Assessment and Follow-up of Chronic Leprosy Wounds with Deep Learning and Smartphone Photography

Karen Sanchez*, Brayan Monroy*, Paula Arguello*, Juan Estupiñan*,
Jorge Bacca*, Claudia V. Correa*, Olinto Mieles†, Henry Arguello*,** and Fernando Rojas*,
*Universidad Industrial de Santander, †Sanatorio de Contratación E.S.E.
**henarfu@uis.edu.co

Abstract

This work presents a novel framework for chronic wound assessment and tracking based on deep learning, which works on RGB images captured with smartphones, avoiding bulky and complicated acquisition setups. The framework integrates mainstream algorithms for medical image processing, including wound detection, segmentation, and quantitative analysis of area and perimeter. Experiments and clinical validations in Colombia leprosy patients demonstrate the validity and accuracy of the proposed framework, with 84.5% *precision. Additionally, this paper provides to the scientific community with a new chronic wounds dataset of 164 images with their respective detection boxes and segmentation maps.*

1. Introduction

Chronic wounds affect 40 million people worldwide. These skin lesions are often caused by type two diabetes, cardiovascular affections, and neglected tropical diseases such as leprosy. Approximately 225,000 new leprosy cases are diagnosed yearly, with 8.9% occurring in children and adolescents [18]. Leprosy is a latent healthcare problem, mainly in tropical developing countries. To cite a particular case, in Colombia, approximately 350-500 new cases of leprosy are reported every year due to the Mycobacterium leprae, and, 383 new cases in 2018 [4].

Chronic wound management involves periodic visual inspection by medical personnel for infection control and moisture balance, where edge and size analysis is used to determine wound progression and make treatment decisions [2]. However, incorrect ulcer management can lead to limb amputations, infection, or even mortality. As occurs in rural areas due to insufficient medical infrastructures and poor rural transportation, which leads to infrequent patient visits, inadequate treatment, and interrupted wound tracking.

This paper was supported by the Vicerrectoría de Investigación y Extensión UIS, project code 3735

Therefore, the development of computational methods for the automatic analysis of medical images is a research trend [3]. Deep learning (DL) algorithms have been employed to detect and produce chronic wound segmentation maps on images [5,7,11]. However, DL methods are mainly designed to work with only one dataset, causing overfitting [17], which can lead to poor performance when these methods are applied to datasets collected by different clinical centers or populations, that may differ significantly in terms of acquisition parameters and protocols. Further, DL-based wound analysis is limited by the lack of publicly available data for training [13]. Specifically, there are few wounds datasets from diabetes patients, mainly from Europe and North America, which means that the model may not perform as well on patients from other regions, including developing countries [9]. This problem is further intensified when considering different types of wounds, such as leprosy, for which no public data is available. Furthermore, a computational framework has not yet been developed for the automatic tracking of leprosy-associated ulcers.

This paper presents "CO2Dnet", a DL framework designed to detect, segment, measure, and track wounds to support medical decision-making. This framework works with RGB images captured by smartphones, outperforming state-of-the-art methods by up to 16% in the F1-score metric. The "CO2Dnet" has been deployed in an online platform where medical staff and patients from Colombia can benefit from temporal wound evolution analysis and a personalized follow-up of each patient's condition. This paper also introduces "CO2Wounds", a labeled dataset of 164 leprosy-associated chronic wound images acquired on patients from the town of Contratación, Santander, Colombia. The full implementation is available at https://github.com/simatec-uis/CO2Dnet.

2. Methods

Captured wound images usually contain unnecessary information around the region of interest, i.e., the actual

wound. This creates a class imbalance scenario, where a significant percentage of image pixels correspond to the background, while only a small portion belong to the wound [8]. We propose a detection-segmentation strategy within the CO2Dnet framework to address this issue, and improve segmentation accuracy. This approach is based on previous research that has demonstrated that working exclusively on the region of interest can enhance the performance of segmentation models [14].

Specifically, the proposed CO2Dnet framework receives RGB images of chronic wounds as input, acquired with commercial smartphones. It consists of four main steps: A) data acquisition, B) wound and calibration pattern detection, C) wound segmentation, and D) area and perimeter calculation. These steps are conducted so that the region of interest that contains the wound is cropped and resized to obtain a standardized centered wound image, that is then segmented to extract the wound. The calibration pattern is fully characterized in advance so that wound area and perimeter can be calculated. Moreover, the estimated metrics from multiple captures at different dates enable a temporal analysis of the evolution of the chronic wound and, allow the medical staff to monitor the patient's condition as well as the efficiency of the prescribed treatment. Figure 1 depicts a general overview of the CO2Dnet framework.

2.1. Step A: Data Acquisition Protocol

The first step of the framework involves the image acquisition process to build a chronic wound dataset from Colombian leprosy patients. A customized calibration pattern was designed, which works as a calibration tool to calculate the area and perimeter of ulcers in the International System of Units. The shapes in the pattern are used in the detection and metrics calculation processes, as detailed in Sections 2.2 and 2.4 (Steps B and D, respectively). Additional characteristics of the calibration pattern, i.e., colored squares were carefully selected, such that the detection network can easily detect the pattern, due to color contrast. The top row includes the RGB and CMYK colors. The bottom row contains six colors corresponding to the Fitzpatrick Skin Color Scale [16]. The squares of the two left columns and the inner right column correspond to 22 colors specially defined and selected to match ulcer tissue representation as in [22]. Finally, the outer right column contains a gray scale, traditionally used in imaging systems to calibrate the white and black sensor responses.

The dataset for wound monitoring can be acquired using smartphones in communities without adequate medical assistance or advanced imaging technology. The images should include a calibration pattern next to the wound. The data set also provides detection and segmentation maps for each image.

2.2. Step B: Wound and Calibration Pattern Detection

This work employs YoloV4 [1] as a detection model to crop the wound region from images and improve segmentation results by reducing the impact of irrelevant information. The YoloV4 network is trained to detect both the chronic wound and the calibration pattern. RGB images from Step A are fed as input to the YoloV4 detection network, which outputs bounding boxes of the regions of interest (wound and calibration pattern) indicated by green and magenta squares in Fig. 1.

This process can be mathematically modeled as follows. Let $\mathcal{X} \in \mathcal{R}^{N\times M\times 3}$ be the input RGB image with $N \times M$ pixels, and $\mathbf{c} = [i_1, j_1, i_2, j_2]$ is the boundary box provided by the YoloV4 network, surrounding the detected object, where (i_1, j_1) and (i_2, j_2) are the top-left and bottom-right coordinates of the box, respectively. Since, in this work, the detection network $\mathcal{D}_\theta(\cdot)$ with parameters θ is adjusted to detect two objects: the wound and the calibration pattern, the actual output of the YoloV4 network is

$$\mathbf{c}_w, \mathbf{c}_p = \mathcal{D}_\theta(\mathcal{X}), \tag{1}$$

where, $\mathbf{c}_w$ and $\mathbf{c}_p$ are the coordinates of the boundary boxes for the wound region and calibration pattern, respectively. Then, the boundary boxes are used to crop the image $\mathcal{X}$, so as to obtain two independent sub-images: the wound $\mathcal{X}_w \in \mathbb{R}^{H\times W\times 3}$ and the calibration pattern $\mathcal{X}_p \in \mathbb{R}^{H\times W\times 3}$ as

$$\mathcal{X}_w = \text{CropResize}(\mathcal{X}, \mathbf{c}_w), \tag{2a}$$

$$\mathcal{X}_p = \text{Crop}(\mathcal{X}, \mathbf{c}_p), \tag{2b}$$

where CropResize$(\cdot)$ crops an image according to the bounding box coordinates, and provides a resized output of $H \times W$ pixels, with $H = W = 320$. Further, considering that task-relevant information varies for different image resolutions [19], the size parameters $H \times W$ for which the wound image is resized must match the size of the images used in the segmentation model training.

2.3. Step C: Wound Segmentation

The segmentation step can be seen as a classification task at pixel level of the wound image $\mathcal{X}_w \in \mathbb{R}^{H\times W\times 3}$, where each pixel in the image is associated with a class (wound/background), yielding a binary map of the wound region. Given a segmentation network $\mathcal{S}_\phi(\cdot)$ and its networks parameters ϕ, the estimated segmentation map can be mathematically modeled as

$$\hat{Y} = \mathcal{S}_\phi(\mathcal{X}_w), \tag{3}$$

where $\hat{Y} \in \{0, 1\}^{H\times W}$ is the estimated segmentation map, with one values indicating that a pixel corresponds to the wound region, and zero otherwise. Based on our previous analysis [14], the segmentation model used in this work follows a U-net architecture.

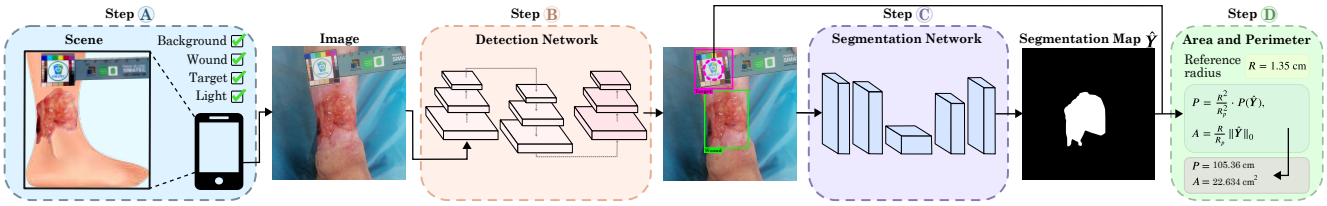

Figure 1. Overview of the proposed CO2Dnet deep learning-based framework for automatic segmentation and measurement of chronic wounds in RGB images acquired with traditional smartphones.

2.4. Step D: Calculation of Wound Area and Perimeter

The estimated binary segmentation map $\hat{Y}$ of the wound region and the calibration pattern region $\mathcal{X}_p$ are used to calculate the area and perimeter of the wound. This process requires the conversion factor from image pixels to centimeters, so as to estimate the wound area (A_w) and perimeter (P_w). Since the dimensions of the calibration pattern are known a priori, as well as the input image size, the relation between camera pixels and centimeters can be estimated via cross-multiplication. Specifically, we first employ a circle detection algorithm to measure (in pixels) the radius R_p of the circle in the detected calibration pattern $\mathcal{X}_p$, as $R_p = \text{calculateRadius}(\mathcal{X}_p)$. Then, since the actual radius of the circle is known, i.e., $R = 1.35$ cm, the conversion factor can be calculated as $C_f = \frac{R}{R_p}$. Thus, the conversion factor is used to calculate the area A_w and perimeter P_w of the wounds from the segmented images $\hat{\mathbf{Y}}$ as

$$A_w = C_f \|\hat{Y}\|_0, \quad P_w = C_f^2 \cdot P(\hat{Y}), \tag{4}$$

where, $\|\cdot\|_0$ is the ℓ_0 norm, namely, the amount of non-zero values, and $P(\cdot)$ denotes the contour perimeter of the binary segmentation map. To compute the contour perimeter of a closed shape, we employ the arcLength$(\cdot)$ function included in the OpenCV libray. The input parameters of arcLength$(\cdot)$ function are an input vector of 2D points, that describes the shape to analyze, and a flag indicating whether the shape is closed or not.

3. Results and Discussion

This section evaluates the performance of the CO2Dnet proposed framework for segmentation and tracking of chronic wound on an RGB image data set acquired from leprosy patients. All simulations were implemented in Python with Tensorflow 2.3 and conducted on an Nvidia Quadro Tesla T4 GPU with 16 GB of memory. Image data sets and evaluation metrics employed in these experiments are described below. The framework performance was evaluated using three metrics: precision, recall, and F1-score. We refer the interested reader to [12] for a detailed explanation on the calculation of these metrics.

3.1. Datasets and Metrics

The performance of the proposed framework was evaluated using two chronic wounds datasets: first, the new "CO2Wounds" dataset constructed, manually labelled, and published in this project; second, the public database "Chronic wounds" (CW-DB) [10]. Specifically, the CO2Wounds, whose images contain the designed calibration pattern, were used to evaluate all steps of the proposed framework, i.e., detection, segmentation, as well as area and perimeter estimation. While the CW-DB was used to evaluate the segmentation step. In the following, a detailed description of both datasets is provided.

3.1.1 Datasets

The ***new CO2Wounds dataset*** contains 164 chronic wound RGB images from leprosy patients, compiled in this study, following the acquisition process described in Step A of the proposed framework (Section 2.1), along with their binary segmentation maps. The images were acquired by the Leprosy program control from the Sanatorio de Contratación in Colombia. Images from 69 consenting patients were collected by medical staff during 7 months (November 2021 to June 2022), using different smartphone models and the provided calibration pattern. The public dataset is available[1], and it is constantly updated.

The public ***Chronic wounds database*** (CW-DB) contains 188 RGB images of diabetes patients from Poland and their corresponding wound segmentation maps.

3.2. Framework Evaluation

To provide a visual representation of the results obtained by each step of the proposed framework, Fig. 2 illustrates the images resulting from Steps A, B, C, and D for six different wound images in the CO2Wounds data set. Specifically, the first column presents the wound images acquired following the data protocol proposed in Step A (Section 2.1), which are the input to the proposed framework. The second column illustrates the results of Step B (Section 2.2), where the wound and calibration pattern are detected; specifically, the calibration pattern

[1]CO2Wounds database: https://doi.org/10.17632/nkw5gx57hw.

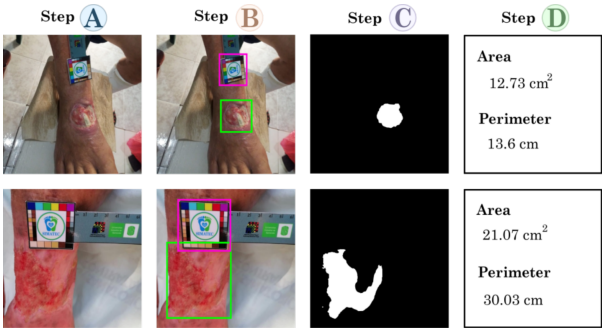

Figure 2. Step-by-step visual results of the proposed framework for six images from the CO2Wounds data set (rows). Each column corresponds to one step of the framework.

is highlighted by the magenta box, while the wound is enclosed by the lime green box. The third column depicts the resulting binary wound segmentation maps from Step C (Section 2.3). Finally, the fourth column presents the numerical results for wound area and perimeter estimation from Step D (Section 2.4). These results show that the proposed framework is able to work on images of different wound severity, and locations within the lower limbs. Also, in all cases, both, the wound and the calibration pattern were appropriately detected. Further, when more than one wound is present in the image, as in the last row of Fig. 2, the proposed framework is able to detect and segment both of them.

In addition, the performance of the proposed framework was compared with respect to state-of-the-art segmentation methods for the analysis of chronic wounds. Specifically, the following segmentation methods were considered: VGG16 [6], SegNet [21], MobileNetV2 [20], and Unet [15]. All methods were implemented and tested using the CO2Wounds data set acquired in this work, following the network configurations suggested by the authors of each work. Average results for 10 runs of each model are reported in Table 1, where it can be seen that the proposed framework outperforms its counterparts by at least 16%, 1.8%, and 11.2% in F1-score, precision, and recall, respectively.

Table 1. Segmentation comparison with state-of-the-art and deep-learning-based segmentation methods on the CO2Wounds data set.

Methods	(↑) **Metrics** (mean±std)		
	F1-Score	**Precision**	**Recall**
VGG16 [6]	50.6	68.7	53.9
SegNet [21]	46.4	68.4	51.5
MobileNetV2 [20]	54.1	65.9	70.7
Unet [15]	64.3	82.7	66.7
CO2Dnet (Ours)	**80.3**	**84.5**	**81.9**

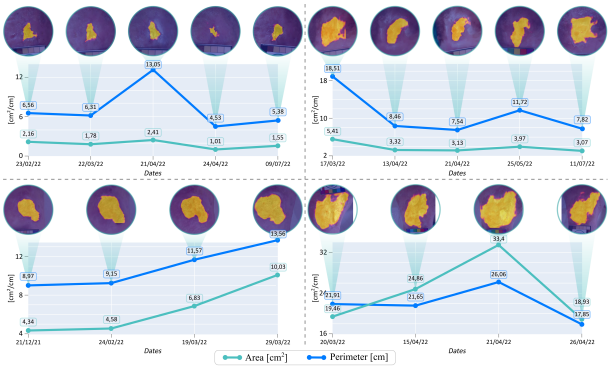

Figure 3. Tracking results of four leprosy chronic wounds using the proposed framework.

3.3. Clinical Tracking Validation

Figure 3 illustrates examples of the wound tacking results with the proposed framework deployed online, used by nurse practitioners from the Leprosy control program of Sanatorio de Contratacion, in the Colombian town of Contratacion, Santander, which are the main users of this platform. In particular, the evolution of four different ulcers is presented in terms of the area and perimeter, in cm^2 and cm, respectively. In this way, medical staff can make informed decisions related to the treatment. Also, the segmented wound image is included at the top for each date, corresponding to the activation maps of each chronic wound image in the last convolutional layer of the segmentation neural network (step C in Section 2.3). Since image acquisition and uploading dates are independent of the algorithm, the top two cases have five records each, while the cases at the bottom have four records each.

4. Conclusions

The CO2Dnet, a deep learning-based framework for chronic wounds assessment and tracking has been presented. The proposed framework includes four steps that include: image acquisition, wound and calibration pattern detection, wound segmentation, and wound area and perimeter estimation, where detection and segmentation employ deep neural networks with transfer learning and data augmentation schemes to avoid overfitting and improve performance. Additionally, we presented a new leprosy chronic wound dataset (CO2Wounds) to provide more variability in chronic wound data sets to enable future works in this context. Experimental results validate the proposed framework, with 80.3% F1 score, 84.5% precision, and 81.9% recall over the CO2Wounds dataset. These results overcome those from state-of-the-art segmentation methods (VGG16, SegNet, MobileNetV2, and U-net). Moreover, temporal analysis of the wounds is allowed by the proposed calibration pattern, which enables wound area and perimeter estimation. In this way, wounds can be monitored over time.

References

[1] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020. 2

[2] Steven Bowers and Eginia Franco. Chronic wounds: evaluation and management. *American family physician*, 101(3):159–166, 2020. 1

[3] Camilo Calderón, Karen Sanchez, Sergio Castillo, and Henry Arguello. Bilsk: A bilinear convolutional neural network approach for skin lesion classification. *Computer Methods and Programs in Biomedicine Update*, 1:100036, 2021. 1

[4] Nora Cardona-Castro. Leprosy in colombia. *Current Tropical Medicine Reports*, 5(2):85–90, 2018. 1

[5] Sawrawit Chairat, Tulaya Dissaneewate, Piyanun Wangkulangkul, Laliphat Kongpanichakul, and Sitthichok Chaichulee. Non-contact chronic wound analysis using deep learning. In *2021 13th Biomedical Engineering International Conference (BMEiCON)*, pages 1–5. IEEE, 2021. 1

[6] Manu Goyal, Moi Hoon Yap, Neil D Reeves, Satyan Rajbhandari, and Jennifer Spragg. Fully convolutional networks for diabetic foot ulcer segmentation. In *2017 IEEE international conference on systems, man, and cybernetics (SMC)*, pages 618–623. IEEE, 2017. 4

[7] Jui-Tse Hsu, Yung-Wei Chen, Te-Wei Ho, Hao-Chih Tai, Jin-Ming Wu, Hsin-Yun Sun, Chi-Sheng Hung, Yi-Chong Zeng, Sy-Yen Kuo, and Feipei Lai. Chronic wound assessment and infection detection method. *BMC medical informatics and decision making*, 19(1):1–20, 2019. 1

[8] Justin M Johnson and Taghi M Khoshgoftaar. Survey on deep learning with class imbalance. *Journal of Big Data*, 6(1):1–54, 2019. 2

[9] Michał Krecichwost, Joanna Czajkowska, Agata Wijata, Jan Juszczyk, Bartłomiej Pyciński, Marta Biesok, Marcin Rudzki, Jakub Majewski, Jacek Kostecki, and Ewa Pietka. Chronic wounds multimodal image database. *Computerized Medical Imaging and Graphics*, 88:101844, 2021. 1

[10] Michał Krecichwost, Joanna Czajkowska, Agata Wijata, Jan Juszczyk, Bartłomiej Pyciński, Marta Biesok, Marcin Rudzki, Jakub Majewski, Jacek Kostecki, and Ewa Pietka. Chronic wounds multimodal image database. *Computerized Medical Imaging and Graphics*, 88:101844, 2021. 3

[11] BK Shreyamsha Kumar, KC Anandakrishan, Manish Sumant, and Srinivasan Jayaraman. Wound care: Wound management system. *IEEE Access*, 2023. 1

[12] Peter Lukac, Robert Hudec, Miroslav Benco, Patrik Kamencay, Zuzana Dubcova, and Martina Zachariasova. Simple comparison of image segmentation algorithms based on evaluation criterion. In *Proceedings of 21st International Conference Radioelektronika 2011*, pages 1–4. IEEE, 2011. 3

[13] Domagoj Marijanović and Damir Filko. A systematic overview of recent methods for non-contact chronic wound analysis. *Applied Sciences*, 10(21):7613, 2020. 1

[14] Brayan Monroy, Jorge Bacca, Karen Sanchez, Henry Arguello, and Sergio Castillo. Two-step deep learning framework for chronic wounds detection and segmentation: A case study in colombia. In *2021 XXIII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, pages 1–6. IEEE, 2021. 2

[15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. 4

[16] Silonie Sachdeva et al. Fitzpatrick skin typing: Applications in dermatology. *Indian journal of dermatology, venereology and leprology*, 75(1):93, 2009. 2

[17] Karen Sanchez, Carlos Hinojosa, Henry Arguello, Denis Kouamé, Olivier Meyrignac, and Adrian Basarab. Cx-dagan: Domain adaptation for pneumonia diagnosis on a small chest x-ray dataset. *IEEE Transactions on Medical Imaging*, 2022. 1

[18] Héctor Serrano-Coll, Hugo Rene Mora, Juan Camilo Beltrán, Malcolm S Duthie, and Nora Cardona-Castro. Social and environmental conditions related to mycobacterium leprae infection in children and adolescents from three leprosy endemic regions of colombia. *BMC infectious diseases*, 19(1):1–10, 2019. 1

[19] Nanne Van Noord and Eric Postma. Learning scale-variant and scale-invariant features for deep image classification. *Pattern Recognition*, 61:583–592, 2017. 2

[20] Chuanbo Wang, DM Anisuzzaman, Victor Williamson, Mrinal Kanti Dhar, Behrouz Rostami, Jeffrey Niezgoda, Sandeep Gopalakrishnan, and Zeyun Yu. Fully automatic wound segmentation with deep convolutional neural networks. *Scientific Reports*, 10(1):1–9, 2020. 4

[21] Changhan Wang, Xinchen Yan, Max Smith, Kanika Kochhar, Marcie Rubin, Stephen M Warren, James Wrobel, and Honglak Lee. A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks. In *2015 37th annual international conference of the ieee engineering in medicine and biology society (EMBC)*, pages 2415–2418. IEEE, 2015. 4

[22] Hazem Wannous, Yves Lucas, Sylvie Treuillet, Alamin Mansouri, and Yvon Voisin. Improving color correction across camera and illumination changes by contextual sample selection. *Journal of Electronic Imaging*, 21(2):023015, 2012. 2