



Less is More: Sample Selection and Label Conditioning Improve Skin Lesion Segmentation

Vinícius Ribeiro¹, Sandra Avila², Eduardo Valle¹

¹RECOD Lab., DCA, FEEC, University of Campinas (UNICAMP), Brazil

²RECOD Lab., IC, University of Campinas (UNICAMP), Brazil

Handling the Inter-Annotator Agreement for Automated Skin Lesion Segmentation

Available at arXiv e-prints

The Impact of Annotation Quality on Deep Learning for Skin Lesion Segmentation

Available at Unicamp's Scientific and Intellectual Production Repository

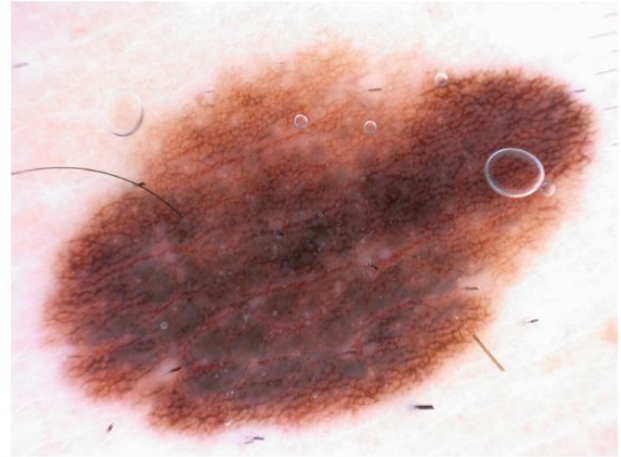
Links at

<https://vribeiro1.github.io/publications/>

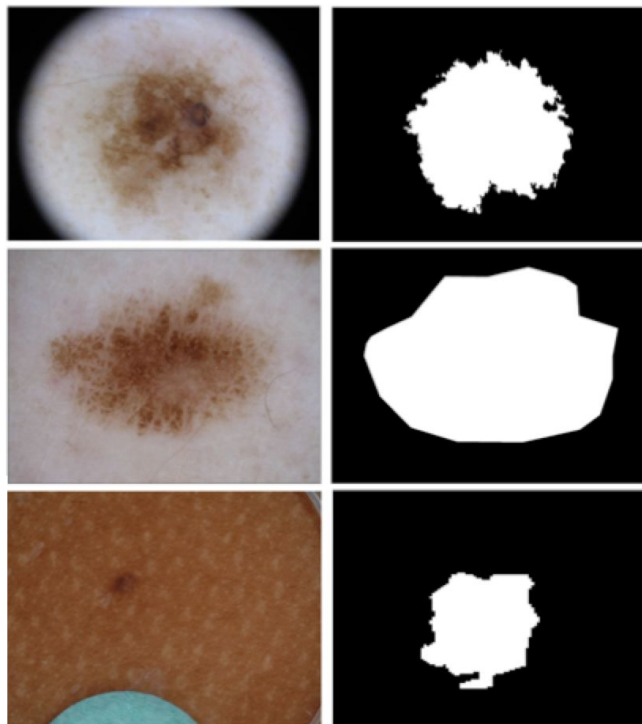


Sample Selection & Inter-annotator Agreement

Lesion Attributes



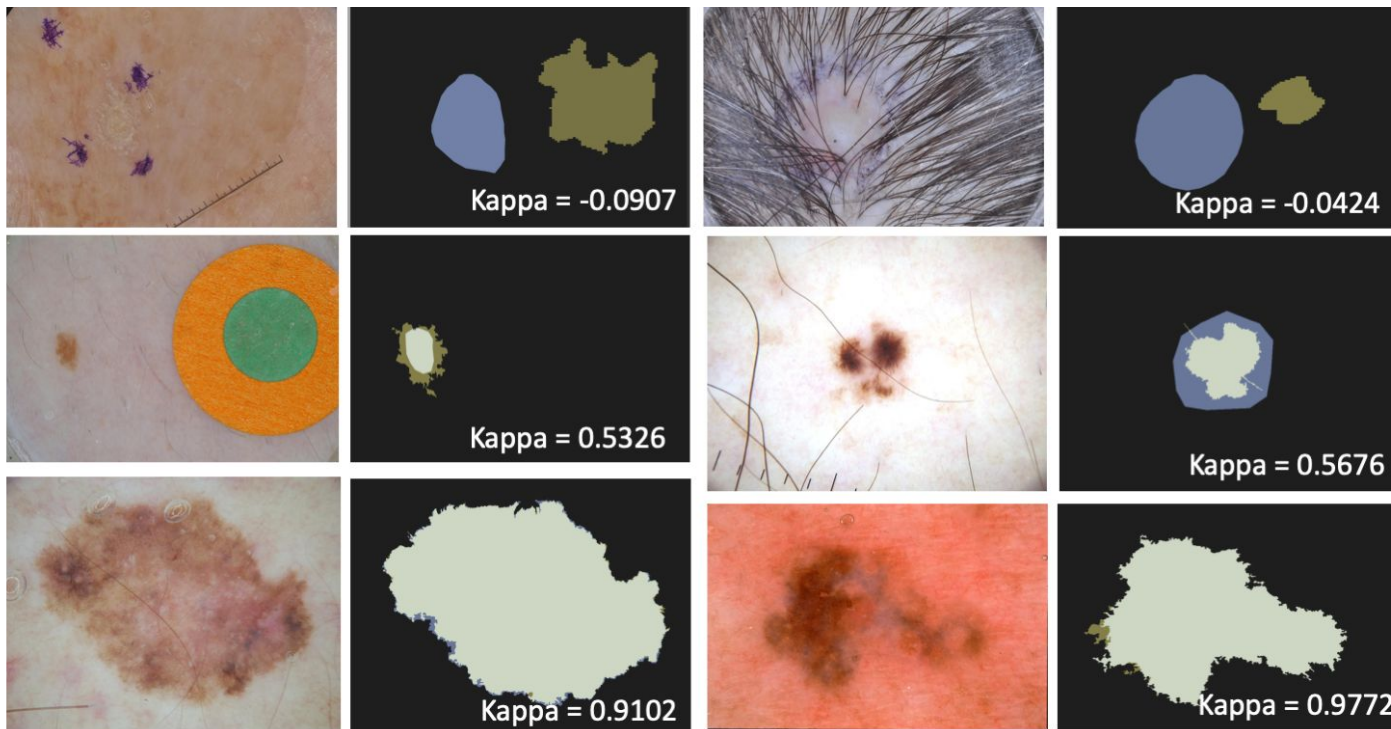
Different Levels of Annotation



Different levels of granularity for image annotation

1. Flood fill controlled by the annotator
2. Manual polygon tracing
3. Fully automated validated by human

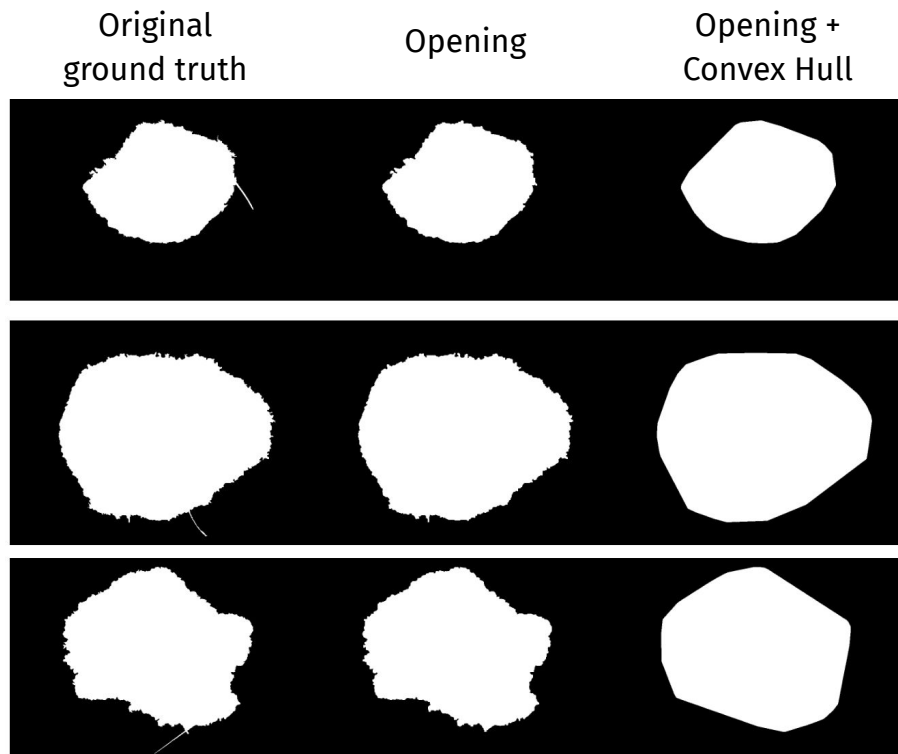
Different Opinions by Physicians



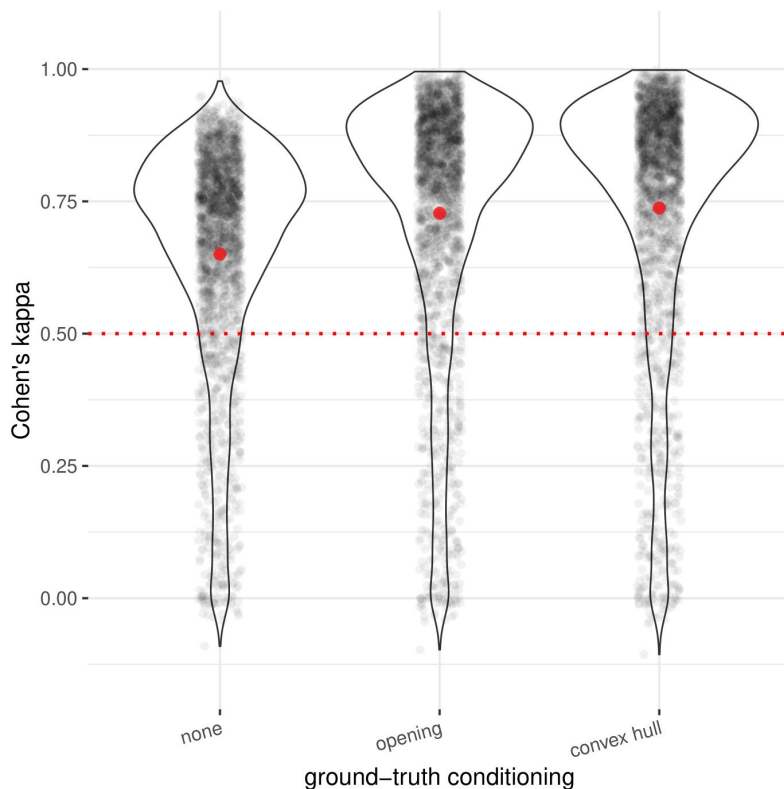
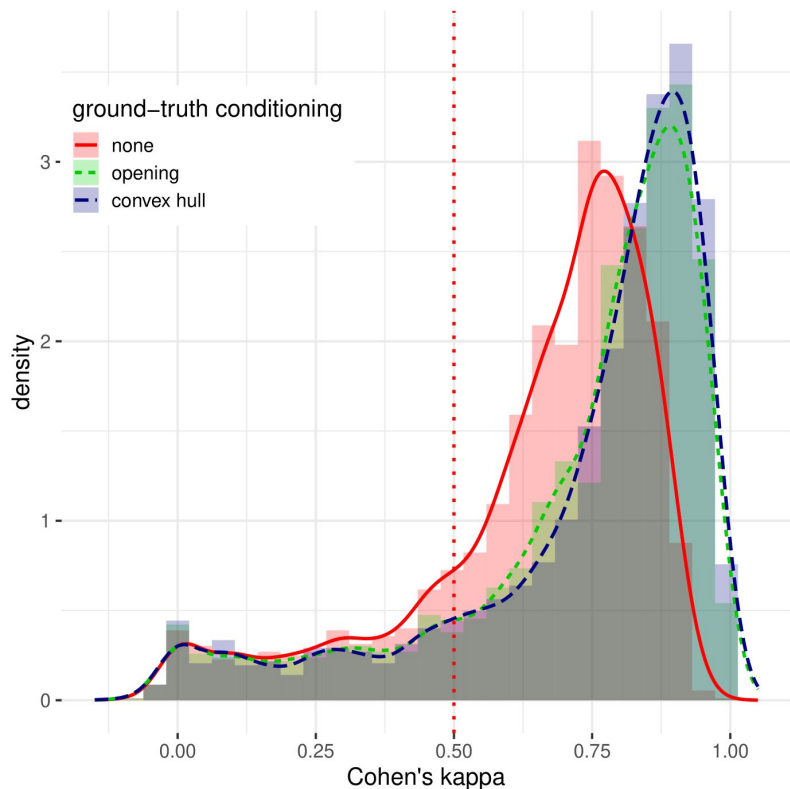


Data Elimination with Label Conditioning

Sample Conditioning



Sample Selection





Materials & Methods

Dataset

ISIC Archive

13,779 images with segmentation

2,233 images with multiple segmentations

Two train datasets:

All samples - 2,233 images

Best samples - 1,808 images ($\kappa \geq 0.5$)

Three test datasets:

ISIC Archive - 2,000 images with single segmentation

PH² - 200 images

Dermofit - 1,300 images

Dataset

ISIC Archive

13,779 images with segmentation

2,233 images with multiple segmentations

Two train datasets:

All samples - 2,233 images

Best samples - 1,808 images ($\kappa \geq 0.5$)

Three test datasets:

ISIC Archive - 2,000 images with single segmentation

PH² - 200 images

Dermofit - 1,300 images

Dataset

ISIC Archive

13,779 images with segmentation

2,233 images with multiple segmentations

Two train datasets:

All samples - 2,233 images

Best samples - 1,808 images ($\kappa \geq 0.5$)

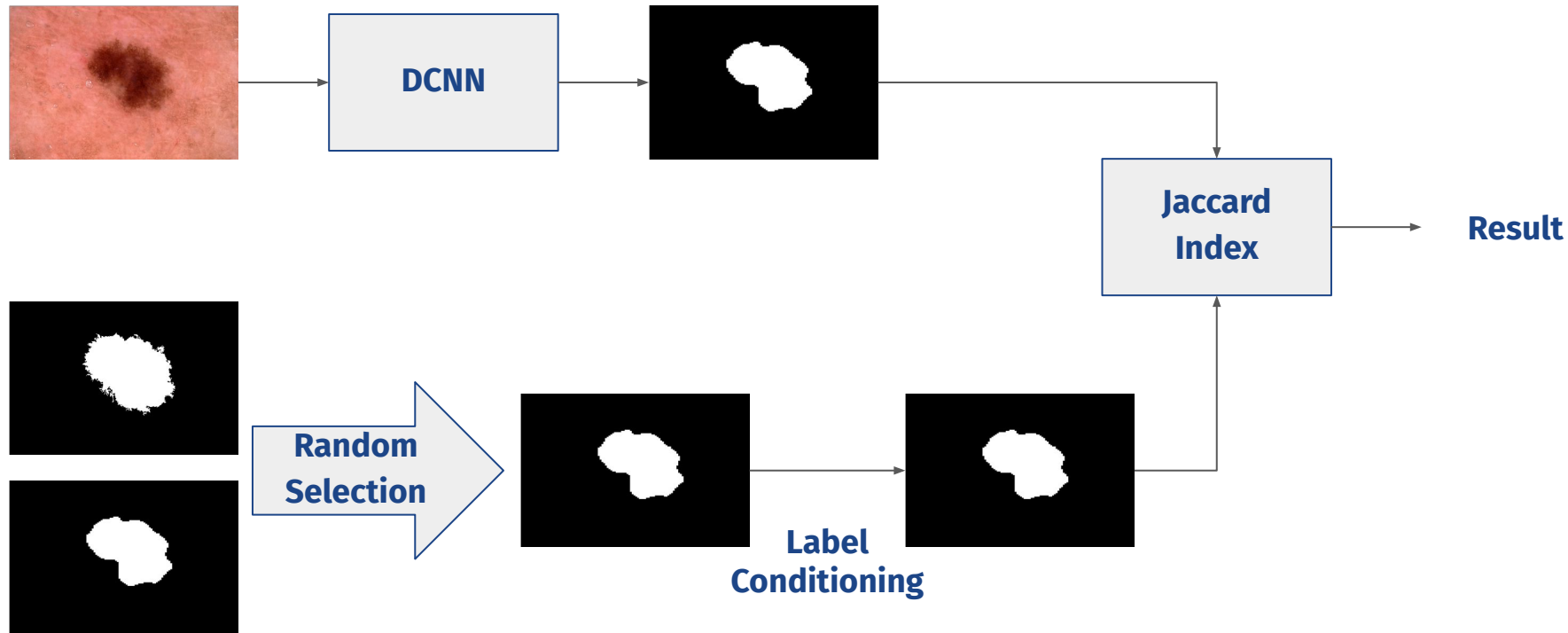
Three test datasets:

ISIC Archive - 2,000 images with single segmentation

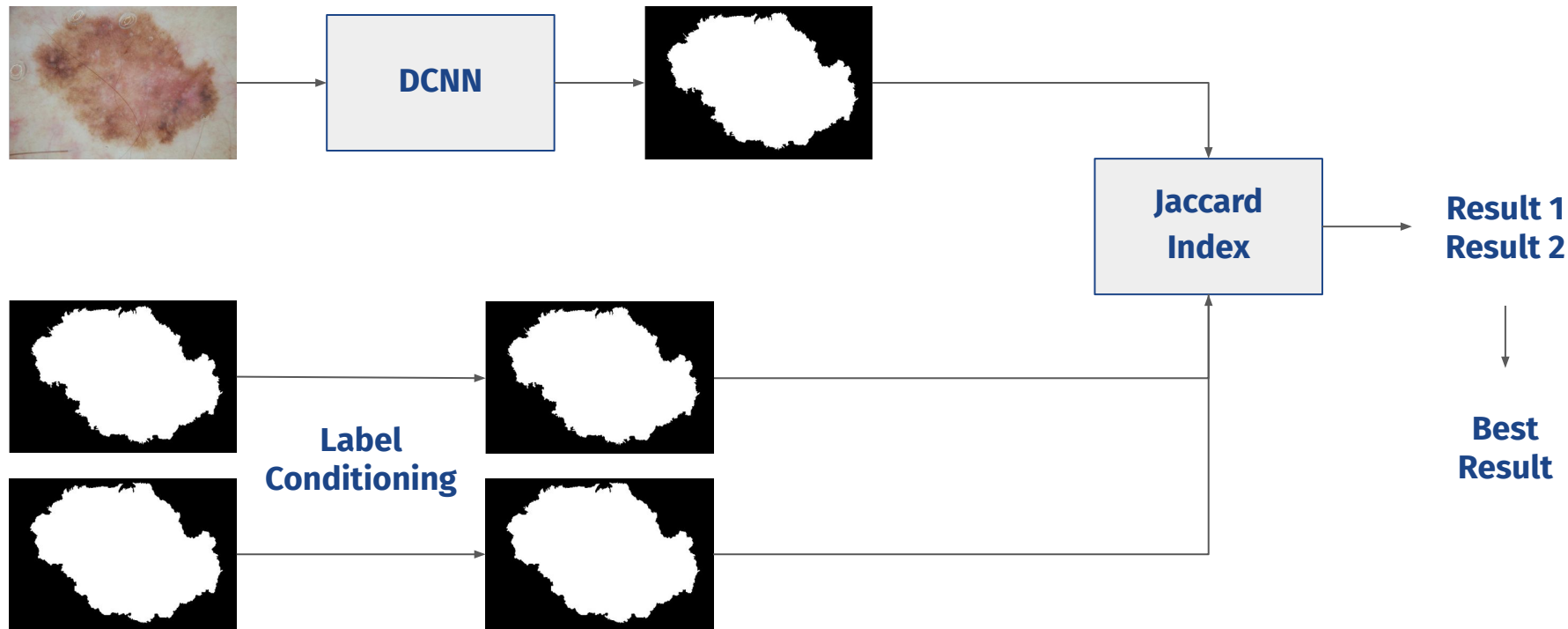
PH² - 200 images

Dermofit - 1,300 images

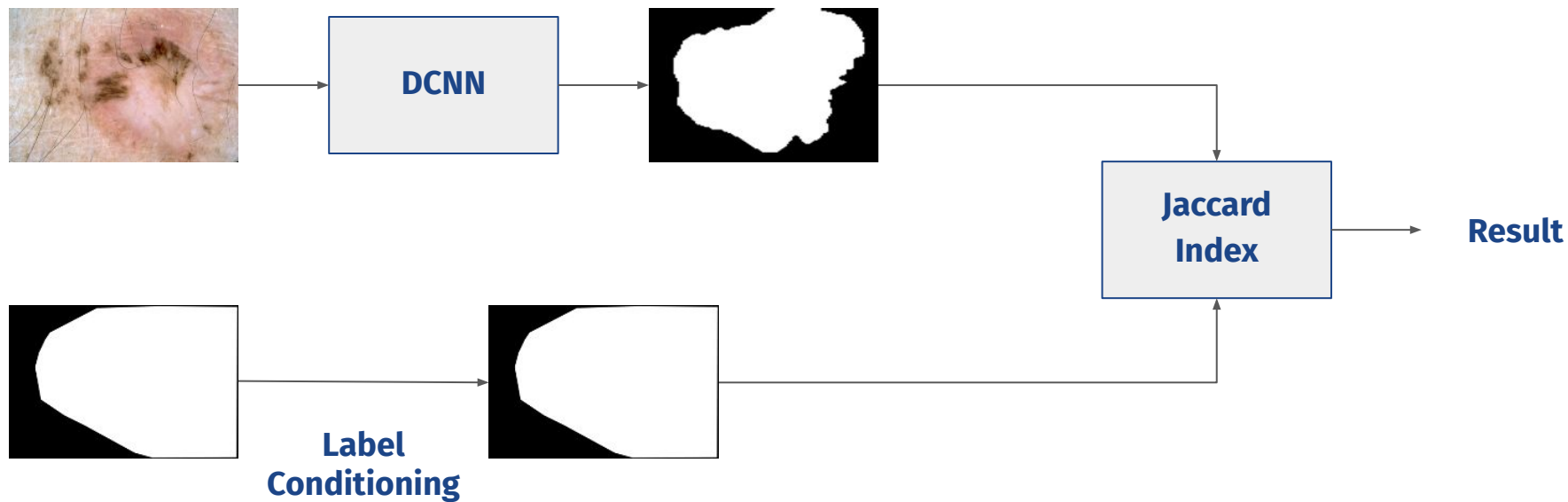
Training



Validation

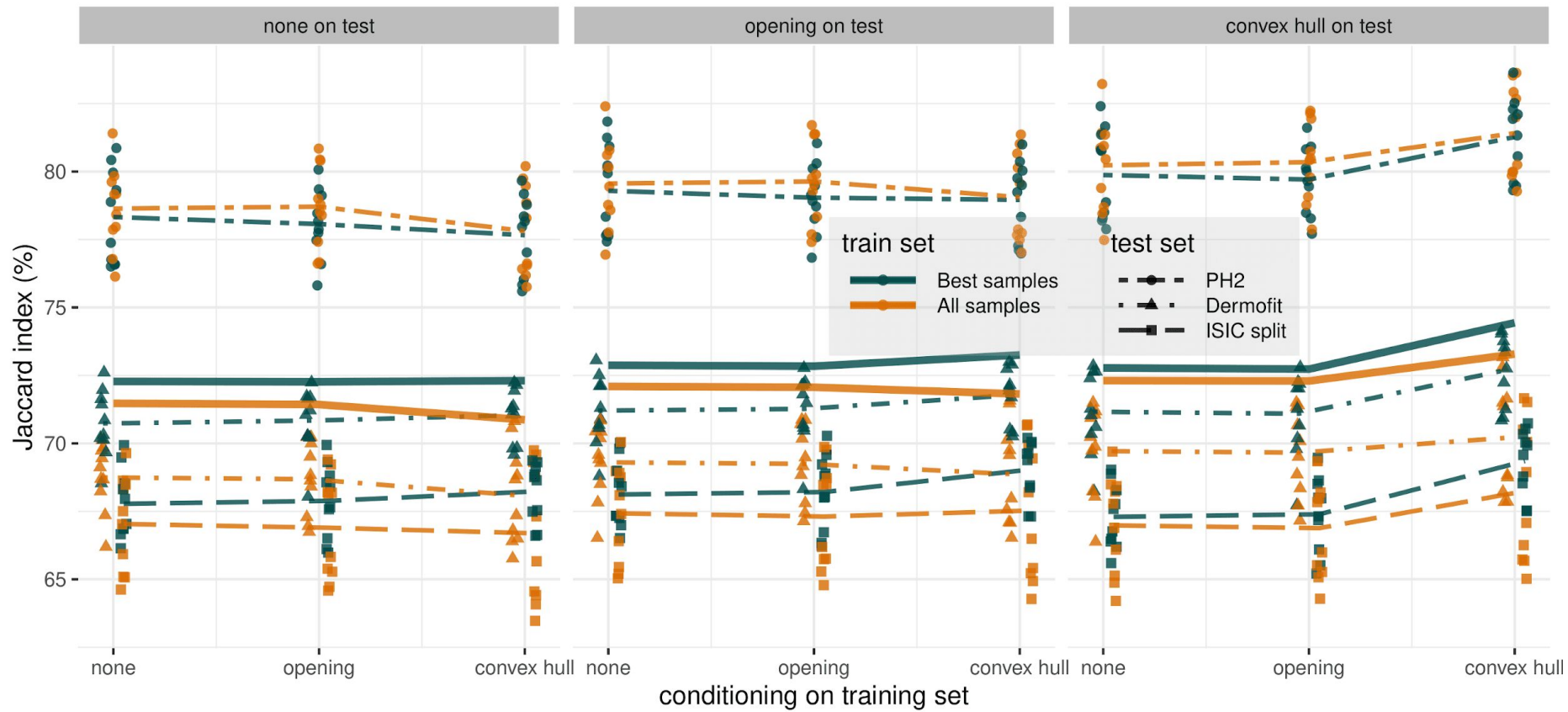


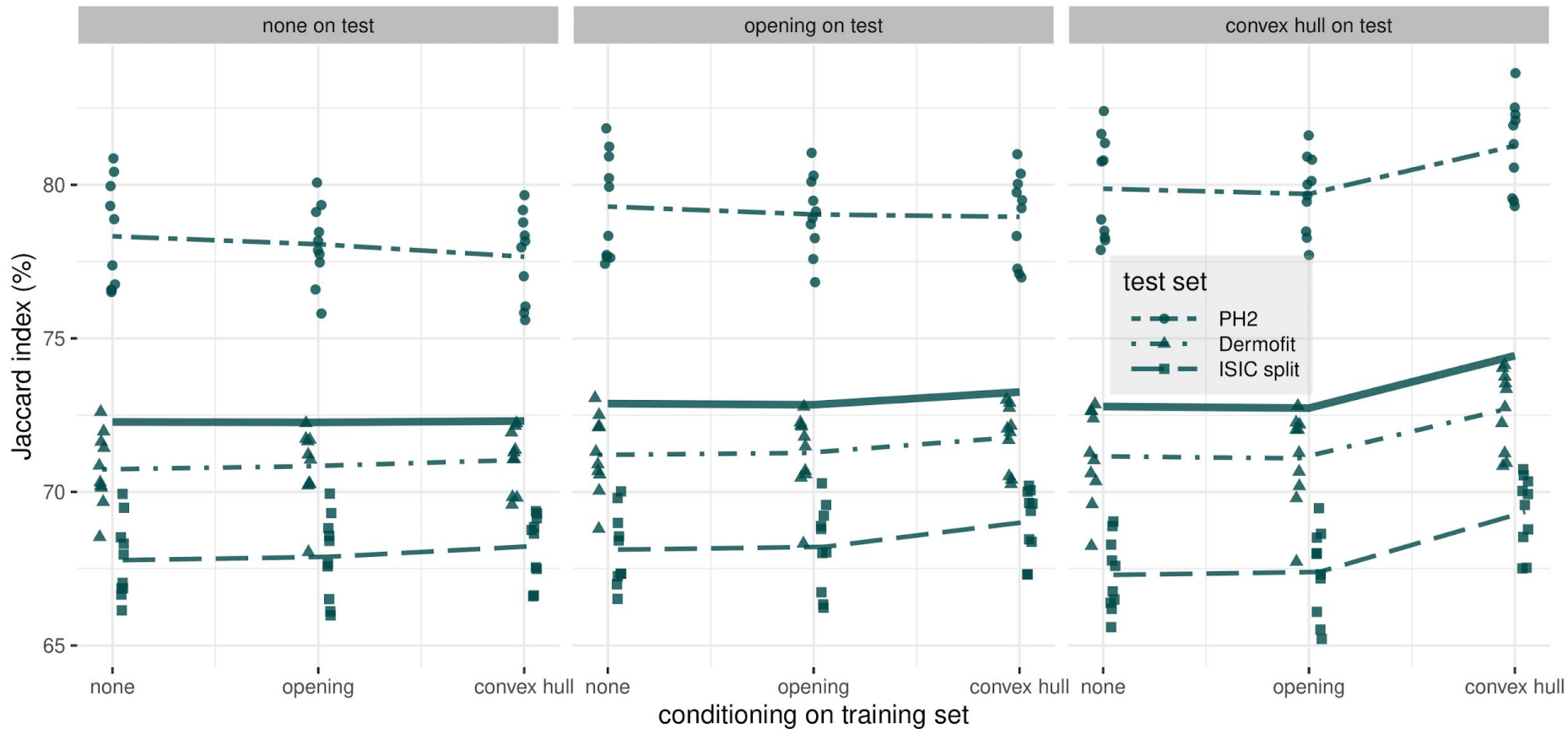
Testing





Results







Conclusion

Conclusion

- Segmentation ground-truths for skin lesion imaging present substantial **inter-annotator disagreement**.
- Withholding **samples with the largest disagreement** may result in significantly improved performance.
- **Removing details** from the segmentation masks may improve the results.
- In the future, we hope to extend our findings to samples with single ground-truth masks, increasing its applicability.

All the skin lesion images and segmentation masks from this presentation were extracted from the ISIC Archive dataset, available at <https://isic-archive.com/>

Code & Data:

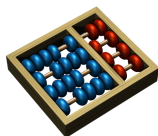


Thank you!

Vinícius Ribeiro vinicius.ribeiro1@gmail.com

Sandra Avila [f](#) [@](#) [@sandraavilabr](#)

Eduardo Valle dovalle@dca.fee.unicamp.br



recod

