

AUTOMATIC GRADING OF ACNE VULGARIS USING DEEP LEARNING

Melina A.¹, Ngo Dinh N.², Tafuri B.³, De Vitis S.⁴, Schipani G.⁵, Nisticò S.⁵, Cosentino C.¹, Amato F.¹, and Cherubini A.^{2,3}

¹Faculty of Biomedical Engineering, University "Magna Graecia", Viale Europa, Catanzaro, Italy

²Linkverse, Rome, Italy

³Institute of Molecular Bioimaging and Physiology (CNR-IBFM), Catanzaro, Italy

⁴Numens, Catanzaro, Italy

⁵Department of Dermatology, University "Magna Graecia", Viale Europa, Catanzaro, Italy

Abstract

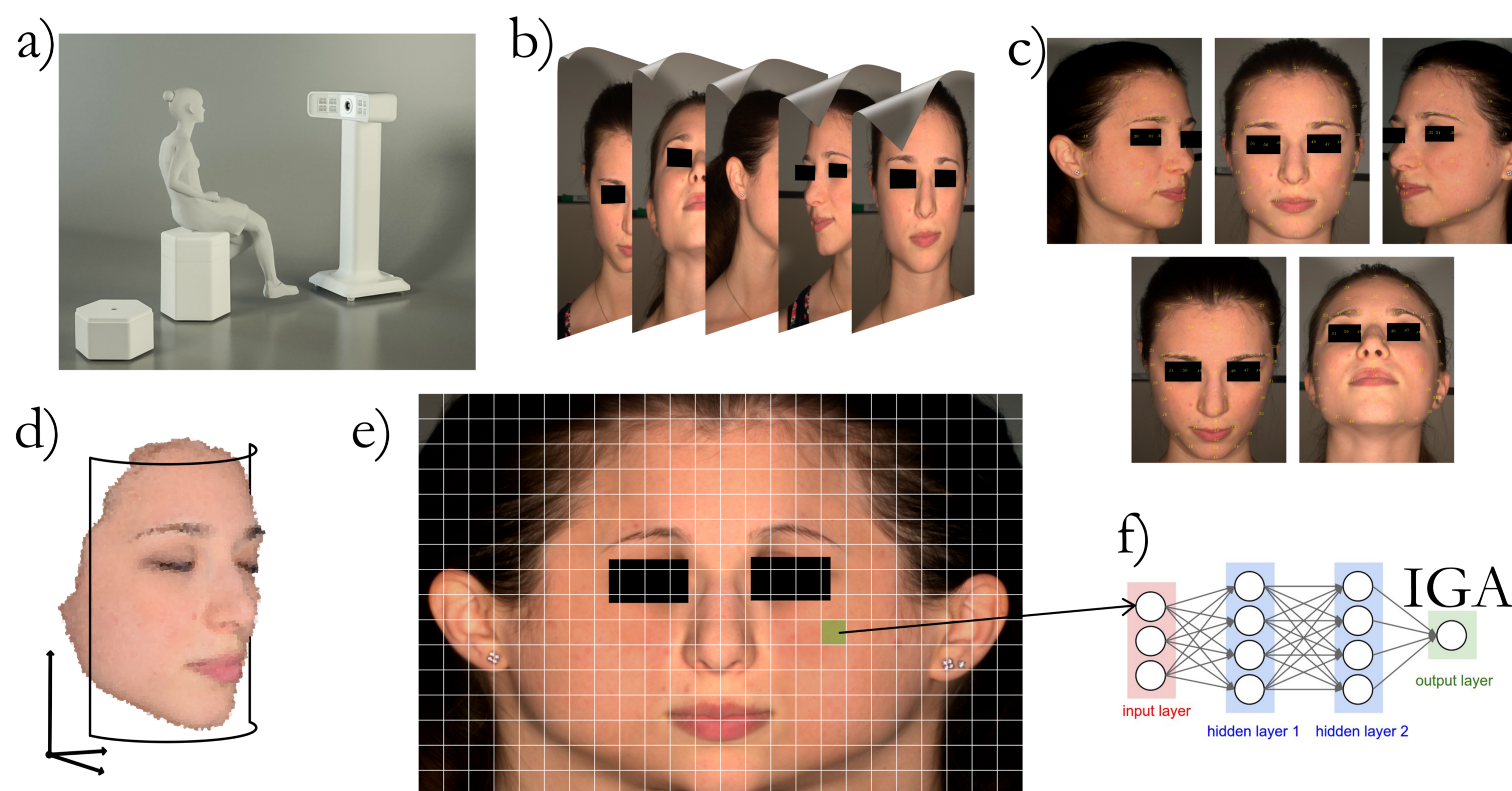
Introduction. The evaluation of Acne grading using ordinal scales reflect the clinical perception of severity but have shown low reproducibility both intra- and inter-rater. In this study, we investigated if Artificial Intelligence trained on images of Acne patients could perform acne grading with high accuracy and reliabilities superior to those of expert physicians.

Methods. 479 patients with acne grading ranging from clear to severe and sampled from three ethnic groups participated in this study. Multi-polarization images of facial skin of each patient were acquired from five different angles using the visible spectrum. An Artificial Intelligence was trained using the acquired images to output automatically a measure of Acne severity in the 0-4 numerical range.

Results. The Artificial Intelligence recognized the Investigator Global Assessment of a patient with an accuracy of 0.854 and a correlation between manual and automatized evaluation of $r=0.958$ ($p<.001$).

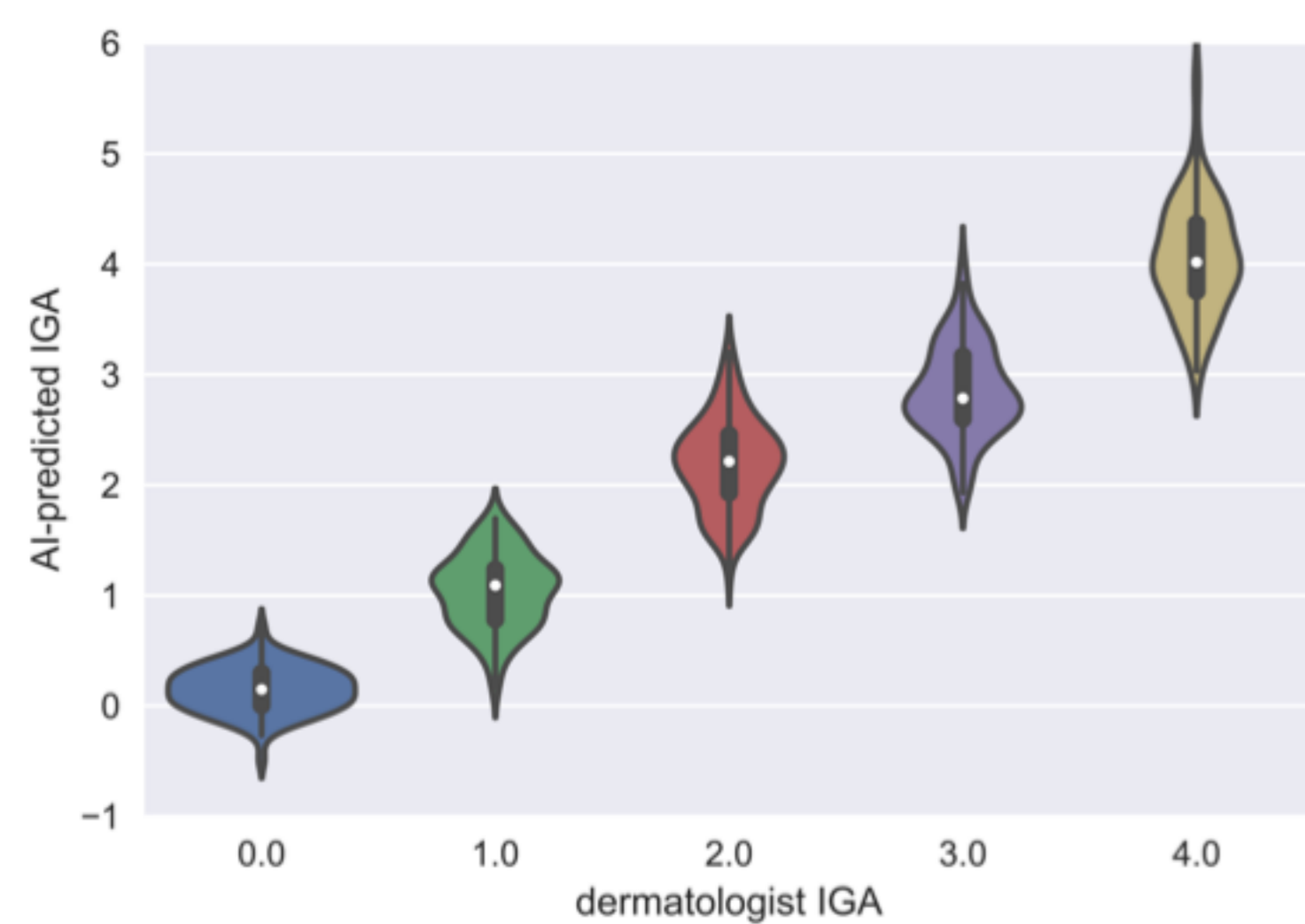
Discussion. This is the first work where an Artificial Intelligence was able to directly classify acne patients according to an Investigator Global Assessment ordinal scale with high accuracy, no human intervention and no need to count lesions.

Materials and Methods



The subjects, seated in front of the device (a), were asked to rotate their head in five different poses (b) with a 30° tilt respect to the lens axis. For each pose, lateral flash bulbs were lighted to collect a first reference image and a second image with cross-polarized filters to reduce reflections in the images. Anatomical landmarks (c) were identified on each view and subsequently used to reconstruct the three-dimensional shape (d) of the individual face. The skin of the patient was then projected upon a cylindrical surface in order to generate a "flattened" image (e) showing the entire facial skin in a single bi-dimensional facial template. In this planar reconstruction, the same spatial region (green squared patch) on different faces corresponded to the same anatomical region on all subjects. The planar images were converted into a hemoglobin/melanin color space, and from each patch and each color channel, several statistical measures were calculated: mean, standard deviation, skewness, kurtosis, and local binary pattern. Thus, at the end of this process, each patient image was converted into a series of about ten thousand ordered parameters that conveyed all information regarding anatomical skin pattern and texture useful for acne grade classification. These numbers were then used to train the deep neural network for automatic acne grading (f).

Results and Discussion



The distribution of patients according to the IGA expressed by the dermatologists (horizontal axis) versus the IGA evaluated by artificial intelligence (vertical axis). The AI showed an accuracy of 0.85 in predicting the IGA based only on images of the patient (Spearman $r=0.96$, $p<.001$). In this work we obtained a remarkable result for three reasons. In fact, this is the first time that an AI was able to directly classify acne patients according to an IGA ordinal scale with no human intervention and no need to count lesions. Second, the accuracy observed is comparable to those reported for skilled physicians. Third, the algorithm was tested on a large dataset comprising subjects belonging to three different ethnic groups, suggesting that the observed results have a high generalizability. Future work will focus on further increasing the accuracy obtained in this work by expanding the dataset and introducing images acquired with multiple devices.

References

- U.S. Department of Health and Human Services Food and Drug Administration. Guidance for Industry Acne Vulgaris: Developing Drugs for Treatment. Center for Drug Evaluation and Research (CDER), 2005.
- Lucky AW, Barber BL, Girman CJ et al. A multirater validation study to assess the reliability of acne lesion counting. Journal of the American Academy of Dermatology, Volume 35, Issue 4, 559 – 565, 1996.
- Di Huang, Caifeng Shan, Mohsen Ardebilian, Yunhong Wang, and Liming Chen. Local Binary Patterns and Its Application to Facial Image Analysis: A Survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C; Volume: 41, Issue: 6, Nov. 2011.
- Zhao Liu, Josiane Zerubia. Melanin and Hemoglobin Identification for Skin Disease Analysis. Asian Conference on Pattern Recognition (ACPR), Nov 2013, Okinawa, Japan. IEEE, 2013.