Skin Cancer Classification using Soft Attention and Metadata Concatenance

Hoang Khoi Do

khoi.dh200322@sis.hust.edu.vn

Ha Noi University of Science and Technology

March 17, 2022

Abstract

Recent estimates are that about 150 million children under five years of age are stunted, with substantial negative consequences for their schooling, cognitive skills, health, and economic productivity. Therefore, understanding what determines such growth retardation is significant for designing public policies that aim to address this issue. We build a model for nutritional choices and health with reference-dependent preferences. Parents care about the health of their children relative to some reference population. In our empirical model, we use height as the health outcome that parents target. Reference height is an equilibrium object determined by earlier cohorts' parents' nutritional choices in the same village. We explore the exogenous variation in reference height produced by a protein-supplementation experiment in Guatemala to estimate our model's parameters. We use our model to decompose the impact of the protein intervention on height into price and reference-point effects. We find that the changes in reference points account for 65% of the height difference between two-year-old children in experimental and control villages in the sixth annual cohort born after the initiation of the intervention.

Keywords: AI-enabled computer-aid diagnosis, Diagnosis, Skin Sancer, Skin Lesion Classification, Artificial Intelligence, Deep Learning, Machine Learning

1   Introduction

Skin cancer is one of the most common cancer leading causes of death worldwide. Every day, more than 9500[1] people in the United States are diagnosed with skin cancer. Otherwise, 3.6[1] million people are diagnosed with basal cell skin cancer each year. According to the Skin Cancer Foundation, the global incidence of skin cancer continues to increase[2]. In 2019, it is estimated that 192,310 cases of melanoma will be diagnosed in the United States[2]. On the other hand, if patients are early diagnosed, the survival rate is correlated with 99%. However, once disease progresses beyond the skin, survival is poor[2]. Moreover, with the increasing incidence of skin cancers, low awareness among a growing population, and a lack of adequate clinical expertise and services, there is a need of effective solution.

Recently, deep learning particularly, and machine learning in generally algorithms have emerged to achieve excellent performance on various tasks, especially in skin disease diagnosis tasks. AI-enabled computer-aided diagnostics (CAD) has solutions in three main categories: Diagnosis, Prognosis and Medical Treatment. Medical imaging, including ultrasound, computed tomography, and magnetic resonance imaging, and X-ray image is used extensively in clinical practice. In Diagnosis, Artificial Intelligence (AI) algorithms are applied for disease detection to save progress execution before these diagnosed results are considered by a doctor. In Prognosis, AI algorithms are used to predict the survival rate of a patient based on his/her history medical data. In Medical Treatment, AI models are applied for building solution for a specific disease, medicine revolutionize is an example. In various studies, AI algorithms has provided various end-to-end solutions in the detection of abnormalities such as breast cancer, brain tumors, lung

cancer, esophageal cancer, skin lesions, and foot ulcers across multiple image modalities of medical imaging[2].

In order to adapt the increase in skin cancer case, AI algorithms over the last decade has a great performance. Some typical models that can be mentioned are DenseNet[3], EfficientNet[4], Inception[5], MobileNets[6][7][8], ResNet[9][10], VGG[11], and Xception[12]. Some of these models have been used as a backbone model in other studies that I will discuss more in the Related Work section.

2 Related Work

3 Objective

In this paper, my objectives are:

- Building a model that has ability to classify in a balanced way between classes instead of specifying almost classes except one or two classes.

- Investigate the performance of the combination of different backbone model and SoftAttention.

- Identifying whether the metadata can make the performance of model better.

4 Implementation

4.1 Data

The dataset used in this paper is the HAM10000 dataset published by Havard University Dataverse[13]. There are total 7 classes in this dataset containing Actinic keratoses and intraepithelial carcinoma or Bowen's disease (AKIEC), Basal cell Carcinoma (BCC), benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, BKL), dermatofibroma (DF), melanoma (MEL), melanocytic nevi (NV), and vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, VASC). The distribution of the dataset is shown in the table below:

Class	AKIEC	BCC	BKL	DF	MEL	NV	VASC	Total
No. Sample	327	514	1099	115	1113	6705	142	10015

More than 50 percent of lesions are confirmed through histopathology (HISTO), the ground truth for the rest of the cases is either follow-up examination (FOLLOWUP), expert consensus (CONSENSUS), or confirmation by in-vivo confocal microscopy (CONFOCAL).

4.2 Model Schema

4.2.1 Input Schema

4.2.2 Backbone Model

4.2.3 SoftAttention

4.2.4 Output Schema

4.2.5 Model

4.3 Training

4.3.1 Data Division

4.3.2 Loss Function

The loss function used in this paper is categorical cross-entropy. Consider $X = [x_1, x_2, \ldots, x_n]$ as the input feature, $W = [w_1, w_2, \ldots, w_n]$. Let n is the number of class in the dataset. The categorical cross-entropy loss is presented as:

$$L(w, x_i) = -\sum_{i=1}^{n} y_i \log(\hat{y}_i)$$

where $\hat{y}_i$ is the output of model and y_i is the target that the model should return.

Since the dataset face the imbalanced problem then I applied the class weight for the loss. This formula below is used to calculate the class weight:

$$CLassWeight = N \odot D$$

3

$$D = [\frac{1}{C \times N_1} \frac{1}{C \times N_2} \cdots \frac{1}{C \times N_n}]$$

where N is the number of sample in dataset, C is the number of class in the dataset, N_i is the number of sample in each class i. D is the matrix contain the inverse of $C \times N_i$.

4.3.3 Optimizer

5 Results

6 Conclusion

[14]

References

[1] Soumyya Kanti Datta, Mohammad Abuzar Shaikh, Sargur N. Srihari, and Mingchen Gao. Soft-attention improves skin cancer classification performance. Available at https://arxiv.org/abs/2105.03358, 4 Jun 2021.

[2] Manu Goyal, Thomas Knackstedt, Shaofeng Yan, and Saeed Hassanpour. Artificial intelligence-based image classification for diagnosis of skin cancer: Challenges and opportunities. Available at https://arxiv.org/abs/1911.11872, 20 Jun 2020.

[3] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Densely connected convolutional networks. Available at https://arxiv.org/abs/1608.06993, 28 Jan 2018.

[4] Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. Available at https://arxiv.org/abs/1905.11946, 11 Sep 2020.

[5] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. Available at https://arxiv.org/abs/1512.00567, 11 Dec 2015.

[6] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. Available at https://arxiv.org/abs/1704.04861, 17 Apr 2017.

[7] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. Available at https://arxiv.org/abs/1801.04381, 13 Jan 2018.

[8] Andrew Howard, Mark Sandler, Grace Chu, Liang-Chieh Chen, Bo Chen, Mingxing Tan, Weijun Wang, Yukun Zhu, Ruoming Pang, Vijay Vasudevan, Quoc V. Le, and Hartwig Adam. Searching for mobilenetv3. Available at https://arxiv.org/abs/1905.02244, 20 Nov 2019.

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. Available at https://arxiv.org/abs/1512.03385, 10 Dec 2015.

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. Available at https://arxiv.org/abs/1603.05027, 25 Jul 2016.

[11] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. Available at https://arxiv.org/abs/1409.1556, 10 Apr 2015.

[12] François Chollet. Xception: Deep learning with depthwise separable convolutions. Available at https://arxiv.org/abs/1610.02357, 4 Apr 2017.

[13] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Available at https://arxiv.org/abs/1803.10417, 25 Nov 2018.

[14] Katherine M. Li and Evelyn C. Li. Skin lesion analysis towards melanoma detection via end-to-end deep learning of convolutional neural networks. Available at https://arxiv.org/abs/1807.08332, 22 June 2018.