Supporting Skin Lesion Diagnosis with Content-Based Image Retrieval

Stefano Allegretti, Federico Bolelli, Federico Pollastri, Sabrina Longhitano, Giovanni Pellacani, and Costantino Grana

Università degli Studi di Modena e Reggio Emilia, Italy
Skin Cancer

• Skin cancer is one of the most common forms of human cancer worldwide

• If caught early, it is usually curable

• Distinguishing skin cancer from other kinds of skin lesion is a difficult task
Computer Aided Diagnosis with CNN

- Convolutional Neural Networks have been widely employed for skin lesion classification
- Classification CNNs have pros and cons:
  - **Good classification accuracy**, comparable to expert dermatologists
  - **Low interpretability**: Scarce diagnostic aid for physicians
- How can interpretability be improved?
Content-Based Image Retrieval (CBIR)

• Given a new lesion, retrieve **similar cases** from a labeled database

• How to define image similarity?

• Past works:
  • Euclidean or Bhattacharyya distance between handcrafted features
  • Hamming distance between hash codes, computed with a modified classification CNN (AlexNet)
Proposed CBIR system

• **ResNet Feature Extractor:** ResNet, except for the last FC layer

• **Embedding Network:** 2 FC layers, built on ResNet-extracted features

• **Cosine similarity** between image embeddings

• Embedding training with **triplet loss** function

• **ISIC dataset** - 20K images, 8 classes
Quantitative Results

4 variations of the proposed model:

- ResNet FE trained for classification
- ResNet FE + EmbNet, end-to-end trained with triplet loss
- ResNet FE pretrained + EmbNet, only EmbNet trained with triplet loss
- ResNet FE pretrained + EmbNet, end-to-end trained with triplet loss
Qualitative Results

• Dermatologists classified 100 lesions two times
  • **Task 1** – without aid
  • **Task 2** – with the 5 most similar labeled images

<table>
<thead>
<tr>
<th>Dermatologist</th>
<th>Task 1</th>
<th>Task 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>75%</td>
<td>79%</td>
</tr>
<tr>
<td>#2</td>
<td>64%</td>
<td>80%</td>
</tr>
<tr>
<td>#3</td>
<td>69%</td>
<td>71%</td>
</tr>
<tr>
<td>#4</td>
<td>68%</td>
<td>82%</td>
</tr>
<tr>
<td>#5</td>
<td>61%</td>
<td>71%</td>
</tr>
</tbody>
</table>

• Average accuracy improvement of 9%