2024 IEEE International Conference on Image Processing (ICIP) | 979-8-3503-4939-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICIP51287.2024.10647789

BuRnSNet: BURN REGION SEGMENTATION NETWORK FROM COLOR IMAGES WITH TWO-WAY CNN

Joohi Chauhan[1,3], Paul L. Rosin[2], Puneet Goyal[3,4]

[1]Motilal Nehru National Institute of Technology Allahabad, India. [2]Cardiff University, UK.
[3]Indian Institute of Technology Ropar, India. [4]NIET, NIMS University, India.

ABSTRACT

Burn injury is a serious health issue leading to several thousands of annual fatalities. The color image-based automated burns diagnostic and assessment methods hold the potential for timely diagnosis and treatment. However, the research is limited in this domain which remains a major challenge. In this work, we explore and address the complex task of burn region segmentation in color images of burn patients. We present a semantic segmentation network that has two parallel sub-networks: a spatial-stream network for extracting low-level features and a contextual-stream network for generating a larger receptive field. Our network utilizes the pre-trained ResNet101 network, global average pooling, and instance normalization for better encoding and fusion of the network outputs. This dual-stream approach optimizes the performance in situations where data scarcity poses a challenge, facilitating robust semantic segmentation despite limited training samples. We prepared a pixel-wise labeled dataset for burn region segmentation and the experimental results on this dataset show that our proposed network outperforms several state-of-the-art semantic segmentation methods. Our method achieved mIOU and Matthews' correlation coefficient (MCC) of 74.3% and 81.7%, respectively, approximately 4.5% higher than the second-best performing method. The *Extended Burn Image Segmentation (EBIS)* dataset and our model are available at *https://github.com/VEDAs-Lab/EBIS*

Index Terms— Burn injury, Semantic image segmentation, Convolutional neural network, Instance Normalization, Matthews' correlation coefficient.

1. INTRODUCTION

Burns are amongst the most life-threatening injuries globally. As per the WHO 2018 Burns report, approximately 180,000 deaths occur annually due to burns, making it a leading cause of mortality and physical disabilities in many countries. In the United States, emergency departments care for over 400,000 patients with burn-related injuries annually, and approximately 3275 of these result in death [1]. In the UK, a concerning statistic reveals that 7335 children experienced burns or scalds in 2022. Also, in January 2023, 631 children suffered from burn injuries or scalds. Among them, 256 were in the 0-2 age group, with hot drinks being the primary source of burns [3]. Some reports in India estimate that there are around 7 million burn injury cases annually, with around 70% of cases in the most productive age group (15-35 years), particularly those belonging to poor socio-economic strata [2]. These reports highlight the severe consequences and challenges associated with burn injuries.

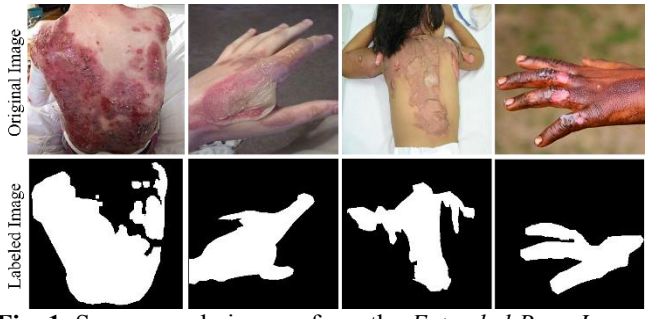

Fig. 1. Some sample images from the *Extended Burn Image Segmentation (EBIS)* dataset. The first row presents the original images, and the second row shows the corresponding labeled images.

Burn injuries can generally result from various sources such as fires, scalding, electricity, radiation, and chemicals. Automated burn diagnosis has the potential to expedite treatment and improve outcomes for burn injuries, addressing challenges associated with the limited availability of burn units and the prevalent use of manual methods in burn care. In this work, we study the complex task of efficiently performing burn region segmentation in color images of burn patients, aiming to contribute to developing low-cost, easily deployable computer-aided burn diagnosis systems. Considering the remarkable proven performance and widespread adoption of Artificial Intelligence (AI) techniques in various computer vision and medical applications, we explore deep learning-based pixel-level segmentation approaches to address the burn region segmentation challenges. While active interest exists in semantic pixel-level segmentation research [4-19] with some methods exhibiting impressive performance across various benchmark datasets [4-6], these methods have not

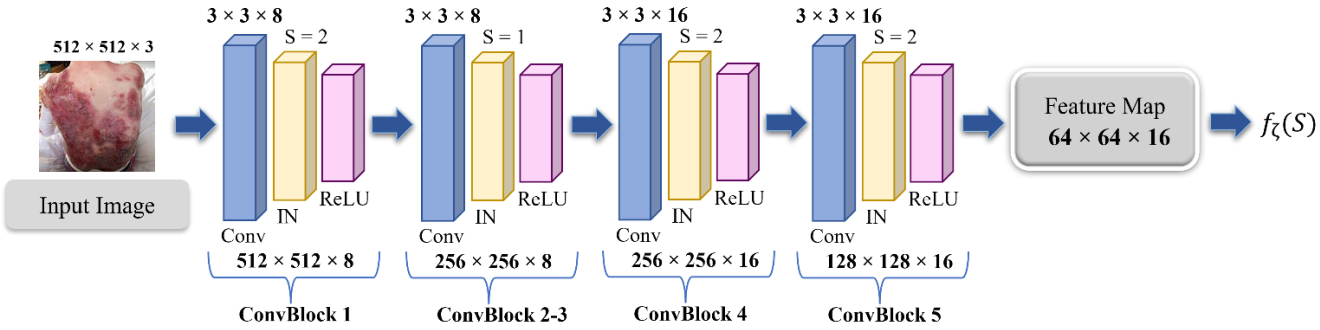

Fig. 2. Overview of *Spatial Stream Network;* A shallow network comprises of 5 convolutional blocks, where each block has a convolutional layer (Conv), Instance Normalization (IN), and ReLU activation function.

been extensively explored in the medical domain. This research aims to bridge this gap and leverage the potential of AI in improving burn diagnosis and assessment. To summarize, our main contributions are as follows:

- We prepared a labeled dataset of 600 burn images, where each pixel is assigned to either the burn or non-burn class; we refer to this dataset as the *Extended Burn Images Segmentation (EBIS)* dataset and is now made publicly available. Fig. 1 shows some of the sample images and their annotations from the dataset.
- We examine the performance of several state-of-the-art semantic image segmentation methods for burn region segmentation.
- We present a dual-stream network that captures and fuses contextual information with encoded rich spatial information. Our method outperforms the state-of-the-art CNNs on the *EBIS* dataset.

2. BACKGROUND

Segmentation methods are designed to differentiate between burn regions and normal regions in an RGB image of a burn patient. Achieving precise segmentation of the burn region from the normal one is crucial for accurately assessing total burn surface area, burn depth, and other related burn diagnosis parameters. Unfortunately, visual assessments by burn specialists yield only 40-70% accuracy in estimating these parameters during the first few days following a burn [20, 21]. Also, the accurate assessment of injury during the initial days is much needed to avoid the delay in the treatment of the patient.

Different imaging techniques, such as laser Doppler imaging (LDI), thermal imaging, and photographic imaging have been proposed for developing automated burn diagnosis tools [22-31]. Research shows that compared to thermal imaging, LDI achieves better diagnostic accuracy, probably because of the susceptibility of thermal imaging to environmental factors. Several research studies support LDI to be the most reliable technique for burn depth assessment, but unfortunately, its real-world use is limited due to its higher costs, associated delays, and limited portability [22-24].

In recent decades, many automatic image analysis algorithms have been developed for healthcare professionals to provide precise clinical decisions and assessment ability through medical images [32, 33]. Varying skin color/characteristics, varying background/illumination conditions, varying burn degrees and different causes of burns make burn color image segmentation problems quite challenging. Earlier explored image processing and traditional ML-based approaches on color images for burn diagnosis parameters estimations have achieved very limited accuracy [25-30]. Acha et al. [25-28] contributed to the domain of burn depth assessment methods using color images, and they also demonstrated a psychophysical experiment and multidimensional scaling analysis to determine the physical characteristics employed by physicians for the diagnosis of burn depth [28]. Cirillo et al. [30] recently proposed a burn images segmentation approach using tensor decomposition followed by GLCM for feature extraction and fuzzy C-means for cluster analysis and found this to be better than PCA and ICA-based methods. They also explored SegNet [7] but unsurprisingly noted it to be ineffective given the limited training data of 11 images. However, not much research has been done on burn image segmentation using deep learning approaches. A 2017 summary paper discussing over 300 deep learning works for medical imaging did not include burn diagnosis [34]. Some of the possible reasons are the lack of any publicly available labeled burn images dataset, limited awareness, and the challenging aspects of this problem. Jiao et al. [35] proposed a deep learning segmentation framework that is based on Mask R-CNN (regions with convolutional neural network). In their work, they used ResNet-101 with atrous convolution in a feature pyramid network as the backbone and tested it on the burn image dataset they prepared using the smartphone camera. That labeled dataset is however not publicly available. A recent work by Cirillo et al. [44] created a dataset of 100 polarized

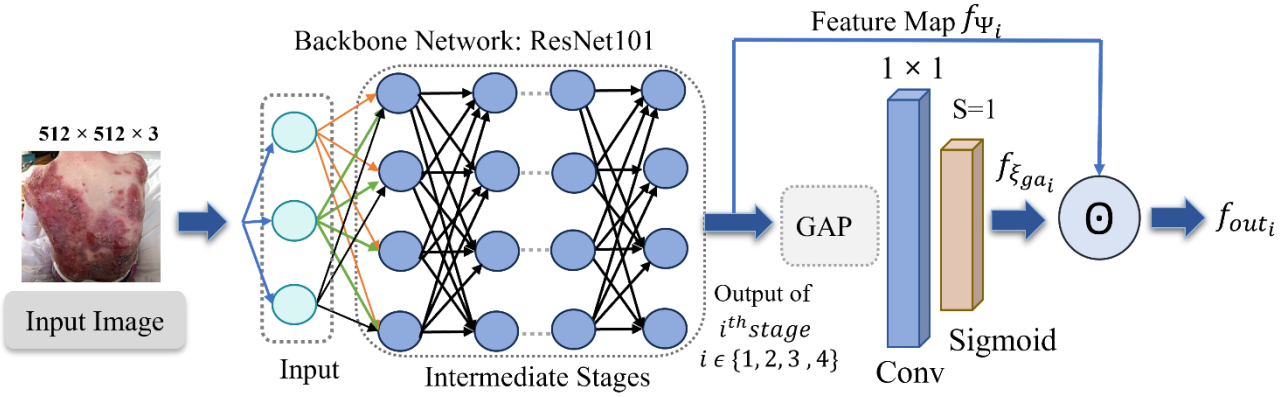

Fig. 3. Illustration of the *Contextual Stream Network* (CSN) that utilizes the output of the intermediate layers of ResNet101; a global average pooling (GAP) followed by a 1×1 Conv layer and sigmoid activation is applied to the output of each stage.

high-performance light camera images of burn wounds of age group less than 4 years old, they utilized the U-Net model for the burn depth assessment based on semantic segmentation, though the dataset is not publicly released.

3. BURN REGION SEGMENTATION NETWORK FRAMEWORK

Pixel-level image segmentation involves assigning a unique label to each pixel in an image from a set of K classes. Given an input color image x of size $M \times N \times 3$, the output of the segmentation network is defined as $y \in \mathbb{R}^{M \times N \times K}$. At the pixel location (i, j) the class probability distribution output by the Convolutional Neural Network (CNN) is represented as a vector $y(i, j) \in \mathbb{R}^K$.

The proposed framework, BuRnSNet, employs a multi-path architecture that is capable of effectively encoding rich spatial and contextual information. It combines these feature maps to generate high-resolution segmentation predictions. The spatial stream network focuses on spatial information, while the contextual stream network extracts high-level features that help in category recognition.

Spatial stream network (SSN) learns visual information such as texture, shape, detailed boundaries, etc. from an input image. The network takes an image size of 512×512×3 as input, this network utilizes the principle of convolutional neural network and comprises of 5 convolution blocks (ConvBlock). Each ConvBlock contains a convolutional layer with a kernel size of 3 × 3, followed by a Rectified Linear Unit (ReLU) activation function [36] and an instance normalization layer. Instance Normalization [37] is employed to normalize each channel independently across the spatial dimensions, introducing invariance to contrast shifts and intensity variations, thus making the feature vector adaptable for each image. Moreover, Instance Normalization alleviates the problem that occurs due to images with varying intensities. Further, to preserve the majority of spatial information, the stacked ConvBlocks are set to generate a feature map that is 1/8 of the original input image resolution. This large size of the feature map ensures that rich spatial details are retained, thus enhancing the ability of the network to capture a larger set of visual features. An overview of this stream of the network is presented in Fig. 2.

Contextual stream network (CSN) is designed to provide a sufficient receptive field and focus on global and local information. Theoretically, it is proven that the receptive field of ResNet [38] is already larger than the input image. However, Zhuo et al. [39] presented that the *empirical* receptive field of a CNN is smaller than the *theoretical* value. Therefore, for generating a large receptive field and maintaining a computationally efficient network, we use pre-trained CNN (ResNet101) as the backbone C_p that down samples the feature map of each i^{th} stage and obtains a large receptive field R_i. The feature map generated from the adopted approach is denoted as $f_\zeta(R)$. We believe that the features from the middle and high-level stages are of great importance for encoding global contextual information. So, we add a global average pooling layer ξ_{ga} at the end of each convolution block of middle and high-level stages of C_p for obtaining maximal receptive field R_{max} having the global context information. ResNet network has generally 4 stages, each comprising multiple residual blocks. For the i^{th} stage, the feature map f_{Ψ_i} is element-wise multiplied with the feature map of the corresponding ξ_{ga} denoted as $f_{\xi_{ga_i}}$, for computing the output feature map f_{out_i} of that stage, illustrated in Fig. 3.

$$f_{out_i} = f_{\Psi_i} \odot f_{\xi_{ga_i}} \tag{1}$$

Finally, the feature maps of the up-sampled output of the convolution blocks are concatenated, shown in Fig. 4. The adopted approach aims to refine the final output

features of each convolution block and extract the global context information at high resolution without excessive computational cost.

$$f_{\zeta}(R) = f^{s}_{out_2} \oplus f^{s}_{out_4} \quad (2)$$

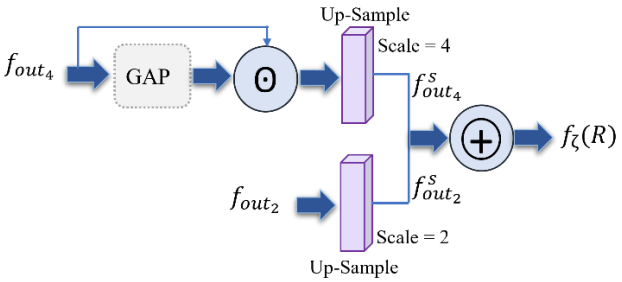

Fig. 4. *Contextual Stream Network (CSN):* Concatenation of the 2^{nd} and 4^{th} stage feature maps obtained from eq. 1.

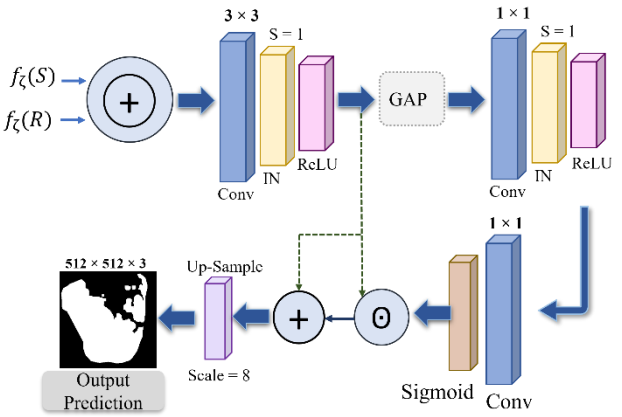

Fig. 5. Fusion module of the proposed network, which performs fusion of the outputs from SSN and CSN modules.

In this network, we use the concept of an incomplete U-shape structure. The hierarchical features of the pre-trained network are fused along with the features obtained from global average pooling. The U-shape structure [40] increases the spatial resolution of the feature map and recovers most of the missing details. Both the network outputs are finally fused together into a high-resolution feature map, as shown in Fig. 5. The final output F_s of the combined network is obtained by upscaling with a factor $S_c = 8$. We use bilinear interpolation $\mathcal{BI}(z, s)$ (where z is the input and s be the scale factor) for resizing the feature map.

4. EXPERIMENTS

4.1. Dataset Description

Due to the non-availability of open access pixel-wise labeled datasets of burn images, we developed a new dataset comprising high-resolution burn images with varied illumination conditions, different source cameras, and varying distances between the patient and the camera. This dataset was created using semi-automated scripts on the Google search engine with queries such as 'skin burn', 'burn wounds', 'burn injuries', 'full thickness burns', 'partial thickness burns', 'deep burn', etc. The images include burns on the back, hand, face, and inner forearm body parts. Further, we manually annotated these images pixel-wise, classifying each pixel as either burn or non-burn. The annotations are done under the consultation of a medical expert who has over three decades of experience in burns diagnosis and treatment. The sample images from the prepared dataset and their corresponding labeled images are illustrated in Fig. 1. The dataset consists of 316 images in the training set, 76 images in the validation set, and 208 images in the test set.

4.2. Implementation Details

We utilize the ResNet-101 [38] network pre-trained on the ImageNet [41] dataset to downsample the feature map. We executed our proposed approach on a system with an Intel® Core i7-8700 CPU @ 3.20 GHz, 32 GB DDR4 RAM, and an NVIDIA GTX1080 (8 GB) Graphics card. The resolution of the output feature map generated by the network is the same as that of the input image and due to the limited GPU memory, input images are resized to 512×512 during the training phase. In our model, a convolutional layer with 1×1 kernel size is employed to predict the label feature map, followed by bilinear interpolation-based up-sampling of output logits by a factor of 8 to facilitate the final computation and prediction task. We adopt the Adam optimizer [42] with an initial learning rate of 1.0×10^{-4} and weight decay of 1.0×10^{-5}. The learning rate follows a polynomial decay policy, where the initial learning rate is multiplied by $\left(1-\frac{epoch}{maxEpoches}\right)^{0.9}$. Additionally, we employ data preprocessing techniques to deal with the problem of overfitting. As part of this preprocessing step, basic data augmentation techniques are implemented, including random horizontal flipping, vertical flipping, and random intensity jittering on the training dataset.

For evaluating the performance of the proposed method and some of the existing state-of-the-art semantic image segmentation methods, we consider the following metrics:

$$PPV = \frac{1}{r}\sum_{i=1}^{r}\left(\frac{\sum_{\epsilon} TP_{\epsilon}}{\sum_{\epsilon}(TP_{\epsilon}+FP_{\epsilon})}\right) \quad (3)$$

$$TPR = \frac{1}{r}\sum_{i=1}^{r}\left(\sum_{\epsilon}\frac{TP_{\epsilon}}{(TP_{\epsilon}+FN_{\epsilon})}\right) \quad (4)$$

$$DSC = \frac{1}{r}\sum_{i=1}^{r}\left(\frac{1}{K}\sum_{\epsilon}\frac{2TP_{\epsilon}}{(2TP_{\epsilon}+FP_{\epsilon}+FN_{\epsilon})}\right) \quad (5)$$

$$mIoU = \frac{1}{r}\sum_{i=1}^{r}\left(\frac{1}{K}\sum_{\epsilon}\frac{TP_{\epsilon}}{(TP_{\epsilon}+FP_{\epsilon}+FN_{\epsilon})}\right) \quad (6)$$

where *PPV* refers to Positive Prediction Value, *TPR:* True Positive Rate, DSC: Dice Similarity Coefficient, and *mIoU*

is the mean Intersection over Union. TP, FP, TN, and FN denote True Positive, False Positive, True Negative, and False Negative, respectively. We additionally employ Matthews' correlation coefficient [43] (MCC) as a metric, which measures the correlation between the ground truth and the predicted binary mask with a value ranging from *-1 to +1*. It is defined as:

$$MCC = \frac{1}{r}\sum_{i=1}^{r}\frac{TP_i \times TN_i - FP_i \times FN_i}{\sqrt{(TP_i + FP_i)(TP_i + FN_i)(TN_i + FP_i)(TN_i + FN_i)}} \quad (7)$$

where r is the number of samples in the set on which the performance of the network is to be evaluated.

4.3. Experimental Results and Analysis

In this section, we present the results achieved by our proposed segmentation frameworks on the *EBIS* test set, along with a comparative analysis of its performance against existing state-of-the-art semantic segmentation methods. To ensure a fair comparison, all the methods are trained on the same dataset, i.e., the *EBIS* training set, and their performance is subsequently evaluated on the *EBIS* test set. The detailed experimental results, including the average performance metrics of our proposed methods and the other methods, are presented in Table 1.

BuRnSNet outperforms other state-of-the-art methods with an average performance improvement of 9.6%, 12.2%, and 10.7% in DSC, *mIoU*, and MCC, respectively, in comparison to RefineNet [18]. Furthermore, this method shows a significant improvement of around 3.8%, 5.4%, and 4.5% in DSC, *mIoU*, and MCC, respectively, in comparison to the MSAC-SNet [45].

Table 1. Performance (in %) of the proposed methods and other state-of-the-art methods on *EBIS test set*. All results are achieved by training only on Burn Image training set.

Method	PPV	TPR	DSC	mIoU	MCC
DenseASPP [8]	61.4	85.4	66.2	52.1	61.4
AdapNet [5]	65.3	80.0	68.1	53.1	62.7
SegNet [7]	64.4	79.7	68.2	53.3	62.5
BiSeNet [9]	73.8	69.7	68.2	53.3	63.3
PSPNet [10]	64.8	83.9	69.2	54.9	64.7
DeepLabV3 [4]	70.2	83.4	74.1	60.3	68.9
UNet [46]	79.3	74.5	73.4	61.0	69.5
DeepLabV3+[11]	67.7	89.3	74.7	61.5	70.5
RefineNet [18]	66.7	92.5	75.2	62.1	71.0
MSAC-SNet [45]	82.4	82.2	81.0	68.9	77.2
BuRnSNet (ours)	**82.8**	**88.7**	**84.8**	**74.3**	**81.7**

Fig. 6 and 7 show the distribution comparison of our proposed methods with other state-of-the-art segmentation methods, on the *EBIS* test set, in terms of *mIoU* and DSC, respectively. The BuRnSNet achieves *mIoU* of 80% or higher for around 39% of test set images, while the RefineNet and MSAC-SNet yield that for only around 13% of the total images. With respect to DSC also, the proposed network BuRnSNet yields a value of 0.9 or above for 35% of the images and the other methods best achieve so for only 13% of the images (by UNet). The analysis of the distribution of *mIoU* and DSC indicates the robustness and potential applicability of our proposed method BuRnSNet for burn region segmentation, as the segmentation results with high *mIoU* and DSC are obtained for the majority of *EBIS* dataset by using BuRnSNet, with significantly better performance than the other methods considered.

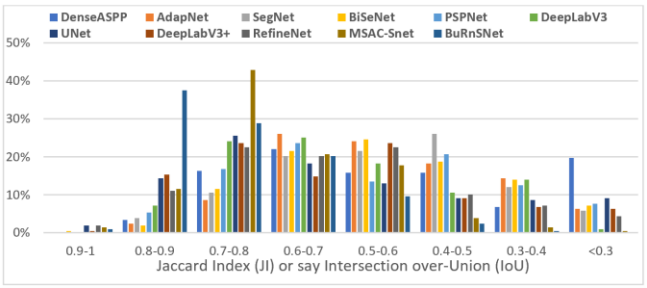

Fig. 6. Distribution comparison of different segmentation methods (on *EBIS test set* images) in terms of Intersection over Union (IoU).

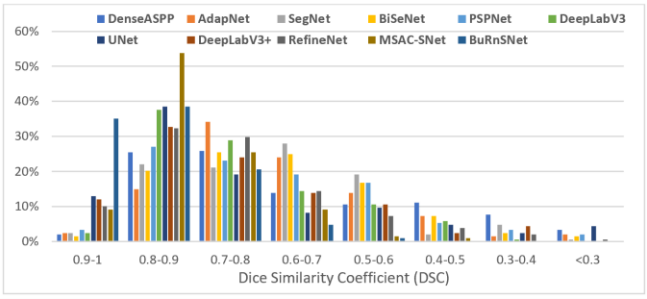

Fig. 7. Distribution comparison of different segmentation methods (on *EBIS test set* images) in terms of Dice similarity coefficient (DSC).

Our proposed method effectively predicts burn segments from color images of burn patients across diverse conditions, including patchy burn regions, irregular boundaries, and variations in the color and texture of burns. These conditions pose significant challenges for existing semantic segmentation methods, as evident from the analysis of original and predicted sample images presented in Fig. 8. Upon analyzing images generated by other best-performing methods, such as RefineNet [18] and DeepLabV3+ [11], we observe that while these models excel in identifying true positives, they also tend to produce an increased number of false positives. Despite their efforts to effectively encode semantic information and consider low-level features, these methods struggle to capture detailed information related to object boundaries. This

limitation compromises the discriminatory ability of these networks, particularly at the boundaries of objects. As shown in Fig. 8, MSAC-SNet [45], inspired by DeepLab which uses the atrous convolution, improvises the output prediction and reduces the number of false positives in comparison to DeepLabV3+ but it still misses similar burn regions. For example, in the third row of Fig. 8, blister in the image is predicted as background by both these methods.

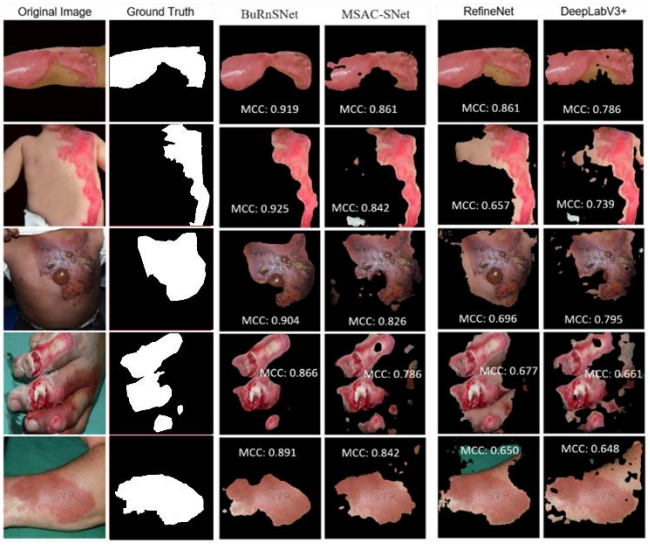

Fig. 8. Some qualitative results as obtained by the proposed segmentation method BuRnSNet and other best performing existing methods on *EBIS test set*.

4.4. Ablation Study

Table 2. Performance (in %) of the proposed network with different network settings.

Method	PPV	TPR	DSC	mIoU	MCC
SSN_only	76.2	74.9	72.4	58.5	68.1
CSN_f4_only	81.7	79.9	78.4	66.5	75.1
CSN (f3+f4)	78.6	87.9	80.4	68.5	77.6
SSN+CSN_f3_only	81.1	84.5	80.8	69.0	77.4
SSN+CSN_f4_only	81.2	85.3	81.8	71.0	78.7
Our	**82.8**	**88.7**	**84.8**	**74.3**	**81.7**

For a more detailed evaluation and interpretation of the proposed network, we conducted multiple experiments with different network settings. The evaluation results on the test set of the *EBIS* dataset are summarized in Table 2. In our experiments, we utilized ResNet101 as the backbone network. Overall, we observe that the Contextual Stream Network (CSN) outperforms the Spatial Stream Network (SSN) by 10% and 9.5% in terms of *mIoU* and MCC, respectively. However, combining both sub-networks: SSN and CSN, with CSN considering output from both stages 3 and 4 (i.e. both mid-level and high-level features) helps in achieving the best performance.

5. CONCLUSIONS

This paper presents a new framework, BuRnSNet, designed specifically for pixel-level segmentation of burn regions from the color images of burn patients. The primary motivation behind this work and the proposed network was the limited exploration in this domain and the need for an efficient segmentation network that can better assist in the burn diagnosis and assessment process. We evaluated the performance of BuRnSNet and compared it with other state-of-the-art models on the prepared *EBIS* dataset. Extensive experimental results showcase that our proposed model significantly outperforms existing state-of-the-art models. BuRnSNet effectively encodes and fuses spatial and contextual information, resulting in superior segmentation accuracy. To ensure reproducibility and facilitate further research, both the codes and the labeled dataset of this work are made publicly available. In the future, we plan to expand the dataset and explore customized approaches to enhance performance even further. We hope that this work will serve as a benchmark for future research in developing advanced visual image-based automated burns diagnostic systems.

6. REFERENCES

[1] Association AB, "National burn awareness week 2017 fact sheet," 2017.

[2] K. Agrawal, Practical Handbook of Burns Management. NPPMRBI.https://dghs.gov.in/WriteReadData/userfiles/file/Practical_handbook-revised_Karoon.pdf

[3] "Causes of burns and preventions: hot drinks," [Online]. Available: https://cbtrust.org.uk/get-informed/causes-of-burns-and-preventions/hot-drinks/. Accessed: Feb. 3, 2024.

[4] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," arXiv preprint arXiv:1706.05587, 2017.

[5] A. Valada, J. Vertens, A. Dhall, and W. Burgard, "AdapNet: Adaptive semantic segmentation in adverse environment conditions," in Proc. ICRA, Singapore, 2017.

[6] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834–848, 2017.

[7] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 12, pp. 2481–2495, 2017.

[8] M. Yang, K. Yu, C. Zhang, Z. Li, and K. Yang, "DenseASPP for semantic segmentation in street scenes," in Proc. CVPR, pp. 3684-3692, 2018.

[9] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, and N. Sang, "BiSeNet: Bilateral segmentation network for semantic image segmentation," in Proc. ECCV, pp. 325-341, 2018.

[10] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in Proc. CVPR, pp. 2881-2890, 2017.
[11] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in Proc. ECCV, 2018.
[12] D. Eigen and R. Fergus, "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture," in Proc. ICCV, 2015.
[13] R. Vemulapalli, O. Tuzel, M. Liu, and R. Chellappa, "Gaussian conditional random field network for semantic segmentation," in Proc. CVPR, Las Vegas, USA, 2016.
[14] L. Chen, Y. Yang, J. Wang, W. Xu, and A. Yuille, "Attention to scale: Scale aware semantic image segmentation," in Proc. CVPR, USA, pp. 3640-3649, 2016.
[15] G. Lin, C. Shen, I. D. Reid, and A. R. Dick, "Efficient piecewise training of deep structured models for semantic segmentation," in Proc. CVPR, Las Vegas, USA, 2016.
[16] V. Jampani, M. Kiefel, and P. V. Gehler, "Learning sparse high dimensional filters: Image filtering, dense CRFs and bilateral neural networks," in Proc. CVPR, USA, 2016.
[17] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," in Proc. ICLR, 2016.
[18] G. Lin, F. Liu, A. Milan, C. Shen, and I. Reid, "RefineNet: multi-path refinement networks for dense prediction," IEEE Trans. Pattern Anal. Mach. Intell., 2019.
[19] P. Wang, P. Chen, Y. Yuan, D. Liu, Z. Huang, X. Hou, and G. Cottrell, "Understanding convolution for semantic segmentation," in Proc. WACV, 2017.
[20] H. Hoeksema, K. Van-de-Sijpe, T. Tondu, M. Hamdi, V. K. Landuyt, P. Blondeel, and S. Monstrey, "Accuracy of early depth assessment by laser Doppler imaging on different days post burn," Burns, vol. 35, no. 1, pp. 36-35, 2009.
[21] S. Monstrey, H. Hoeksema, J. Verbelen, A. Pirayesh, and P. Blondeel, "Assessment of burn depth and burn wound healing potential," Burns, vol. 34, no. 6, pp. 761-769, 2008.
[22] A. Pape, S. Skouras, and P. O. Bryne, "An audit of the use of laser Doppler imaging (LDI) in the assessment of burns of intermediate depth," Burns, vol. 27, pp. 233-239, 2001.
[23] T. Resch, M. R. Drake, D. S. Helmer, D. G. Jost, and S. J. Osland, "Estimation of burn depth at burn centers in the United States: A survey," J. Burn Care Res., vol. 35, no. 6, pp. 491-497, 2014.
[24] C. Wearn, C. K. Lee, J. Hardwicke, A. Allouni, A. Bamford, P. Nightingale, and N. Moiemen, "Prospective comparative evaluation study of laser Doppler imaging and thermal imaging in the assessment of burn depth," Burns, vol. 44, no. 1, pp. 124-133, 2018.
[25] B. Acha, C. Serrano, and L. Roa, "Segmentation and classification of burn color images," in Proc. IEEE Eng. Med. Biol. Soc., Turkey, 2001.
[26] B. Acha, C. Serrano, I. J. Acha, and M. L. Roa, "Segmentation and classification of burn images by color and texture information," J. Biomed. Opt., vol. 10, no. 3, 034014, 2005.
[27] B. Acha, C. Serrano, I. Fondon, and T. Gomez-Cia, "Burn depth analysis using multidimensional scaling applied to psychophysical experiment data," IEEE Trans. Med. Imaging, vol. 32, no. 6, pp. 1111-1120, 2013.
[28] S. Carmen, R. Blanco, T. Gomez-Cia, and B. Acha, "Features identification for automatic burn classification," Burns, vol. 41, no. 8, pp. 1883-1890, 2015.
[29] U. Sevik, E. Karakullakcu, T. Berber, Y. Akbas, and S. Turkylmaz, "Automatic classification of skin burn color images using texture-based feature extraction," IET Image Process., vol. 13, no. 11, pp. 2018-2028, 2019.
[30] D. M. Cirillo, R. Mirdell, F. Sjoberg, and T. D. Pham, "Tensor decomposition for color image segmentation of burn wounds," Nature Sci. Rep., 2019.
[31] M. Badea, C. Vertan, C. Florea, L. Florea, and S. Bădoiu, "Severe burns assessment by joint color-thermal imagery and ensemble methods," in Proc. IEEE HealthCom, Germany, 2016.
[32] E. Iglesias and R. Sabuncu, "Multi-atlas segmentation of biomedical images: A survey," Med. Image Anal., vol. 24, no. 1, pp. 205-219, 2015.
[33] M. Viergever, J. Maintz, S. Klein, K. Murphy, M. Staring, and P. Pluim, "A survey of medical image registration-under review," Med. Image Anal., vol. 33, pp. 140-144, 2016.
[34] G. Litjens et al., "A survey on deep learning in medical image analysis," Med. Image Anal., vol. 42, pp. 60-88, 2017.
[35] C. Jiao, K. Su, W. Xie, and Z. Ye, "Burn image segmentation based on mask regions with convolutional neural network deep learning framework: more accurate and more convenient," Burns Trauma, vol. 7, p. 6, 2019.
[36] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in Proc. ICML, 2010.
[37] D. Ulyanov and A. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," arXiv preprint arXiv:1607.08022, 2016.
[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. CVPR, USA, 2016.
[39] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Object detectors emerge in deep scene CNNs," arXiv preprint arXiv:1412.6856, 2014.
[40] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proc. CVPR, 2015.
[41] J. Deng, W. Dong, R. Socher, L.-J. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in Proc. CVPR, Miami, USA, 2009.
[42] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2017.
[43] M. D. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," J. Mach. Learn. Technol., vol. 2, no. 1, pp. 37-63, 2011.
[44] M. D. Cirillo, R. Mirdell, F. Sjoberg, and T. D. Pham, "Improving burn depth assessment for pediatric scalds by AI based on semantic segmentation of polarized light photography images," Burns, vol. 47, no. 7, pp. 1586-1593, 2021.
[45] J. Chauhan and P. Goyal, "Convolution neural network for effective burn region segmentation of color images," Burns, vol. 47, no. 4, pp. 854-862, 2021.
[46] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," MICCAI, vol. 9351, 2015.