Automatic Information Extraction from Neurological Assessment Records for Leprosy Cases Using AI

Jamilly Braga Melo[1], **Anthony Alexandre Morais Militão**[2], **Patricia Takako Endo**[2], **Hilson Gomes Vilar de Andrade**[1,2]

[1]Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco (IFPE) Recife – PE – Brazil

[2]Programa de Pós-Graduação em Engenharia da Computação (PPGEC), Universidade de Pernambuco (UPE), Caruaru – PE – Brazil

jbm2@discente.ifpe.edu.br, aamm2@ecomp.poli.br,

{patricia.endo, hilson.andrade}@upe.br

Abstract. *Leprosy is a chronic infectious disease that continues to be present in more than 120 countries. In 2023, over 20,000 new cases were reported in Brazil, making it the second most endemic country in the world. The development of new information systems that utilize available clinical records is essential to support decision-making during the treatment of this highly stigmatizing disease. In this study, we present preliminary results demonstrating the feasibility of using YOLO for the automatic recognition of non-textual data from neurological assessments of patients undergoing treatment. Our approach achieved an accuracy of 97.5% in recognizing sensory assessment records used in the Brazilian healthcare system.*

1. Introduction

Leprosy is an infectious disease caused by the bacterium *Mycobacterium leprae* and is one of the most common causes of non-traumatic peripheral neuropathy worldwide [WHO 2023]. The clinical consequences of nerve damage include sensory changes that can lead to repeated trauma and skin ulcers, vascular and glandular alterations, as well as muscle weakness that can result in physical deformities [Slacel 2000].

Despite advancements in molecular biology and serological techniques, the diagnosis of leprosy remains primarily clinical. A thorough assessment, including a careful examination of skin lesions and peripheral nerves, will, in most cases, be sufficient for the diagnostic definition of leprosy. In Brazil, the country responsible for 92% of new leprosy cases in the Americas in 2023 [Organization et al.], the Simplified Neurological Assessment (from Portuguese, *Avaliação Neurológica Simplificada*, ANS) is a mandatory clinical examination that aims to monitor the patient's neural function, checking for autonomic changes, impaired sensitivity or decreased muscle strength as a result of neural damage resulting from leprosy [Secretaria de Vigilância em Saúde 2022]. Currently, the information obtained from this assessment is recorded manually and on paper forms, stored locally in the health units.

Current literature employing Artificial Intelligence (AI) has primarily focused on diagnosing the condition by analyzing image of skin lesions [De Souza et al. 2021,

Steyve et al. 2022, Yotsu et al. 2023, Beesetty et al. 2023]. However, we did not identify any prior studies employing AI to facilitate the digitization of existing clinical data in health units, such as the sensitive assessments available in the ANS.

This work presents preliminary results on YOLO (You Only Look Once) [Redmon 2016] performance for automatic recognition of non-textual data present in the ANS form. Facilitating the extraction of non-textual information from neurological assessment forms of patients with leprosy will aid in data collection and the development of new information systems to support decision-making in leprosy treatment.

2. Materials and Method

2.1. ANS form

The ANS form includes sociodemographic and clinical data of patients, encompassing the assessment of the face, upper limbs, strength evaluation, nerve palpation and the inspection and sensory evaluation of the hands and feet. Semmes-Weinstein monofilaments (SWM) are employed to evaluate and monitor tactile sensation in specific territories of the nerve trunks of the hands and feet [Frade et al. 2022]. The standard esthesiometer kit recommended by Brazil's Ministry of Health consists of six nylon monofilaments, each 38 mm in length and varying in diameter, designed to exert specific forces ranging from 0.07 to 300 gram-force (gf) [Guimarães et al. 2013].

Figure 1 shows the fields from the ANS where tactile sensitivity is recorded during the examination conducted using the SWM. Each specific point, marked with a circle, is individually assessed. The examination begins with the contact of the smallest diameter filament (green - 0.07 gf) and progresses to the largest diameter filament (pink - 300 gf) if the patient does not report sensitivity to the touch of the filaments, totaling six different filaments (and colors). In cases of severe neural impairment, where the patient does not indicate sensitivity to the touch of the largest diameter filament, the black color is used for marking. Unlike the other fields in the ANS form, which utilize tabular data, the records related to sensory assessment are presented in a graphical format, making it challenging to convert this information into text that can be interpreted by computerized systems.

The ANS form is composed of many other fields, but the image showed in Figure 1 is our focus point for model learning. The objective of our model is to detect and classify the colored markings within this section, which correspond to neurological sensitivity assessments. By analyzing these markings, the model aims to automate the identification of evaluated points, enabling the digitization of data.

2.2. Dataset

We utilized 262 images (130 hands and 132 feet) obtained from scanned ANS forms collected at a Hospital Otávio de Freitas, Recife, Brazil. To create our labeled dataset, we followed the steps presented in Figure 2, annotating only the areas containing color markings, as they are our target.

To enhance the accuracy of the YOLO model in detecting sensitivity test points, a preprocessing script was developed to isolate the regions corresponding to the hands and feet in the images used for the sensitivity test. The low image quality, combined with the small size of the points of interest, poses a significant challenge for detection, leading to

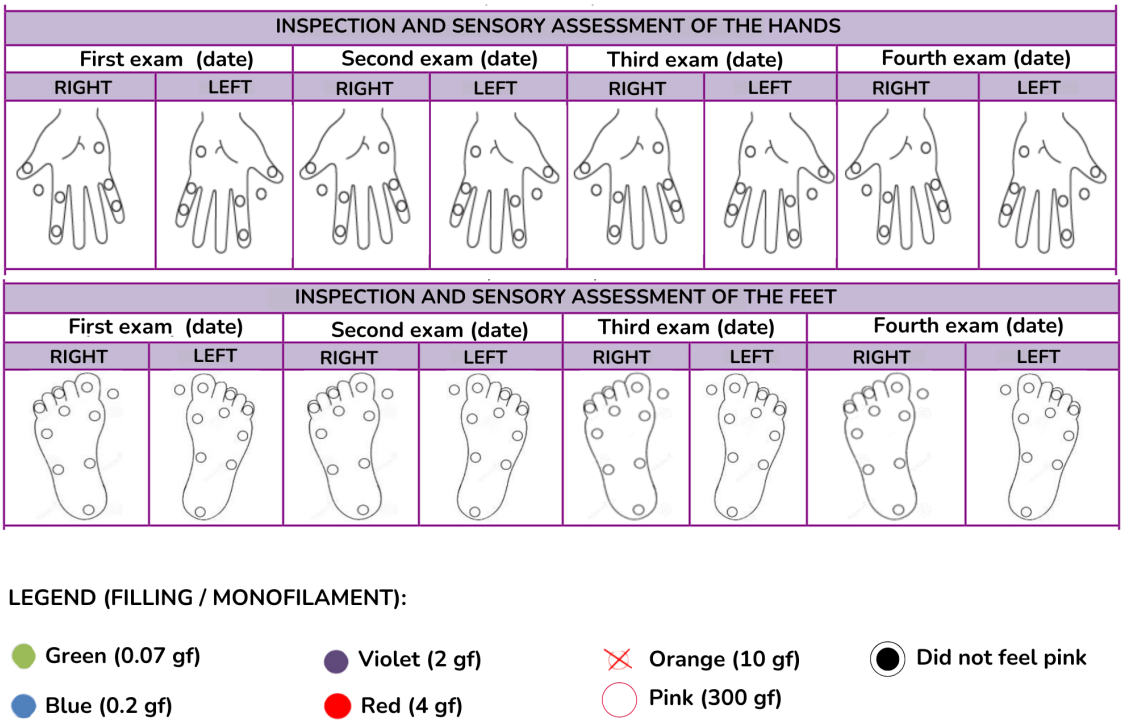

Figura 1. Table for recording the sensitive assessment in the ANS form

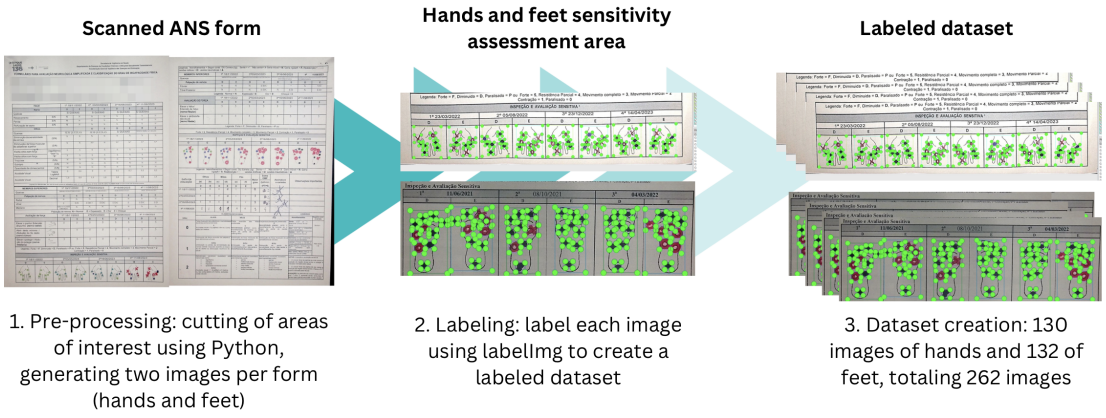

Figura 2. Process of creating the labeled dataset from ANS forms.

low confidence scores. To address this issue, the script segments these regions based on predefined coordinates, generating two separate images (hands and feet). This procedure enables the YOLO model to analyze the images more precisely, optimizing the training process and improving detection results.

To ensure effective training and minimize the risk of overfitting, the dataset was split into 70% for training and 30% for testing, allowing the model to generalize properly to new samples.

2.3. Model inference

The model was developed utilizing YOLOv8[1] - a computer vision model architecture for detection, segmentation, pose, estimation, tracking and classification, based on AI - with its primary objective being the detection of two specific categories of bounding boxes: (a) hand or foot bounding boxes, which delineate the overall area encompassing the detected hand or foot within the image, and (b) color bounding boxes, smaller regions within the hand or foot area that correspond to colored zones, each representing a distinct level of neurological sensitivity, see Figure 3.

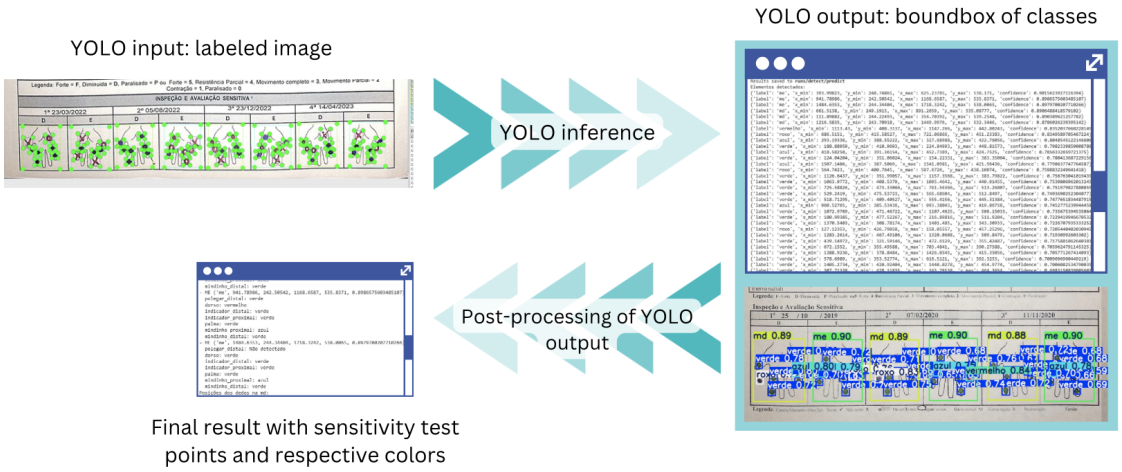

Figura 3. Model inference process

Thus, the output of the YOLO will consist of the inference of a multiclass classification, where each class corresponds to the identification of the location of the point where sensitivity was tested (right hand, left hand, right foot, or left foot) and the respective color (green, blue, purple, red, orange, pink, or black). The hyperparameters were empirically fine-tuned to optimize the model's performance, with training conducted over 100 epochs. Additionally, thresholds were set at 0.05 for bounding box loss, improving object localization accuracy, and 0.5 for classification loss, ensuring reliable categorization of detected instances.

As the model's output summarizes the results by reporting only the detected classes and the number of occurrences for each class, a post-processing step was incorporated to identify the location of each class occurrence. This step aimed to generate an output that maps the location of each sensitive point, as detailed in the following section.

2.4. Mapping of the positions of sensitivity points

Based on the model outputs (color bounding box, and hands or feet bounding box), we developed an algorithm to process these detections hierarchically, verifying which color bounding boxes were contained within the larger bounding boxes corresponding to the hands or feet.

This ordering was performed by considering the y-coordinate to identify elements from top to bottom and the x-coordinate to differentiate elements appearing side by side.

[1] https://yolov8.com

The position of colors within the hand and foot regions was associated with predefined anatomical points:

- **Hands:** The color correspondence was established for seven anatomical points: distal phalanx of the thumb, dorsal hand, distal phalanx of the second finger, proximal phalanx of the second finger, medial border of the palm, proximal phalanx of the fifth finger, and distal phalanx of the fifth finger.
- **Feet:** The mapping of the feet followed a more detailed structure, considering both plantar and dorsal anatomy. The defined points included: distal phalanx of the fifth toe, fifth toe, lateral border of the plantar surface, calcaneus, medial border of the plantar surface, third toe, distal phalanx of the third toe, head of the metatarsophalangeal joint of the hallux, distal phalanx of the hallux, dorsal surface of the foot

To ensure the correct identification of anatomical points, a mirroring criterion was adopted between the left and right sides. The left foot, for example, was treated as a mirrored image of the right foot, allowing the ordering logic to be applied symmetrically to both limbs. With this approach, the system was able to generate a structured output associating each detected color with a specific anatomical point, enabling a detailed analysis of the distribution of neurological sensitivity areas in the hands and feet.

2.5. Evaluation metrics

The main metrics used for evaluation are precision, recall, and Mean Average Precision (mAP) at various Intersection over Union (IoU) thresholds.

Precision assesses the accuracy of the model's predictions by indicating the proportion of true positive detections among all positive detections. Recall, in contrast, evaluates the model's ability to identify all relevant instances, reflecting how effectively it avoids missing positive cases. The mAP serves as a comprehensive metric that accounts for the trade-off between precision and recall across different IoU thresholds. It provides an overall evaluation of the model's performance across various classes and thresholds. In this analysis, mAP is computed at the standard IoU threshold of 0.5 (mAP50) and over a range of thresholds from 0.5 to 0.95 (mAP50-95).

3. Preliminary Results

According Table 1, the overall performance of the model across all classes combined is quite promising. It achieved a precision of 97.5% and a recall of 96.5%, indicating that the proposed model can accurately identify the color associated with tactile sensitivity point indicated on the ANS form. The mAP at the standard IoU threshold of 0.5 is 98.1%, which is considered a strong performance, and the more comprehensive mAP across IoU thresholds from 0.5 to 0.95 is 68.1%. Examining the performance of individual classes reveals variations in the model's capability to detect different markings, though all classes demonstrate strong results across the evaluated metrics.

As illustrated in Figure 4, the F1-confidence curve highlights the model's promising performance in detecting our points of interest. Reaching an F1-score of 0.97 at a confidence threshold of 0.271 reflects an effective balance between precision and recall for identifying markings across all classes. The curve indicates that the model can accurately detect points while minimizing false positives, making it well-suited for extracting information from sensitive assessments in ANS forms.

Classes	Precision	Recall	mAP50	mAP50-95
all	0.975	0.965	0.981	0.681
green	0.963	0.944	0.962	0.527
blue	0.974	0.912	0.945	0.573
purple	0.977	0.993	0.989	0.632
red	0.949	0.934	0.965	0.594
orange	0.936	0.974	0.984	0.631
pink	0.961	0.997	0.994	0.635
black	0.986	0.994	0.994	0.599
right hand	0.993	0.979	0.985	0.844
left hand	0.981	0.979	0.982	0.846
right foot	1.000	0.972	0.993	0.791
left food	1.000	0.943	0.995	0.821

Tabela 1. YOLO performance based on evaluation metrics

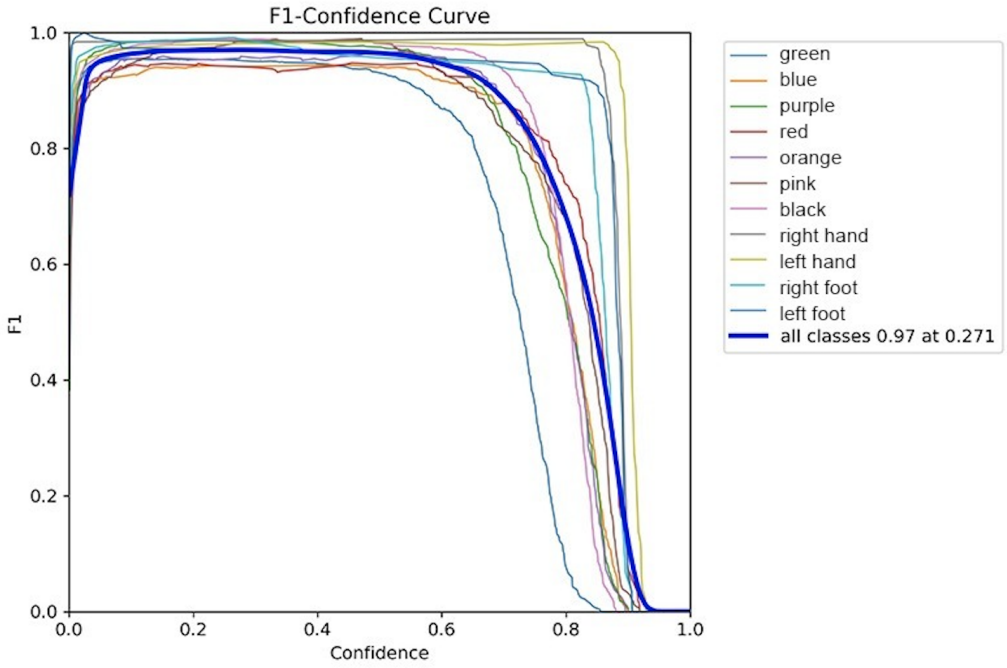

Figura 4. F1-confidence curve of the proposed model

4. Conclusions and Future Works

This study successfully developed a methodology for detecting and mapping neurological sensitivity points in the hands and feet using YOLOv8. The approach efficiently associates detected colors with predefined anatomical points, enabling a structured analysis. Future work will focus on evaluating the performance of other models in comparison to YOLOv8 and developing a software system to automate data extraction from assessment forms. This system will integrate computer vision techniques to streamline data collection and analysis. Such automation will enhance accuracy, reduce manual processing time, and facilitate large-scale studies on neurological sensitivity patterns, enabling data-driven analysis and the discovery of new patterns.

Referências

Beesetty, R., Reddy, S., Modali, S., Sunkara, G., Dalal, J., Damagathla, J., Banerjee, D., and Venkatachalapathy, M. (2023). Leprosy skin lesion detection: An ai approach using few shot learning in a small clinical dataset. *Indian J Lepr*, 95:89–102.

De Souza, M. L. M., Lopes, G. A., Branco, A. C., Fairley, J. K., and Fraga, L. A. D. O. (2021). Leprosy screening based on artificial intelligence: Development of a cross-platform app. *JMIR mHealth and uHealth*, 9(4):e23718.

Frade, M. A. C., Bernardes Filho, F., Silva, C. M. L., Voltan, G., Lima, F. R., Abi-Rached, T. L. C., and de Paula, N. A. (2022). Evaluation of altered patterns of tactile sensation in the diagnosis and monitoring of leprosy using the semmes-weinstein monofilaments. *Plos one*, 17(8):e0272151.

Guimarães, L. d. S. et al. (2013). Incapacidade física em pessoas afetadas pela hanseníase: estudo após alta medicamentosa.

Organization, W. H. et al. Global leprosy (hansen disease) update, 2023: Elimination of leprosy disease is possible–time to act!. 2024.

Redmon, J. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Secretaria de Vigilância em Saúde, D. d. D. d. C. C. e. I. S. T. (2022). Protocolo clínico e diretrizes terapêuticas da hanseníase.

Slacel, M. (2000). The diagnosis of leprosy among patients with symptoms of peripheral neuropathy without cutaneous lesions. *Arq neuropsiquiatr*, 58:800–807.

Steyve, N., Steve, P., Ghislain, M., Ndjakomo, S., et al. (2022). Optimized real-time diagnosis of neglected tropical diseases by automatic recognition of skin lesions. *Informatics in Medicine Unlocked*, 33:101078.

WHO (2023). Leprosy: Key facts.

Yotsu, R. R., Ding, Z., Hamm, J., and Blanton, R. E. (2023). Deep learning for ai-based diagnosis of skin-related neglected tropical diseases: A pilot study. *PLOS Neglected Tropical Diseases*, 17(8):e0011230.