Melanoma Recognition via Visual Attention

Yiqi Yan Jeremy Kawahara Ghassan Hamarneh

Medical Image Analysis Lab, School of Computing Science, Simon Fraser University, Canada

2019 HK 50th Anniversary IPMI

1 Introduction

Motivation

- ❑ Melanoma is a leading cause of skin cancer related deaths
- ❑ Encourage the CNN to attend to the skin lesion region rather than irrelevant objects when diagnosing melanoma.
- ❑ Develop a more flexible approach to utilize pixel-level prior information.

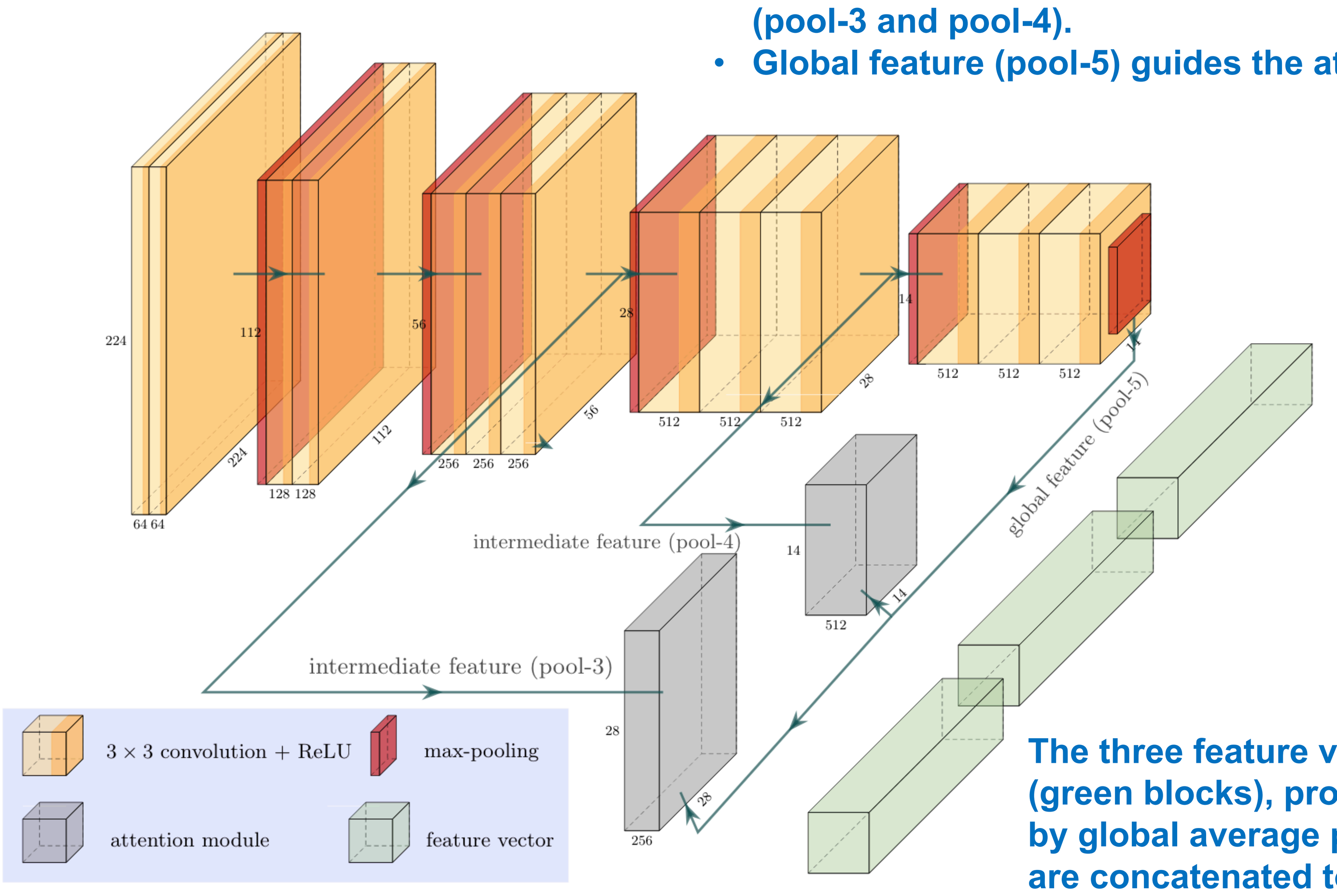
Overview

- ❑ Learnable attention modules → Higher accuracy, better interpretability
- ❑ Attention map regularization → Use prior information (if available)

2 Method

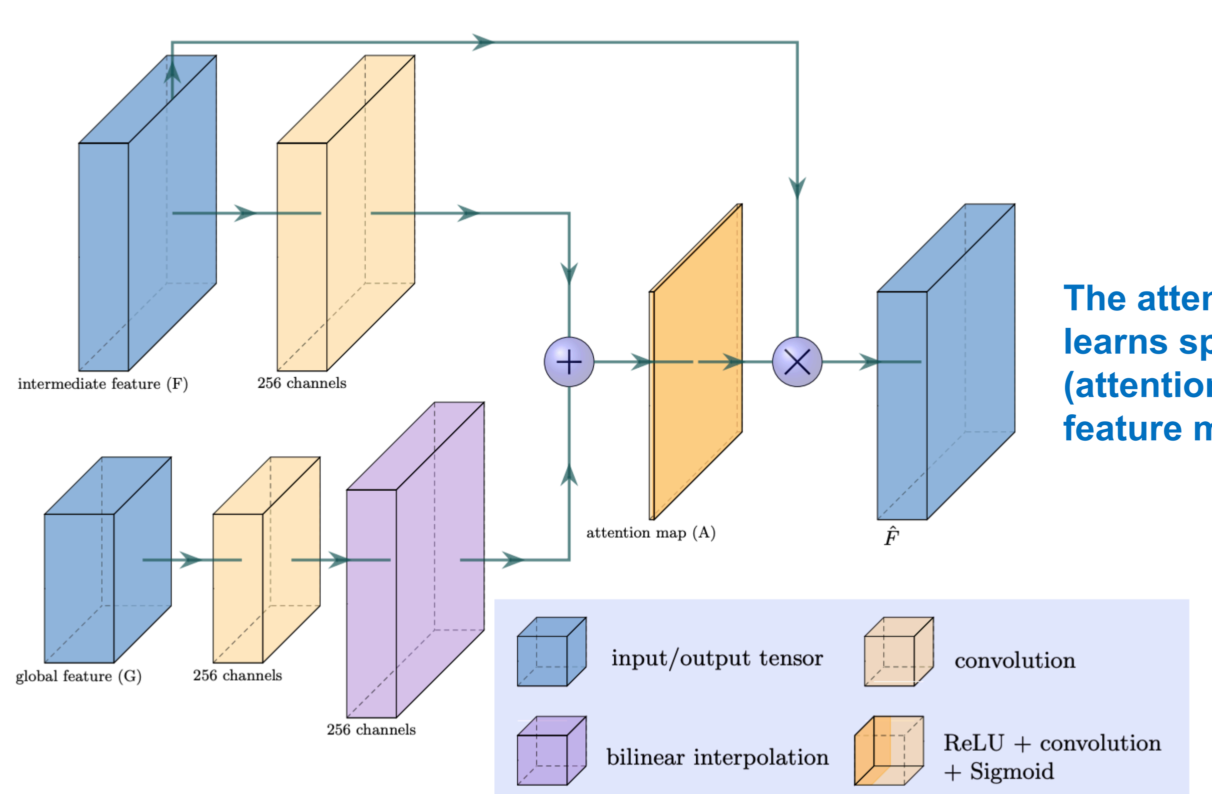
Network Architecture

- Apply attention to intermediate features (pool-3 and pool-4).
- Global feature (pool-5) guides the attention.



The three feature vectors (green blocks), produced by global average pooling, are concatenated together as the input of the final classification layer.

Learnable Attention Module

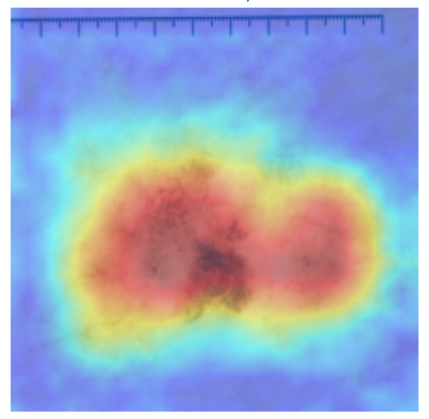


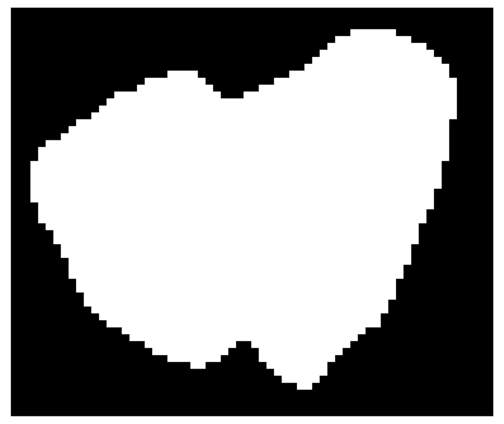
The attention model learns spatial weights (attention map $\mathcal{A}$) for the feature map.

Attention Map Regularization

- Use negative Sørensen-Dice-F1 loss to regularize the attention map. a_i represents the i-th pixel.

$$\mathcal{L}_D(\mathcal{A}, \bar{\mathcal{A}}) = 1 - D(\mathcal{A}, \bar{\mathcal{A}}) = 1 - \frac{2 \cdot \sum_{i=1}^{n} (a_i \cdot \bar{a}_i)}{\sum_{i=1}^{n} (a_i + \bar{a}_i)}$$

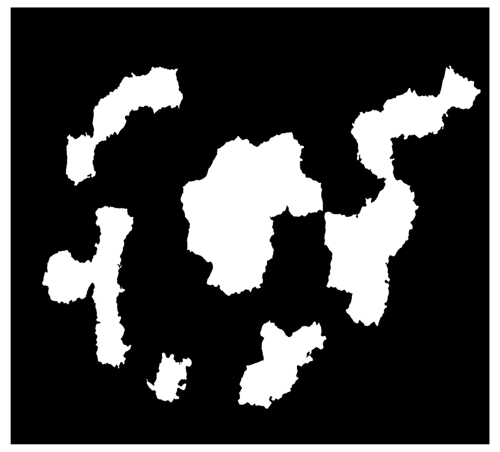




Lesion Segmentation

or

Dermoscopic Features



- Use focal-loss[1] as the main classification term.
- Regularize attentions from both pool-3 and pool-4 layers, yielding two regularization terms.
- When pixel-level annotations are unavailable, set $\lambda_1 = \lambda_2 = 0$; otherwise set the value empirically $\lambda_1 = 0.001$, $\lambda_2 = 0.01$.

$$\mathcal{L} = \mathcal{L}_{focal} + \lambda_1 \mathcal{L}_D(\mathcal{A}^{(3)}, \bar{\mathcal{A}}^{(3)}) + \lambda_2 \mathcal{L}_D(\mathcal{A}^{(4)}, \bar{\mathcal{A}}^{(4)})$$

3 Results

Ablation & Comparison Study

- VGG-16: the original VGG architecture
- VGG-16-GAP: replace the dense layers with global average pooling
- Mel-CNN: use intermediate features, but without attention

- Attn-Mel-CNN: proposed model, trained without regularization
- Attn-Mel-CNN-Dermo: proposed model, regularized with dermoscopic features
- Attn-Mel-CNN-Lesion: proposed model, regularized with lesion segmentation

- Attn-Mel-CNN-Bkg (counterexample): proposed model, regularized with background (inverse of segmentation)

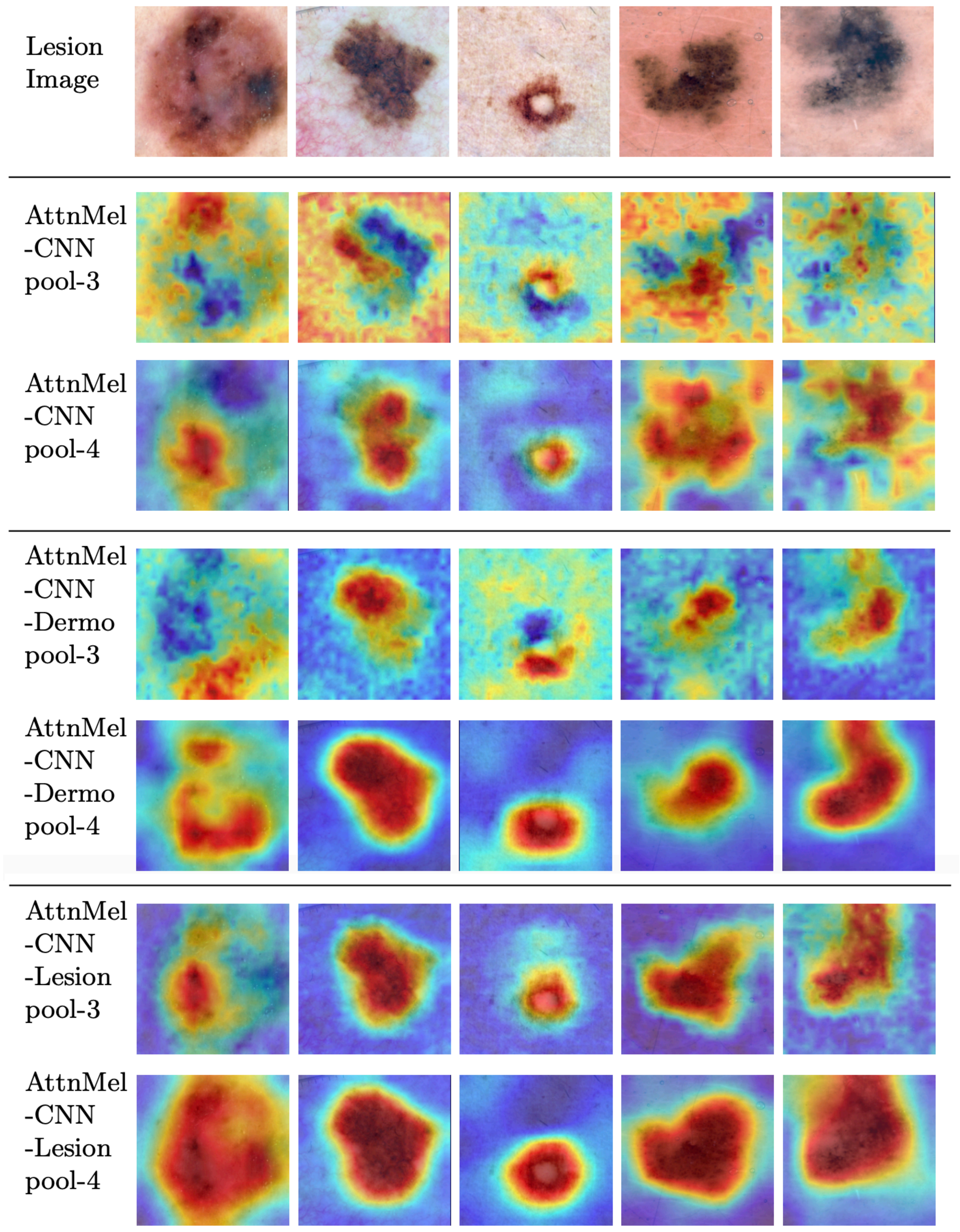
ISIC 2016[2]

	Average precision	Area under the the ROC curve	Whole lesion prior	Interpretable	Ensemble
#1 Lequan et al.	0.637	0.804	✓	✗	✗
#2 Codella et al.	0.596	0.808	✗	✗	✓
#3 Yu et al.	*0.685*	**0.852**	✗	✗	✓
#4 VGG-16	0.602	0.806	✗	✗	✗
#5 VGG-16-GAP	0.635	0.815	✗	✓	✗
#6 Mel-CNN	0.664	*0.844*	✗	✗	✗
#7 **AttnMel-CNN**	**0.693**	**0.852**	✗	✓	✗

ISIC 2017[3]

	Average precision	Area under the the ROC curve	Whole lesion prior	Dermoscopic feature prior	Interpretable	Ensemble	External data
#1 ISIC 2017 Winner 1	–	0.868	✗	✗	✗	✓	✓
#2 ISIC 2017 Winner 2	–	0.856	✓	✓	✗	✗	✓
#3 ISIC 2017 Winner 3	–	*0.874*	✗	✗	✗	✓	✓
#4 Harangi et al.	–	0.836	✗	✗	✗	✓	✗
#5 Mahbod et al.	–	*0.873*	✗	✗	✗	✓	✓
#6 VGG-16	0.600	0.824	✗	✗	✗	✗	✗
#7 VGG-16-GAP	0.627	0.834	✗	✗	✓	✗	✗
#8 Mel-CNN	0.653	0.854	✗	✗	✗	✗	✗
#9 **AttnMel-CNN**	0.655	0.872	✗	✗	✓	✗	✗
#10 **AttnMel-CNN-Dermo**	*0.665*	0.864	✗	✓	✓	✗	✗
#11 **AttnMel-CNN-Lesion**	**0.672**	**0.883**	✓	✗	✓	✗	✗
#12 AttnMel-CNN-Bkg	0.647	0.849	✓	✗	✓	✗	✗

Attention Map Visualization



- The shallower layer (pool-3) tends to focus on more general and diffused areas, while the deeper layer (pool-4) focus more on the lesion and avoids irrelevant objects.
- Regularization makes the learned attention more concentrated.

4 Conclusion & Future Work

- ❑ Proposed an attention-based network for melanoma recognition with attention map regularization.
- ❑ Achieved *state-of-the-art* and *interpretable* performance for melanoma classification on two public datasets.
- ❑ Future work would explore visual attention in more general multi-class skin lesion classification problems.

REFERENCES

[1] Lin et al. ICCV (2017)
[2] Gutman et al. arXiv:1605.01397 (2016)
[3] Codella et al. ISBI (2018)

Mitacs NSERC CRSNG NVIDIA

contact: yiqiy@sfu.ca **code:** https://github.com/SaoYan/IPMI2019-AttnMel