
BURNNET: AN EFFICIENT DEEP LEARNING FRAMEWORK FOR ACCURATE DERMAL BURN CLASSIFICATION

Rohan Bhansali
 Loudoun Academy of Science
 Leesburg, VA 20175
rohanbhansali@connectai.org

Rahul Kumar*
 Loudoun Academy of Science
 Leesburg, VA 20175
rahulkumar@icloud.com

January 30, 2021

ABSTRACT

Burns are the fourth most prevalent unintentional injury around the world, and when left untreated can become permanent and sometimes fatal. An important aspect of treating burn injuries is accurate and efficient diagnosis. Classifying the three primary types of burns – superficial dermal, deep dermal, and full thickness – is essential in determining the necessity of surgery, which is often critical to the afflicted patient’s survival. Unfortunately, reconstructive burn surgeons and dermatologists are merely able to diagnose these types of burns with approximately 50-75% accuracy. As a result, we propose the use of an eight-layer convolutional neural network, BurnNet, for rapid and precise burn classification with 99.87% accuracy. We applied affine transformations to artificially augment our dataset and found that our model attained near perfect metrics across the board, demonstrating the high propensity of deep learning architectures in burn classification.

1 Introduction

Burns are a form of human injury caused by the application of high amounts of heat to the skin. Burns are extremely painful as they result in the rapid death of skin cells that can result in permanent damage through the loss of sensation or functionality. They are a prevalent public health issue, ranking as the fourth source of unintentional injury [3].

1.1 Dermal Burn Diagnosis

When burns are treated, a determination must be made as to the type and location of the burn for an adequate treatment to occur. The three primary types of burns are superficial dermal burns, deep dermal burns, and full-thickness burns [4]. Superficial burns primarily harm the outer layer of the skin, otherwise known as the epidermis. Typically, they can be described as appearing to be red in color and relatively dry. Blisters may also potentially be present. Deep dermal burns primarily harm the dermis and epidermis layers of the skin and are similar in appearance to superficial burns. Full-thickness burns destroy the dermis, epidermis, and subcutaneous tissue, and generally cause permanent damage. They appear distinctly white and charred. Properly distinguishing between these burns is extremely important in providing pertinent and rapid treatment to burn victims. Unfortunately, in many parts of the world access to specialists is limited and burns are often incorrectly classified, which leads to inadequate treatment measures. In some cases, specialists can take days to be available for consultation which is especially detrimental when surgery is required for treatment. Even when specialists are present, their accuracy can range anywhere from 50% to 76% based on their experience [7].

*Both authors contributed equally.

1.2 Previous Deep Learning Research

Convolutional neural networks (CNNs) have been used in the past for burn classification. In one study, CNNs were used to produce feature maps that were fed to a support vector machine (SVM) for classification and an accuracy of 82.43% was achieved [8]. Others have also experimented with CNN architectures for burn degree determination. One research group utilized deep CNNs with local binary operators in order to classify burns as first, second, third, or fourth degree. Notably, it was recorded that the primary features used in burn classifications (by CNNs and surgeons) were color and texture of the burned area [4].

1.3 Automated Classification Model

In order to alleviate the problem, we propose BurnNet, a deep convolutional neural network designed to rapidly and accurately classify skin burns. Our research looks to explore the ability of affine transformations to improve the functionality of a deep CNN in burn type classification presented with a small dataset. Specifically, we applied affine transformations in order to artificially magnify our dataset and achieve a training set that is considered adequate in size for deep CNNs. In previous medical imaging work, affine transformations have been shown to have significantly improve results across all metrics, specifically when small datasets were utilized [1]. Additionally, we look to provide clinical viability for burn classification through CNNs by achieving validation accuracy and F1-score above .85, which would be significantly more accurate than most experienced burn surgeons [5].

2 Methodology

The dataset that was used in this study was the BIP_US database, containing 94 images of full-thickness, deep dermal, and superficial dermal burns. It was produced by the Signal Theory and Communications Department [2]. The burned area has been cropped and refitted to a gray background.



Figure 1: Sample full-thickness, deep dermal, and superficial dermal burn images from BIP_US database

2.1 Data Processing

First, all of the images were resized to 512x512 through the anti-aliasing technique. It is among the most sophisticated and robust techniques used in resolution manipulation. It is primarily advantageous because it removes high-frequency noise artifacts prior to resolution manipulation and thus minimizes distortion and black and white noise. Second, we applied affine transformations to all images in the dataset. We applied 72 transformations per image, going in increments of 5 degrees. Examples of the transformed images can be seen below. Next, we split the dataset into a training and testing subset; 75% of images were used for training, with the remainder being used for testing.

2.2 BurnNet

The burn classification task is a tertiary classification problem, where the input is an image of a burn and the output is a tertiary label indicating a superficial burn, deep dermal burn, or full-thickness burn, respectively. To accomplish this task, an eight-layer CNN was utilized that we dubbed BurnNet. This model contained an input layer, two convolution layers, two pooling layers, a flattening layer, and two fully connected dense layers. The dense layers alleviate the vanishing gradient problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters. Within our model, we utilized Adaptive Moment Estimation (Adam), an adaptive learning rate optimization algorithm, with an initial learning rate of 0.00001. Our activation functions for the intermediate layers was the rectifier function, with the final layer using the softmax function. A batch size of ten was used over twenty epochs,

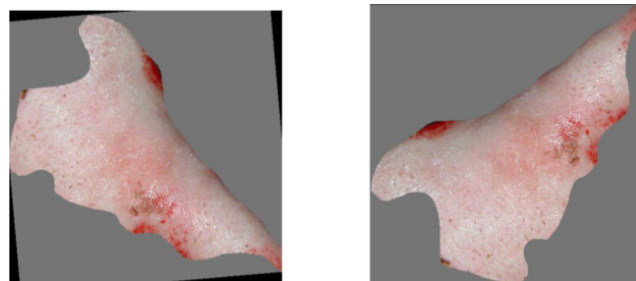


Figure 2: Burn images with affine transformations performed

after each of which the model outputted accuracy, performance metrics (precision, recall, and F1 score), and loss, given by the binary cross-entropy loss function:

$$H_p(q) = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \quad (1)$$

In order to prevent the networks from overfitting, early stopping was performed by saving the network after every epoch and choosing the saved network with the lowest loss on the tuning set.

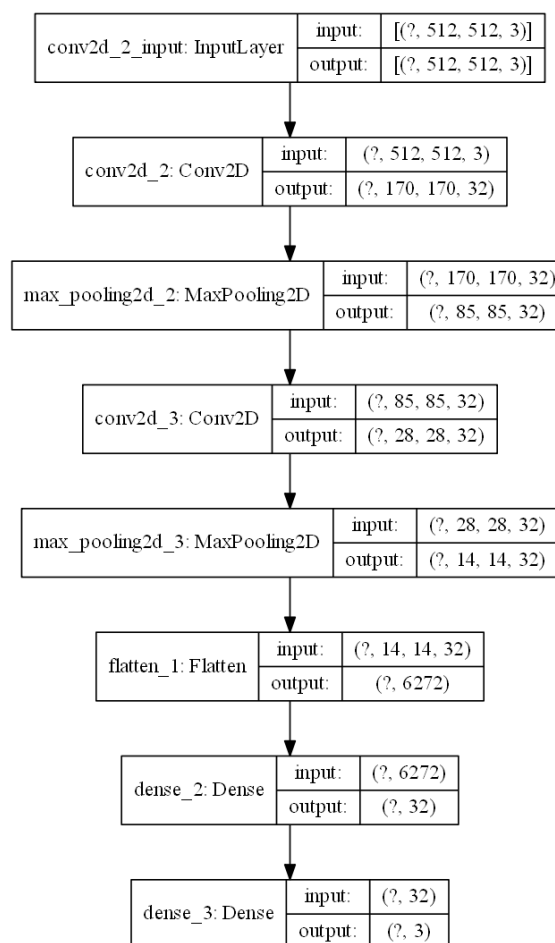


Figure 3: BurnNet Architecture

3 Results

Our model attained near-perfect performance metrics across the board, reaching an accuracy, F1, sensitivity, and specificity of 0.9987. We performed early stopping after fifteen epochs to prevent overfitting due to the training loss becoming stagnant.

Table 1: Performance Metrics

Epoch	5	10	15	20
Accuracy	0.8692	0.9590	0.9987	0.9987
F1	0.8667	0.9592	0.9987	0.9987
Sensitivity	0.8639	0.9592	0.9987	0.9987
Specificity	0.8698	0.9592	0.9987	0.9987

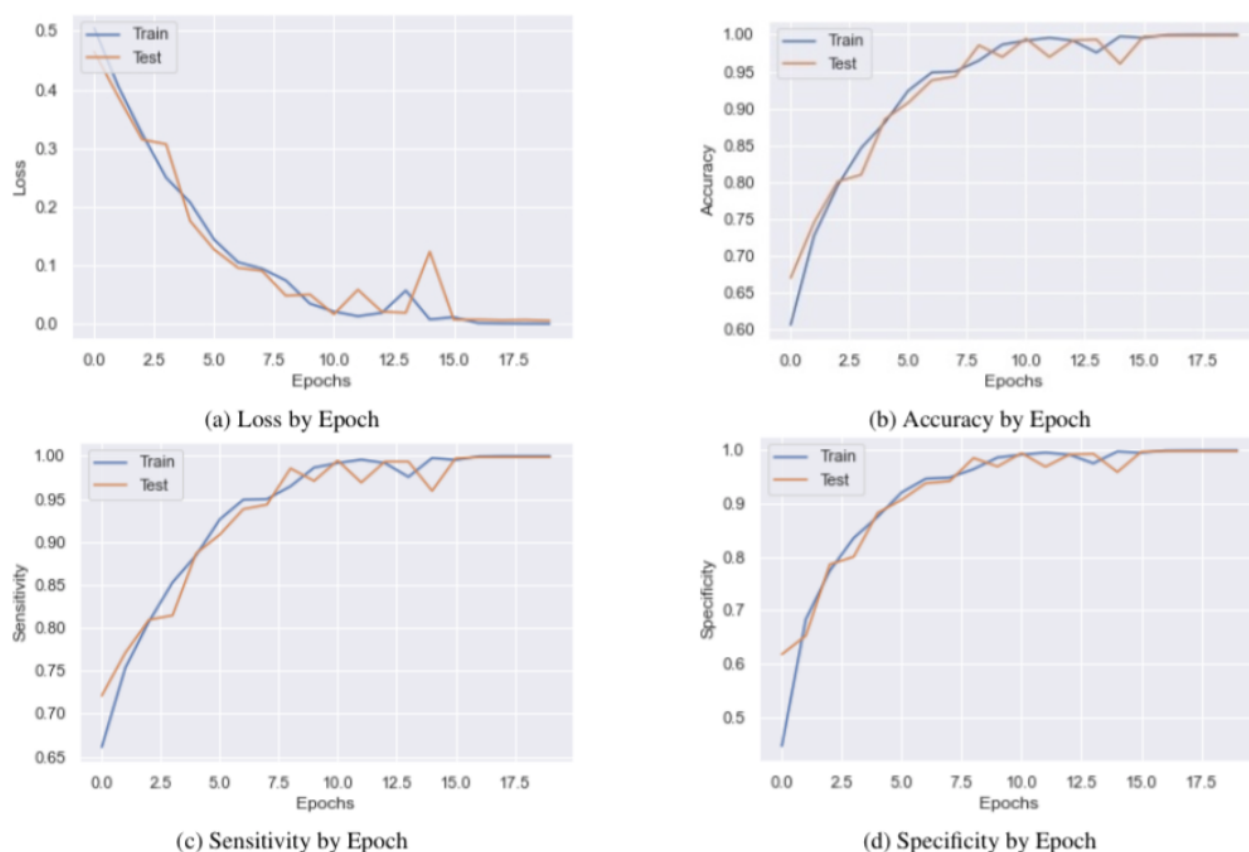


Figure 4: Training Logs

4 Conclusion

We utilized affine transformations to artificially augment our image dataset to determine its effect on the performance of our deep convolutional neural network algorithm, BurnNet, to accurately classify images of burned skin as full-thickness, deep dermal, or superficial dermal. Specifically, we applied 72 affine transformations per image in the original dataset and subsequently trained the model for twenty epochs. We found that our model diagnosed dermal burns with 99.87% accuracy, significantly exceeding the efficiency and performance of trained professionals. Our study not only supports the efficacy of deep learning frameworks for dermal burn classification but also provides evidence for the application of

data augmentation techniques for improved performance. We hope that our study will provide a major step forward in the automation of medical diagnosis and lower healthcare costs on a global scale.

References

- [1] Bhansali, R., Kumar, R., Writer, D. (2020). CoronaNet: A Novel Deep Learning Model for COVID-19 Detection in CT Scans. *Journal of Student Research*, 9(2). doi:10.47611/jsrshs.v9i2.1246
- [2] Biomedical Image Processing (BIP) Group From the Signal Theory and Communications Department (University of Seville, SPAIN) and Virgen del Rocío Hospital (Seville, Spain). Burns BIP_US Database. Accessed: Nov. 1, 2018. [Online]. Available: http://personal.us.es/rbolloix/Burns_BIP_US_database.zip
- [3] King, D. R., Li, W., Squiers, J. J., Mohan, R., Sellke, E., Mo, W., Zhang, X., Fan, W., Michael DiMaio, J., Thatcher, J. E. (2015). Surgical wound debridement sequentially characterized in a porcine burn model with multispectral imaging. *Burns*, 41(7), 1478-1487. doi:10.1016/j.burns.2015.05.009
- [4] Li, H. (2012). Skin Burns Degree Determined by Computer Image Processing Method. *Physics Procedia*, 33(7), 758-764. doi:10.1016/j.phpro.2012.05.132
- [5] Sabeena, B. and Raj Kumar, P. (2012). Diagnosis and Detection of Skin Burn Analysis Segmentation in Colour Skin Images *International Journal of Advanced Research in Computer and Communication Engineering ISO Certified* vol. 6, no. 2, pp. 369-374, 2017.
- [6] Sato, R., Iwamoto, Y., Cho, K., Kang, D.Y., and Chen, Y. (2020). Accurate BAPL Score Classification of Brain PET Images Based on Convolutional Neural Networks with a Joint Discriminative Loss Function. *Applied Sciences*, 10(3), 965. doi:10.3390/app10030965
- [7] Vertan, C., Badea, M. S., Florea, C., Florea, L., and Bădoiu, S. (2017). MPEG-7 visual descriptors selection for burn characterization by multidimensional scaling match. *2017 E-Health and Bioengineering Conference, EHB 2017* 253-256. doi:10.1109/EHB.2017.7995409
- [8] Yadav, D. P., Sharma, A., Singh, M., and Goyal, A. (2019). Feature Extraction Based Machine Learning for Human Burn Diagnosis From Burn Images. *IEEE Journal of Translational Engineering in Health and Medicine*, 7, 1-7. doi:10.1109/jtehm.2019.2923628