Textural descriptors for quantification of mitochondrial states

Matúš Kočalka
Supervised by: Jiří Hladůvka

Faculty of Mathematics, Physics and Informatics Comenius University Bratislava / Slovakia

Abstract

Mitochondria morphology reflects cellular health and is linked to diseases like cancer and diabetes. While neural networks achieve high accuracy across diverse computer vision tasks, they often lack interpretability and demand substantial computational resources, as well as large datasets. We propose a machine learning approach using texture and fractal descriptors to classify yeast mitochondria states. Our method provides an energy-efficient, interpretable alternative that achieves performance comparable to standard neural networks.

Keywords: texture features, mitochondria state, feature selection, neural networks, classification

1 Methodology

Our methodology extracts interpretable classifications from raw microscopy images through these stages:

- Segmentation: We perform instance segmentation of individual cells using Cellpose [6].
- Feature Extraction: We extract IBSI-compliant [3] texture descriptors [5] (such as GLCM, GLRLM, and GLSZM) alongside fractal features, including multifractal analysis and lacunarity.
- Feature Selection: We prioritize model simplicity by narrowing the descriptor set in three distinct phases. First, we eliminate redundancy by grouping correlated features and retaining the representative with the highest mutual information relative to the target variable. The remaining features are then ranked and filtered using Recursive XGBoost [1], followed by an exhaustive search to identify the optimal subset of *n* features that maximizes the performance of a k-nearest neighbor classifier [2].

2 Results and Comparison

We benchmarked our method against a ResNet model [4], trained on the same dataset of isolated yeast cells.

We evaluate the performance using the Macro F1-score, which calculates the mean of the F1-scores for each class to ensure that every mitochondrial state is weighted equally, regardless of its sample size.

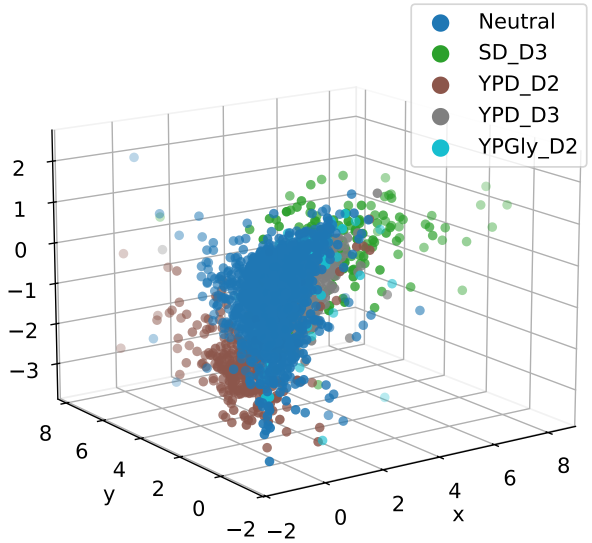

Figure 1: Plot of the dataset using extracted texture features: GLCM cluster prominence (x), GLCM difference variance (y), and intensity volume histogram (z).

Our proposed method achieved a Macro F1-score of 67%, approaching the 77% produced by the ResNet model. The interpretable representation of the classification is illustrated in Figure 1.

3 Conclusion

In our experiments, our proposed method was able to approximate the results of a standard neural network model, showing that for isolated classification tasks, traditional machine learning approaches and descriptors remain viable alternatives, with potential advantages in terms of energy efficiency and explainability.

Acknowledgments: This work was supported by the European Union's NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project 09I03-03-V04-00363.

References

[1] Tianqi Chen and Carlos Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794. ACM, August 2016.

[2] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theor.*, 13(1):21–27, September 2006.

[3] Zwanenburg Alex et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*, 295(2):328–338, May 2020.

[4] Kaiming He, X. Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2015.

[5] Anne Humeau-Heurtier. Texture feature extraction methods: A survey. *IEEE Access*, 7:8975–9000, 2019.

[6] Carsen Stringer, Tim Wang, Michalis Michaelos, and Marius Pachitariu. Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*, 18:100 – 106, 2020.