Self-supervised Methods for Ugly Duckling Detection in Wide Field Images

Vullnet Useini[*1,2] VUSEINI@ETHZ.CH
Nicolaus Andratschke[2] NICOLAUS.ANDRATSCHKE@USZ.CH
Stephanie Tanadini-Lang[2] STEPHANIE.TANADINI-LANG@USZ.CH
Quentin Lohmeyer[1] QLOHMEYER@ETHZ.CH
Ralph P. Braun[3] RALPH.BRAUN@USZ.CH
Javier Barranco García[*2] JAVIER.BARRANCOGARCIA@USZ.CH
[1] *Department of Mechanical and Process Engineering, ETH Zurich, Zurich, Switzerland.*
[2] *Department of Radiation Oncology, USZ and UZH, Rämistrasse 100, 8091, Zurich, Switzerland.*
[3] *Department of Dermatology, USZ and UZH, Rämistrasse 100, 8091, Zurich, Switzerland.*

Editors: Under Review for MIDL 2022

Abstract

Screening skin lesions is a very time-consuming process in which the dermatologist examines hundreds of lesions all over the patient's body in a limited period of time. The decision as to which lesions should be further examined is made based on the "ugly duckling" sign. The dermatologist compares all lesions on the same patient and identifies those that are different from the average-looking lesions. Deep learning algorithms have been shown to be efficient tools for detecting outliers in large image datasets. In this study, we propose a self-supervised approach for lesion clustering and outlier detection to identify and suggest lesions of interest for each individual patient.

Keywords: Self-supervised, melanoma, ugly duckling, outlier

1. Introduction

During a routine dermatology consultation, a dermatologist sometimes has to screen hundreds of lesions in an average time frame of 20 minutes. Experienced dermatologists would identify those of interest for further consideration leaving around 2 minutes for a final decision per lesion. In recent studies (Soenksen et al., 2021), deep learning models have been proposed to automate the detection and ranking of "ugly ducklings" using real wide field images. However, one limitation of this approach is the fact that the embedding is performed based on the features extracted from a classifier trained on previously labeled data. Labeling data in a clinical context is very time consuming, and in many cases there is no clear consensus on the ground truth. To avoid such limitations, we propose the use of recent self-supervised algorithms that learn from unlabeled data and rank the different instances according to their similarity to the average-looking lesion from the same patient.

[*] Contributed equally

2. Methods

An end-to-end process that includes image acquisition, lesion detection, and self-supervised clustering is proposed next.

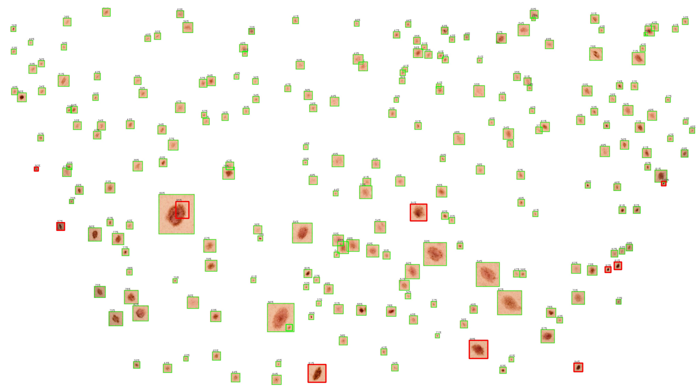

Figure 1: Real space distribution of the lesions in the back of the patient as detected by YOLOR. The top 10 highest scoring lesions are shown with red bounding boxes.

2.1. Dataset

The USZ Dermatology Clinic has acquired FotoFinder© (FotoFinder) devices, which semi-automatically capture high-quality polarized whole-body images. Data from a total of 91 patients have been acquired at the USZ dermatology clinic to date. For the purpose of this study, we limited ourselves to the assessment of lesions in the dorsal region of patients.

2.2. Lesion Detection

The methodology for lesion detection consists of two steps. In the first step, we semi-automatically label all lesions on a few selected patient images with the help of a blob detector and then train a state-of-the-art object detection algorithm, namely YOLOR (Wang et al., 2021), in a supervised manner. This model is then used to detect and extract lesions from unseen patients.

2.3. Self-supervised Learning

Diagnosis in dermatology depends heavily on the context of the patient. The "ugly duckling" sign compares how similar a particular lesion is to other lesions in the same patient. To avoid bias from using pre-labeled images of other patients, we propose to apply self-supervised learning to our real-world dataset. We will rely on the DINO libraries (Caron et al., 2021) for this purpose. In our case, ResNet18 was used as the backbone for DINO. The final clustering is done by calculating a weighted and normalized distance composed of a normalized cosine distance between the embeddings of all lesions and their median embedding, and a normalized euclidean distance between the average RGB values of the whole image and the minimum RGB values of all lesions. The latter helps to remove unwanted outlier lesions such as pimples or red spots. This distance is then proposed as the ugly duckling score for the lesions.