Use of the Smartphone Camera to Monitor Adherence to Inhaled Therapy

Sofia Ferraz[1], Rute Almeida[2], Pedro Vieira-Marques[2], Nuno Escudeiro[3]

[1] INESC TEC - Institute for Systems and Computer Engineering, Technology and Science, Porto, Portugal
sofia.a.ferraz@inesctec.pt

[2] CINTESIS - Center for Health Technology and Services Research, Faculty of Medicine, University of Porto, Porto, Portugal
{rutealmeida, pmarques}@med.up.pt

[3] Department of Computer Science, Polytechnic of Porto - School of Engineering, Porto, Portugal
nfe@isep.ipp.pt

Abstract. *Adherence to inhaled controller medications is crucial for patients with chronic respiratory illnesses to achieve favorable clinical outcomes. Self-management measures have been shown to enhance health outcomes, decrease unnecessary interventions, and improve disease control. However, compliance evaluations have difficulties in establishing a high level of trustworthiness, as patient's self-reported high compliance rates are frequently regarded unreliable. A mobile application module to objectively verify inhalation usage using image snapshots of the inhalation counter and optic character recognition has shown to be promising, but insufficient for some inhaler models. In this paper a model specific approach was explored to enable reliable adherence measurement. To achieve this, a machine learning model was trained on an inhaler image dataset. The trained model had an average accuracy of 88% in recognizing the digits on the dose counter of an inhaler model. These results show the potential to gain additional evidence for inhaler compliance.*

Keywords. medication adherence, mHealth, remote monitoring, optical character recognition, inhaled therapy

University of Maribor Press

DOI https://doi.org/10.18690/um.3.2022.26
ISBN 978-961-286-658-7

1 Introduction

Asthma affects about 300 million people globally and accounts for 1 in every 250 deaths. In Europe alone, approximately 30 million people have asthma and 15,000 people die yearly from this disease [1]. Although asthma exacerbations occurrences can be reduced with appropriate regular therapy and patient education, treatment adherence is generally low among patients with asthma. As a matter of fact, some studies show that adherence is less than 50% in children and as low as 30% in adults [2]. This low adherence may be due in part to misinformation or confusion regarding complicated treatment regimens. Poor medication adherence is concerning, since it is known to increase risk of asthma exacerbations, leading to higher mortality, greater financial burden for the patient and health system, as well as decreased quality of life [2]. Numerous adherence-improvement interventions have been introduced, but most have been only moderately successful with little evidence of long-term sustainability or reduction of health care utilization and cost [3].

Mobile Health (mHealth) technologies can improve disease outcomes and may be an especially powerful tool to deliver effective behavioral health interventions that are dynamic, user-centric, and continuously adapted [4]. Medication-use monitoring can provide important information for patients, researchers, and health professionals, with the aim of facilitating improved adherence and of improving treatment prescribing. Patient self-report and clinician assessments of medication adherence are notoriously unreliable [5].

Regarding inhaled medication, current mHealth applications require the user to manually enter the readings from the dose counters of these medical devices. This process is slow and prone to error. As the internet becomes more embedded into medical monitors through Wi-Fi and Bluetooth technologies, more sophisticated systems transmit the values from the connected devices to the smartphone. However, this adds costs to the manufacturing of the device and brings connectivity issues [6]. People who cannot afford to upgrade to these expensive devices will fail to receive the benefits [7].

In the United States, smartphones are owned and regularly carried by approximately 50% of 12–17 year-olds and 75% of adults ages 30–49 [5]. The advantages of smartphones over other devices is not only the fact that they are affordable, but also that they are very powerful, with most models nowadays integrating several cores in their main processor. They are also standalone devices with a camera, a battery, and audio output and an Internet connection [10]. Therefore, these devices show high potential to be explored as a relevant mHealth tool.

2 Methods

Optical character recognition (OCR) is a tool that converts scanned images of typewritten or hand-written text into machine-readable text [8]. Despite recent technology advancements, the available OCR approaches still present several limitations (e.g. the dependency on the quality of input images), and are still not able to compete with human reading capabilities with desired accuracy levels [7][9].

Previous research [10] used a mobile application module to objectively verify inhalation usage through image snapshots of the inhaler counter and OCR. Although this research demonstrated encouraging results in the use of OCR to address non-adherence to inhaled medicine, the results indicate that the approach used is less effective for particular inhaler models (*e.g. Seretaide* and *Twisthaler*). As such, we hope to build on prior work by developing a machine learning algorithm trained on an image database of the *Seretaide* model and optimize the OCR performance for this specific inhaler. To accomplish the proposed objectives, annotated and pre-processed images of *Seretaide* inhaler devices were used to train and validate a machine learning model.

Dataset. The database consists of 354 images of the *Seretaide* inhaler model in PNG format, with dimensions of 640 x 360 pixels, collected under ideal conditions using a LG-V700 (Android) camera app. These images show a wide representation of digits in the dose counter; additionally, the background varies between black, white, and multi-color, and the lighting source of the photo varies between natural light and artificial light. For the purpose of training a machine learning model on inhaler pictures, a manual annotation process was performed on all images present in the dataset. This task was accomplished with the help of the VGG Image Annotator (VIA) tool [11].

With the aid of this software, it was possible to collect annotations regarding the position and dimensions of the dose counter and the corresponding digits that appear on the display.

Image Pre-Processing. All images suffered four pre-processing steps: rotating, cropping, convert to grayscale and resizing. The images in the dataset were rotated 90º right, since the photos were acquired horizontally, and cropped according to the dimensions of the dose counter indicated on the annotations file. Since the cropping measurements were not consistent across the dataset, the cropped-out images did not have the same size; hence, all images were resized to standardize the data.

Model Architecture. With an OCR approach in mind, a neural network (NN) consisting of convolutional layers (CNN) to extract a sequence of features and recurrent layers (RNN) to propagate information through this sequence, was developed. Additionally, it instantiates a new "endpoint layer" for implementing CTC loss. The former enables using unsegmented pairs of images and corresponding text transcriptions to train the model without any character/frame-level alignment [12]. More specifically, the architecture of the NN consists of an input layer, two convolutional layers each followed by a pooling layer, two bidirectional layers, a CTC layer and finally an output layer (Figure 1).

Model Training. In later stages, a training and validation dataset were generated. For this purpose, the dataset was shuffled randomly, so that each time the dataset is split a new training and validation dataset are created. Additionally, the dataset split into 90 % training set (318 images) and 10% validation set (36 images). This project runs using the Google Colab environment and the network was built using TensorFlow 2. 6.. The model was trained in 200 epochs and a batch size of 5.

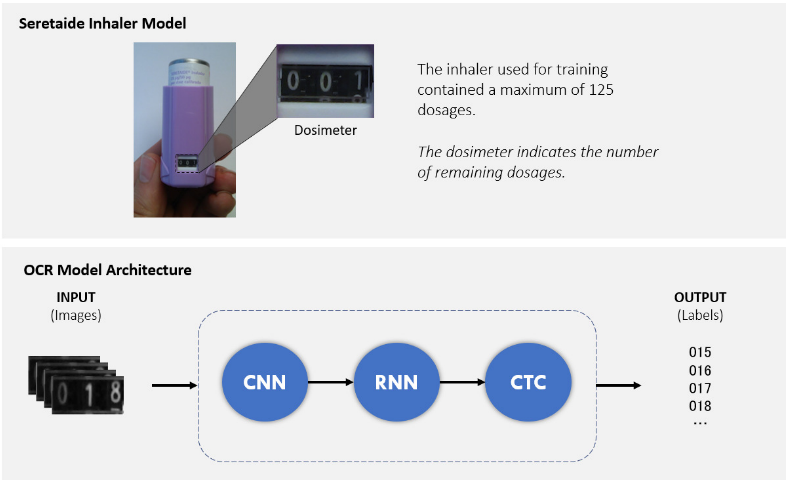

Figure 1. Information regarding the inhaler used in the dataset, along side a photo of a *Seretaide* inhaler (above); Schematics of the optical character recognition model's architecture (bellow).

3 Results

To evaluate the model, several metrics were taken into consideration. The findings of this examination are summarized in Table 1, which includes the average values for the metric following five tests, as well as the standard deviation.

Table 1. Evaluation Metrics for the Trained Model.

Metrics	Average (%)	Standard Deviation
Accuracy	87,2	0,04
Exact Match Ratio	87,2	0,04
Hamming Loss	0,13	0,04
Recall	88,3	0,06
Precision	87,2	0,04
F1-measure	87,1	0,04

4 Discussion

The exact match ratio can be considered a challenging metric since it doesn't support the notion of being partially correct.

As it can be seen in Table 1, the exact match ratio is 87,2%, which indicates that a large part of the predicted results were entirely correct, and consequently reflects a good model performance.

The Hamming Loss considers the incorrect label predictions and the relevant labels not predicted, over the total number of labels. In this case, the computed hamming loss is

0,13 %, which is a significantly low value and indicates a good performance of the learning algorithm.

For this model, the calculated recall was 88,3 % (Table 1), which indicates that a large number of the actual labels were predicted. On the other hand, precision is the ratio of how much of the predicted is correct, i.e., it only considers the positive predicted results. In this case, the precision equals the exact match ratio (87,2 %).

Furthermore, the F1 measure is the harmonic mean of Precision and Recall and gives a better measure of the incorrectly classified cases than the Accuracy Metric. The F1 measure reaches 87,1%, which is an indication of both good precision and good recall.

The purpose of this effort was to enhance the *Seretaide* model's OCR performance. In comparison to prior study, the model's accuracy has increased significantly (from 38% to 87%). It is important to highlight, however, that the datasets utilized to evaluate these methodologies are not equal, and hence a direct comparison cannot be drawn. Nevertheless, the advancements in this research might be seen as an indication of significant improvement.

5 Conclusions and Future Work

The purpose of this paper was to develop a text recognizer for an inhaler model (*Seretaide*). This was done by building a machine learning model, trained on a database of inhaler images, compatible with mobile applications.

To the best of our knowledge, there are not many approaches in the literature that help to reduce the unreliability of patient compliance and self-reporting by making use of OCR to record effective dosage in inhaler dose counters. Furthermore, the proposed work explores the potential of a customized OCR approach to enhance the performance of already existing algorithms, thus making this work relevant to help mitigating the patient's unreliable, self-reported adherence.

Nevertheless, further improvements are still needed to enhance the detection performance. It will be critical in future work to compare the findings of this model to those of the general model in the previous research, on the same collection of photos (independent of the image set used to train the model). Additionally, this model could be applied to additional inhaler models in order to assess the impact of employing customized machine learning models on other inhalers, thus, determining if performance is maintained. Furthermore, an object detector-like algorithm can be implemented to detect the dose counter, thus avoiding the cropping stage in the image pre-processing.

Acknowledgments

This research was part of the project "mINSPIRE—mHealth to measure and improve adherence to medication in chronic obstructive respiratory diseases - generalization and evaluation of gamification, peer support and advanced image processing technologies", POCI-01-0145-FEDER-029130, supported by the ERDF (European Regional Development Fund), COMPETE2020 (Programa Operacional Competitividade e Internacionalização), Portugal

2020 and by Portuguese Funds through FCT (Fundação para a Ciência e a Tecnologia).

We thank PaCeIT for providing the *Seretaide* image dataset, without which this work would not have been feasible.

References

[1] J. E. Fergeson, S. S. Patel, and R. F. Lockey, "Acute asthma, prognosis, and treatment," *Journal of Allergy and Clinical Immunology*, vol. 139, no. 2, pp. 438–447, 2017, ISSN: 10976825. DOI: 10.1016/j.jaci.2016.06.054. [Online]. Available: http://dx.doi.org/10.1016/j.jaci.2016.06.054.

[2] R. Jeminiwa, L. Hohmann, J. Qian, K. Garza, R. Hansen, and B. I. Fox, "Impact of eHealth on medication adherence among patients with asthma: A systematic review and meta-analysis," *Respiratory Medicine*, vol. 149, no. February, pp. 59–68, 2019, ISSN: 15323064. DOI: 10.1016/j.rmed.2019.02.011. [Online]. Available: https://doi.org/10.1016/j.rmed.2019.02.011.

[3] B. G. Bender, "Technology Interventions for Nonadherence: New Approaches to an Old Problem," *Journal of Allergy and Clinical Immunology: In Practice*, vol. 6, no. 3, pp. 794–800, 2018, ISSN: 22132198. DOI: 10.1016/j.jaip.2017.10.029. [Online]. Available: https://doi.org/10.1016/j.jaip.2017.10.029.

[4] D. A. Fedele, A. McConville, J. Graham Thomas, *et al.*, "Applying Interactive Mobile health to Asthma Care in Teens (AIM2ACT): Development and design of a randomized controlled trial," *Contemporary Clinical Trials*, vol. 64, no. March, pp. 230–237, 2018, ISSN: 15592030. DOI: 10.1016/j.cct.2017.09.007. [Online]. Available: http://dx.doi.org/10.1016/j.cct.2017.09.007.

[5] A. H. Y. Chan, H. K. Reddel, A. Apter, M. Eakin, K. Riekert, and J. M. Foster, "Adherence Monitoring and E-Health: How Clinicians and Researchers Can Use Technology to Promote Inhaler Adherence for Asthma," *Journal of Allergy and Clinical Immunology: In Practice*, vol. 1, no. 5, pp. 446–454, 2013, ISSN: 22132198. DOI: 10.1016/j.jaip.2013.06.015. [Online]. Available: http://dx.doi.org/10.1016/j.jaip.2013.06.015.

[6] E. Finnegan, M. Villarroel, C. Velardo, and L. Tarassenko, "Automated method for detecting and reading seven-segment digits from images of blood glucose metres and blood pressure monitors," *Journal of Medical Engineering and Technology*, vol. 43, no. 6, pp. 341–355, 2019, ISSN: 1464522X. DOI: 10.1080/03091902.2019.1673844. [Online]. Available: https://doi.org/10.1080/03091902.2019.1673844.

[7] D. Tsiktsiris, K. Kechagias, M. Dasygenis, and P. Angelidis, "Accelerated Seven Segment Optical Character Recognition Algorithm," *5th Panhellenic Conference on Electronics and Telecommunications, PACET 2019*, pp. 1–5, 2019. DOI: 10.1109/PACET48583.2019.8956283.

[8] K. Kanagarathinam and K. Sekar, "Text detection and recognition in raw image dataset of seven segment digital energy meter display," *Energy Reports*, vol. 5, pp. 842–852, 2019, ISSN: 23524847. DOI: 10.1016/j.egyr.2019.07.004.

[9] A. Chaudhuri, K. Mandaviya, P. Badelia, and S. K. Ghosh, *Optical character recognition systems*. 2017, vol. 352, pp. 9–41, ISBN: 9783319502526. DOI: 10.1007/978-3-319-50252-6_2.

[10] P. Vieira-Marques, R. Almeida, J. F. Teixeira, *et al.*, "InspirerMundi-Remote Monitoring of Inhaled Medication Adherence through Objective Verification Based on Combined Image Processing Techniques," *Methods of Information in Medicine*, 2021, ISSN: 2511705X. DOI: 10.1055/s-0041-1726277.

[11] A. Dutta and A. Zisserman, "The VIA annotation software for images, audio and video," in *Proceedings of the 27th ACM International Conference on Multimedia*, ser. MM '19, Nice, France: ACM, 2019, ISBN: 978-1-4503-6889-6/19/10. DOI: 10.1145/3343031.3350535. [Online]. Available: https://doi.org/10.1145/3343031.3350535.

[12] M. Yousef, K. F. Hussain, and U. S. Mohammed, "Accurate, data-efficient, unconstrained text recognition with convolutional neural networks," *Pattern Recognition*, vol. 108, p. 107 482, Dec. 2020, ISSN: 00313203. DOI: 10.1016/j.patcog.2020.107482. arXiv: 1812.11894.