Major Project Proposal

Melanoma Detection: An Automated Approach

Aishwarya Rajan, 13CO109
Ankit Tripathi, 13CO113
Tejaswi Ravu, 13CO248

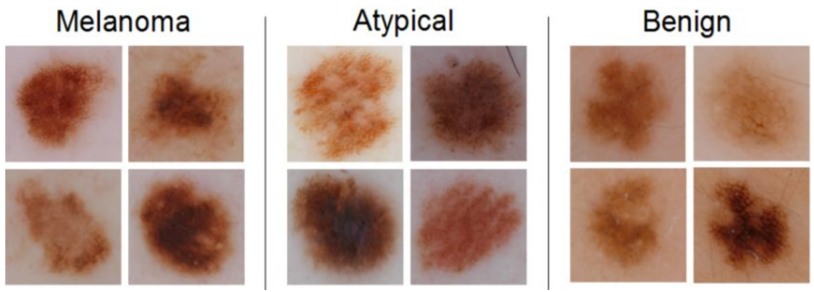
Table of Contents

Significance

1. Early melanoma diagnosis is critical to combatting the disease
2. Lack of experts and increasing melanoma incidences
3. An automation technique that can easily be adapted for other image based medical diagnostics
4. Explore recent algorithms in CV for image classification

Codella - using CNN and SVM (Transfer Learning)

- **Dataset** : International Skin Imaging Collaboration (ISIC) dataset
- Uses pre-trained Caffenet model from the ILSVRC
  - Concept detector layer, FC8 (1000 dimensions)
  - Fully connected layer, FC6 (4096 dimensions)
- The preprocessing step involves resizing the image and Subtract the model's input mean image to "centralize"
- **Feature normalization** : Sigmoid
- **Classifier** : Non-linear SVM using a histogram intersection kernel
- SVM score averaged FUSION



Drawbacks

- Melanoma vs Atypical and Benign has 92% accuracy while Melanoma vs. Atypical has a max of 72%.
- The network has been optimized for natural photographs of real-world objects.
- Based on stochastic learning with an annealing learning rate - later samples in the dataset have lower effect.
- Model is highly depended on the training dataset to give high level of accuracy and doesn’t adjust with new dataset, easily.
- Requires large datasets while most medical imaging problems have considerably small datasets

Codella - Sparse Coding and SVM

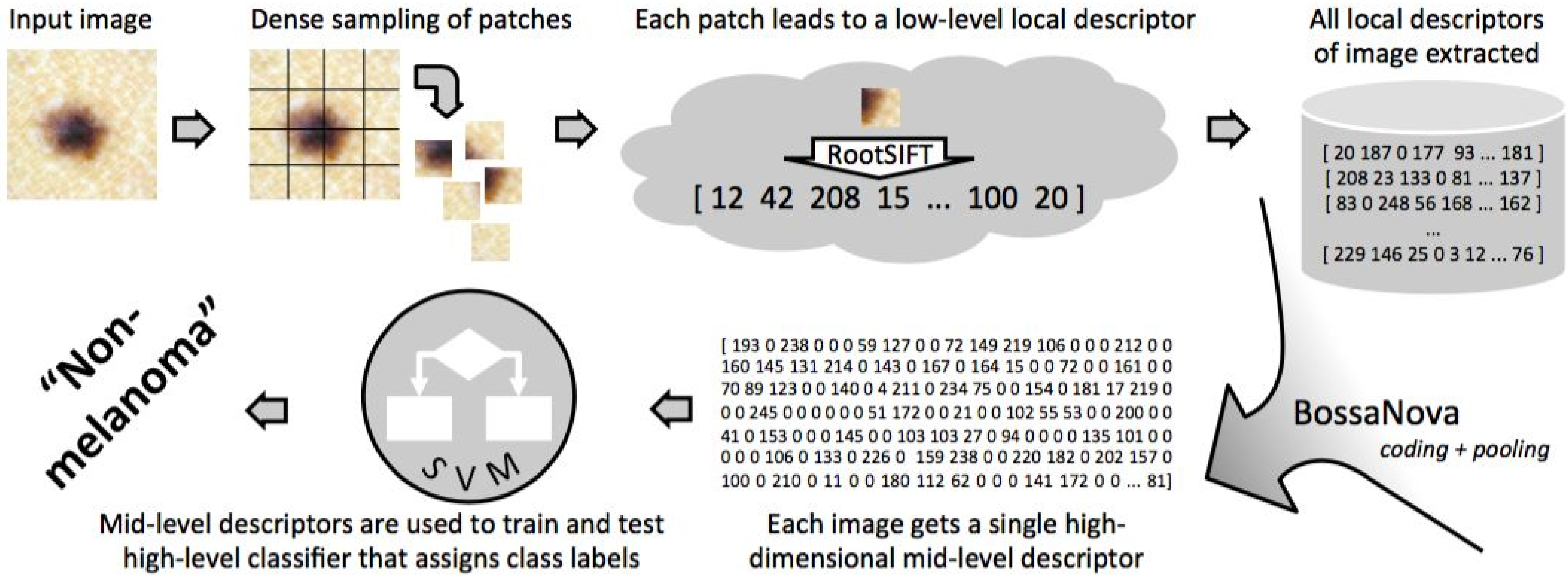
- **Dataset** : International Skin Imaging Collaboration (ISIC) dataset
- Unsupervised methods - learns a dictionary of sparse codes
- SPAMS sparse coding dictionary learning algorithm- based on stochastic approximations.
- Images are rescaled to 128x128 pixel dimensions before extraction of 8x8 patches, to learn dictionaries of 1024 elements.
- Two dictionaries are constructed in color (RGB) and grayscale color spaces.
- **Classifier** : Non-linear SVM using a histogram intersection kernel

Drawbacks

- We performed the experiment and found that specificity is in accordance with the results in the paper but the accuracy depends upon the dataset
- The feature vector is multi-dimensional and hence, converting the feature vector to one dimension results in loss of spatial information
- It involves solving a non-convex optimization problem and hence, finding the global solution takes long time and sometimes results in dead end (as observed in the experiment)

BossaNova

- **Dataset** : International Skin Imaging Collaboration (ISIC) dataset
- The pipeline
  1. Pre-processing image by resizing using ImageMagick
  2. Low-level local feature extraction using RootSIFT in VLFeat
  3. Sparsification by mapping all dimensions below threshold to zero
  4. Descriptor sampling
  5. Learn PCA matrix (to reduce dimensionality)
  6. Create codebook by random choice via k-means routine
  7. Mid-level features are created
- **Feature normalization** : PCA
- **Classifier** : SVM



Input image
Dense sampling of patches
Each patch leads to a low-level local descriptor
RootSIFT
[12 42 208 15 ... 100 20]
All local descriptors of image extracted
[20 187 0 177 93 ... 181]
[208 23 133 0 81 ... 137]
[83 0 248 56 168 ... 162]
...
[229 146 25 0 3 12 ... 76]
BossaNova
coding + pooling
[193 0 238 0 0 0 59 127 0 0 72 149 219 106 0 0 0 212 0 0
160 145 131 214 0 143 0 167 0 164 15 0 0 72 0 0 161 0 0
70 89 123 0 0 140 0 4 211 0 234 75 0 0 154 0 181 17 219 0
0 0 245 0 0 0 0 0 0 51 172 0 0 21 0 0 102 55 53 0 0 200 0 0
41 0 153 0 0 0 145 0 0 103 103 27 0 94 0 0 0 0 135 101 0 0
0 0 0 106 0 133 0 226 0 159 238 0 0 220 182 0 202 157 0
100 0 210 0 11 0 0 180 112 62 0 0 0 141 172 0 0 ... 81]
Each image gets a single high-dimensional mid-level descriptor
SVM
Mid-level descriptors are used to train and test high-level classifier that assigns class labels
"Non-melanoma"

Drawbacks

- All identified features are treated equally - lacks a hierarchy of characteristics
- Soft coding both saves geometric data but also results in denser feature vectors in the coding step (94 GB)
- The complete pipeline takes days to run => complex feature extraction.

Research Gap

1. Most of the methods proposed are based on empirical dataset and hence, the results are not consistent when tested on different datasets.
2. The topology of the Deep Convolutional Neural Network is decided by the user before training and hence, it is based on heuristics leading to underfitting or overfitting issues.
3. There is no algorithm that deterministically build the network appropriate for the dataset available, though work has been done on Constructive Learning for Multi-Layer Perceptron.

What next?

1. Explore constructive learning for DNN
2. Understand characteristics of DNN by performing experiments
3. Formulate an algorithm to train topologies for image classification
4. A robust model for Melanoma Detection in particular.