Abstract

Melanoma is the deadliest form of skin cancer. Incidence rates of melanoma have been increasing, especially among non-Hispanic white males and females, but survival rates are high if detected early. Due to the costs for dermatologists to screen every patient, there is a need for an automated system to assess a patient’s risk of melanoma using images of their skin lesions captured using a standard digital camera. One challenge in implementing such a system is locating the skin lesion in the digital image. A novel texture-based skin lesion segmentation algorithm is proposed. A set of representative texture distributions are learned from an illumination-corrected photograph and texture distinctiveness metric is calculated for each distribution. Next, regions in the image are classified as normal skin or lesion based on the occurrence of representative texture distributions. The proposed segmentation framework is tested by comparing lesion segmentation results and melanoma classification results to results using other state-of-art algorithms. The proposed framework has higher segmentation accuracy compared to all other tested algorithms.

1. Introduction

Melanoma is the most deadly form of skin cancer, with an estimated 76,690 people being diagnosed with melanoma and 9,480 people dying of melanoma in the United States in 2013. In the United States, the lifetime risk of getting melanoma is 1 in 49. Melanoma accounts for approximately 75% of deaths associated with skin cancer. It is a malignant tumour of the melanocytes and usually occurs on the trunk or lower extremities. Recent trends found that incidence rates for non-Hispanic white
males and females were increasing at an annual rate of approximately 3%. If melanoma is detected early, while it is classified at Stage I, the 5-year survival rate is 96%; however, the 5-year survival rate decreases to 5% if the melanoma is in Stage IV. With the rising incidence rates in certain subsets of the general population, it is beneficial to screen for melanoma in order to detect it early. To reduce costs of screening melanoma in the general population, development of automated melanoma screening algorithms have been proposed.

Early automated melanoma screening systems assess the risk of melanoma using images acquired via a digital dermatoscope. A dermatoscope is a special device for dermatologists to use to look at skin lesions that acts as a filter and magnifier. Images acquired through a digital dermatoscope are referred to as dermoscopy images and have relatively low levels of noise and consistent background illumination. Optional pre-processing algorithms applied to dermatological images include normalizing or enhancing image colors. However, requiring dermatologists to have a dermatoscope impedes the adoption of these systems as only 48% of practicing dermatologists use dermatoscopes. The most common reasons against using the dermatoscope include a lack of training or interest. Recent work with automated melanoma screening algorithms tries to adapt the algorithms to analyze images taken by a standard digital camera. Examples of digital images of melanoma are shown in Fig. 1(a) and (c). There is a need for a segmentation algorithm designed specifically for digital images of skin lesions.

Before extracting features from the skin lesion and classifying the lesion as malignant or benign, the location of the lesion border must be identified using a segmentation algorithm. Finding an accurate estimate of the lesion border is important because of the types of features used for classification. One common set of features is the ABCD scale: asymmetry, border irregularity, color variegation, and diameter.

Figure 1: Uncorrected and corrected skin lesion images. In (a) and (c), examples of the uncorrected skin lesion images are shown. In (b) and (d), the images after being corrected for illumination variation using the MSIM algorithm. Shadows which appear on the left side of the uncorrected images are removed in the corrected images, while the color of the lesion has changed minimally.
In particular, metrics that measure border irregularity may depend heavily on the accuracy of the estimated lesion border. Therefore, it is important that the skin lesion segmentation algorithm is accurate, as the resulting segmentation is used as an input to feature extraction and melanoma classification algorithms. Many segmentation algorithms have been proposed to locate skin lesion in images automatically.

2. Existing System

In the existing system, a segmentation algorithm has been used which is based on texture distinctiveness (TD) in order to locate skin lesions in photographs. This algorithm is referred to as the TD lesion segmentation (TDLS) algorithm. The TDLS algorithm consists of two main steps. First, a set of sparse texture distributions that represent skin and lesion textures are learned. A TD metric is calculated to measure the dissimilarity of a texture distribution from all other texture distributions. Second, the TD metric is used to classify regions in the image as part of the skin class or lesion class. The use of joint statistical information has been introduced to characterize skin and lesion textures as representative texture distributions. Then the regions in the image are classified as being part of the lesion or normal skin. This region classification algorithm incorporates the texture information captured by the TD metric.

Using a sparse texture model allows the image to be stored efficiently and allows for efficient computation of algorithms that involve textures from the image. There are many ways to learn the model, including clustering or by formulating the problem as an optimization problem. The initial over-segmentation algorithm is adapted from the statistical region merging (SRM) algorithm. SRM sorts pixels in an image to determine the order in which pixels are compared, and then merges pairs of pixels into regions based on their similarity. To learn whether each texture distribution belongs to the skin or lesion class, a TD metric is formulated. Textures include smoothness, roughness, or the presence of ridges, bumps or other deformations and are visible by variation in pixel intensities in an area.

3. Proposed System

3.1 Texture Distinctiveness

The TDLS algorithm consists of two main steps. First, a set of sparse texture distributions that represent skin and lesion textures are learned. A TD metric is calculated to measure the dissimilarity of a texture distribution from all other texture distributions. Second, the TD metric is used to classify regions in the image as part of the skin class or lesion class. Here, the first step is described in detail and Fig. 2 illustrates the overall process to learn the representative texture distributions and calculate the TD metric.

Existing sparse texture algorithms use sparse texture models for segmentation or classification of images with different texture patterns. Sparse texture models find a small number of texture representations, such as texture patches, to characterize an entire image. Sparse texture models learn important local texture details present in an image. Using a sparse texture model allows the image to be stored efficiently and allows for efficient computation of algorithms that involve textures from the image. There are many ways to learn the model, including clustering or by formulating the problem as an optimization problem. A common method to learn a sparse texture model is by employing a dictionary-learning algorithm, where a set of texture patches that can best match details in the original image is learned. Incorporating probabilistic information to learn sparse
A continuous $27^{th}$ Edition, Page No.: 1

Reshma S. Nair, Marvin P. Polson:: Skin Lesion Segmentation From Digital Images Using Statistical Texture Distinctiveness

Figure 2: Algorithm flowchart displaying the steps to learn the representative texture distributions and calculate the TD metric.

3.2: Representative Texture Distributions

An existing sparse texture model algorithm is modified to find representative sparse texture distributions from the input photograph. Our proposed sparse texture model algorithm incorporates statistical information. The advantage of using a joint probabilistic sparse model is that the sparse texture distributions can model both local and global texture characteristics. To learn the sparse texture model, a local texture vector is obtained for each pixel in the image. The input image has been corrected for illumination variation, contains $N \times M$ pixels and each pixel has $a$ channels. The texture vector contains pixels in a neighborhood of size $n$ centered on the pixel of interest. Let $s$ be a pixel location $(x, y)$ in the photograph. Then, the vector $t_s$ represent then $\times n \times a$ texture patch centered at pixel $s$. The process of extracting the texture vector for a pixel in a single channel is illustrated in Fig. 3. To account for edge pixels, the borders of image are padded.

Figure 3: Extracting a texture vector. For images with multiple channels, a separate vector is obtained for each channel and concatenated sequentially.

In the case of $a$ multiple channels, $t_{A,s}$ is the texture patch centered at pixel $s$ and corresponding to channel $A$. The texture vector is constructed by concatenating each $t_{A,s}$ corresponding to the same pixel across all channels. For example, if the color image contains three channels $\{R,G,B\}$ for each pixel, three texture vectors, $t_{R,s}$, $t_{G,s}$, and $t_{B,s}$, are extracted and...
concatenated such that \( ts = [tR, s, tG, s, tB, s] \). After extracting the set of texture vectors for an image, we have a set of \( N \times M \) texture vectors is extracted, with each vector of size \( n \times n \times a \):

\[
T = \{ts_j | 1 \leq j \leq N \times M\}.
\] (1)

Using the set of all texture vectors extracted from an image, we find a set of representative texture distributions. By characterizing the sparse model as a set of distributions, we can capture both local and global characteristics in the image. The texture distributions are able to capture the commonly occurring texture patterns found in lesion and normal skin regions. The \( k \)th representative texture distribution is defined as \( T^r_k \). By using a small set \( T^r \) comprised of \( K \) representative texture distributions instead of using all the local texture vectors, the computational complexity and memory requirements are reduced,

\[
T^r = \{ T^r_k | 1 \leq k \leq K \}.
\] (2)

Each texture vector belongs to a single representative texture distribution, which best corresponds with that texture vector. All parameters needed to characterize the \( k \)th texture distribution are contained in \( \theta_k \). Each distribution has its own distinct set of parameters. A mixture model is used to represent the set of texture distributions associated with the input photograph. Texture distributions are chosen to maximize the log-likelihood of the mixture model

\[
T^r = \arg \max \sum_{k=1}^{K} \sum_{t \in C_k} \log \left( P(t_s | T^r_k) \right)
\] (3)

Figure 4: Map of representative texture distributions. In (a) and (d), the original images are shown. In (b) and (e), five representative texture distributions have been learned and each pixel in the image is replaced by one of five colors, depending on which texture distribution that pixel is associated with. In (c) and (f), maps of the texture distinctive metric are constructed. The pixel intensities in (c) and (f) depend on the TD of the texture distribution associated with each pixel.

To find the representative texture distributions and the sets of texture vectors corresponding to each representative distribution, an unsupervised clustering algorithm is used. The set \( C_k \) is
comprised of the texture vectors corresponding to texture distribution $T_k^r$. A Gaussian distribution is assumed, so $\theta_k$ contains the two required parameters to define a multivariate Gaussian distribution. The mean and covariance of the $k$th texture distribution are represented by $T_k^r$ and $\Sigma$, respectively. $P(t_j | T_k^r)$ is the probability of the $j$th texture vector given the parameters of the $k$th texture distribution. The parameters of the texture distributions are chosen to maximize the log-likelihood in (3). Examples of photographs where pixels associated with a set of five representative texture distributions are shown in Fig. 4. Each solid color in Fig. 4(b) and (e) represents pixels belonging to the same representative texture distribution. In Fig. 4(b), the lesion is represented by texture distribution associated with dark blue and in Fig. 4(e), the lesion is represented by texture distributions associated with dark blue and light green.

### 3.3. TD Metric

A TD metric is formulated using the learned sparse texture model. Since we are only interested in two classes, normal skin and lesion, but have learned many texture distributions, multiple texture distributions must represent the same class. To measure similarity of two texture distributions, we first measure the probability that the mean of one texture distribution is a realization of the mean of the other texture distribution, which is defined as $l_{j,k}$ in (4). Because we assume that the texture distributions are Gaussian, $T_j^r$ and $\Sigma_j$ are the mean and covariance of distribution $T_j^r$. The metric $l_{j,k}$ is asymmetric, because when comparing most pairs of distributions, $\Sigma_j \neq \Sigma_k$. The measure of similarity $L_{j,k}$ given in (5) is the average of $l_{j,k}$ and $l_{k,j}$.

After $L_{j,k}$ has been calculated for each pair of texture distributions, they are normalized to be between 0 and 1,

$$l_{j,k} = \frac{1}{\sqrt{(2\pi)^{n \times n \times n} |\Sigma_j|}} \exp \left( -\frac{1}{2} (t_j^r - t_k^r)^T \Sigma_j^{-1} (t_j^r - t_k^r) \right) (4)$$

$$L_{j,k} = \frac{1}{2} (l_{j,k} + l_{k,j}) . \quad (5)$$

Interested in finding distinct texture distributions. For example, lesion texture distributions are both dissimilar from the normal skin texture distributions and also from other texture distributions, due to color variegation and textural patterns found in skin lesions. The probability that a texture distribution is distinct from another texture distribution is given by

$$d_{j,k} = 1 - L_{j,k} \quad (6)$$

Using the texture distributions and probabilities of distinctiveness, a weighted graphical model can be constructed to characterize all pair-wise relationships. The graphical model is defined as $G = \{V, E\}$. $V$ represents the set of vertices for the graphical model, which are the texture distributions associated with each pixel in the image. $E$ represents the set of edges between every pair of texture distributions, which are given a weight based on the probability of distinctiveness, $d_{j,k}$.

A TD metric $D_j$ is used to capture the dissimilarity of texture distribution $T_j^r$ from other texture distributions. The metric is defined in (7) and measures the expected distinctiveness of $T_j^r$ given the photograph $I$, where $P(T_j^r|I)$ is the probability of occurrence of a pixel being associated with a texture distribution $T_j^r$. $P(T_j^r|I)$ is estimated using the histogram of the number of pixels associated with each texture distribution across the entire image,

$$D_j = \sum_{k=1}^{K} d_{j,k} P(T_k^r|I). \quad (7)$$
In the case of normal skin texture distributions, the dissimilarity of one skin texture distribution from other skin texture distributions is very small. The TD metric for skin texture distributions is small overall. Lesion texture distributions are dissimilar from other skin and lesion texture distributions, so the textural distinctiveness metric is large. Fig. 4(c) and (f) give illustrative examples of the TD metric corresponding to each pixel in the images. A brighter pixel corresponds to a higher TD metric. In both figures, the lesion is predominately white, meaning that the lesion texture distributions have higher TD metrics, as expected. In Fig. 4(f), there are two texture distributions that correspond to the lesion class and have high TD. However, in Fig. 4(c), some normal skin pixels to the right of the lesion also have high TD. This can occur when there are unique texture patterns in normal skin areas. This commonly occurs, motivating the region classification step of the TDLS algorithm. The region classification step allows the algorithm to be more robust and minimize misclassification of pixels.

### 3.4 Region Classification

The second main step in the TDLS algorithm is to find and classify regions in the input image as being part of the lesion based on the sparse texture distributions and their associated TD metric. First, the image is over segmented, which results in the image being divided into a large number of regions. Next, each region is independently classified as representing normal skin or lesion based on the textural contents of that region. Finally, post-processing steps refine the lesion segmentation.

#### 3.4.1. Initial Regions

The corrected lesion image is divided into a large number of regions. This initial over segmentation step is incorporated to increase the TDLS algorithm’s robustness to noise. Furthermore, it allows for the use of an efficient and fast classification algorithm to find which regions belong to the skin or lesion class. The initial over segmentation algorithm is adapted from the level set segmentation method. The main difference is that the level set algorithm uses the image in the RGB color space, while the TDLS algorithm converts the photograph to the XYZ color space. The advantages of using the level set algorithm as the initial over segmentation algorithm are that it directly takes into account pixel location, is simple and is computationally efficient.

Level set methods (LSM) are a conceptual framework for using level sets as a tool for numerical analysis of surfaces and shapes. The advantage of the level set model is that one can perform numerical computations involving curves and surfaces on a fixed Cartesian grid without having to parameterize these objects (this is called the Eulerian approach). Also, the level set method makes it very easy to follow shapes that change topology, for example when a shape splits in two, develops holes, or the reverse of these operations. All these make the level set method a great tool for modeling time-varying objects, like inflation of an airbag, or a drop of oil floating in water. The result of the initial over segmentation step is a map of several regions which correspond to the normal skin or lesion classes. To reduce the number of regions, all segments that touch the edges of the photograph are merged into a single region. This is based on the assumption that the lesion is not touching the edges of the photograph, which is reasonable for situations where the photographs are captured in controlled, clinical environments. As such, regions touching the edges are all likely to be part of the normal skin class.
3.4.2. Distinctiveness-based Segment Classification

Following the initial over-segmentation step, each region must be classified as belonging to the normal skin class or lesion class based on a criterion. The classification step is illustrated in (8), where \( y \) is the resulting segmentation map. Each element in \( y \) is either 1 (lesion) or 0 (normal skin), depending on the classification results for that element's corresponding region. The threshold is denoted by \( \tau \) and it represents the decision boundary between the normal skin and lesion class. The feature used to discriminate between the two classes is the regional textural distinctiveness metric \( DR \). This metric is based on the TD across a region,

\[
y(R) = \begin{cases} 
1, & \text{if } D_R \geq \tau \text{ (lesion)} \\
0, & \text{otherwise (normal skin).}
\end{cases}
\]

From Section 2.1, each pixel in the input photograph is associated with a texture distribution. A TD metric \( D \) is calculated for each texture distribution based on the probability of it being similar to other texture distributions. This information is combined with the contents of each region to determine a regional TD metric, \( DR \). \( DR \) represents the average TD across region \( R(9) \), where \( P(T_j^R|R) \) is the probability of a pixel being associated with the \( j \)th texture distribution in region \( R \). Again, \( P(T_j^R|R) \) is estimated using the histogram of the number of pixels associated with each texture distribution across the region \( R \),

\[
D_R = \sum_{j=1}^{K} D_j P(T_j^R|R). 
\]

Finally, a threshold \( \tau \) is defined to divide the set of representative texture distributions into two classes, normal skin and lesion, and is also based on the TD metrics. There are many ways to find two classes from a one-dimensional set of features. In the TDLS algorithm, the threshold is found that divides the set of texture distributions into two classes such that the total intraclass variance of the TD metric for each class is minimized as

\[
\tau = \arg \min_{\tau} \left( \sigma_{C_1(\tau)}^2 P(T^1|C_1(\tau)) + \sigma_{C_2(\tau)}^2 P(T^2|C_2(\tau)) \right). 
\]

The threshold \( \tau \) is used to divide the set of texture distributions into two classes \( C1(\tau) \) and \( C2(\tau) \). The classes depend directly on \( \tau \) because if the distinctiveness metric of the associated texture distribution is above \( \tau \), that texture distribution is in class \( C1(\tau) \). Likewise, if it is below \( \tau \), it is in class \( C2(\tau) \). The probability that a texture distribution is in the class \( C \) for given \( \tau \) is \( P(T|C(\tau)) \) and the variance of the TD based on the elements in the class is \( \sigma_C(\tau) \). This threshold is known as the Otsu’s threshold.

3.4.3. Segmentation Refinement

After the regions are classified as being normal skin or lesion, the following post-processing steps are applied to refine the lesion border: morphological dilation and region selection. First, the morphological dilation operator is applied to fill holes and smooth the border. Morphological dilation is a process that expands binary masks to fill small holes. The shape and amount that the binary mask is expanded is controlled by a structuring element, which is a disc with a radius of 5 pixels in the TDLS algorithm. Next, since multiple noncontiguous regions may have been identified as part of the lesion class, the number of regions is reduced to one. While it is possible to have multiple lesions in a single image, it is necessary to reduce the number of lesions for the feature extraction step. Features proposed by both Celebi et al. and Cavalcanti and Scharcanski assume that
only a single lesion is being analyzed in the image. To eliminate the small regions, the number of pixels in each contiguous region is counted. The contiguous region with the largest number of pixels is assumed to correspond to the lesion class and any other regions are converted to the normal skin class. This gives the final lesion segmentation.

4. Implementation details

4.1 Color Space

In the implementation of the TDLS algorithm, the photograph is in the RGB domain and has three channels ($a = 3$). However, the algorithm can be generalized and expanded to take into account multi- or hyper spectral images of a skin lesion, where $a$ is much greater than three channels. For standard digital images, we convert the image to the XYZ color space to find texture distributions and during the initial over segmentation. Work by Terrillon et al. found that the XYZ color space proved to be an efficient color space in which to segment the skin region of human faces. These color spaces is designed to better model color perception and reduce correlation between the XYZ channels, compared to the standard RGB color space.

4.2 Learning Representative Texture Distributions

In this implementation, a two-step clustering algorithm is used. First, a $k$-means clustering algorithm is run, which is followed by learning a finite mixture model. $K$-means clustering is used as an initial step to increase the robustness and to speed up the number of iterations required for the finite mixture model to converge. $K$-means clustering finds $K$ clusters of texture data points that minimizes the sum of squared error between cluster members and the cluster mean. The optimization function for $k$-means clustering is shown in (11), where $C_k$ is the $k$th set of texture vectors, and $\mu_k$ is defined as the mean vector for the $k$th set. Here, the initial cluster means are randomly assigned. Other methods to initialize the clusters could be used to decrease the sensitivity of $k$-means clustering to initial cluster placement,

$$\hat{C} = \arg \min_C \sum_{k=1}^{K} \sum_{t_{s_j} \in C_k} \|t_{s_j} - \mu_k\|^2. \quad (11)$$

One limitation with $k$-means clustering is that it does not take into account any probabilistic information. Therefore, the second step is to apply finite mixture model clustering. To fit the finite mixture model, the model parameters in the set $\Theta$ are found to maximize the log-likelihood function shown in (12). In this implementation, a Gaussian distribution is assumed for all clusters and the model parameters are the distribution mean $\mu$ and distribution covariance $\Sigma$. $\Theta$ also contains the parameter $\alpha$, which is the mixing proportion. No closed form solution exists for (12) in general, so an expectation-maximization iterative algorithm is used. The expectation-maximization algorithm is initialized using cluster means, co-variances, and mixing proportions based on the results of the $k$-means clustering,

$$\Theta = \arg \min_{\Theta} \sum_{j=1}^{n} \sum_{k=1}^{K} \log (\alpha_k P(t_{s_j} | \mu_k, \Sigma_k))$$
Where $\sum_{k=1}^{K} \alpha_k = 1$ and $\Theta = \{\mu_1, \mu_2, \ldots, \mu_K, \Sigma_1, \Sigma_2, \ldots, \Sigma_K, \alpha_1, \alpha_2, \ldots, \alpha_K\}$. (12)

Expectation-maximization is an iterative algorithm. The initial parameters for the Gaussian mixture model are obtained from the results of the $k$-means clustering. That is, the initial Gaussian means are equal to the $k$-means cluster means:

$$\mu_k = \mu_{c_k}$$ (13)

and the distribution co-variances and mixing proportions are also dependent on the cluster results. The initial estimate of the initial mixing proportion is $P(t_{sj} \in C_k)$. It is calculated by assuming that the clusters found using $k$-means clustering have a Gaussian distribution with mean $\mu_{c_k}$ and covariance $\Sigma_{C_k}$.

$$\Sigma_k = \Sigma_{C_k}$$ (14)

$$\alpha_k = P(t_{sj} \in C_k)$$ (15)

The parameters defining the $K$ representative texture distributions are taken to be the mean and co-variances for the $K$-estimated Gaussian distributions ($t_k = \mu_k$). Furthermore, each texture vector is assigned to belong to the distribution which maximizes the weighted probability $\alpha_k P(t_{sj} | \mu_k, \Sigma_k)$. The number of clusters in $k$-means clustering or distributions in the Gaussian mixture model is 10, which is determined to best model the set of skin and lesion textures.

### 4.3 Summary of the TDLS Segmentation Algorithm

i.) Convert the corrected image to the XYZ color space.

ii.) For each pixel $s$ in image $I$, extract the texture vector $t_s$ to obtain the set of texture vectors $T$.

iii.) Cluster the texture vectors in $T$ to obtain the representative texture distributions.

iv.) Calculate probability that two texture distributions are distinct $d_{j,k}$ using (6) for all possible pairs of texture distributions.

v.) Calculate the textural distinctiveness metric $D_j(7)$ for each texture distribution.

vi.) Apply the level set segmentation algorithm to find the initial regions.

vii.) Calculate the region distinctiveness metric $DR$ for each initial region using (9).

viii.) Calculate the threshold $\tau$ between the normal skin and lesion classes (10).

ix.) Classify each region as normal skin or lesion based on the results of steps 7 and 8.

x.) Apply a morphological dilation operator to the initial lesion classification.

xi.) For each contiguous region in the initial segmentation, count the number of pixels in the region.

xii.) As the final lesion segmentation, return the contiguous region consisting of the most pixels.

### 5. Conclusion

A novel lesion segmentation algorithm using the concept of TD is proposed. A probabilistic TD metric is introduced based on a learned model of normal skin and lesion textures. Representative texture distributions are learned from the image itself and the TD metric captures the dissimilarity between pairs of texture distributions. Then, the image is divided into numerous smaller regions and each of those regions are classified as lesion or skin based on the TD map. The entire proposed framework is tested by using the illumination corrected images as the input to the texture-based segmentation algorithm. It is compared to state-of-art lesion segmentation algorithms, including...
three algorithms designed for lesion images. The proposed framework produces the highest segmentation accuracy using manually segmented images as ground truth. A larger data collection and annotation process, including additional testing on a wide range of images, will be undertaken as future work. While the experimental results show that the proposed method is able to segment the lesion in images of different scales and levels of quality, it is worth conducting a more comprehensive analysis on the impact of image quality and scale on the proposed method.

6. References


