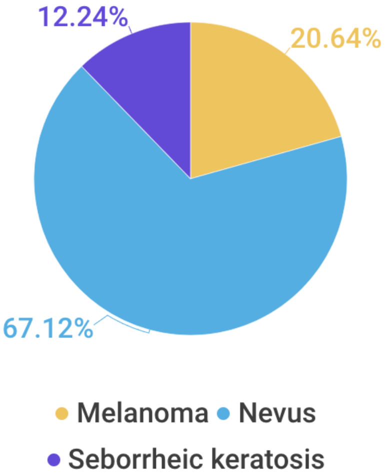
Towards Automated Melanoma Detection with Deep Learning: Data Purification and Augmentation

Devansh Bisla, Anna Choromanska, Russell S. Berman, Jennifer A. Stein, David Polsky

New York University, New York, NY, USA

Code: https://bit.ly/2KFRp5e Paper: https://bit.ly/2FBgOZP

Motivation



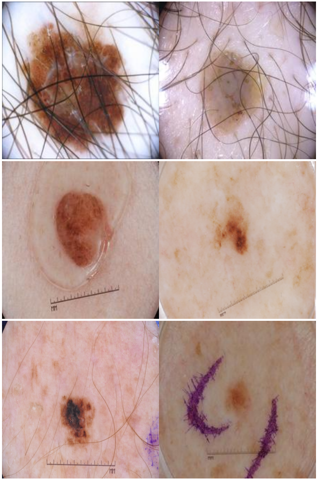
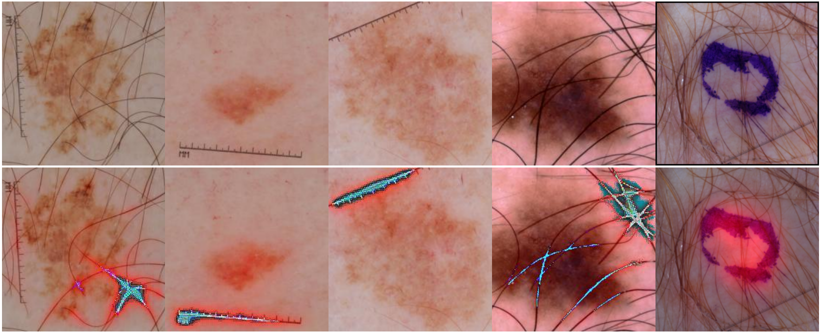


Figure 1: **(Left)** Data Imbalancedness **(Right)** Data Impurities

Existing computational techniques

- Traditional machine learning
  - Hand-crafted extraction of features from the data such as
    - Lesion Symmetry/Asymmetry.
    - Irregular borders.
    - Non-Uniform pigmentation.
    - Lesion size.
  - Problem: not scalable to large data sets.
- Deep Learning
  - Automatically extract features from large sized data.
  - Problem: Needs **large, balanced, and unbiased data**.

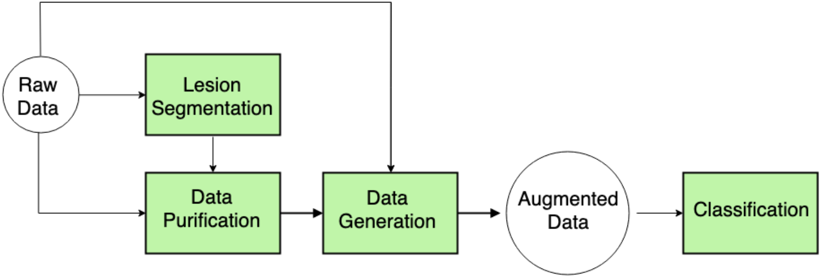
Traditional training



Visualization results for the conventionally-trained model **(Top)**: Original image. **(Bottom)**: Visualization mask overlaid on the original image.

The model overfits to image occlusions such as hairs, rulers and ink marks.

Proposed approach

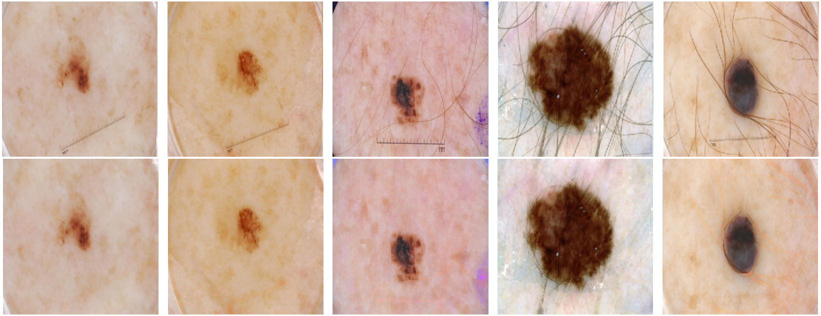


- Data Impurities:
  - Removal of unwanted objects such as hair, rulers etc.
- Data Imbalancedness
  - Synthetic data generation.
  - Data augmentation.

Data purification

- Thresholding in the LUV color space combined with morphological operations. Note that this may also remove dark regions belonging to the lesion itself.[Philippe Schmid-Saugeon et al]
- Overlay the processed image with the segmented lesion obtained from our segmentation algorithm.

Data purification - results



(a) (b) (c) (d) (e)

Figure 2: **Top**: Original images. **Bottom**: Images obtained after a,b) scales, c) hairs and scales, and d,e) hairs removal.

Data generation

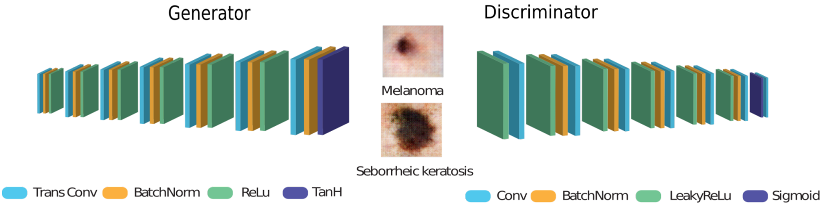


Figure 3: Architecture of Generative Adversarial Network

Main idea:

- Train a generator network to generate images which have similar distribution to the one followed by the training data, but do not appear in the training data set.
- The discriminator provides a feedback on similarity between the two distributions.

We generated 350 **images of melanoma and** 750 **images of seborrheic keratosis.**

Data generation - results

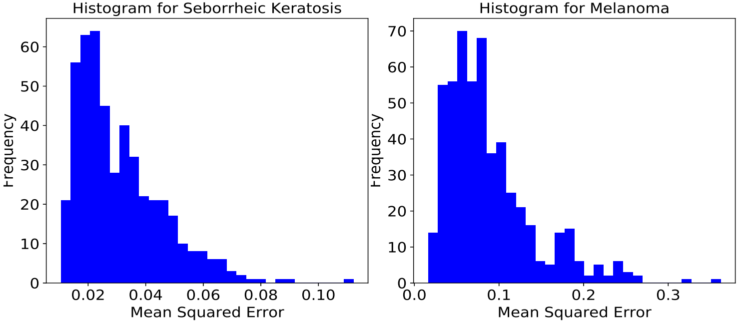
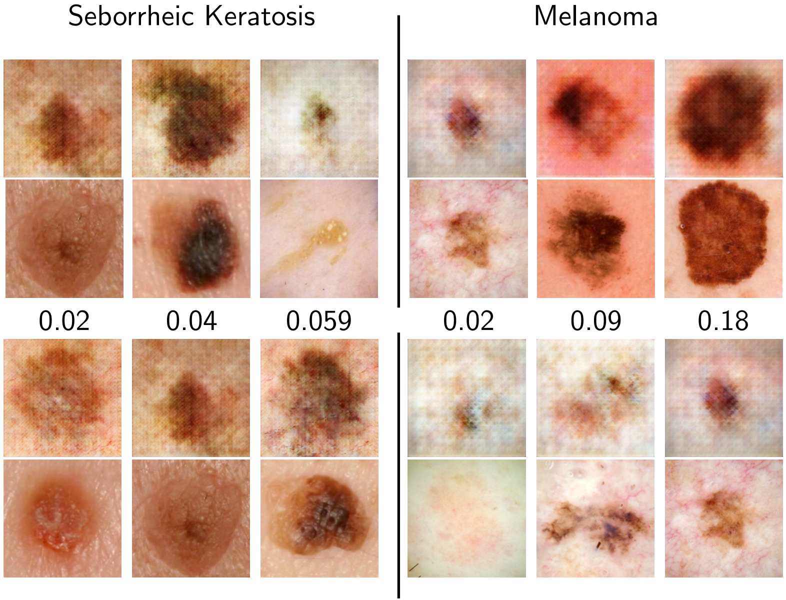


Figure 4: Histograms of the MSE values for (**left**) seborrheic keratosis and (**right**) melanoma.

Data generation - results



Classification results: confusion matrix

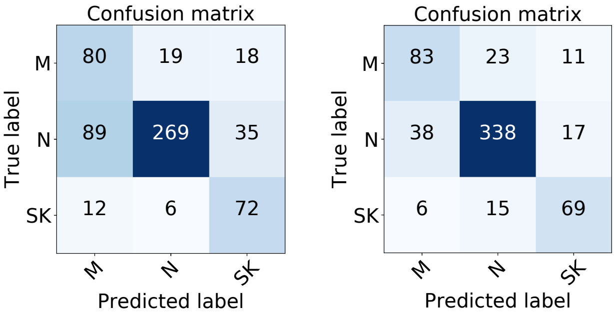


Figure 5: Confusion matrix obtained by traditional baseline (**left**) and proposed model (**right**).

Classification results: ROC-AUC

Mean Value	ROC-AUC
Our Approach	**0.915**
Kazuhisa Matsunaga[K. Matsunaga et al.]	0.911
RECOD Titans[A. Menegola et al.]	0.908

Table 1: Leader-board for melanoma and seborrheic keratosis combined.

Method	82%	89%	95%
Top AVG[K. Matsunaga et al.]	0.729	0.588	0.366
Top SK [I. Gonzalez Diaz et al.]	0.727	0.555	0.404
Top M [A. Menegola et al.]	**0.747**	0.590	0.395
Our Approach	0.697	**0.648**	**0.492**

Table 2: Specificity values at sensitivity levels of 82%/89%/95% for melanoma classification. Top AVG, Top SK, and Top M denote the winning approaches of the ISIC 2017 challenge.

Classification results visualized

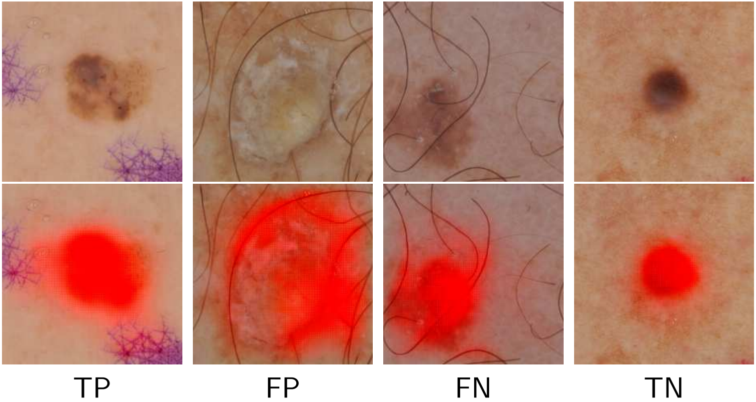


Figure 6: Visualization results for seborrheic keratosis. **Top**: Original image. **Bottom**: Visualization result.

Classification results visualized

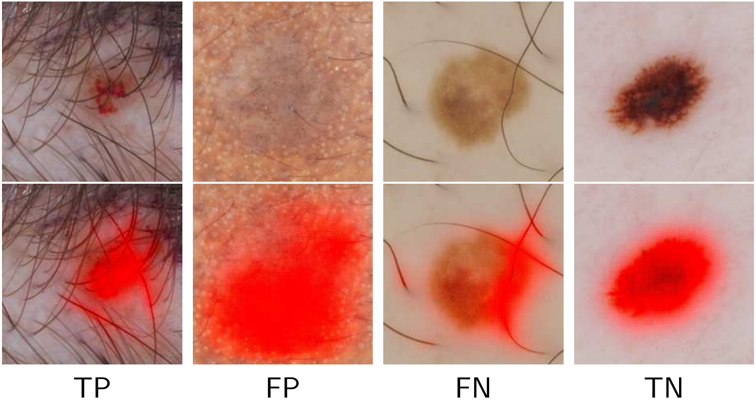


Figure 7: Visualization results for Nevus. **Top**: Original image. **Bottom**: Visualization result.

Conclusion

- Deep learning based methods are the most accurate and scalable, but they require **large**, **pure** and **balanced** training data sets.
- We presented solutions to improve effectiveness of classification systems by data purification (**removal of unwanted objects**) and data augmentation (**synthetic data generation**).