WASEDA University
Graduate School of Information, Production and Systems

Skin Lesion Classification Using Weakly-supervised Fine-grained Method

Present. Xue Xi

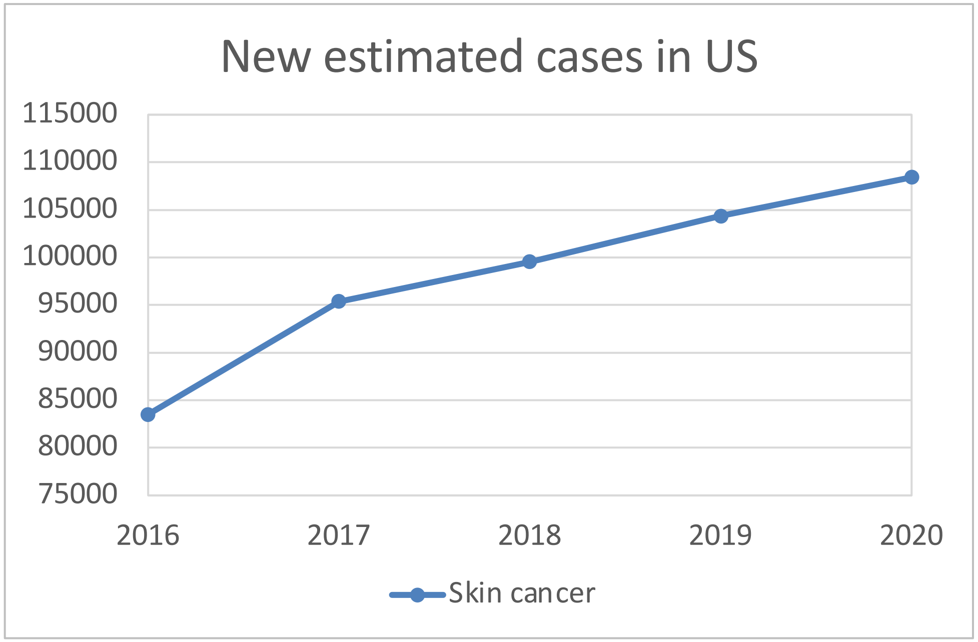
Outline

I. Introduction
II. Proposed Methods
III. Experiments & Discussions
IV. Conclusion & Future Work
V. References

I.　Introduction

Background

- Skin Cancer
  - one of the deadly diseases
  - the number of people who are attacked by skin cancer is increasing by years



Background

Causes of Skin Cancer

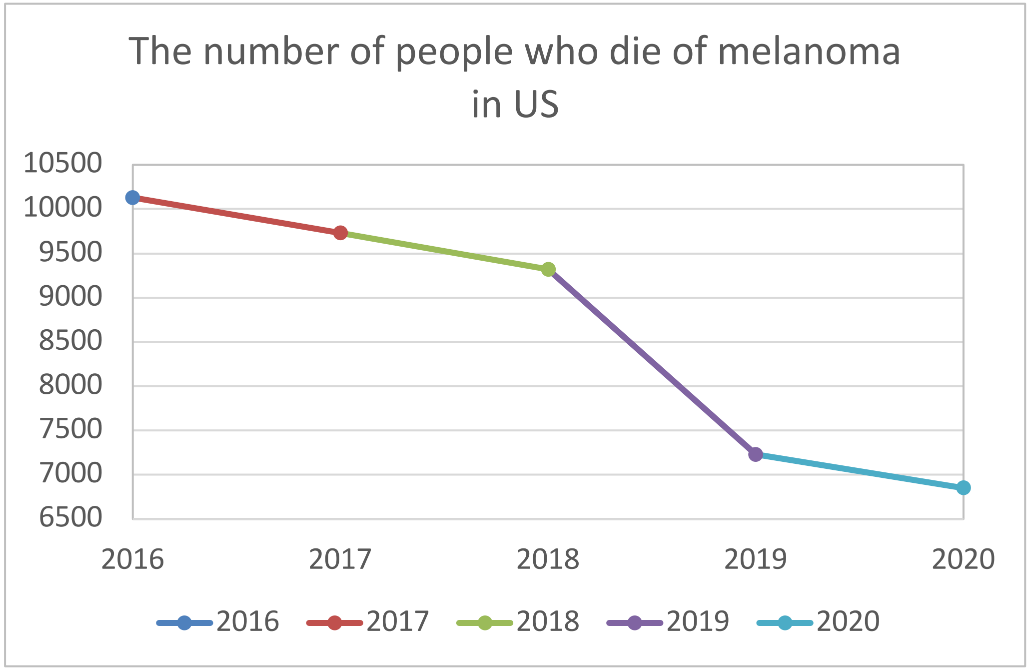
- Skin cancer is caused by mutation occurring in the DNA of skin cells.
- Several factors may risk to cause mutation.

Risk factors

- Genetic factors
- Environmental factors
  - UV exposure (Leading factor)
  - Viral infections
  - Immunosuppression
  - Having light skin

Background

- Melanoma
  - one of the deadliest skin cancers
  - up to 92% of patients can recover if melanoma is diagnosed in early stage



Related Work

Existing Work on Skin Lesion Classification

- methods based on convolutional neural networks (CNNs)
- three widely used methods

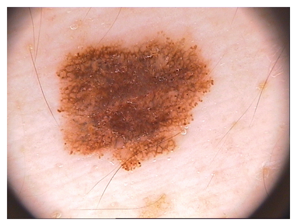
Methods	Year	Publication	Details
Ensemble learning	[1], 2018	CIBEC	often combine three or more networks together
Hierarchical classification	[2], 2019	CVPR	from general categories to more specific classes
Data augmentation	[3], 2018	IEEE	transform one image into several images
Other methods	[4], 2019	IEEE	design a new loss weight formula
	[5], 2019	IEEE	propose a patch-based attention architecture

Motivation

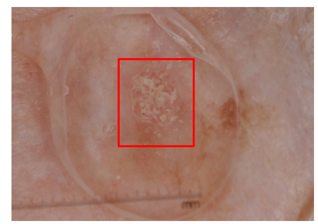
Challenge

- Different skin cancers may look quite similar.
- likely for network to make the wrong judgement on similar skin images

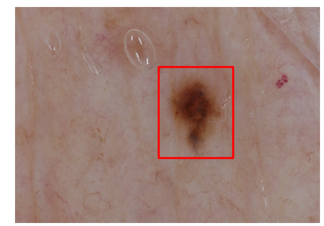
Melanoma



Nevi



Seberrheic Karatosis



Nevi

Motivation

- Problems
  - Noisy artifacts
  - Data Imbalance
  - **Similarity among skin lesion images**

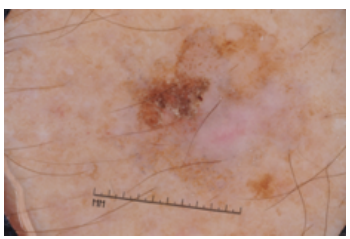
How to effectively solve these three problems and improve the classification performance?

II. Proposed methods

Image Pre-processing(1)

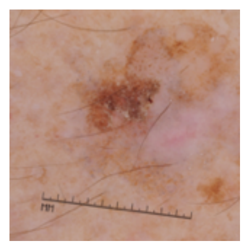
- Problem 1: Noisy artifacts
  - often include hair, ruler, bubbles and other annotations marked by doctors
  - increase difficulties in focusing on the pivotal parts
  - ***Solution:*** Artifact remover

original image



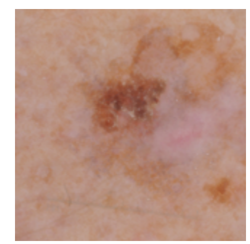
4288×2848

center crop
resize to 1024×1024



2848×2848
↓
1024×1024

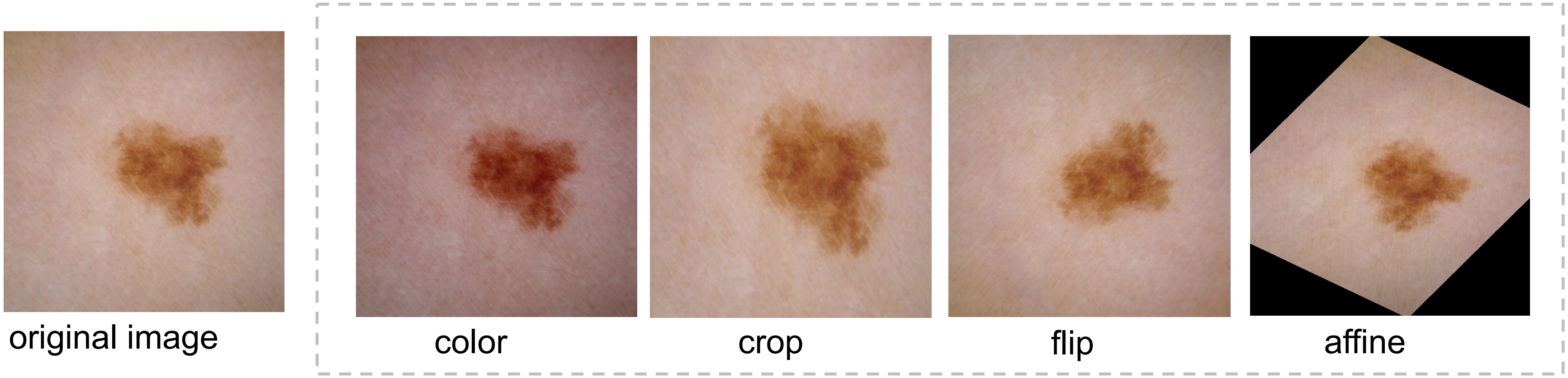
artifacts removal algorithm



1024×1024

Image Pre-processing(2)

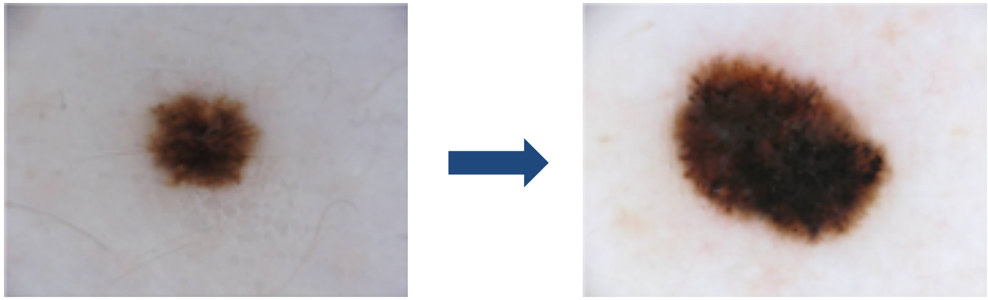
- Problem 2: Data Imbalance
  - poor performance in testing
  - ***Solution:*** Data augmentation
    - **Color:** modify saturation, contrast and brightness by random factors
    - **Crop:** randomly crop original images
    - **Flip:** randomly flip the images horizontally or vertically
    - **Affine:** rotate the image by up to 90°, shear by up to 20° and scale the area by [0.8-1.2]



DRPM

- Problem 3: Similarity among skin lesion images
  - particularly between melanoma and nevi

    about 30% of melanoma mutate from nevi



Nevi   Melanoma

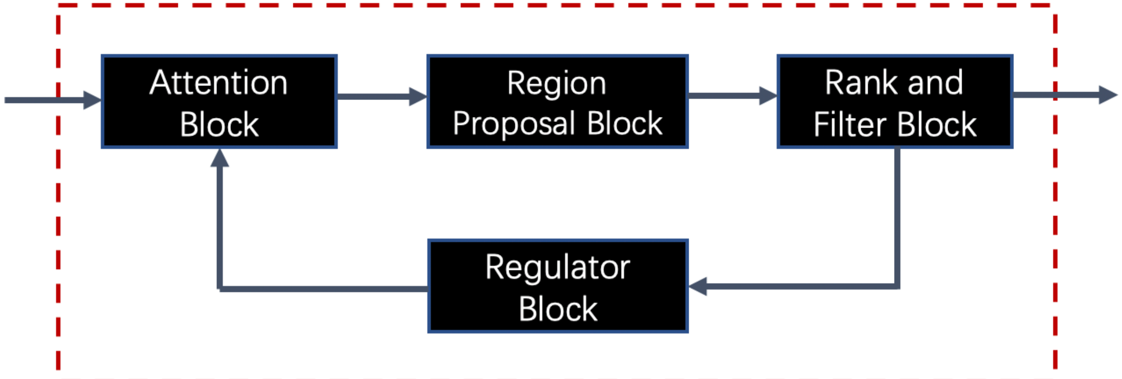
  - hard for network to make a right decision
  - inspired by fine-grained image classification and weakly-supervised object detection task

    Key point: discover subtle and distinguished local areas

    Solution: Distinct Region Proposal Module (DRPM)

DRPM

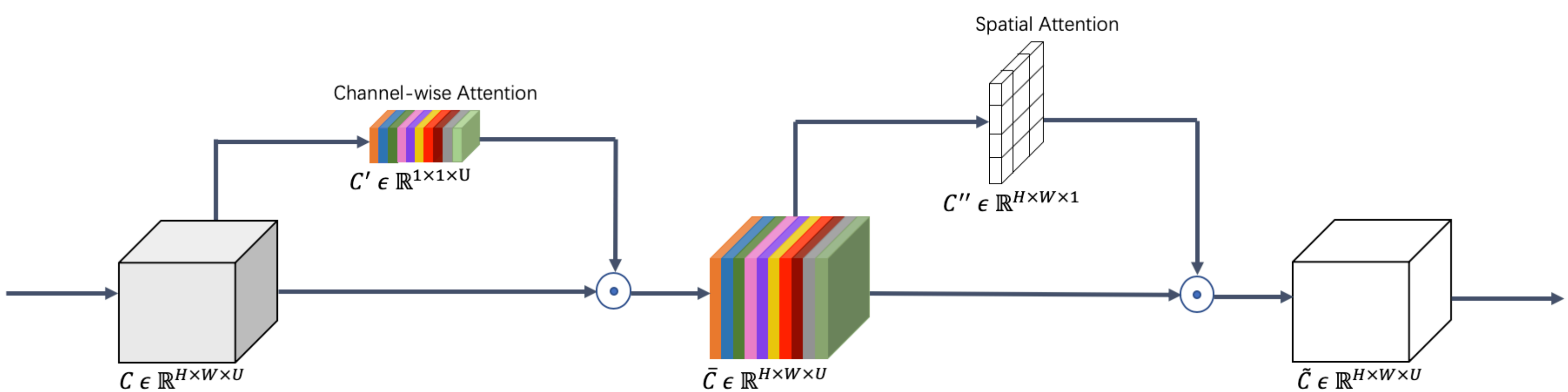
- Distinct Region Proposal Module



- **Target:** localize the most informative and characterized regions
- **Attention Block:** produce the strengthened feature maps and highlight the most informative areas
- **Region Proposal Block:** generate possible regions containing objects
- **Rank and Filter Block:** rank all the selected regions according to their score of informativeness
- **Regulator Block:** help the network to adjust and get the most distinctive parts

Attention Block

- **Channel-wise attention** focuses on one main part of an image and simultaneously considers the semantic information.
- **Spatial attention** generates heat map automatically to emphasize the related features and restrain the unrelated features.



$$\overline{C} = \psi(C) \otimes C = C' \otimes C$$

$\psi(\cdot)$: global average pooling

$$\tilde{C} = \varphi(\overline{C}) \otimes \overline{C} = C'' \otimes \overline{C}$$

$\varphi(\cdot)$: max pooling

DRPM

Region Proposal Block

- Region proposal block uses anchors with three scales of {48, 92, 192} and three ratios of {1:1, 3:2, 2:3} to predict multiple bounding boxes for proposed regions in a weakly-supervised way.

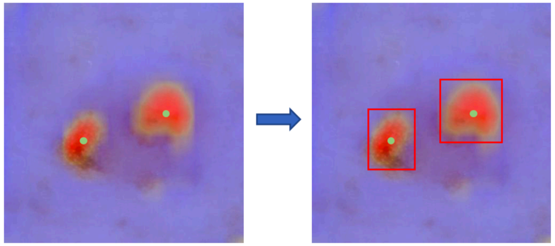
Rank and Filter Block

- Non-maximum suppression is followed to reduce redundancy of regions and get {A1, A2,…, Ay} local images.
- Local images are ranked in a descending order based on their score of informativeness.

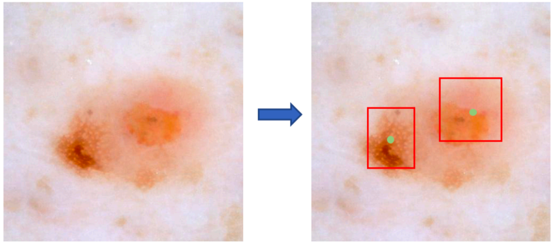
$$S(A_1) \geq S(A_2) \geq \ldots \geq S(A_{y-1}) \geq S(A_y)$$

- A hyper-parameters k is set to chooses top-k regions $\{A_1, A_2 \ldots, A_k\}$.

Here k = 2.



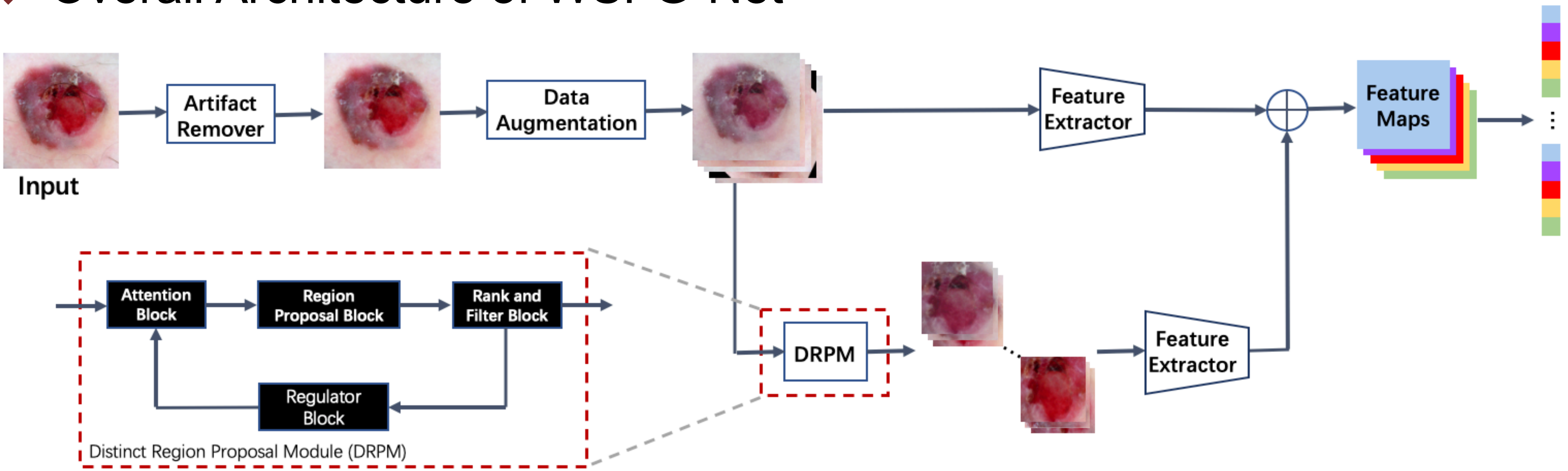
(a)



(b)

Overall Architecture

◆ Overall Architecture of WSFG-Net



❖ Whole process:

- Firstly, an input image conducts center crop and eliminates the artifacts through artifact remover.
- Secondly, for categories with small quantity, each image is implemented on four data augmentation methods.
- Then the structure is split into two branches, extracted features from both global and local images.
- Finally, the features from two branches are concatenated together and sent into classifier.

III. Experiments & Discussions

Experiments

Dataset

Table 1. Details of ISIC 2017 dataset.

	Melanoma (M)	Seborrheic Keratosis (SK)	Nevi (NV)
Train	374	254	1372
Test	117	90	393

- Task: calculate SE, SP and AUC of Melanoma and Seborrheic Keratosis
- Evaluation Metrics
  - Sensitivity (SE) = $\frac{TP}{TP + FN}$
  - Specificity (SP) = $\frac{TN}{TN+FP}$
  - Aera under the curve (AUC)

TP: True positive
TN: True negative
FP: False positive
FN: False negative

Experiments

Comparison Result

Table 2. Comparison of SE, SP and AUC of melanoma and seborrheic keratosis on ISIC 2017.

Method	Extra Data	Ensembles	Melanoma			Seborrheic Keratosis			Average AUC
			SE	SP	AUC	SE	SP	AUC	
[6], 2017	Y	Y	73.5	85.1	86.8	97.8	77.3	95.3	91.1
[7], 2017	Y	Y	42.7	96.3	87.0	58.9	97.6	92.1	89.6
[8], 2019	Y	N	65.8	89.6	87.5	87.8	86.7	95.8	91.7
[9], 2018	N	Y	40.2	71.9	85.1	71.1	85.1	93.0	89.1
[2], 2019	N	Y	73.5	83.8	85.5	61.1	97.2	93.2	89.4
[10], 2020	N	Y	37.6	96.5	89.1	72.2	97.3	93.5	92.6
[8], 2019	N	N	59.0	89.6	85.9	77.8	93.1	95.1	90.5
WSFG-Net	N	N	75.8	85.3	86.6	64.7	98.0	96.2	91.5
WSFG-Net-Ens	N	Y	76.1	88.4	89.5	73.8	98.3	96.9	93.2

The performance of proposed method is better than [8], which both are under the same condition. Two architectures, ResNet-50 and VGG- 16, are ensembled and the performance of average AUC improves from 91.5% to 93.2%, which achieves the state-of-the-art performance.

Experiments

Ablation Study

Table 3. Ablation study on two preprocessing steps.

Artifact Remover	Data Augmentation	M AUC	SK AUC
		81.4	90.9
√		83.7	92.8
√	√	86.6	96.2

This ablation study shows that the two preprocessing steps are conducive to the performance of our model.
In particular, training the network with imbalanced data is adverse so using data augmentation to equilibrate each category is helpful.

Discussions

- Resized Images and Cropped Images

Table 4. Comparison results on performance of using resized images and cropped images.

	Melanoma			Seborrheic Keratosis		
	SE	SP	AUC	SE	SP	AUC
Resized images	66.9	85.1	84.9	59.3	94.5	93.4
Cropped images	75.8	85.4	86.6	64.7	98.0	94.5

The performance of cropped images is better than resized images in all evaluation metrics.

Analyzing: This may be because locations of skin lesions got from dermoscopy equipment often lie in the center of images and using cropping operation can delete parts of background and make the areas of lesions become clear.

Discussions

Artifact Remover and Attention Block

- We think artifact remover and attention block in DRPM are interacted with each other.
  - Without eliminating artifacts in images, the attention mechanism may highlight some artifacts, which makes the latter region proposal block focus on artifacts instead of skin lesions.
  - Without attention block, the network can't concentrate on objects of interest.

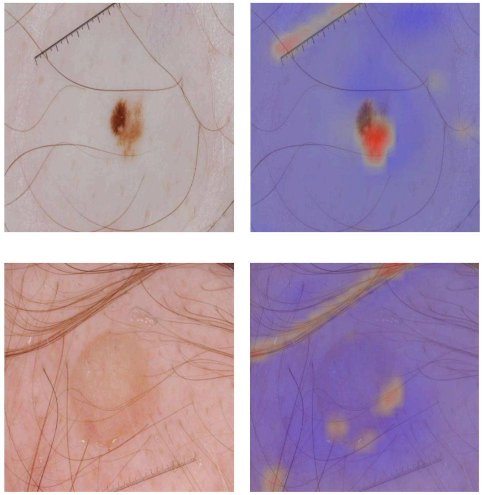


Table 5. Ablation study on the performance of channel-wise attention and spatial attention in attention block.

Channel-wise Attention	Spatial Attention	M AUC	SK AUC
√		85.8	95.3
	√	85.5	95.7
√	√	86.6	96.2

The performance of using either of the attention has no big difference but after combining two attention methods, the performance of AUC on both categories has been improved.

IV. Conclusion & Future Work

Conclusion

- A novel end-to-end skin lesion classification method is proposed to solve hard-recognized skin lesion classification problems. This task is regarded as a fine-grained image classification and the network can be trained in a weakly-supervised way.
- Two pre-processing steps, artifact remover and data augmentation, are designed to remove the noisy artifacts and make each category relatively balanced.
- DRPM is designed in the network to automatically localize the most distinct regions in an image, guided by both channel-wise attention and spatial attention.
- The experiments, conducted on ISIC 2017 dataset, prove that the proposed method can not only classify each category effectively but also be trained efficiently.

Future Work

- Drawbacks
  - Two indexes haven't achieved the state-of-the-art performance.
  - Only one dataset is used in the experiments.

- Future Work
  - need to further improve the performance of skin lesion classification
  - use other datasets to evaluate the proposed network, such as the training data of ISIC 2018 and 2019

References

[1] A.H.Shahin, A.Kamal, and M.A.Elattar. "Deep ensemble learning for skin lesion classification from dermoscopic images," in CIBEC. IEEE, 2018.

[2] C. Barata, J. S. Marques, and M. Emre Celebi, "Deep attention model for the hierarchical diagnosis of skin lesions," in CVPR, 2019.

[3] F. Perez, C. Vasconcelos, S. Avila, and E. Valle, "Data augmentation skin lesion analysis," in Computer Assisted Robotic Endoscopy, IEEE, 2018.

[4] M. A. Albahar, "Skin lesion classification using convolutional neural network with novel regularizer," IEEE Access, 2019.

[5] N.Gessert, etal. "Skin lesion classification using cnns with patch-based attention and diagnosis-guided loss weighting," IEEE Transactions on Biomedical Engineering, 2019.

[6] K. Matsunaga, et al. "Image classification of melanoma, nevus and seborrheic keratosis by deep neural network ensemble," arXiv preprint arXiv:1703.03108, 2017.

[7] L. Bi, J. Kim, E. Ahn, and D. Feng, "Automatic skin lesion analysis using large-scale dermoscopy images and deep residual networks," arXiv preprint arXiv:1703.04197, 2017.

[8] J. Zhang, Y. Xie, Y. Xia, and C. Shen, "Attention residual learning for skin lesion classification," IEEE transactions on medical imaging, 2019.

[9] B. Harangi, "Skin lesion classification with ensembles of deep convolutional neural networks," Journal of biomedical informatics, 2018.

[10] P. Tang, Q. Liang, X. Yan, S. Xiang, and D. Zhang, "Gp-cnn-dtel: Global-part cnn model with data-transformed ensemble learning for skin lesion classification," IEEE Journal of Biomedical and Health Informatics, 2020.

Q & A