A Comparative Study of Features Extracted in the Classification of Human Skin Burn Depth

Kuan Pei Nei[1], Stephanie Chua[1] and Ehfa binti Bujang Safawi[2]
[1]*Faculty of Computer Science and Information Technology Universiti Malaysia Sarawak*
[2]*Faculty of Medicine and Health Science Universiti Malaysia Sarawak*
94300 Kota Samarahan, Sarawak, Malaysia
16020107@siswa.unimas.my

***Abstract*—The first burn treatment provided to patient is usually based on the first evaluation of the skin burn injury by determining the burn depths. In this paper, the objective is to conduct a comparative study of the different set of features extracted and used in the classification of different burn depths by using an image mining approach. Seven sets of global features and 5 local feature descriptors were studied on a skin burn dataset comprising skin burn images categorized into three burn classes by medical experts. The performance of the studied global and local features were evaluated using SMO, JRIP, and J48 on 10-fold cross validation method. The empirical results showed that the best set of features that was able to classify most of the burn depths consisted of mean of lightness, mean of hue, standard deviation of hue, standard deviation of A* component, standard deviation of B* component, and skewness of lightness with an average accuracy of 77.0% whereas the best descriptor in terms of local features for skin burn images was SIFT, with an average accuracy of 74.7%. It can be concluded that a combination of global and local features is able to provide sufficient information for the classification of the skin burn depths.**

***Index Terms*—Skin Burn; Classification; Feature Extraction; Image Mining Approach.**

I. INTRODUCTION

Human skin is made up of three layers as shown in Figure 1, which are: (i) the epidermis, which is the outermost layer of the skin, (ii) the dermis, lay underneath the epidermis layer and is divided into two sub-layers, which are papillary layer (superficial) and reticular layer (deep) and (iii) the subcutaneous layer, which is the inner layer of the skin, constitutes of fat and connective tissue [1]. Generally, burns are classified into: (i) Superficial burn, which involves only the epidermis, (ii) Partial thickness burn, which is further divided into (a) superficial partial thickness burn, involving the entire epidermis and the upper layer of the dermis (papillary layer) and (b) deep partial thickness burn, affecting the entire epidermis and most of the dermis and (iii) Full thickness burn, in which all the layers of the skin are destroyed, and some may extend into muscle and bone [2].

When burn accident happened, patients with burn injuries often consult doctors for treatment. The doctors normally diagnose a burn injury based on visual examination. Sometimes the depth of the burn is not easily defined through visual examination, as there could be mixed depth appearance. Medical practitioners with limited experience may at times be confused with the depth and severity of the burns, especially in non-clear-cut cases. In some places like rural areas, patients may only have access to other healthcare facilities which are Medical Assistant or Nurse-led. A wrong assessment of burn depth can result in inappropriate and inaccurate initial management of the burn injuries. These mistakes will eventually translates into poor healing process, infections, undesirable scars and impaired body functions post burns.

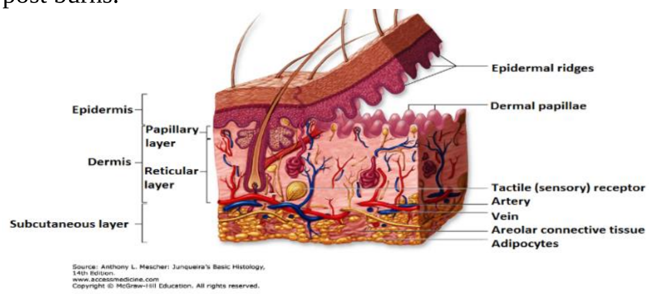

Figure 1: Human skin structure [3]

The current state-of-the-art in burn depth classification is performed by using intelligent classifiers, one of which is deep learned convolutional neural network to identify features that are capable to differentiate between healthy skin and the burn [4]. However, the images used were captured using colour-thermal camera instead of digital camera and the images were manually registered with infrared markings. Besides that, colour-thermal camera are expensive to acquire as compared to regular digital camera.

Automated classification of skin burn depth by using computer vision is still a challenging task especially when the digital images of various burn depths are captured under uncontrolled environment with various lighting and different camera resolution level used. They are two types of feature extraction approaches often used by many object recognition system for classification, which are global and local features [5]. To the best of our knowledge, there are very few or almost no other work extracting local features from the skin burn depths images. Most of the related work were extracting global features by computing the statistical moments such as mean of hue, standard deviation of chroma and so on. In this section, some works related to feature extraction are presented.

According to Lisin et al. [5], global features describe the image as a whole by generalizing the entire image with a single vector whereas local feature represent the image patches that are computed at multiple points in the image and are robust to clutter and occlusion [5]. The global features produce very compact representation of an image but are very sensitive to clutter and occlusion in which a good segmentation of an object from the background image is assumed to be available [5]. The local features, often with variable number of feature vector for each image, may require specialized classification algorithms such as Non-Parametric Density (NPD) [5]. Local features mostly represent the texture in the image patch. Local features are defined as a distinct structure in an image, in which the representation of that feature does not matter [6]. Examples of global features according to Lisin et al. [5] are contour representation, shape descriptors, and texture features [5] whereas the example of local features are blobs, corners, and edge pixels [6]. The examples of descriptors for local features are histogram of oriented gradients (HOG), local binary pattern (LBP), speeded up robust feature (SURF), fast retina keypoint (FREAK), binary robust invariant scalable keypoint (BRISK), and scale-invariant feature transform (SIFT).

Both global and local features are equally important in the classification of skin burn depths. According to Murphy et al. [7], the local features sometimes can be ambigous to a system, especially when the region of interest is relatively small or the image contain variations such as illuminants. This is when the global features are able to help in reducing this ambiguity by acting as an additional source of evidence [7].

In this paper, an image mining approach is used to evaluate the image of the skin burn injury and to classify the burn injury into one of the burn depths. Based on the burn depth classification, suitable treatment can then be recommended. Most of the classification of skin burn depths in previous research works tend to use the global feature extraction approach. Most of them used the colour feature as the main characteristic to differentiate between different burn depths. There are some related work focusing on extracting both colour and texture features. According to Acha et al. [8-10] and Serrano et al. [11], colour and texture are the characteristics observed by experts in order to differentiate the burn depths and give the diagnosis. Thus, in this research work, both colour and texture features are used and compared. The main contribution of this paper is the comparative study of features extracted in the classification of skin burn depths.

II. MATERIALS AND METHODS

This work proposes to use an image mining approach to evaluate the image of a skin burn injury and classify the burn injury into one of the burn depths. Image mining is not just an extension of data mining to image domain. It is an interdisciplinary field with a combination of techniques such as computer vision, image processing, image retrieval, data mining, machine learning, database and artificial intelligence [12]. Figure 2 shows the image mining approach that is used in this work. The image mining approach consists of several processes as described in the following sub-sections.

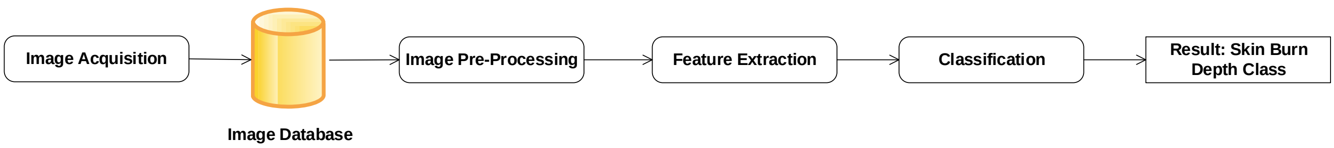

Figure 2: Image Mining Approach

A. *Image Acquisition*

The burn images used in this work were collected by a burn specialist. There is currently no open source dataset available for skin burn depth. The skin burn depths considered in this work are second degree burn and third degree burn. The burn images are categorized into superficial partial thickness (SPT) burn, deep partial thickness (DPT) burn and full thickness (FT) burn. The total images collected are 120 images; 40 for each burn class.

B. *Image Pre-Processing*

All the collected images are standardized to image file types of PNG and with the size of 90*90 pixels. The camera lighting reflected on the burn wound are also being removed. The dataset was standardized because preliminary experiments indicated that better results could be obtained.

C. *Feature Extraction*

After the images had been pre-processed, feature extraction was performed. The features used in this work are based on the features proposed in previous related works. This study is conducted to compare the use of those features in the classification of our own dataset. Table 1 shows the global features used in the related works and adopted in our work whereas Table 2 shows the descriptors for local features extraction adopted, utilizing a toolbox developed by Khosla et al. [13, 14].

Table 1
Details of global feature sets

Set ID	Type	Features extracted
1	Colour & Texture	Mean of lightness, Mean of hue, Mean of chroma, Standard deviation of lightness, Standard deviation of hue, Standard deviation of chroma, Mean A*, Mean B*, Standard deviation of A* component, Standard deviation of B* component, Skewness of lightness, Kurtosis of lightness, Skewness of A*, Kurtosis of A*, Skewness of B* and Kurtosis of B* [8-11]
	Texture	Contrast, Correlation, Energy, Homogeneity, Mean, Standard deviation, Entropy, RMS, Variance, Smoothness, Kurtosis, Skewness and Inverse difference moment (IDM) [15]
2	Colour & Texture	Mean of lightness, Mean of hue, Standard deviation of hue, Standard deviation of A* component, Standard deviation of B* component, and Skewness of lightness [8-11]
	Texture	Contrast, Correlation, Energy, Homogeneity, Mean, Entropy, Smoothness, Kurtosis, Skewness and Inverse difference moment (IDM) [15]
3	Colour & Texture	Mean of lightness, Mean of hue, Mean of chroma, Standard deviation of lightness, Standard deviation of hue, Standard deviation of chroma, Mean A*, Mean B*, Standard deviation of A* component, Standard deviation of B* component, Skewness of lightness, Kurtosis of lightness, Skewness of A*, Kurtosis of A*, Skewness of B* and Kurtosis of B* [8-11]
4	Colour & Texture	Mean of lightness, Mean of hue, Standard deviation of hue, Standard deviation of A* component, Standard deviation of B* component, Skewness of lightness [8-11]

Set ID	Type	Features extracted
5	Colour & Texture	Mean of h-space, Standard deviation of h-space, contrast, homogeneity [16]
6	Colour	Mean and (2,1)th coefficient of Discrete Cosine Transform (DCT) [17-19]
7	Texture	Contrast, Correlation, Energy, Homogeneity, Mean, Standard deviation, Entropy, RMS, Variance, Smoothness, Kurtosis, Skewness and Inverse difference moment (IDM) [15]

The motivation of comparing the different sets of feature, including the use of hybrid sets of feature is to find the set of feature that is able to represent the different classes of burn. Feature selection was also used in Set ID 2 to select the relevant features that are able to contribute to the class value.

Table 2
Descriptor for local feature extraction

Descriptors	Descriptions
Colour	The image is converted to colour names [20, 21] and the dense overlapping patches of multiple size are extracted in the form of a histogram of colour names.
Gist	The spatial envelope of the image is described by Gist descriptor [22].
Dense HOG 3X3	HOG [23] are extracted in a dense manner and are concatenated in 3x3 cells to obtain a descriptor at each grid location.
Local Binary Pattern	Non-uniform local binary pattern [24] descriptor are extracted and are concatenated on 3 levels of spatial pyramid to obtain the final feature vector.
Dense SIFT	SIFT [25] descriptor are extracted in a dense manner at multiple patch sizes.

After feature extraction by each descriptor, the features underwent a pipeline called the Bag-of-Words pipeline. This pipeline worked as follows: (i) The extracted features from various patches were sampled randomly to learn a dictionary by using K-means proposed by Elkan [26], (ii) Locality-constrained linear coding (LLC) proposed by Wang et al. [27] was applied to assign the features to dictionary entries, (iii) Max pooling [26] with spatial pyramid [28] was applied to obtain the final feature vector.

The images collected in this work contain various luminance and were taken by different camera resolution levels in each burn depths classes. Instead of extracting the colour features and map them to a preselected set of colour names, Weijer et al.'s [20] method of learning the colour name from real-world images and Khan et al.'s [21] method of clustering the colour values based on the colour description's discriminative power were adopted in this work. In this work, the input images were first converted to colour names. Then, colour descriptor was used to extract features which were later used as input for classifiers. The colour descriptor was represented in the form of histogram of colour names. Other descriptors like Gist, HOG, LBP and SIFT were also used to extract local features. The features extracted from these descriptors are considered local features because they were extracted from multiple size of patches in an image.

D. Burn Depth Classification

The performance of different features sets under global and local features were compared using a machine learning workbench, the Waikato Environment for Knowledge Analysis (WEKA) [29]. Three classification algorithms were used on the skin burn dataset for this comparative study, which were SMO, JRIP, and J48 using the 10-fold cross validation method. The 10-fold cross validation method takes the average of the different test partitions in the dataset. Hence, the results would be void of bias and more consistent.

III. Results And Discussion

The overall performances for each classifier on the different set of features for all the three burn depths are shown in Table 3 and 4 for global and local features respectively, taking the average accuracies of the three classifiers using the multi-class classification approach.

On closer inspection, there were some DPT burn which were misclassified as FT burn in both global and local features due to the images being taken under dim environment for the DPT burn, causing the image to be dark and thus, mistaken as FT burn. In the misclassification of DPT burn as SPT burn, the DPT burn, which is usually cream or almost white in colour, was surrounded by bright red colour, which usually indicates SPT burn. The classifier eventually classify the burn as SPT burn due to the colour coverage of the burn. Gist had a poor classification result as this descriptor is normally used in scene recognition by finding edges, naturalness of surface and so on. Apparently in skin burn images, the Gist descriptor was unable to find these attributes, resulting in many misclassifications of the burn depths. SIFT descriptor is well known as a shape descriptor in finding the shape of an object in an image [25]. However, in burn images classification, colour and texture features play an important role instead of the shape feature [8-11]. Some burn images in the dataset might not have a distinct shape or structure as the burn covers a whole image.

A. Global Features

Based on Table 3, it can be seen that Feature Set ID 4, which is the selected features of colour and texture proposed by Acha et al. [8-10] and Serrano et al. [11] achieved the best performance with an average accuracy of 77.0% for global features. The average accuracies for Feature Set ID 1, 2 and 3 were quite close to each other, with the value of 75.3%, 76.7% and 76.4% respectively. The lowest average accuracy was resulted from the use of feature from Feature Set ID 7, which was 33.3%.

Table 3
Multi-class classification results for various sets of global features by using 10-fold cross validation

Set ID	Classification Accuracy SMO	JRIP	J48	Average
1	75.8%	75.8%	74.2%	75.3%
2	77.5%	75.8%	76.7%	76.7%
3	75.8%	79.2%	74.2%	76.4%
4	77.5%	76.7%	76.7%	77.0%
5	41.7%	37.5%	41.7%	40.3%
6	65.0%	61.7%	56.7%	61.1%
7	33.3%	33.3%	33.3%	33.3%

B. Local Features

Based on Table 4, the SIFT features descriptor achieved the best performance with an average accuracy of 74.7% for local features. The second highest accuracy was achieved by using colour, with an average accuracy of 74.2%. The lowest average accuracy was yielded by features extracted by GIST descriptor, with the value of 45.6%.

Table 4
Multi-class classification results for various sets of local features by using 10-fold cross validation

Descriptor	Classification Accuracy			
	SMO	JRIP	J48	Average
Colour	79.2%	71.7%	71.7%	74.2%
GIST	51.7%	41.7%	43.3%	45.6%
HOG3X3	54.2%	45.8%	52.5%	50.8%
LBP	63.3%	50.0%	62.5%	58.6%
SIFT	76.7%	71.7%	75.8%	74.7%

IV. Conclusion

A comparative study of the features extracted in the classification of human skin burn depth was conducted using an image mining approach. Different set of features from related works were experimented on our collection of skin burn images. Both global and local features were used in this comparative study. The performance of the extracted features were evaluated using SMO, JRIP and J48 classification algorithms. The best feature set in terms of global features were the mean of lightness, mean of hue, standard deviation of hue, standard deviation of A* component, standard deviation of B* component, and skewness of lightness proposed by Acha et al. [8-10] and Serrano et al. [11] with an average accuracy of 77.0% whereas the best descriptor in terms of local features for skin burn images was SIFT, with an average accuracy of 74.7%. The performance in classification of different skin burn depths is largely dependent on the features extracted in terms of whether the features extracted from a particular image is able to represent the burn class of that image. In future work, the combination global and local features will be studied for the skin burn depths images. Besides that, to solve the current issues in misclassifications, extracting local features for each specific burn depths classes by studying the characteristics of each burn depths will also be conducted in the future.

Acknowledgments

I would like to thank Universiti Malaysia Sarawak for providing the funding to present this research in the UNIMAS Silver Jubilee Conference.

References

[1] "Boundless Anatomy and Physiology in Structure of the Skin: Dermis," 2016. [Online]. Available: https://www.boundless.com/physiology/textbooks/boundless-anatomy-and-physiology-textbook/integumentary-system-5/the-skin-64/structure-of-the-skin-dermis-395-7489/.

[2] "Burn Classification," *UNM hospitals*. [Online]. Available: http://hospitals.unm.edu/burn/classification.shtml.

[3] A. L. Mescher, *Junqueira's Basic Histology*, 14th ed. McGraw-Hill Education, 2016.

[4] M. S. Badea, C. Vertan, C. Florea, L. Florea, and S. Badoiu, "Automatic burn area identification in colour images," *2016 International Conference on Communications (COMM)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 65–68, 2016.

[5] D. A. Lisin, M. A. Mattar, M. B. Blaschko, M. C. Benfield, and E. G. Learned-miller, "Combining Local and Global Image Features for Object Class Recognition," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005.

[6] "Local Feature Detection and Extraction," *MathWorks*. [Online]. Available: https://www.mathworks.com/help/vision/ug/local-feature-detection-and-extraction.html.

[7] K. Murphy, A. Torralba, D. Eaton, and W. Freeman, "Object detection and localization using local and global features," *Toward Category-Level Object Recognition - Lecture Notes in Computer Science*, vol. 4170, pp. 382–400, 2006.

[8] B. Acha, C. Serrano, and J. I. Acha, "Segmentation of burn images using the L*u*v* space and classification of their depths by colour and texture imformation," *Medical Imaging 2002: Image Processing*, pp. 1508–1515, 2002.

[9] B. Acha, C. Serrano, J. I. Acha, and L. M. Roa, "CAD tool for burn diagnosis," *Biennial International Conference on Information Processing in Medical Imaging*, pp. 294–305, 2003.

[10] B. Acha, C. Serrano, J. I. Acha, and L. M. Roa, "Segmentation and classification of burn images by colour and texture information," *Journal of Biomedical Optics*, vol. 10, no. 3, pp. 34014–3401411, 2005.

[11] C. Serrano, B. Acha, T. Gómez-Cía, J. I. Acha, and L. M. Roa, "A computer assisted diagnosis tool for the classification of burns by depth of injury," *Burns*, vol. 31, no. 3, pp. 275–281, 2005.

[12] R. Sudhir, "A survey on image mining techniques: Theory and applications," *Computer Engineering and Intelligent Systems*, vol. 2, no. 6, pp. 44–52, Oct. 2011.

[13] A. Khosla, T. Zhou, T. Malisiewicz, A. Efros, and A. Torralba, "Undoing the damage of dataset bias," Oct. 2012.

[14] A. Khosla, J. Xiao, A. Torralba, and A. Oliva, "Memorability of Image Regions," *Advances in Neural Information Processing Systems*, no. 1, pp. 296–304, 2012.

[15] B. N. Manu, "Brain MRI Tumor Detection and Classification," *MathWorks*. 2016.

[16] K. Wantanajittikul, S. Auephanwiriyakul, N. Theera-Umpon, and T. Koanantakool, "Automatic segmentation and degree identification in burn colour images," *The 4th 2011 Biomedical Engineering International Conference (BMEiCON)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 169–173, 2012.

[17] L. Deepak, J. Antony, and C. Niranjan U, "Hardware Co-Simulation of skin burn image analysis," *19th IEEE International Conference in High Performance Computing (HiPC-2012): Student Research Symposium. Pune, India.* 2012.

[18] M. Suvarna, S. Kumar, and N. U C, "Classification Methods of Skin Burn Images," *International Journal of Computer Science and Information Technology*, vol. 5, no. 1, pp. 109–118, 2013.

[19] M. Suvarna, K. Kumar, Sivakumar, and N. U. C, "Diagnosis of burn images using template matching, k-nearest neighbor and artificial neural network," *International Journal of Image Processing (IJIP)*, vol. 7, no. 2, 2013.

[20] J. van de Weijer, C. Schmid, J. Verbeek, and D. Larlus, "Learning colour names for real-world applications," *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1512–1523, 2009.

[21] R. Khan, J. Van De Weijer, F. S. Khan, D. Muselet, C. Ducottet, and C. Barat, "Discriminative colour descriptors," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2866–2873, 2013.

[22] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *International Journal of Computer Vision*, vol. 42, no. 3, pp. 145–175, 2001.

[23] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, vol. I, pp. 886–893, 2005.

[24] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns," 2002.

[25] D. G. LOWE, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[26] C. Elkan, "Using the Triangle Inequality to Accelerate k-Means," *Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)*, pp. 147–153, 2003.

[27] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained Linear Coding for Image Classification," 2010.

[28] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features : spatial pyramid matching for recognizing natural scene categories," in *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 2006.

[29] E. Frank, M. A. Hall, and I. H. Witten, *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*, Morgan Kaufmann, Fourth Edition, 2016.