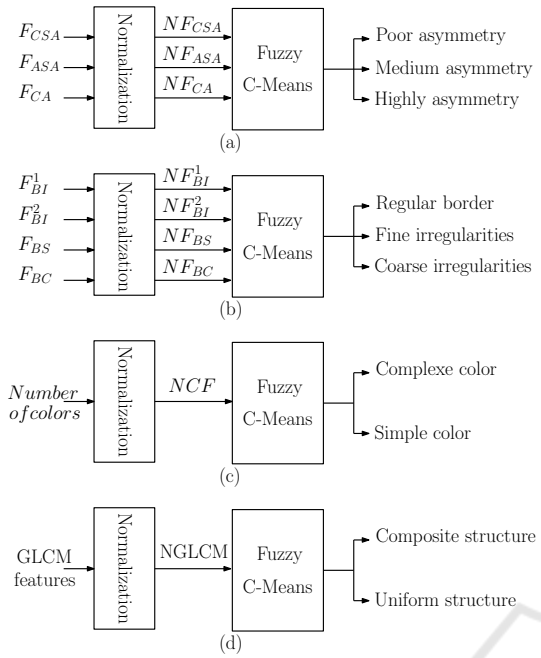


Figure 4: The modeling step from the extracted features. (a) Asymmetric modeling (b) Border modeling (c) Color modeling (d) Differential structure modeling.

plexity in decision making. Recently, Deep Neural Network (DNN) has developed as an active research sub-area of AI and it has been significantly increased performance on various pattern-recognition tasks. CAD system based on DNN has been reported to be an accurate tool for the evaluation of pigmented skin lesions (Binder et al., 1994; Piccolo et al., 2002; Blum et al., 2004). DNNs are considered as an efficient classifier. It is a computational paradigm based on mathematical models that unlike traditional computing have a structure and operation that imitates the mammal brain (Sordo, 2002). The input features of DNNs are processed in parallel in interconnected processors. Moreover, the ability of adaptation by learning algorithms adjusts the connectivity between the nodes of DNNs. DNNs can be trained by specific domain samples, therefore, they obtain their "knowledge" about appropriate processing in order to extract pertinent information from those samples.

In our work, we apply the DNN process in order to make the decision about the severity of the lesion. Indeed, an input layer of nodes represents external data which are the membership degrees of the lesion to each BoW, whereas an output layer represents the identity of classes which are the benign and the melanoma classes, and two layers of hidden nodes were included to represent possible interactions among the input variables as it is shown in Figure.5. We apply the hyperbolic tangent sigmoid as a non-linearity function and gradient descent as an optimization algorithm in this neural networks.

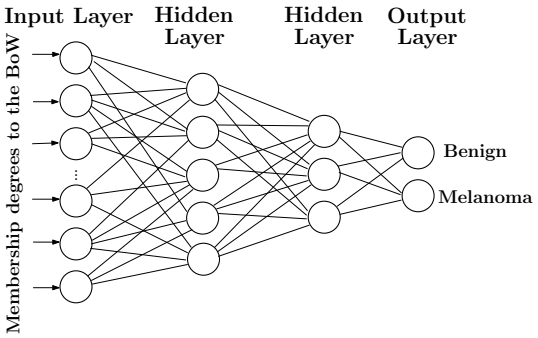


Figure 5: A multilayered deep neural network for decision making based on BoW memberships.

6 EXPERIMENTAL RESULTS

In this section, we show an experimental evaluation of the proposed CAD system. Firstly, we represent the experimental data set. Then, we define the evaluation metrics as well as the performance evaluation of the proposed CAD system. Based on performance metrics, results are compared to the existing approaches.

6.1 Experimental Dataset

In order to validate our work, we use 206 images of skin lesion, in which 119 are melanoma and 87 are benign lesions. These skin lesion images were obtained using standard cameras in varying environmental conditions. Each image has a single region of interest. This dataset was extracted from the two online public databases DermQuest (Der, 2012) and Dermatology Information System(DIS, 2012). Figure. 6 shows some samples of our experimental dataset.

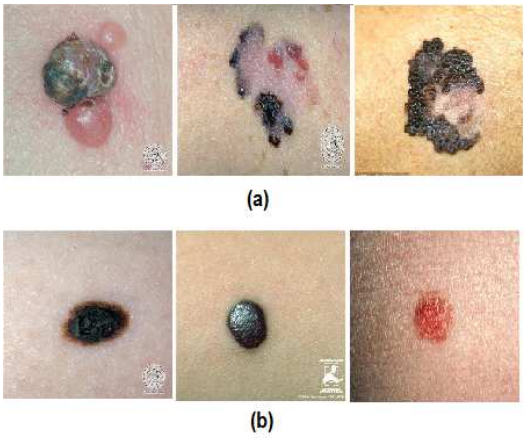


Figure 6: Samples of skin lesion images acquired by an optical camera: (a) Melanoma cases and (b) Benign lesions.

6.2 Evaluation Metrics

To evaluate the error rate of the diagnostic results, some kind of performance metrics is determined in terms of the true and false negatives as well as true and false positives. Using these terms, the performance of the classification is described by the sensitivity, specificity and accuracy metrics. Specificity determines the proportion of negatives that are correctly identified. In our work, it presents the percentage of benign lesions which are correctly identified as benign. Sensitivity defines the proportion of positives that are correctly identified. Here, it presents the percentage of melanomas which are correctly identified as melanoma. Whereas accuracy determines the proportion of negatives and positives that are correctly identified to the total number of benign lesions and melanomas. Thus, the perfect diagnostic test would correctly classify skin lesions with 100% of sensitivity, of specificity and of accuracy. These evaluation metrics are expressed in the following equations:

$$Sensitivity = \frac{TP}{TP+FN} \times 100 \quad (9)$$

$$Specificity = \frac{TN}{TN+FP} \times 100 \quad (10)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (11)$$

where TP is the number of true positives, FP is the number of false positives, FN is the number of false negatives and TN is the number of true negatives. Therefore, TP represents a correct malignant prediction, FP is the incorrect malignant prediction, FN defines the incorrect benign prediction and TN describes the correct benign prediction. Thus, FP and FN should be minimized using the proposed CAD system.

6.3 Performance Evaluation

The proposed CAD system allows decreasing FP value, which presents the most dangerous cases. Sensitivity, specificity, and accuracy are the evaluation metrics of our CAD system. They are calculated in order to evaluate the diagnostic performance. Accordingly, our system yields a good sensitivity of 90.1%, an acceptable specificity of 84.4% and a good accuracy of 87.5% on our database.

6.4 Comparative Study

In this section, we define a comparative study of our approach with five recent studies from the state-of-the-art (Celebi and Zornberg, 2014; Codella et al., 2017; Lopez et al., 2017; Yu et al., 2017; Quang et al., 2017). We are interested in comparing the results of these different diagnostic approaches to situate our approach among the state of the art methods. These previous works are considering dermoscopy images. Each work develops automated systems of image analysis in order to detect malignant melanoma from skin lesion images.

About the work of (Celebi and Zornberg, 2014), authors adopt a k-means clustering algorithm and a symbolic regression for classification. The classification rates are nearly 75.8% for specificity, 71.7% for accuracy, and 61.6% for sensitivity. In (Codella et al., 2017), authors combine feature extractors coded by hand, sparse-coding approaches, and Support Vector Machines (SVM), with deep residual networks and fully convolutional neural networks for decision making about melanoma. This work yields an accuracy of 85%, specificity of 93% and sensitivity of 54%. The work of Lopez et al. (Lopez et al., 2017) is based on an existing CNN architecture. Firstly, it trains the CNN from scratch. Then, it uses the transfer learning method to leverage features from a VGGNet model pre-trained on a larger data set. Finally, it keeps the transfer learning method and fine-tuning the existing CNN architecture. The classification rates are 81% for accuracy and 78% for sensitivity. In (Yu et al., 2017), Yu et al. develop deep networks, which have more than 50 layers, for both segmentation and classification steps to acquire more relevant features for more precise recognition. They use the residual learning methods. Indeed, they develop a fully convolutional residual network (FCRN) using a multi-scale contextual information integration scheme. The evaluation of this work yields a sensitivity of 54%, a specificity of 93% and an accuracy of 85%. Quang et al. (Quang et al., 2017) use two different CNN architectures to improve classification accuracy. The first method adopts a simple CNN and it trains the dataset from scratch. In the second method, they are based on VGG-16 with fine-tuning approaches in order to make use of pre-trained VGG-16 on ImageNet dataset. The classification rates of the first method are 49% for accuracy, 83% for sensitivity and 41% for specificity. The second method achieves 79% for accuracy, 34% for sensitivity and 90% for specificity.

We illustrate the performance metrics (Specificity, sensitivity, and accuracy) of the results of these studies in Table 2. We note that most of the previous studies are poor in term of sensitivity. Poor sensitivity is due to high false negative rates, i.e. the rate of melanoma detection as a benign lesion is high. Furthermore, we remark that the previous studies based on CNN architecture are working on unbalanced da-

Table 2: Comparative Study.

Performance		Sensitivity (%)	Specificity (%)	Accuracy (%)
(Celebi and Zornberg, 2014)		61.6	75.8	71.7
(Codella et al., 2017)		54	93	85
(Lopez et al., 2017)		78	-	81
(Yu et al., 2017)		51	94	85
(Quang et al., 2017)	Method 1	83	41	49
	Method 2	34	90	79
Our work		**90.1**	84.4	**87.5**

taset between two classes (80% benign lesion class and 20% melanoma class). Thus, the lesions are diagnosed as benign in most cases, and this leads to increasing the specificity result rates. Therefore, we are trying to improve the sensitivity in our study by keeping an acceptable result of specificity. Indeed, our CAD system achieves a good sensitivity of 90.1% and an acceptable specificity of 84.4%. This is due to the relevant extracted features, the fuzzification of feature modeling, and the deep neural network based classification. In conclusion, classification using fuzzy feature modeling exhibits the best sensitivity, while reaching a better accuracy compared to all previous studies.

7 CONCLUSIONS

In this work, we propose an approach using fuzzy feature modeling and DNN in order to help dermatologists and improve the accuracy of melanoma diagnosis based on standard optical images. Firstly, we present a framework for modeling the extracted features using FCM method and proposes BoW for quantifying skin lesion characteristics for melanoma detection. This BoW is modeled to describe some human-observable characteristics. The output of this feature modeling step is the membership degrees of the lesion to the different label in BoW. Then, these membership degrees are applied as input to the DNN classifier. The proposed approach achieves promising results in a classification problem. It yields the best result in term of accuracy (87.5%)and of sensitivity (90.1%). However, this approach has some inconveniences that can be avoided in the future. In future work, we will improve the specificity rates by adopted the convolutional neural networks.

REFERENCES

(2012). Dermatology information system. *www.dermis.net. from DermIs.net*.

(2012). Dermquest. *www.dermquest.com. from University of California, San Francisco, UCSF*.

Abbes, W. and Sellami, D. (2016). High-level features for automatic skin lesions neural network based classification. In *Image Processing, Applications and Systems (IPAS), 2016 International*, pages 1–7. IEEE.

Abbes, W. and Sellami, D. (2017). Automatic skin lesions classification using ontology-based semantic analysis of optical standard images. *Procedia Computer Science*, 112:2096–2105.

Argenziano, G., Fabbrocini, G., Carli, P., De Giorgi, V., Sammarco, E., and Delfino, M. (1998). Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: comparison of the abcd rule of dermatoscopy and a new 7-point checklist based on pattern analysis. *Archives of dermatology*, 134(12):1563–1570.

Bickers, D. R., Lim, H. W., Margolis, D., Weinstock, M. A., Goodman, C., Faulkner, E., Gould, C., Gemmen, E., and Dall, T. (2006). The burden of skin diseases: 2004: A joint project of the american academy of dermatology association and the society for investigative dermatology. *Journal of the American Academy of Dermatology*, 55(3):490–500.

Binder, M., Steiner, A., Schwarz, M., Knollmayer, S., Wolff, K., and Pehamberger, H. (1994). Application of an artificial neural network in epiluminescence microscopy pattern analysis of pigmented skin lesions: a pilot study. *British Journal of Dermatology*, 130(4):460–465.

Blum, A., Luedtke, H., Ellwanger, U., Schwabe, R., Rassner, G., and Garbe, C. (2004). Digital image analysis for diagnosis of cutaneous melanoma. development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions. *British Journal of Dermatology*, 151(5):1029–1038.

Celebi, M. and Zornberg, A. (2014). Automated quantification of clinically significant colors in dermoscopy images and its application to skin lesionclassification. *IEEE systems journal*, 8(3):980–4.

Chang, C.-C., Hsiao, J.-Y., and Hsieh, C.-P. (2008). An adaptive median filter for image denoising. In *Intelligent Information Technology Application, 2008. IITA'08. Second International Symposium on*, volume 2, pages 346–350. IEEE.

Codella, N. C., Nguyen, Q.-B., Pankanti, S., Gutman, D., Helba, B., Halpern, A., and Smith, J. R. (2017). Deep learning ensembles for melanoma recognition in der-

moscopy images. *IBM Journal of Research and Development*, 61(4):5–1.

Dunn, J. C. (1973). A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters.

Engasser, H. C. and Warshaw, E. M. (2010). Dermatoscopy use by us dermatologists: a cross-sectional survey. *Journal of the American Academy of Dermatology*, 63(3):412–419.

Facts, C. and Figures (2018). American cancer society. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2018/cancer-facts-and-figures-2018.pdf.

Ferlay, J., Shin, H.-R., Bray, F., Forman, D., Mathers, C., and Parkin, D. M. (2010). Estimates of worldwide burden of cancer in 2008: Globocan 2008. *International journal of cancer*, 127(12):2893–2917.

Haralick, R. M., Shanmugam, K., et al. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621.

Henning, J. S., Dusza, S. W., Wang, S. Q., Marghoob, A. A., Rabinovitz, H. S., Polsky, D., and Kopf, A. W. (2007). The cash (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. *Journal of the American Academy of Dermatology*, 56(1):45–52.

Kaliyadan, F., Ashique, K. T., Jagadeesan, S., et al. (2018). A survey on the pattern of dermoscopy use among dermatologists in india. *Indian Journal of Dermatology, Venereology, and Leprology*, 84(1):120.

Kuo, Y.-W., Chang, Y.-J., Wang, S.-H., Lu, P.-H., Su, Y.-L., Chu, T. W., and Chu, G.-Y. (2015). Survey of dermoscopy use by taiwanese dermatologists. *Dermatologica Sinica*, 33(4):215–219.

Lopez, A. R., Giro-i Nieto, X., Burdick, J., and Marques, O. (2017). Skin lesion classification from dermoscopic images using deep learning techniques. In *Biomedical Engineering (BioMed), 2017 13th IASTED International Conference on*, pages 49–54. IEEE.

Marchetti, M. A., Codella, N. C., Dusza, S. W., Gutman, D. A., Helba, B., Kalloo, A., Mishra, N., Carrera, C., Celebi, M. E., DeFazio, J. L., et al. (2018). Results of the 2016 international skin imaging collaboration international symposium on biomedical imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. *Journal of the American Academy of Dermatology*, 78(2):270–277.

Menzies, S. W., Ingvar, C., Crotty, K. A., and McCarthy, W. H. (1996). Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features. *Archives of Dermatology*, 132(10):1178–1182.

Peizhuang, W. (1983). Pattern recognition with fuzzy objective function algorithms (james c. bezdek). *SIAM Review*, 25(3):442.

Piccolo, D., Ferrari, A., Peris, K., Daidone, R., Ruggeri, B., and Chimenti, S. (2002). Dermoscopic diagnosis by a trained clinician vs. a clinician with minimal dermoscopy training vs. computer-aided diagnosis of 341 pigmented skin lesions: a comparative study. *British Journal of Dermatology*, 147(3):481–486.

Quang, N. H. et al. (2017). Automatic skin lesion analysis towards melanoma detection. In *Intelligent and Evolutionary Systems (IES), 2017 21st Asia Pacific Symposium on*, pages 106–111. IEEE.

Ries, L. A., Harkins, D., Krapcho, M., Mariotto, A., Miller, B., Feuer, E. J., Clegg, L. X., Eisner, M., Horner, M.-J., Howlader, N., et al. (2006). Seer cancer statistics review, 1975-2003.

Sheha, M., Mabrouk, M., and Sharawy, A. (2012). Automatic detection of melanoma skin cancer using texture analysis. *International Journal of Computer Applications*, 42(20):22–26.

Smith, A. R. (1978). Color gamut transform pairs. *ACM Siggraph Computer Graphics*, 12(3):12–19.

Sordo, M. (2002). Introduction to neural networks in healthcare. *Open Clinical: Knowledge Management for Medical Care*.

Stolz, W. (1994). Abcd rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. *Eur. J. Dermatol.*, 4:521–527.

Yu, L., Chen, H., Dou, Q., Qin, J., and Heng, P.-A. (2017). Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE transactions on medical imaging*, 36(4):994–1004.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS