Non-invasive Assessment and Classification of Human Skin Burns Using Images of Caucasian and African Patients

Aliyu Abubakar [a*], Hassan Ugail [a], Ali Maina Bukar [a] ,

[a]Centre for Visual Computing, Faculty of Engineering and Informatics, University of Bradford, Bradford BD7 1DP, UK

Abstract. Early medical intervention to address burn injuries in hospitals is an important step towards relieving the burden on patients and the patient's family at large. However, in most developing countries, the medical centres have major obstacles including but not limited to inadequate workforce, poor diagnostic facilities, and high maintenance/operational costs. Hence, the aforementioned issues have become a bottleneck to majority of people living in the third world. Towards this end, there is a need to develop an automatic machine learning algorithm to non-invasively identify skin burns; this will operate with little or no human intervention thereby acting as an affordable substitute to human expertise. Here, we leverage the weights of pre-trained deep neural networks for image description, subsequently, the extracted image features are fed into a Support Vector Machine (SVM) for classification. To the best of our knowledge, this is the first study that investigates black African skin data. Interestingly, the proposed algorithm achieves state of the art classification accuracy on both Caucasian and African datasets.

Keywords: Burns, Deep Neural Network, Image Descriptors, Support Vector Machine, Classification.

*****Aliyu Abubakar**,** E-mail: a.abubakar6@bradford.ac.uk

I. INTRODUCTION

Burn injuries are devastating to affected patients as well as the victim's family and consequently, to the nation in general [1]. Affected individuals face life challenges including but not limited to discrimination in the society, disfigurement due to scars, low self-esteem and other psychological problems. These injuries are reported by [2-5] to be the most cause of morbidity and mortality in the world with over 90% ruinous burns related injuries in low- and middle-income countries (LMIC) such as those in Africa and Asia. According to World Health Organisation (WHO) as reported in 2014 [6], almost 11 million people were affected by burn injuries severe enough for medical attention. South-east Asian countries such as Indonesia, India, Bangladesh, Maldives, Nepal and Myanmar have the highest burn incidences with mortality rate of at least 8.3 deaths per 100,000 affected individuals. Africa is the second with highest mortality rate of approximately 5.5 deaths per 100,000 people. On the other hand, the report showed that high-income countries (HIC) such as Americas have an average death rate of 1.3 per 100,000 people. Burns are caused by fire flames, chemicals such as concentrated acid which mostly occur in work places (industries) as a result of accidental spillage, ultraviolet sun radiation and friction. Reports have shown that about 85% of the total global population is in low-and middle-income countries where high-income countries constitute only 15% of the global population [7]. Unfortunately, about three-quarter (75%) of the population residing in LMIC die before the age of 70 compared to those in

HIC as a result of infections and non-communicable diseases. Furthermore, these calamities were associated with inadequate healthcare systems, and lack of experienced health care workforce [7]. According to a publication of the Global Burden of Disease in 1996 [8], it was projected that by the year 2020, non-communicable diseases are expected to rise and account to about seven out of ten deaths in poor countries. Similarly, Lack of specialized hospital for burn victims and lack of epidemiological data by the concerned agency has resulted to unavailability of policy that could help to strategize preventive measures in LMICs such as Ghana [9]. Moreover, in a study by [10], it was observed that up to 4 million women suffer burns injuries worldwide almost same number as those found with HIV/AIDS but not much attention is given to burn injuries. The authors further reported that burn injuries in pregnant women is not well documented. They further lamented that the consequences of the injury are felt by both the mother and the unborn child. Complications such as precipitated labour due to septicaemia, and catabolic state is associated to be the consequence experienced by the pregnant women with burn injuries.

Researchers at Stanford University [11] found that over 33.5 million thermal burn injuries were recorded in 2013 mostly in low-and middle-income countries (about 90% of the global incidences), which substantially contributed to high mortality rate [12]. This was mostly influenced by socioeconomic condition in addition to culture and life style. According to authors in [3], socioeconomic condition is the biggest factor as such this incidence is highly reported in low- and middle-income countries while lesser occurrences in high-income countries (HIC). However, views have been expressed that preventive measures in order to minimise the morbidity and mortality rates should be a primary concern rather than focusing on burn care improvement measures. This conform to the expression made by Keswani [13] in 1986 that "100 per cent successful care of the burn injury does not provide the fully needed outcome but the successful preventive measure that will ensure none occurrence of the incidences". Nevertheless, preventive measures are found to be much more effective in HIC [6, 14] than in LMIC due to the socioeconomic and life style differences.

While this calamity is tearing people apart, specialised burns centres are lacking in most crowded communities living in extreme poverty, access to medical centres in many parts of the globe is hard, available burn centres in some parts of the world are overcrowded with no good facilities and such condition results to diagnosis delay and ineffective assessment. In 2015, Boissin and other scholars from South-Africa [15] also lamented that most of the improvements in burn prevention and care is highly recorded in HIC, with low recorded improvement in LMIC. They similarly further associated the imbalance improvement of care received by patients in poor countries with lack of medical access and in extreme cases due to poverty that hinders majority from utilising or affording good medical treatment. However, this disproportion can be improved using alternative approaches which can provide timely and low-cost effective service delivery for burn patients. We believe our work is novel because of the following contributions: we have collected the largest dataset so far as compared to the literature (to the best of our knowledge); we intend to make it publicly available; using the weights of deep neural

networks we have attained results that are extremely good which we hope will form the baseline for future research; our rigorous evaluation protocol proves the efficiency of the proposed technique
The rest of the paper is organized as follows: section two presents an excerpt from the related literature; section three provides an overview of deep neural networks; method such as image acquisition, feature extraction and classification techniques are presented in section four; section five presents the experiment and results; and finally the research is concluded in section six.

II. RELATED LITERATURES

Attempts to automate the process of burn identification have been proposed in few studies such as [16], where five burn wounds were obtained from the department of medical science in the ministry of public health in Thailand. These acquired images were artificially augmented by experienced surgeons who identified second degree and third degree regions. Afterwards SVM was trained using 34 sub-images, through a 4-fold cross-validation, they achieved classification accuracy of 89.29%.

Automatic classification of burns based on observable features by experienced burn surgeons was also proposed in [17]. Here, a total of 20 images were analysed, where the surgeons rated images based on the images similarities. Specifically, grouping of the images were based on severity of the burn wounds; such as reddish appearance for the superficial burns and pink-whitish colour for the deep dermal burns. The detected features were subsequently classified using support vector machine; they reported an accuracy of 80% with specificity as well as sensitivity of 60% and 97% respectively. Obviously the number of images used is too few to draw conclusion. Moreover, none of the aforementioned authors has made his data public.

In [18], 611 images obtained from 53 pediatric patients were augmented and used to train a convolutional neural network (CNN) to distinguish skin burns from healthy skins. The algorithm's reported sensitivity and specificity were 82.96% and 75.91% respectively. Unfortunately, the dataset used here is too small to produce the best from CNN. Rather than training from scratch, transfer learning will have been a better alternative. The dataset is also inaccessible.

It is worth noting that none of the studies considered classifying burns in African patients. To the best of our knowledge, researchers only used images of Caucasian patients, including the most recent study [19] who proposed segmentation of burn wounds. Furthermore, the data used by the aforementioned researchers is not publicly available, thus the algorithm proposed in this study cannot be compared to any of the reviewed literature.

As such, this study aims to investigate the use of deep neural networks for automatic recognition of burns to infer fast, accurate, reliable and cost-effective diagnostic process. Here we do not intend to train the model from scratch. Instead, the weights of pre-trained models shall be used for image feature extraction. Additionally, the research investigates the efficacy of the algorithm when applied on images of both Caucasian and African patients. We believe that the proposed pipeline could aid tremendously when deployed to complement the diagnostic reliability of medical experts thereby enhancing service delivery in remote locations for different ethnicities. As an additional benefit to the research community, and to ensure ease of algorithm comparison, the dataset used in this study shall be made publicly available to researchers.

III. CONVOLUTIONAL NEURAL NETWORKS

Since the advent of machine learning (ML), many aspects of human activities such as commerce [20], healthcare [21] and police detective operations [22] have been revolutionized. As a branch of ML, deep learning uses stacks of multiple hidden neural network layers arranged in a hierarchical passion to learn deep representations of data and to subsequently classify them in accordance to the categories for which they were trained to identify. Different architectures of convolutional neural network have been utilised recently in diverse fields such as face recognition [23], speech emotion recognition [24], and medical image analysis [25-27]. The most popular pre-trained CNN models are those used for the ImageNet competition [28], including but not limited to AlexNet [29], VGG-16 and VGG-19 [30]. Here, we consider three different VGG models due to their simplicity and reported accuracy [30].

A. VGG-16

This is a deep neural network developed by Visual Geometry Group (VGG) at the University of Oxford. VGG-16 has a total of 37 layers out of which there are 13 convolution layers and 3 FC layers and all the convolution layers are equipped with 3 x 3 size filters as depicted in figure 1. The remaining layers are activation and pooling layers and the last layer is the decision layer. This deep neural network was trained on ImageNet database in 2014 [30]. In 2015, the VGG-16 model was retrained on a dataset of 2.6 million human faces [31, 32], this they named the VGG-Face; it has been reported to have achieved state of the art accuracy in face recognition.

Figure 1. Architectural illustration of VGG-16 model

B. VGG-19

VGG-19 architecture is a deeper updated version of the VGG-16, the depth of the network was increased reaching up to 43 layers comprising of 16 convolution layers, 3 FC layers, and the rest are interweaved activation and pooling layers.

IV. METHOD

Our method in this paper is simple, the weights of VGG-Face as well as both VGG-16 and VGG-19 deep neural networks are used to extract discriminatory features from burn images, subsequently, three sets of features are used to train support vector machine classification algorithms. The rationale for the proposed pipeline is to achieve efficient burn identification and also to assess whether the training datasets initially used to train the (compared) models has an impact on burn identification accuracy. We hope this comparison will further give us more insight on to the effect of neural network depth. Since, VGG-16 and VGG-19 were trained on database of 1000 different categories of objects while VGG-face model was trained on a database of human faces (VGG-Face). Intuitively, VGG-Face has closer relationship to our skin data, but will this provide significantly stronger features?

Additionally, we believe this approach of using a pre-trained model is simple and well suited for our task, especially due to the fact that our dataset has insufficient images; any attempt to retrain the whole neural network or a few of its layers will most likely lead to over-fitting.

A. *Data Collection and Ethical Consideration*

- Caucasian patients: The datasets were captured with the patient's full consent in a hospital from the city of Bradford, United Kingdom; sample images contained in this dataset are presented in figure 2.

Figure 2. Samples of healthy and burnt skin from Caucasian patients

- African patients: Burn images from patients with dark (black) skin were obtained from Federal Teaching Hospital Gombe (FTHG) in the North-Eastern Nigeria. The hospital is situated in a city of over 3 million people and it is the only hospital with a burn unit. The dataset was released upon the successful approval of the submitted application/request by the teaching hospital's research and ethics committee. Figure 3 shows samples of the African burn images. The quality of images contained in this dataset is very poor, having low resolution as well as poor contrast.

Figure 3. Samples of healthy and burnt skin from African patients

The skin in Caucasian and African people differ in appearance due to a colouring pigment called melanin. This pigment is present in all human races however, the secretion of these pigments varies in accordance to geographic location. Melanin pigment is produced by melanocyte which is contained in both white and black skin [33]. Melanin in black skin is brown-black pigment called eumelanin and yellow-reddish pigment in white skin called Pheomelanin. The secretion of these pigments depends on the ultraviolent (UV) sun radiation in a region. People living in region with high UV light produce more melanin to shield them against the negative effect of too much exposure to sun heat. Thus, people with high rate of melanin secretion possess black skin, for instance people living in sub-Saharan Africa. Another distinguishing element between white and black people found in the epidermal layer of skin is the outermost protective barrier that sits on top of the epidermis. This layer is called stratum Corneum; it aids water retention as well as hydration [34] and has 15 or more flattened corneocytes. According to [16] there are 20 corneocytes in black skin compared to 16 corneocytes in white skin, however the thickness of the stratum Corneum in both races is the same but they are more compacted in black skin.

B. Data Preparation

The collected datasets contain burn images of different complexity levels (i.e., different burn degrees and sizes), also the patients are of varying age group; adults and infants. Furthermore, the affected burn locations are from different regions of the body (head, torso to limbs). We harvested 1360 rectangular patches from 32 Caucasian subjects. Extraction of patches from the images was done carefully to include varying degrees of burn information. Similarly, 700 regions were cropped out from the images of 60 black patients. All cropped regions were resized to 224 by 224 pixels to conform to the input of the VGG models.

C. Feature Extraction

As stated earlier, three sets of features were extracted from colour images using the pre-trained models (VGG-16, VGG-19 and VGG-Face). Assuming input images are given as I_0 and is represented as $I_0 \varepsilon V^{H \times W \times C}$, the parameters H,W and C stand for height of the images, width of the image and colour channel respectively. Each input image passes through different layers of the pre-trained model which is composed of series of learning functions $F_L = f^1 \gg f^2 \gg f^3 \ldots f^n$, where output of each layer is the input of the immediate layer. Specifically, layer FC7 of the three models are utilized. Hence each image feature is given by a vector of 4096 elements.

D. Classification

Following successful feature extraction, a linear support vector machine (SVM) classifier is trained for each set of features. Specifically, for each feature set, three SVMs were trained on; Caucasian images, African images, and a combination of the two. This we have done to get an insight on how skin colour affects the overall performance of the classifier. Hence, in total 9 SVMs were trained.

Skin-burn discrimination is a binary classification problem, where the test data are either healthy skin or burnt skin. Given a training set (x_i, y_i) for i=1,…,n with x ε R^N and y_i ε {-1 , +1} a classifier is learned such that

$$f(x_i) = \begin{cases} \geq 0 \; y_i = +1 \\ < 0 \; y_i = -1 \end{cases} \qquad (1)$$

Here, +1 and -1 denotes skin burn and healthy skin. The goal of the SVM is to find the most favourable separating hyperplane that best divides the two classes. It functions by projecting the training samples via a function Ø into an infinite dimensional space F. Then the optimal separating hyperplane is obtained in F by solving an optimization problem. However, the mapping from input space X to the feature space

F is not done explicitly; rather it is done via the kernel trick, which computes the inner dot products of the training data. Reader is referred to [35] for more detail on SVM. In this paper linear kernel SVM is utilised, given by:

$$K(x_{i,}x_j) = x_i^T x_j \quad (2)$$

V. EXPERIMENT AND RESULTS

Here, results of experiments performed using the proposed pipeline are presented. Unfortunately, we cannot make comparison to other works highlighted in section II. This is because the dataset they all used are private and every attempt to get access to them have been futile. We however hope to make our dataset publicly available, thus the results presented in this section can act as a baseline for future research.

All nine SVMs are evaluated through a 10 fold cross-validation technique. Hence, in each of the 9 scenarios, the dataset say $\boldsymbol{S}$ is randomly partitioned into 10 mutually exclusive parts/folds $\boldsymbol{S}_1, \boldsymbol{S}_2, \ldots, \boldsymbol{S}_{10}$ of same size. The SVM was trained 10-times where in each iteration 1-fold is held-out and the classifier is trained using 9 folds and tested on the held out split. Thus, after 10 iterations all data samples will have been used for testing and training at some point in time. Classification accuracies for all 9 scenarios is presented in table 1. Figure 4 depicts the experiment procedure.

Figure 4. Experimental setup

Table 1. Classification accuracy

CNN Models	Caucasian	African	Hybrid
VGG-16	99.286%	98.869%	98.750%
VGG-19	98.333%	97.500%	97.560%
VGG-Face	96.310%	97.202%	95.208%

As can be seen in table 1, all 9 scenarios achieved excellent classification accuracies. It is also obvious that skin burn classification is easier in Caucasian images, this can be attributed to the fact that the burns have a clear contrast as compared to the normal skin colour. Moreover, the Caucasian dataset is bigger and has better image quality as compared to its African counterpart. It can also be observed that combining both white and dark skin images into a single global dataset further introduces confusion, thereby deflecting the accuracy even further. Table 1 also highlights that out of the three VGG models, the two (VGG-16 and 19) that were trained on ImageNet outperform VGG-Face which was trained on faces. This highlights to us the importance of training data and the diversity of data categories. The features of VGG-Face seem to have been fine-tuned so much toward face images, hence the slight under performance when used to describe non-face images. Additionally, we have seen that VGG-16 outperforms VGG-19, here is a scenario where increase in depth only made the performance worst, it can be argued that this is due to the fact that as depth increases in CNNs, accuracy saturates and subsequently degrades; this has been mentioned in the literature [36], as one of the reasons why residual neural networks were introduced. Figure 5, graphically summarizes the result of Table 1. Lastly, it can be noted that the result in VGG-Face shows that the model represents dark skins better than the lighter skins, despite that fact, the overall accuracy is still below the other two models.

Figure 5. Classification accuracy

The accuracy of a classification algorithm tells the overall correctness of the trained model on the dataset in question. However, there is need to know which of the instances or class samples were less/highly misclassified. This information can be obtained if proportion of the positive samples were determined, likewise for the negative samples. Since the study deals with binary classification, let consider a situation represented by a 2 x 2 dimensional table known as confusion matrix or contingency table as shown in table 2

Table 2. Confusion matrix

Predicted Classes	**Actual Classes**			
		Negative	**Positive**	
	True (T)	Healthy Skin: True Negative (TN)	Misclassified TP: False Negative (FN)	False(F)
	False (F)	Misclassified TN: False Positive (FP)	Abnormal Skin: True Positive (TP)	True (T)

The contingency table shown in Table 2 shows there can be only 4 possible outcomes when the test is conducted. The abnormal class is represented by the cell labelled TP which indicates percentage of correctly classified positive (burned skins) instances and the normal or healthy instances that were correctly classified are represented by TN which indicates the actual normal negative. For simplicity, instances that falls within the shaded cells are the correct prediction made by the classification model. However, those that fall into the unshaded cells are the misclassified instances; FN represents a percentage of burned-skins that were falsely classified as healthy, on the other hand, FP shows percentage of normal skins that were flagged as burnt.

The confusion matrix of our 9 models, are represented in Tables 3, 4 and 5. For each of the tables the a) b) and c) sub-tables represent VGG-16, VGG-19 and VGG-Face respectively. Figures 6, 7 and 8 further summarize as well as compare the specificities and sensitivities of the three features under the 3 scenarios.

Table 3a. Classification of burns in Caucasian using VGG-16 features

Predicted Classes	Target Classes		
		Burns	**Healthy**
	Burns	837	9
	Healthy	3	831

Table 3b. Classification of burns in black/African using VGG-16 features

Predicted Classes	Target Classes		
		Burns	**Healthy**
	Burns	834	13
	Healthy	6	827

Table 3c. Classification of burns in both Caucasian & African using VGG-16 features

Predicted Classes	Target Classes		
		Burns	**Healthy**
	Burns	1664	26
	Healthy	16	1654

Table 4a. Classification of burns in Caucasian using VGG-19 features

Predicted Classes	Target Classes		
		Burns	**Healthy**
	Burns	830	18
	Healthy	10	822

Table 4b. Classification of burns in African using VGG-19 features

Predicted Classes	Target Classes		
		Burns	**Healthy**
	Burns	820	28
	Healthy	20	813

Table 4c. Classification of burns in both Caucasian & African using VGG-19 features

Predicted Classes	Target Classes		
		Burns	**Healthy**
	Burns	1646	48
	Healthy	34	1632

Table 5a. Classification of burns in Caucasian using VGG-Face features

		Target Classes	
Predicted Classes		**Burns**	**Healthy**
	Burns	818	40
	Healthy	22	800

Table 5b. Classification of burns in African using VGG-19 features

		Target Classes	
Predicted Classes		**Burns**	**Healthy**
	Burns	820	27
	Healthy	20	813

Table 5c. Classification of burns in both Caucasian & African using VGG-19 features

		Target Classes	
Predicted Classes		**Burns**	**Healthy**
	Burns	1606	87
	Healthy	74	1593

Figure 6. Sensitivity and specificity from the classification of burns in Caucasian patients

Figure 7. Sensitivity and specificity from the classification of burns in African patients

Figure 8. Sensitivity and specificity from the classification of burns in both Caucasian and African patients

The results of the confusion matrix presented above, further corroborate the output of the accuracies presented in Table 1. In all scenarios VGG-16 weights have proven to have the strongest burn features. The confusion (FN and FP) are consistently minimal when training is done on isolated ethnicities. The global classifier seems to have more confusion, this suggests that it will be better to encode some ethnicity information into the classifier in the feature, that way it is hoped that the combined classifier having prior information about the ethnicity will tend to perform better.

Figure 9 shows some misclassified images, this failure of the algorithm can be attributed to waxy, white and leathery appearances of full-thickness burns. Poor resolution and low contrast are also additional causes of misclassification

Figure 9. Images (a) and (b) are full-thickness burns, (d) and (e) are healthy skin images from Caucasian patient while (c) and (f) are burn images from African patients

Finally, the results of our 9 experiments are represented using receiver operating characteristics (ROC). This is a tool that gives a graphical representation of sensitivity of the diagnostic test against the corresponding specificity. In summary, the computed area under the (AUC) curve gives a summary of which test/algorithm has lesser confusion and hence better overall performance. It is worth nothing that the values of AUC range between 0 and 1, with 1 been the ideal state.

Figures 10, 11 and 12 show ROC curves generated for all three scenarios. Again, the performance of our pipeline is close to the ideal case (i.e. AUC = 1). Furthermore, VGG-16 features outperform the other two, and again the global classification shows a slight drop in performance. Figure 13 summarizes the AUC and compare the performances under three scenarios.

Figure 10. ROC curve from the classification of burns in Caucasian patients

Figure 114. ROC curve from the classification of burns in African patients

Figure 12. ROC curve from the classification of burns in both Caucasian and African patients

Figure 14. AUC comparison

VI. Conclusion

In this paper, we presented a straightforward but very effective means of recognizing human skin burns. To the best of our knowledge, this is the first work that used machine learning to discriminate burn skin injuries from healthy skins for patients of different ethnicities.

Burn images obtained from Caucasian patients as well as those from black-African patients living in two different geographic regions were used to train various machine learning algorithms to discriminate healthy and burnt human skins. Weights of pre-trained deep neural networks were utilized for image feature extraction and subsequently linear kernel SVM was used for classification.

Besides achieving extremely good classification accuracies under nine scenarios, we were able to observe that deep learning models trained on multiple data categories have strong generic information that can be used on the fly for image feature representation.

Our work has also thoroughly investigated the effect of race/ethnicity on the overall performance of the skin-burns identification algorithm. We have thus observed that training local models for each ethnic group (or race) tends to be more robust than a single global model for all skin colors. As such, in the future, we shall investigate the development of a single global model with a priori ethnicity information encoded into the model.

We have further observed that the performance of CNNs saturates and subsequently degrade as they go deeper. Hence, in the future we will also like to investigate the weights of extremely deep residual neural networks since they were invented with a view to tackling problems associated with the conventional CNNs. Additionally, a multi-class problem of categorizing different degrees of skin burn shall be studied.

REFERENCES

[1] J. Bayuo, P. Agbenorku, and R. Amankwa, "Study on acute burn injury survivors and the associated issues," *Journal of Acute Disease*, vol. 5, no. 3, pp. 206-209, 2016.

[2] N.-T. Dai, H.-I. Chang, Y.-W. Wang, K.-Y. Fu, T.-C. Huang, N.-C. Huang, J.-K. Li, P.-S. Hsieh, L.-G. Dai, and C.-K. Hsu, "Restoration of skin pigmentation after deep partial or full-thickness burn injury," *Advanced drug delivery reviews*, vol. 123, no. pp. 155-164, 2018.

[3] W. Cheng, S. Wang, C. Shen, D. Zhao, D. Li, and Y. Shang, "Epidemiology of hospitalized burn patients in China: A systematic review," *Burns Open*, vol. 2, no. 1, pp. 8-16, 2018.

[4] M. M. Rybarczyk, J. M. Schafer, C. M. Elm, S. Sarvepalli, P. A. Vaswani, K. S. Balhara, L. C. Carlson, and G. A. Jacquet, "Prevention of burn injuries in low-and middle-income countries: a systematic review," *Burns*, vol. 42, no. 6, pp. 1183-1192, 2016.

[5] A. Abubakar and H. Ugail, "Discrimination of Human Skin Burns Using Machine Learning," in *Intelligent Computing-Proceedings of the Computing Conference*, 2019, pp. 641-647.

[6] J. P. Heard, B. A. Latenser, and J. Liao, "Burn prevention in Zambia: a targeted epidemiological approach," *Journal of Burn Care & Research*, vol. 34, no. 1, pp. 65-73, 2013.

[7] K. Hofman, A. Primack, G. Keusch, and S. Hrynkow, "Addressing the growing burden of trauma and injury in low-and middle-income countries," *American journal of public health*, vol. 95, no. 1, pp. 13-17, 2005.

[8] C. J. Murray, A. D. Lopez, and W. H. Organization, "The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020: summary," no., 1996.

[9] J. Bayuo, F. B. Agyei, and P. K. Baffour, "Burns in the middle belt of Ghana: A systematic review," *Burns Open*, vol. 2, no. 2, pp. 85-89, 2018.

[10] C. J. Ogbogu, A. Uduezue, W. I. Anetekhai, and C. C. Agunwa, "Burn injuries in pregnancy in a regional burns center in Nigeria: Presentation, maternal and fetal outcome," *Burns Open*, vol. 2, no. 1, pp. 53-58, 2018.

[11] J. A. Newberry, C. B. Bills, E. A. Pirrotta, M. Barry, G. V. R. Rao, S. V. Mahadevan, and M. C. Strehlow, "Timely access to care for patients with critical burns in India: a prehospital prospective observational study," *Emerg Med J*, vol. 36, no. 3, pp. 176-182, 2019.

[12] D. R. Davé, N. Nagarjan, J. K. Canner, A. L. Kushner, B. T. Stewart, and S. R. Group, "Rethinking burns for low & middle-income countries: differing patterns of burn epidemiology, care seeking behavior, and outcomes across four countries," *Burns*, vol. 44, no. 5, pp. 1228-1234, 2018.

[13] M. Keswani, "The prevention of burning injury," *Burns*, vol. 12, no. 8, pp. 533-539, 1986.

[14] P. Hodgins, T. Potokar, and P. Price, "Comparing rich and poor: burn prevention in Wales, Pakistan, India, Botswana and Zambia," *Burns*, vol. 37, no. 8, pp. 1354-1359, 2011.

[15] C. Boissin, L. Laflamme, L. Wallis, J. Fleming, and M. Hasselberg, "Photograph-based diagnosis of burns in patients with dark-skin types: the importance of case and assessor characteristics," *Burns*, vol. 41, no. 6, pp. 1253-1260, 2015.

[16] K. Wantanajittikul, S. Auephanwiriyakul, N. Theera-Umpon, and T. Koanantakool, "Automatic segmentation and degree identification in burn color images," in *The 4th 2011 Biomedical Engineering International Conference*, 2012, pp. 169-173.

[17] C. Serrano, R. Boloix-Tortosa, T. Gómez-Cía, and B. Acha, "Features identification for automatic burn classification," *Burns*, vol. 41, no. 8, pp. 1883-1890, 2015.

[18] M.-S. Badea, C. Vertan, C. Florea, L. Florea, and S. Bădoiu, "Automatic burn area identification in color images," in *Communications (COMM), 2016 International Conference on*, 2016, pp. 65-68.

[19] M. D. Cirillo, R. Mirdell, F. Sjöberg, and T. D. Pham, "Tensor Decomposition for Colour Image Segmentation of Burn Wounds," *Scientific reports*, vol. 9, no. 1, p. 3291, 2019.

[20] G. Liu, T. T. Nguyen, G. Zhao, W. Zha, J. Yang, J. Cao, M. Wu, P. Zhao, and W. Chen, "Repeat buyer prediction for e-commerce," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 155-164.

[21] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease prediction by machine learning over big data from healthcare communities," *Ieee Access*, vol. 5, no. pp. 8869-8879, 2017.

[22] P. J. Phillips, A. N. Yates, Y. Hu, C. A. Hahn, E. Noyes, K. Jackson, J. G. Cavazos, G. Jeckeln, R. Ranjan, and S. Sankaranarayanan, "Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms," *Proceedings of the National Academy of Sciences*, vol. 115, no. 24, pp. 6171-6176, 2018.

[23] P. Tambi, S. Jain, and D. K. Mishra, "Person-Dependent Face Recognition Using Histogram of Oriented Gradients (HOG) and Convolution Neural Network (CNN)," in *International Conference on Advanced Computing Networking and Informatics*, 2019, pp. 35-40.

[24] H. M. Fayek, M. Lech, and L. Cavedon, "Evaluating deep learning architectures for Speech Emotion Recognition," *Neural Networks*, vol. 92, no. pp. 60-68, 2017.

[25] T.-C. Chiang, Y.-S. Huang, R.-T. Chen, C.-S. Huang, and R.-F. Chang, "Tumor detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation," *IEEE transactions on medical imaging*, vol. 38, no. 1, pp. 240-249, 2019.

[26] X. Li, S. Zhang, Q. Zhang, X. Wei, Y. Pan, J. Zhao, X. Xin, C. Qin, X. Wang, and J. Li, "Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images: a retrospective, multicohort, diagnostic study," *The Lancet Oncology*, vol. 20, no. 2, pp. 193-201, 2019.

[27] K.-H. Yu, F. Wang, G. J. Berry, C. Re, R. B. Altman, M. Snyder, and I. S. Kohane, "Classifying Non-Small Cell Lung Cancer Histopathology Types and Transcriptomic Subtypes using Convolutional Neural Networks," *bioRxiv*, no. p. 530360, 2019.

[28] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, and M. Bernstein, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211-252, 2015.

[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.

[30] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, no., 2014.

[31] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *BMVC*, 2015, vol. 1, p. 6.

[32] A. Elmahmudi and H. Ugail, "Deep face recognition using imperfect facial data," *Future Generation Computer Systems*, no., 2019.

[33] Z. Yin, Y. Ge, H. Ning, Y. Zhu, L. Chen, S. Zhang, X. Xia, X. Wang, L. Wang, and Q. Pang, "Expression and tissue distribution analysis of Angiotensin II in sheep (Ovis aries) skins associated with white and black coat colors," *Acta histochemica*, no., 2019.

[34] M. B. Murphrey and P. M. Zito, "Histology, Stratum Corneum," in *StatPearls [Internet]*, vol.: StatPearls Publishing, 2018.

[35] V. Jakkula, "Tutorial on support vector machine (svm)," *School of EECS, Washington State University*, vol. 37, no., 2006.

[36] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

Aliyu Abubakar has received his first class BSc Honors degree in Computer Science from Gombe State University, Nigeria, in 2012 and MSc degree in Cyber Security from Faculty of Engineering and Informatics, University of Bradford, UK in 2015. He is currently working towards his PhD at the Centre for Visual Computing, University of Bradford, UK. His research interest includes machine learning, computer vision and image processing.

Hassan Ugail is the director of the Centre for Visual Computing at the Faculty of Engineering and Informatics at University of Bradford. He has a first class BSc Honours degree in Mathematics from King's College London and a PhD in the field of Geometric Design from the School of Mathematics at University of Leeds. Professor Ugail's research interests include computer based geometric and functional design, image analysis and machine learning.

Ali Maina Bukar received a B.Eng. Electrical and Electronics Engineering degree with Second Class Upper Honours in 2006, and a Distinction in MSc Computing in 2010 from Robert Gordon University, Aberdeen, UK. Furthermore, he obtained his PhD in Computer Vision and Machine Learning from the University of Bradford, UK, in 2017. His research interests include Computer Vision, Pattern recognition and Machine Learning.