An auxiliary tool for preliminary tests of skin cancer

A self-modifying meta-learning method for clean and noisy data

Yifeng Cao †
Automation
North China University of Technology
Beijing, China
immorta1s@126.com

Yuefan Wu †
Gifted Young
University of Science and Technology of China
Hefei, China
wyf8899@mail.ustc.edu.cn

Zhenyu Tian*, †
Industrial Technology
XianJiaotong-liverpool University
Suzhou, China
zhenyu.tian18@student.xjtlu.edu.cn

Xuan Yu †
Computer Science
East China Normal University
Shanghai, China
10185102157@stu.ecnu.edu.cn

† These authors contributed equally.

Abstract—**Deep learning is a popular method when it comes to disease detection problems. As for skin cancer, a rather common kind of disease, however, unexpected rare cases often occur with few written records and referential materials, resulting in a disadvantageous situation for a usual neural network to learn. Hence, J in this paper, we propose a self-modifying meta-learning model which combines the idea of meta-learning with curriculum learning. Applying this mechanism, our model will first train on data of common diseases and then adapt the model to rare disease classification. Moreover, despite the existence of natural noise in data, like manually mistaken labels, our model can still handle which. We evaluate our algorithm on ISIC 2018 skin lesion classification dataset. Employing only 5 samples from each class, we achieve our accuracy up to 79.2%. Apart from that, when predicting data with a 20% noisy rate, our model can also adapt to classify unseen classes by accuracy of 76.2%. The further utilization for our model can be limited to skin cancer detection and diagnosis and extend to be applied to all kinds of diseases, serving as helpful assistants for medical workers, which would be a win-win for both patients and doctors.**

Keywords-Skin-cancer; Deep Learning; Meta Learning; Noise Robustness

I. INTRODUCTION

Statistics show that cancer has become the second leading cause of death, responsible for approximately 1 in 6 deaths every year [1]. Skin cancer, however, has become the most common cancer worldwide. There are 5.4 million new cases of basal cell carcinoma and squamous cell carcinoma in the United States every year [2], including 76380 cases of malignant melanoma. While in Australia, cases worsen, considering nearly two-thirds may be diagnosed as skin cancer before the age of 70 [3].

Even though it is not that frequent among persons with skin color, their illnesses are often associated with greater morbidity and mortality. Unlike our common sense, ultraviolet radiation is not an important etiologic factor for it, except for basal cell carcinoma. Thus, skin cancer can still pose significant risks in the skin of color [4]. Scary though it may sound, if early screening and popularization of skin cancer-related knowledge are carried out, the mortality rate caused by skin cancer will be greatly reduced. A screening study conducted in a German state from 2003 to 2004 shows that through the popularization of skin cancer knowledge, the mortality rate of melanoma in this area can be reduced by 48% by 2009[5]. Also, if detected early, the 5-year survival rate for melanoma is 99 percent [1].

However, many problems exist between a person falling ill and getting appropriate medications and other professional treatments in a standard hospital. Most people won't be aware of small changes in their body in the first place, especially in skin cancer cases. Let alone some cancerous are easily misleading. Under circumstances of early stages of melanoma diagnosis, particularly when the diameter of the pigmentary lesions is below 6 millimeters, clinical diagnosis criteria can fail to offer credible help. This is when misjudgements occur: people will mistakenly take melanoma as ordinary moles [6]. Without professional knowledge and means of detections related to this field, one can have difficulty making judgements about abnormality. In the next stage, if people feel skeptical about themselves, they may turn to the Internet for help. According to the NBC-news data, people have this tendency to search the information about diseases on the Internet, of which 63% look for information about specific diseases or medical problems, and 47% look for answers about specific medical treatment or procedures [7]. Hence it won't be too surprising to find out that about 35% of adults in the United States use the Internet to make diagnoses by themselves [8]. On the contrary, adults around 50 years old fail to believe that the Internet can provide correct diagnosis and treatment information [9]. At the same time, they are more likely to believe the suggestions and solutions provided by doctors [10, 11]. Even if the accuracy of information provided by anonymousness is not to blame, another study clarifies communication barriers, which will cause an individual difference in recognition and misunderstandings while searching for related information on the Internet [12-15]. In addition, without professional

978-1-6654-2709-8/21/$31.00 ©2021 IEEE
DOI 10.1109/ICBASE53849.2021.00040

knowledge, one can have a hard time considering which disease do his symptoms fit. Especially considering when some serious diseases share the same symptoms with the mild ones, leading to people's panic and unnecessary detours to the right diagnosis. The final barrier would possibly be tedious and expensive examinations one needs to go through, which may discourage some families from giving up treatment. Skin cancer has become the fifth most expensive cancer in the United States medical insurance, and it costs more than 8 billion dollars to treat which every year [16]. Lack of in-depth researches about skin cancer and proper medical conditions may be some objective factors contributing to the high diagnosis rate and mortality.

There is a growing trend that researchers try to combine advanced technology like Artificial Intelligence with medical issues [17]. Existing websites like the Skin Cancer Foundation can just offer self-detection tips to assist early discovery. In addition, some papers about skin cancer detection have been published. [18] uses a deep CNN model, trained on nearly 130 thousand clinical images to classify them into over 2000 categories. A massive attempt like this may fail to emphasize small details within a specific disease and requires huge devotion of time and energy to collect and tag the dataset. [19] performed well with pre-trained deep learning and transfer learning models to distinguish three different types of skin lesion, which may not make it to pathology requirements. While in [20], recent technology such as feature matching comes into use. With a method based on Generative Adversarial Networks (GANs) to train a segmentation model, they achieve a better performance than traditional adversarial training methods. Other previous works present excellent results when only a small amount of training datasets are available with meta learning. Methods like MAML [21] and Meta-Weight-Net [22] prove their advanced performances and high-efficiency on classification tasks of few shots.

However, some existing deficiencies of these methods stop them from being used in a large scope. Thus, to make actual easily accessible and available classification systems, we develop a model that can classify common skin cancer and detect some rare skin diseases or even find new ones that have not been discovered yet. In these cases, people can have greater chance to accept proper disease-detection and treatments services in time. Few shot is a kind of method that only requires a small amount of data to train, thus eliminating the cost for data collection and labeling, further reducing calculation amount conspicuously [23]. Here, we propose a self-modifying meta-learning model which can have great performance in skin cancer detection. It can also deal with the learning gap between different kinds of samples with self-modification.

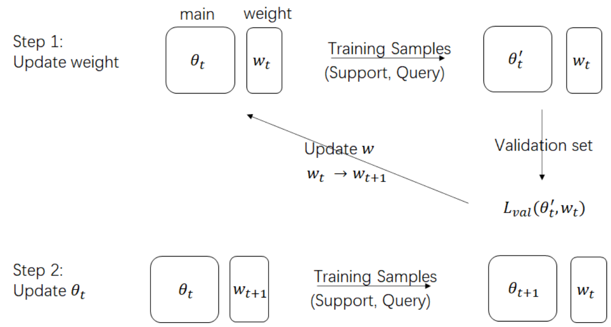

Fig. 1: The pipeline of our proposed self-modifying meta-learning system.

II. Methods

This section aims to provide a new meta-learning framework that can generate excellent results on rare skin cancer cases within a few steps of gradient descent, using only a small number of pictures of common diseases as meta-training data. Fig 1 illustrates the procedure of our system. We have a major model with its parameter θ and a weighted network with parameter w. First, sample some images from the dataset as a buffer. Then sample few images from buffer to train few-shot learning, on which the model will be trained, and value θ' will be available. With this newly acquired parameter, the validation set can apply the same method before calculating losses. In this way, the parameters in the weighted network can be updated based on the back propagation of its loss. θ However, the major model can then be updated with both the newly generated weighted network parameter w and its original value. Detailed steps are identical to what was illustrated before when calculating the temporary weight θ'.

A. Data Collection

We employ the ISIC 2018 Skin Lesion Analysis Towards Melanoma Detection Dataset [21,24], which has a total of 10,015 skin lesion images from seven skin diseases, including melanocytic nevus (6705), melanoma (1113), benign keratosis (1099), basal cell carcinoma (514), actinic keratosis (327), vascular lesion (142) and dermatofibroma (115). We simulate the categorizing problem by utilizing the three classes with the largest amount of cases as common diseases (i.e., meta-train dataset D_{tr}), two classes with a modest number of cases as a validation set D_{val}, and the least two classes as rare diseases (i.e., meta-test dataset D_{te}).

Task instance $\mathcal{T}_i$ is randomly sampled from a distribution over tasks $p(\mathcal{T})$ and $D_{tr}, D_{val}, D_{te} \in p(\mathcal{T})$. During the meta-train stage, learning tasks $\mathcal{T}_i$ are binary classification tasks, and each task consists of two random classes with k samples per class in D_{tr}. The same is true for the validation set D_{val}. During the meta-test stage, each test task instance is sampled from D_{te}.

Meanwhile, we also add noise to the dataset to test the robustness of our model and its ability to deal with the existing noise in real datasets. To widen the gap between different batches and make the noise have a greater influence, there is a

50% chance that the batch will be added noise. Then in each batch, every piece of data has the possibility of p(i.e., noisy rate) to be added noise by randomly choosing a label.

B. Algorithms

Current model-agnostic meta-learning algorithms only apply one-order optimization during the learning process, which may contribute to the instability of the model and the great variance of the results. Additionally, with simply one-order optimization, the model may not learn that effectively from data. Our algorithm achieves a harmonious combination of supervised learning and meta-learning, presenting a two-order optimization meta-learning. This means we obtain a stable and satisfying result and make excellent contributions to handling datasets with noise. The core of our algorithm is described in Algorithm 1.

After splitting the dataset into D_{tr} D_{te}, we obtain a model parameter θ based on some learning tasks D_{tr}. First, we randomly sample $\mathcal{T}_i$ instances according to a probability distribution $p(\mathcal{T})$, which can help us get the updated parameter θ' with the cross-entropy we acquired before serving as a binary classification task that contains k samples in each random-selected class. Then take θ' and apply our model to learning tasks sampled from D_{val}, getting the validation loss L_{val} that can be used to update parameters w in our weighted network. Finally, renew the original model's parameter θ with the new weighted network.

After updating weight, data has been soundly explored by our model; thus, the model can comprehend exactly which piece of data needs to be put extra attention on. To be more specific, this self-modifying weighted model can automatically tell which batches to emphasize on and which don't work according to the losses it generated. Through this mechanism, our model has quite good performances on not only clean but also noisy datasets.

The algorithm of Self-Modified Meta Learning Framework

Algorithm 1: Self-Modified Meta Learning Framework

Initialize the main model parameter θ
Initialize weight net parameter w
for *session* $= 1$ *to* M **do**
  Sample several data and put into meta buffer
  Randomly select k samples from meta buffer as meta train data
  Add noise to train data is necessary
  Train main model , update parameter $\theta_t' \rightarrow \theta_t + \eta\nabla\theta_t$
  Sample k_{val} pieces of data as validation set the way train data selected
  Calculate loss on validation set, L_{val}
  Update weight net with L_{val}, $w_{t+1} \rightarrow w_t + \eta\nabla w_t$
  Update θ with new $w_{t+1}, \theta_{t+1} \rightarrow \theta_t + \eta\nabla\theta_t$
end

III. Results & Discussion

We employ accuracy on the test dataset as our criterion to tell whether a model performs well or not. The results indicate that we can only train our model for at most 10 thousand iterations, or it will lead to overfitting. Hence, each figure only contains results up to 10 thousand iterations. First, we make comparisons with other methods that train on clean ISIC 2018 datasets. Then we will demonstrate the robustness of our method on data with different noise levels.

A. Results on Clean Data

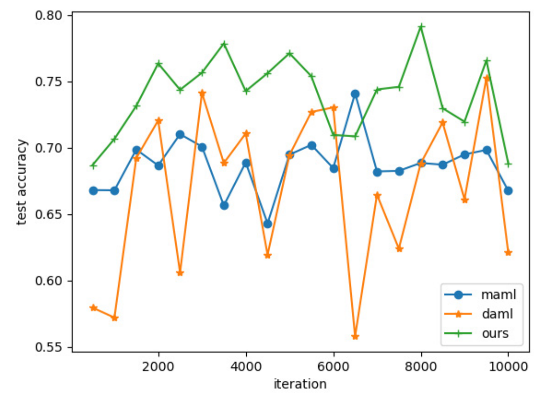

Fig. 2: Accuracy on the test set every 500 iterations

As can be seen from the results shown in Fig. 2, our method gets the best accuracy and achieves a satisfying accuracy within 2000 iterations, which outperforms other models a lot.

Table 1: Results on clean data

	Acc
MAML	74.1 $\pm$ 0.9%
DAML	75.2 $\pm$ 1.0%
Ours	**79.1 $\pm$ 0.8%**

B. The result on Noisy Data

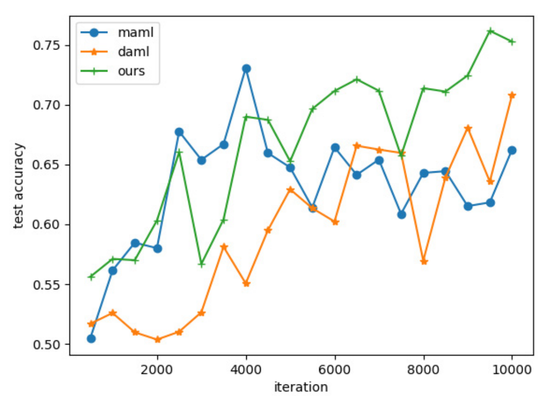

Fig 3: Accuracy on the test set with 0.4 noise rate

First, each batch will have a probability of 50% to decide whether it will get some noise. Second, according to the noisy rate we set, the meta training data from the support set will randomly choose a label from the label set. Results are shown in Fig 3 and Fig 4. The accuracy of our method is much higher than two baselines with several iterations. If the noise rate is 0.4, the expectation of noisy data rate in the train data set is 0.2, which means that it will have about 20% randomly chosen labels on average.

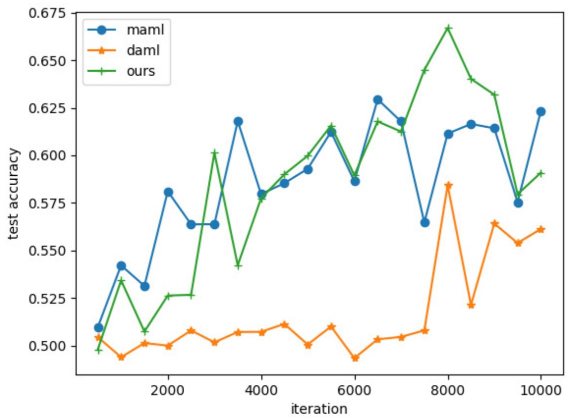

Fig 4: Accuracy on the test set with 0.8 noise rate

Table 2: Results of noisy data

	Noisy Rate	Acc
MAML	0.4	73.1±1.3%
DAML	0.4	70.8±1.3%
Ours	0.4	**76.2±1.5%**
MAML	0.8	62.9±1.3%
DAML	0.8	58.4±1.1%
Ours	0.8	**66.7±1.5%**

C. Discussion

The model we raise in this paper applies meta-learning to CNN and the method based on which can be used in early-stage skin cancer preliminary tests. Though other researches involve adopting CNN to skin cancer or other kinds of diseases, they need a huge amount of data for training to support their algorithm. By contrast, the few shot methods we adopt can free the model from massive datasets with better applicability and less time and achieve excellent results on rare or even undiscovered diseases. Apart from that, we run tests on our method's robustness. By adding noise, an unavoidable factor, to different degrees to the dataset, we imitate possible manual labeling faults in real life to test our method's response. The results illustrate that our model maintains fine performances under different circumstances.

In addition to the virtues mentioned before about our model, our research emphasizes preliminary tests of skin cancer. Most of the former researches are discussions about methods and models but fail to consider the patients' and hospitals' sides. We regard preprocessing as a crucial point of our work, making it more easily available and convenient to use. Our model can offer a tool to help ordinary users have some basic knowledge and recognition about their current situations while saving a lot of time and energy. As for hospitals, doctors won't bother to answer some simple questions that our model is fully capable of, saving time for patients with more urgent cases. This will not only increase the efficiency of the hospital but also relieve the doctors' pressure.

However, there still exist some flaws for us to make further improvements. Though our model can achieve excellent results with few clean or noisy data, considering the choice of few-shot learning, we inevitably encounter the problem of insufficient training. Compared to traditional deep learning models, the new one we raise can't acquire adequate information on given datasets, while the accuracy of diagnosis is inferior. Suppose we are to carry out some further study. In that case, we will maintain its generalization as much as possible, then increase the accuracy by multiple-dataset training and modifying and improving our model to better carry out the task of preliminary diagnosis. At the same time, we will employ more medical-related materials of multi-aspects, such as use images of cells or pathological sections, to serve as auxiliary means.

IV. Conclusion

This paper presents a convolutional neural network to make preliminary diagnoses on skin cancer by categorizing them to their belonging classifications. Normal CNN model requires datasets with huge amounts of images to train intended for specific diseases. Hence it will fail to give satisfying results in unseen situations. Moreover, due to the lack of an adversarial noise mechanism, it can't predict correctly when natural noises occur. After tests and adjustments, we propose a self-modifying meta-learning neural network, which provides excellent results in low-data regime problems and achieves excellent performances. This model is capable of helping medical practitioners to carry out routine and basic examinations more easily, like dermatoscopy and melanoma classification. We also contribute to future research that may be carried out on skin cancer with an effective classification method and that the diseases they study have rare cases and insufficient data and records. In addition, experiments show that our method has great robustness when dealing with truly existing noise on images. Compared to other skin cancer classification work, our method achieves more practicability. Compared to other skin cancer classification work, our method achieves more generalization, which means it can perform quite well in unseen datasets using only a few iterations. We admit that this clinical pre-diagnosis only focuses on the aspect of skin cancer, still having a lot of work to do if it is truly to be applied to real work categorizing tasks. For example, we will train it on more data and improve the model structure intended for higher accuracy on unseen data. In our future research, we hope to implement our model on a wider range, serving as an auxiliary tool for later stages of skin cancer, even intend for other diseases. In addition, as for our method, preprocessed data with super-resolution can help our model learn more detailed information. Considering the performances are imperfect, ensemble learning would make a fine combination with our model, and we believe it will increase our model's ability to make more accurate classifications.

References

[1] Skin Cancer Facts & Statistics, skin cancer foundation, https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/

[2] Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012. JAMA Dermatol. 2015;151(10):1081–1086.

[3] Information from https://www.cancercouncil.com.au/skin-cancer/

[4] Gloster Jr H M, Neal K. Skin cancer in skin of color[J]. Journal of the American Academy of Dermatology, 2006, 55(5): 741-760.

[5] Systematic skin cancer screening in Northern Germany. Breitbart EW, Waldmann A, Nolte S, Capellaro M, Greinert R, Volkmer B, Katalinic AJ Am Acad Dermatol. 2012 Feb; 66(2):201-11.

[6] V. de Giorgi, I. Savarese, S. Rossari et al., “Features of small melanocytic lesions: does small mean benign? A clinical-dermoscopic study,” Melanoma Research, vol. 22, no. 3, pp. 252–256, 2012.

[7] Information from https://www.nbcnews.com/id/wbna3077086

[8] Fox S, Duggan M. Pew Internet & American Life Project. 2013. [2014-12-22]. webcite Health Online 2013 http://www.pewinternet.org/files/old-media//Files/Reports/PIP_HealthOnline.pdf.

[9] Hirji F. Freedom or Folly? Canadians and the Consumption of Online Health Information. Information, Communication & Society. 2004 Jan;7(4):445–465. doi: 10.1080/1369118042000305593.

[10] Zulman DM, Kirch M, Zheng K, An LC. Trust in the internet as a health resource among older adults: analysis of data from a nationally representative survey. J Med Internet Res. 2011;13(1): e19. doi: 10.2196/jmir.1552. http://www.jmir.org/2011/1/e19/

[11] Price EL, Bereknyei S, Kuby A, Levinson W, Braddock CH. New elements for informed decision making: a qualitative study of older adults' views. Patient Educ Couns. 2012 Mar;86(3):335–41. doi: 10.1016/j.pec.2011.06.006. http://europepmc.org/abstract/MED/21757315.

[12] Robertson N, Polonsky M, McQuilken L. Are my symptoms serious Dr Google? A resource-based typology of value co-destruction in online self-diagnosis. Australasian Marketing Journal. 2014;22(3):246–256. doi: 10.1016/j.ausmj.2014.08.009.

[13] Begum S, Hasan M, Sharmeen A, Sultana S. Self-Medication of Anti-microbial Agents (AMAs) and Over-the-Counter Practices: A Study in Mymensingh Sadar Area. Comm Based Med J. 2013 Mar 17;2(1):15–20. doi: 10.3329/cbmj.v2i1.14176.

[14] Lau F, Price M, Keshavjee K. From benefits evaluation to clinical adoption: Making sense of health information system success in Canada. hcq. 2011 Jan 27;14(1):39–45. doi: 10.12927/hcq.2011.22157.

[15] Lewis N, Gray SW, Freres DR, Hornik RC. Examining cross-source engagement with cancer-related information and its impact on doctor-patient relations. Health Commun. 2009 Dec;24(8):723–34. doi: 10.1080/10410230903264030. http://europepmc.org/abstract/MED/20183381.

[16] Surveillance, Epidemiology, and End Results Program, National Cancer Institute. [Accessed July 14,2016]; SEER stat fact sheets: melanoma of the skin. http://seer.cancer.gov/statfacts/html/melan.html.

[17] Guoguang Rong, Arnaldo Mendez, Elie Bou Assi, Bo Zhao, Mohamad Sawan,Artificial Intelligence in Healthcare: Review and Prediction Case Studies, Engineering,Volume 6, Issue 3, 2020, Pages 291-301, ISSN 2095-8099,

[18] Esteva, A., Kuprel, B., Novoa, R. et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115–118 (2017)

[19] K. M. Hosny, M. A. Kassem and M. M. Foaud, "Skin Cancer Classification using Deep Learning and Transfer Learning," 2018 9th Cairo International Biomedical Engineering Conference (CIBEC), Cairo, Egypt, 2018, pp. 90-93, doi: 10.1109/CIBEC.2018.8641762.

[20] Mondal, A.K., Dolz, J., Desrosiers, C.: Few-shot 3d multi-modal medical image segmentation using generative adversarial learning. arXiv preprint arXiv:1810.12241 (2018)

[21] Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks, Chelsea Finn, Pieter Abbeel, Sergey Levine, ICML 2017

[22] Meta-Weight-Net: Learning an Explicit Mapping For Sample Weighting, Jun Shu, Qi Xie, Lixuan Yi, Qian Zhao, Sanping Zhou, Zongben Xu, Deyu Meng, NeurIPS 2019

[23] Information from https://research.aimultiple.com/few-shot-learning/

[24] Information from https://canceratlas.cancer.org/the-burden/europe/