A Benchmark for Automatic Visual Classification of Clinical Skin Disease Images
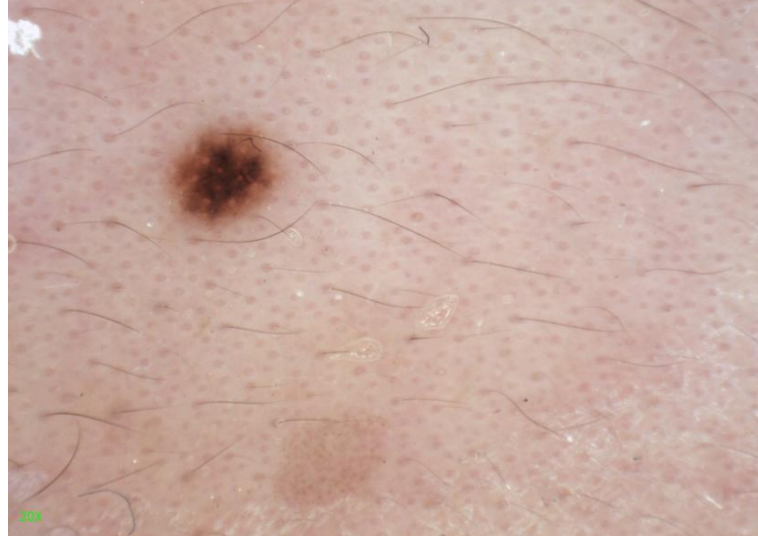
Xiaoxiao Sun, Jufeng Yang*, Ming Sun, Kai Wang

Nankai University

2019/1/15

预警！！！！！！！

Motivation

- Low contrast between lesion and surrounding skin

- Irregular and fuzzy borders, fragmentation

- Variegated coloring inside the lesion

Motivation

- Small datasets (few species)
- Not publicly available
- The absence of benchmark datasets is a barrier to a more dynamic development of this research area.

Main Contribution

- A large scale benchmark dataset for skin disease image recognition.

       SD-128:128 categories, more than 20 images per category

       SD-198:6584 images,198 categories

- Evaluation of skin disease classification

Main Contribution

Related Work

- Razeghi et al(Medical Image Understanding and Analysis (MIUA2012)): two subsets containing 90 and 706 images from 3 and 7 different skin diseases

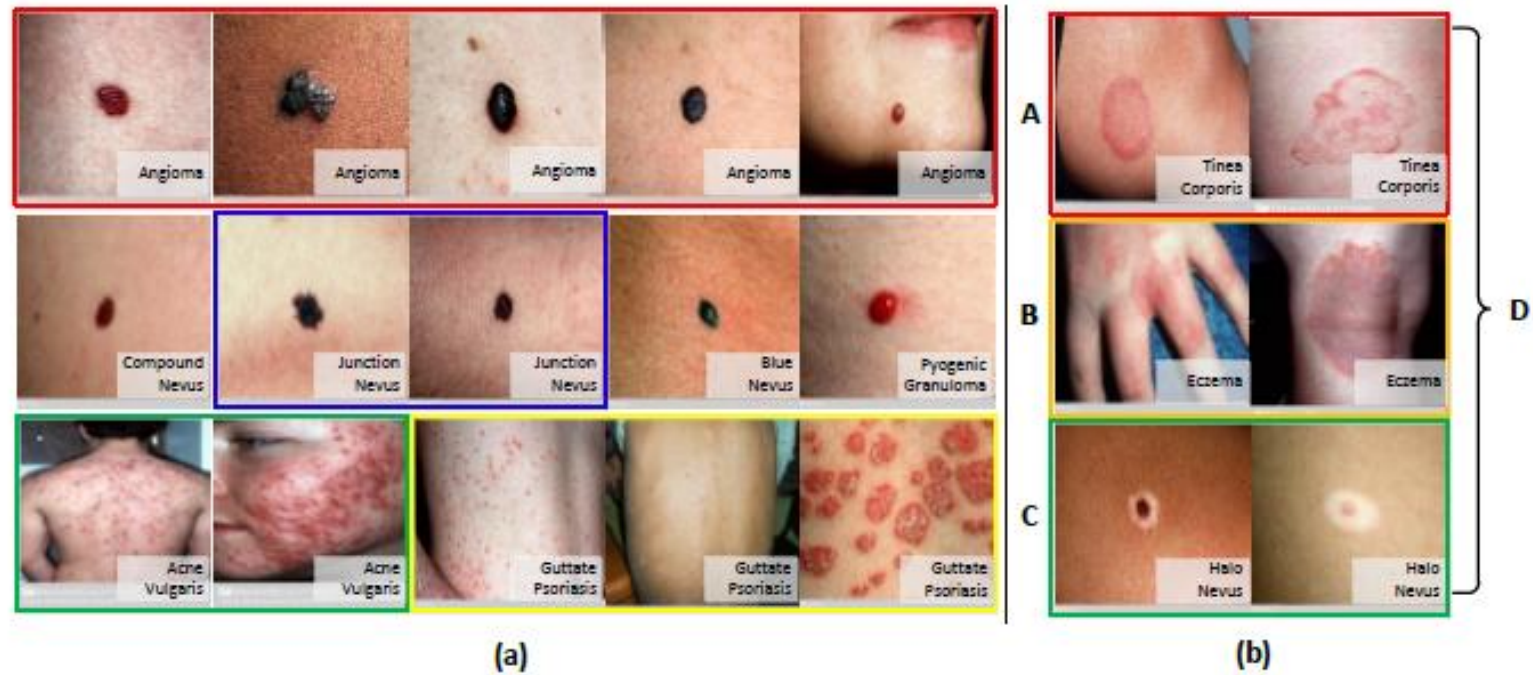- Razeghi et al(ICME2013): 2,309 images,44 categories

- Not available

Analysis

Image Collection and Annotation

- Https://www.dermquest.com
- 10000 to 6584(balance)
- Ground truth annotations

Properties

- Scale:3-10 times
- Diversity:

    Species

Properties

- Scale:3-10 times
- Diversity:

    Species

    Appearance

    Attribute:

    age

    sex

    disease color of skin

    different periods of lesions

Challenge and lack

- Low color contrast of foreground(lesion) and background(health skin)
- Details of dermatosis marks
- Imbalance among different categories

Experiments

Clinical Skin Disease Classification

• Hand-Crafted Features based Classification

    texture:

        SIFT, HOG, LBP, Gist, Gabor

    color:

        Color Histogram, Color Names(CN)

    SIFT+CN+SVM

    Spatial Pyramid

Clinical Skin Disease Classification

Num	Features	features dimension	Classifier	SD-198 %	SD-128 %
1	SIFT	21000	SVM	25.85	29.40
2	HOG	12400	SVM	12.78	14.17
3	LBP	23200	SVM	15.46	17.09
4	Color Histogram	768	SVM	4.19	5.59
5	Color Names(CN)	21000	SVM	20.20	20.32
6	Gist	512	SVM	16.49	17.52
7	Gabor	4000	SVM	10.14	11.37
Num	Methods	features dimension	methods or features	SD-198 %	SD-128 %
8	[49]	21000	SIFT+CN+SVM	52.19	53.29
9	[51]	4200	Spatial Pyramid	22.45	24.45

Clinical Skin Disease Classification

- Deep features based Classification
  - CaffeNet(pre-trained)
  - VGGNet(per-trained)

Method	SD-198[%]	SD-128[%]
CaffeNet	42.31	42.83
CaffeNe+ft	46.69	47.38
VGG	37.91	39.27
VGG+ft	50.27	52.15

Discussion

- Traditional features: cleaner background
- CNNs: structure and semantic information

My own thinking

- Quality is poor: Image Quality Assessment, Image Augmentation, Image compression(for good-quality but space-taking image)
- Scale can be promoted: Cooperate
- Algorithm is too simple: Classification