Buds morphometrics – How to distinguish and predict tree species with images of buds

Felix Nößler

Bioimage Analysis and Extended Phenotyping, Dr. Christian Kappel

March 13, 2021

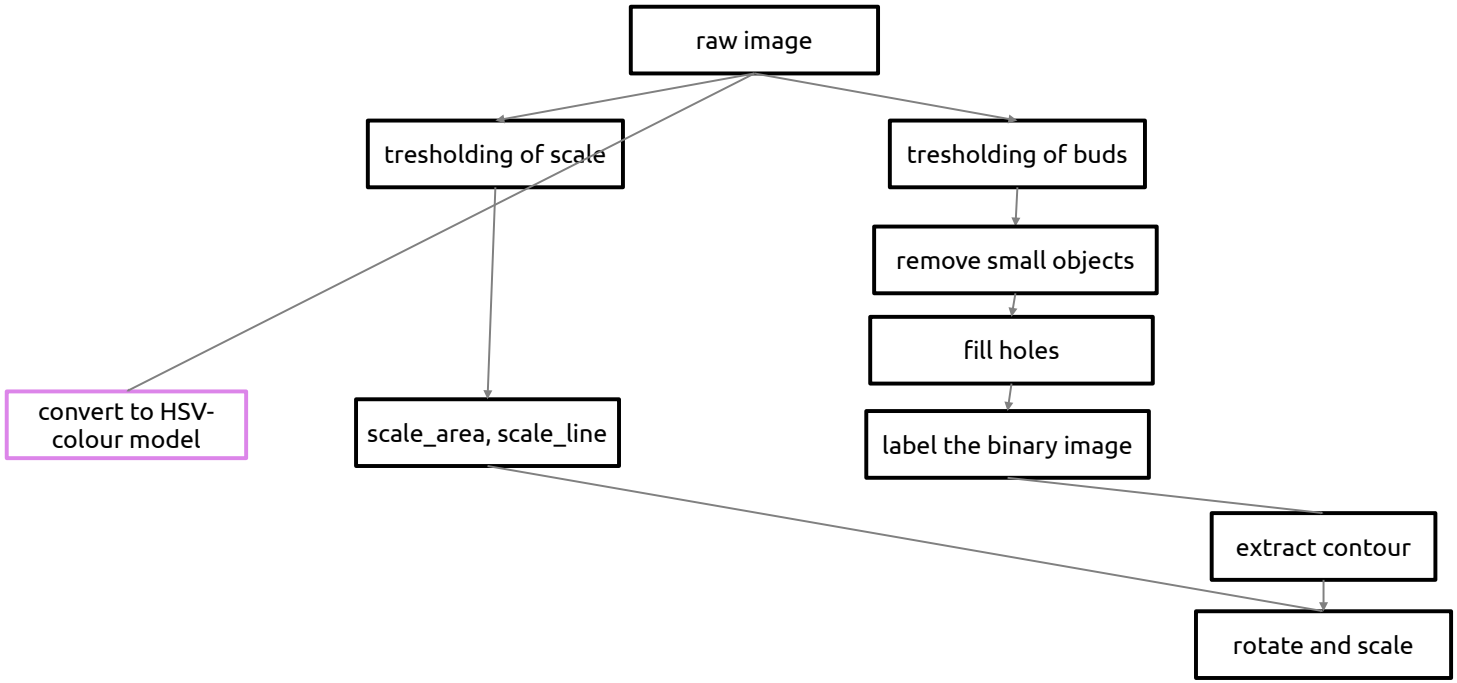
Overview

Introduction

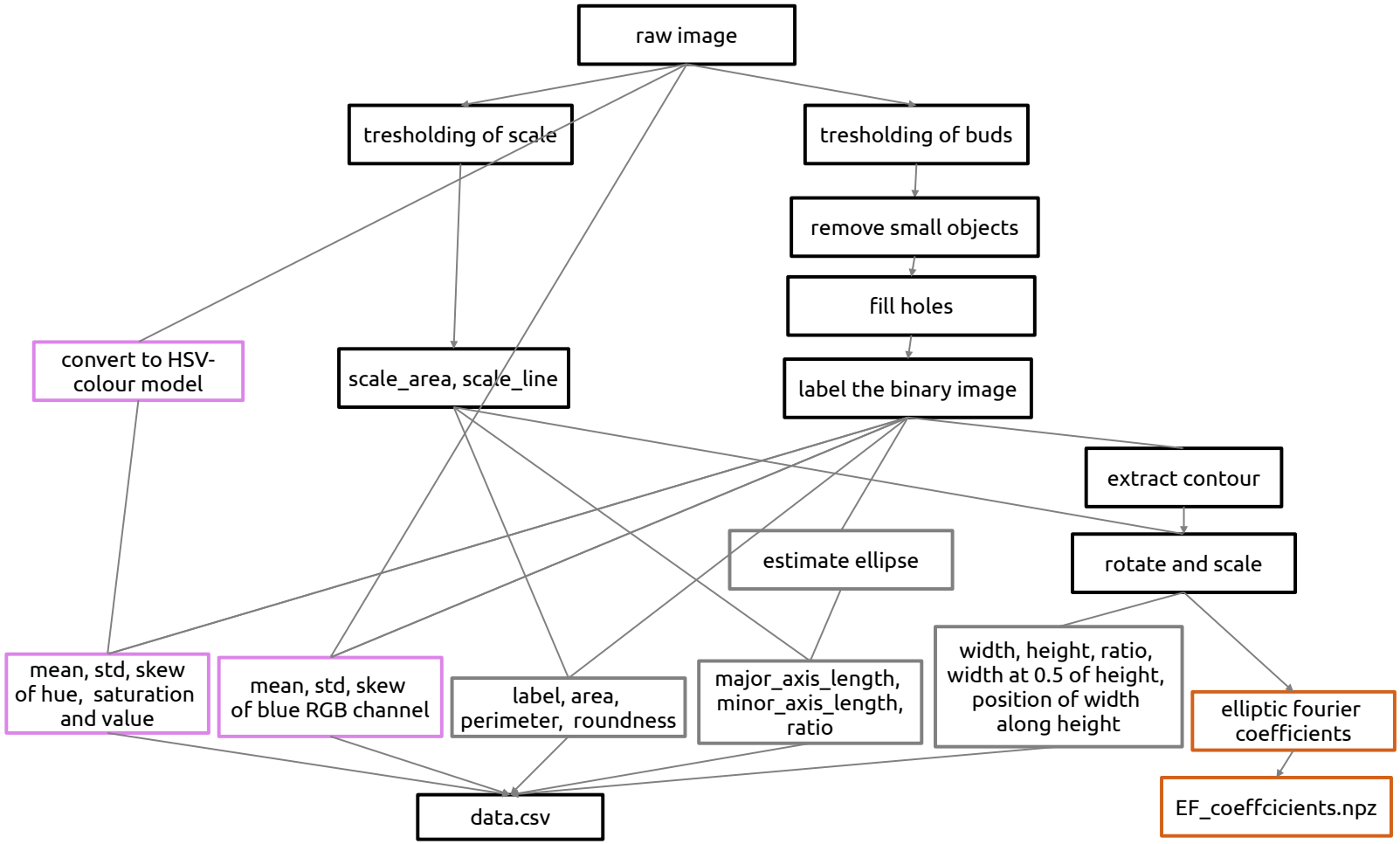
Figure: Setup for taking images

- Goal was to describe buds with image descriptors and make species predictions
- 46 pictures of 1422 buds

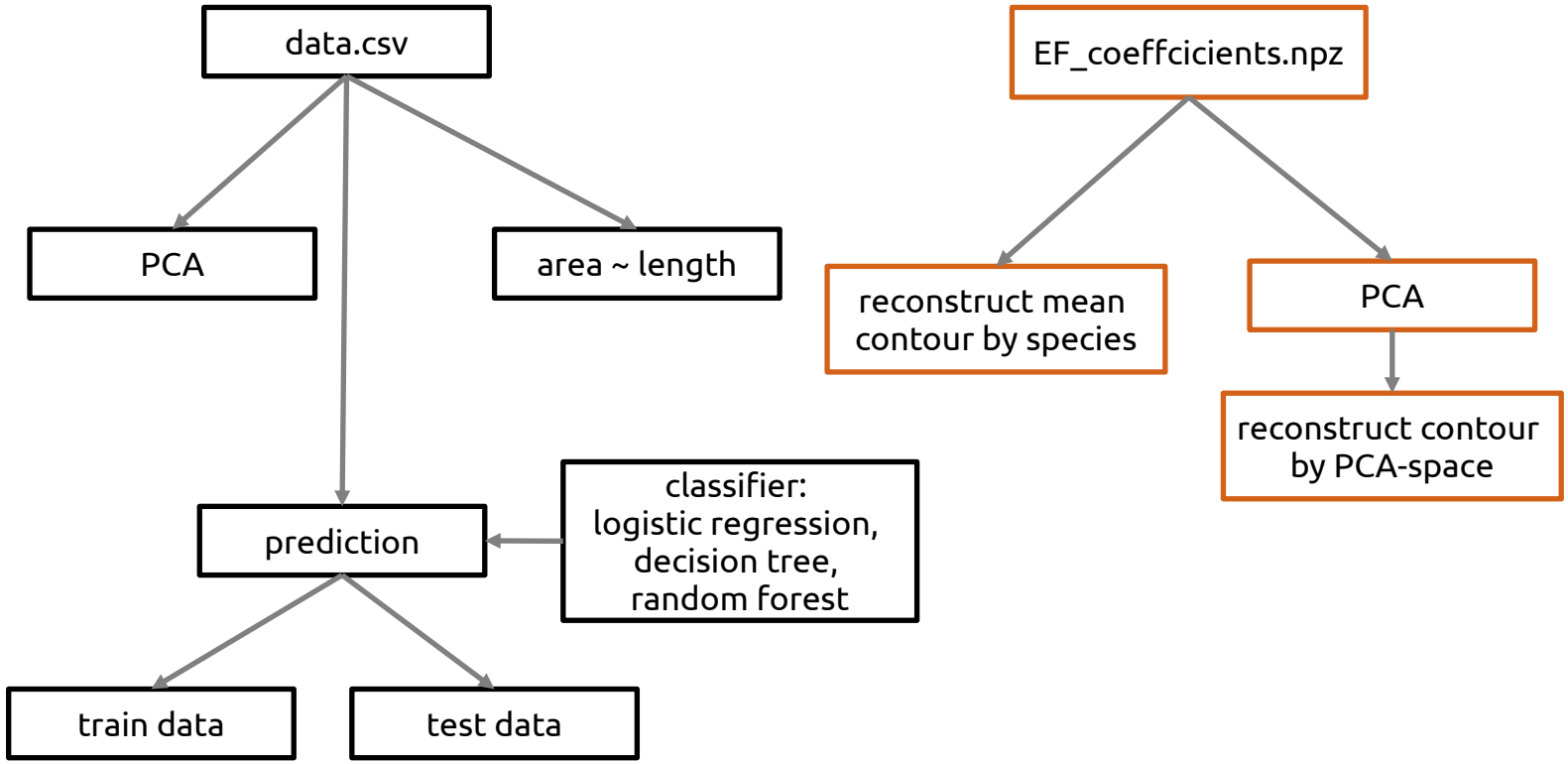
Analysis workflow



Analysis workflow



Analysis workflow



Results

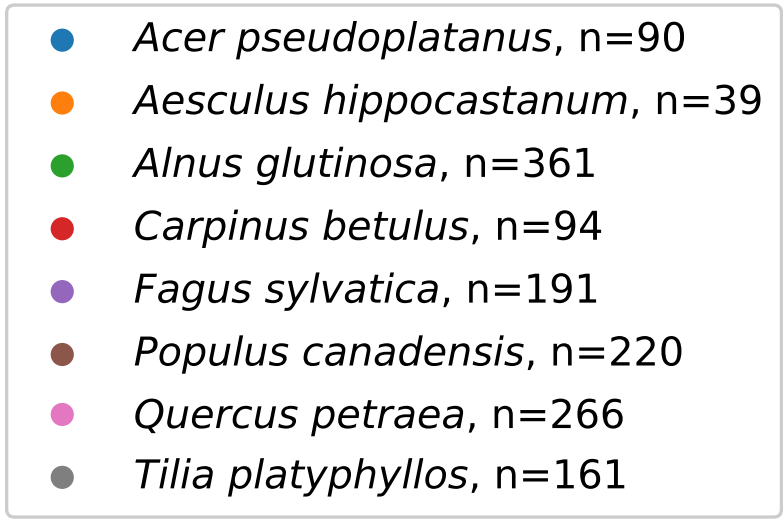


Figure: Colours used in all folowing images

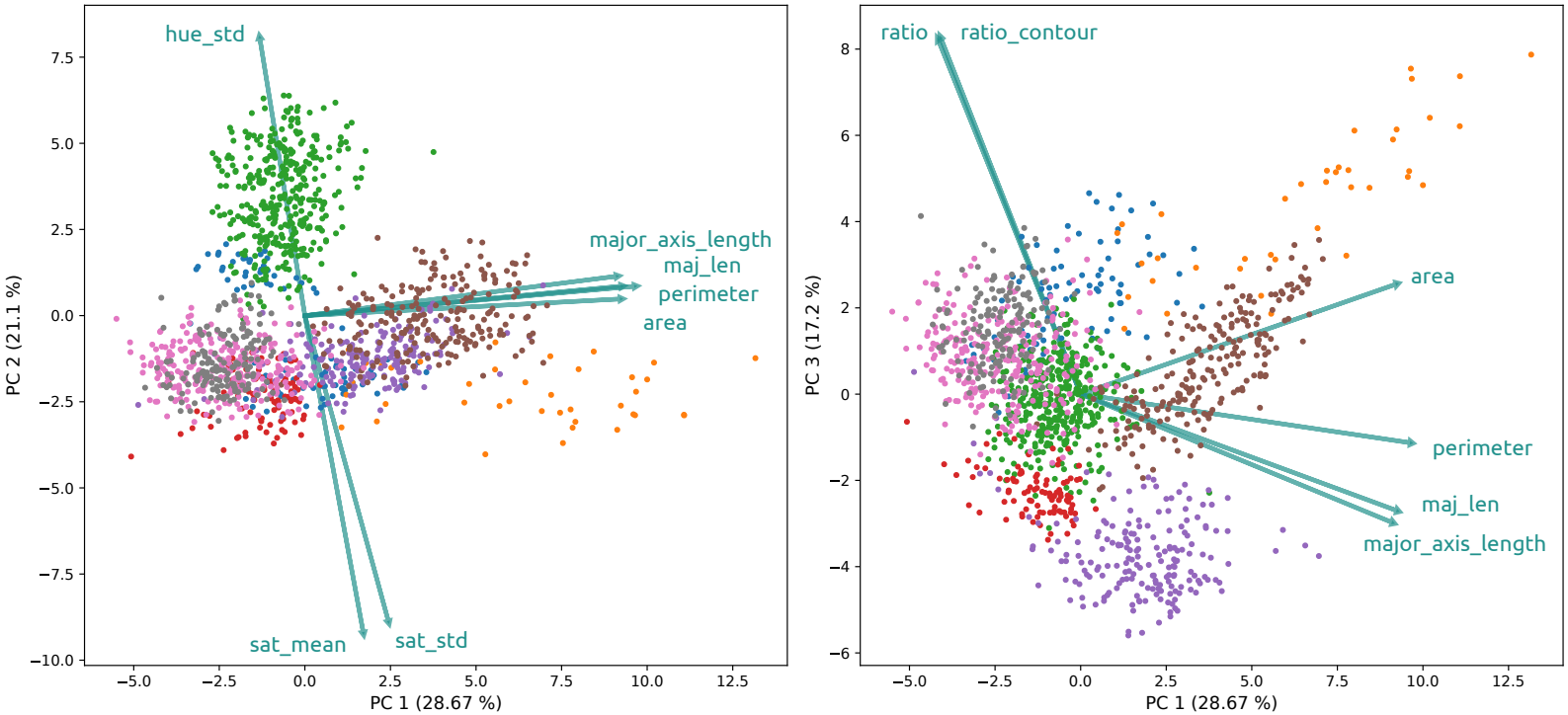


Figure: Principal component analysis of the complete data set

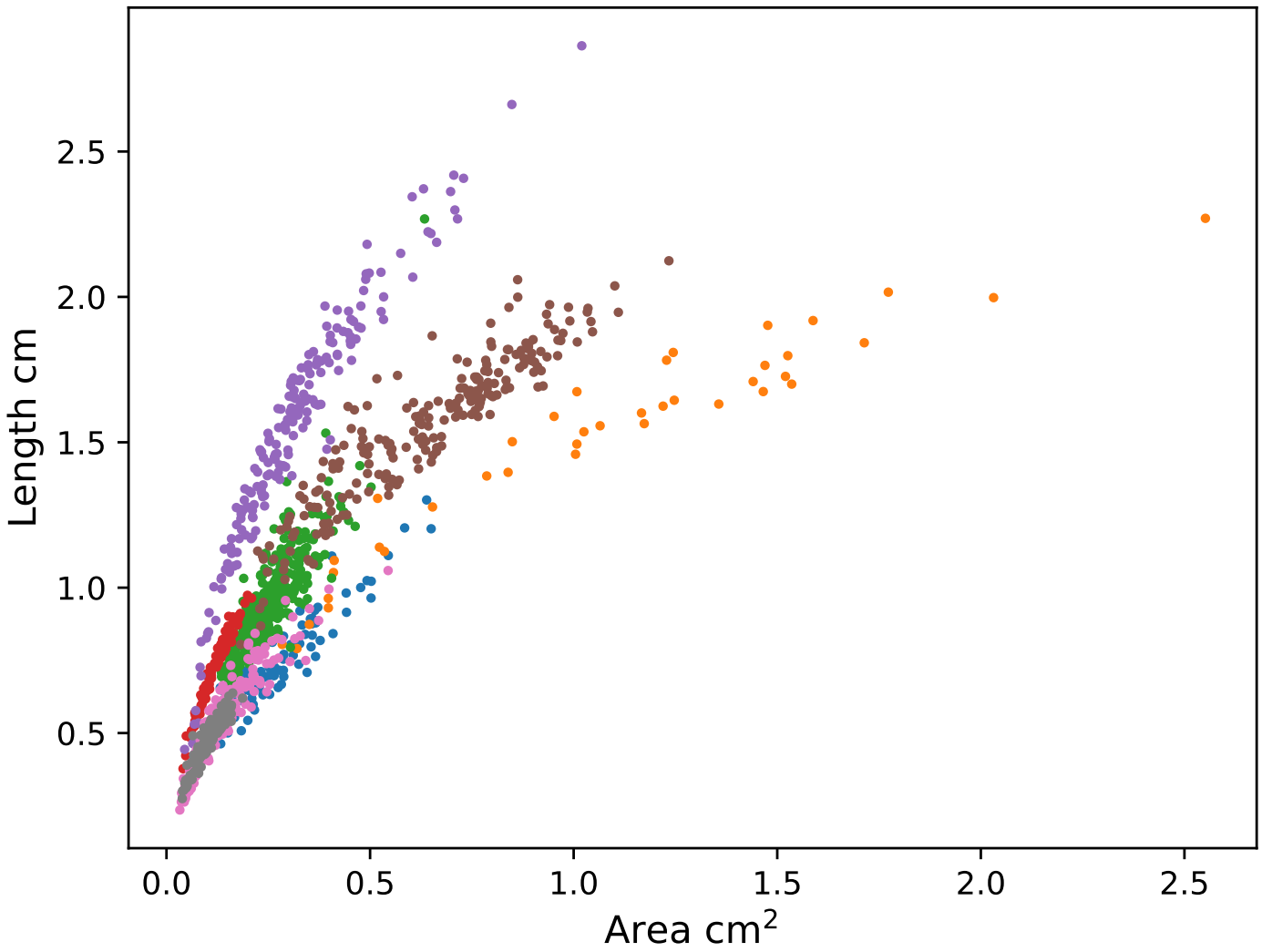


Figure: Area and major axis length of estimated ellipse of all buds

Outline analysis

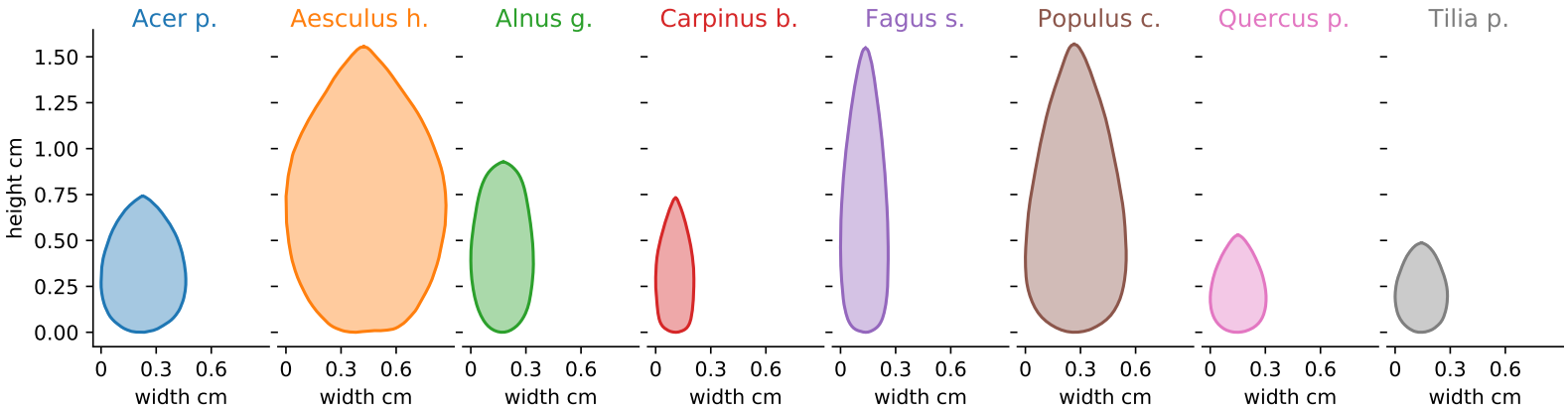


Figure: Mean contours of all species reconstructed from elliptic fourier coefficients

Outline analysis

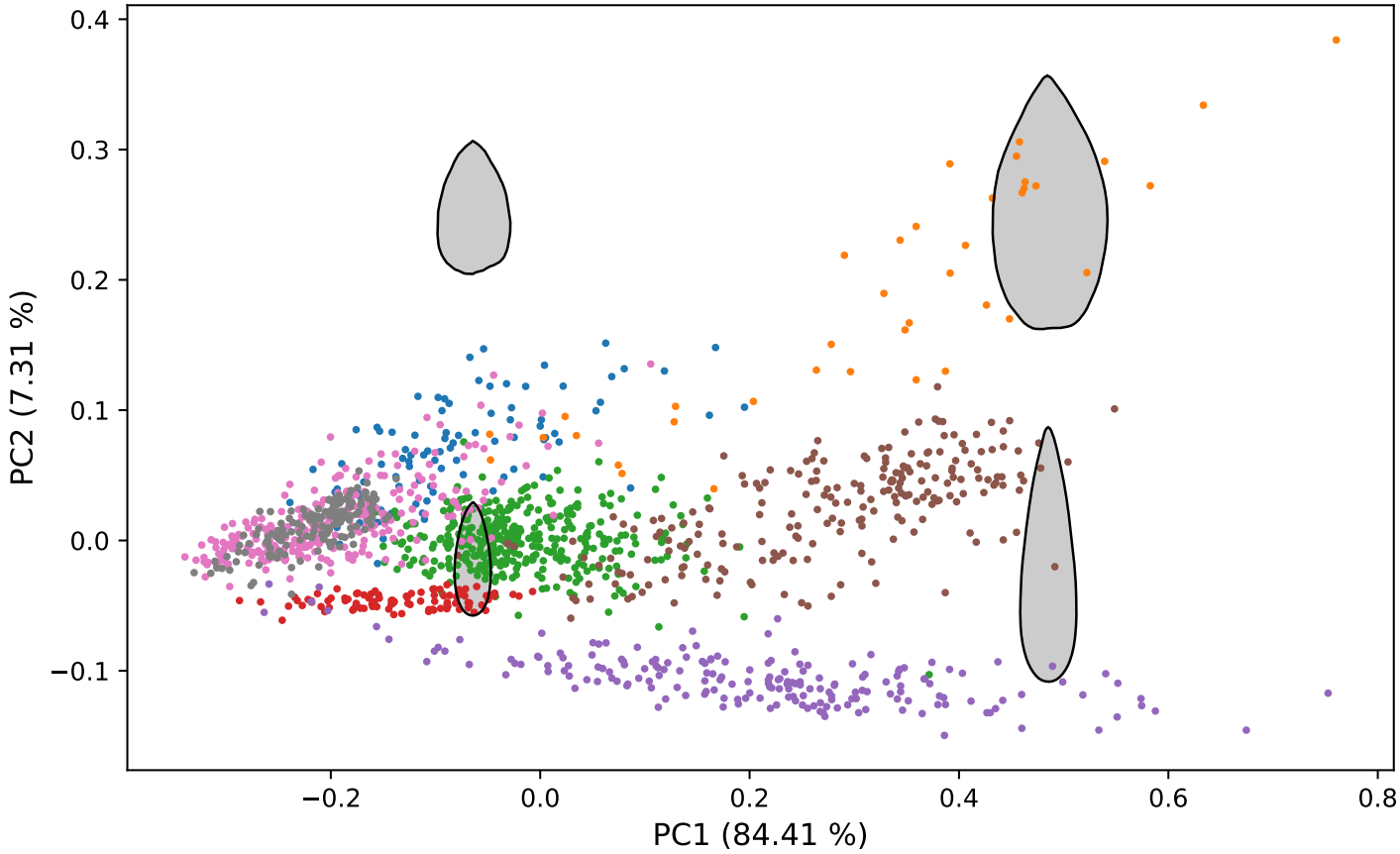


Figure: Principal component analysis of the elliptic fourier coefficients of the shape of buds

Prediction of species

Table: Result of 200 times predicting the species with the three methods, numbers for the correct assignment are given in percent ± standard deviation

Method	Train data	Test data
Logistic regression	97.4 ± 0.3	95.9 ± 0.8
Decision Tree	99.3 ± 0.3	92.8 ± 1.3
Random Forest	99.9 ± 0.1	96.2 ± 0.9

Prediction of species

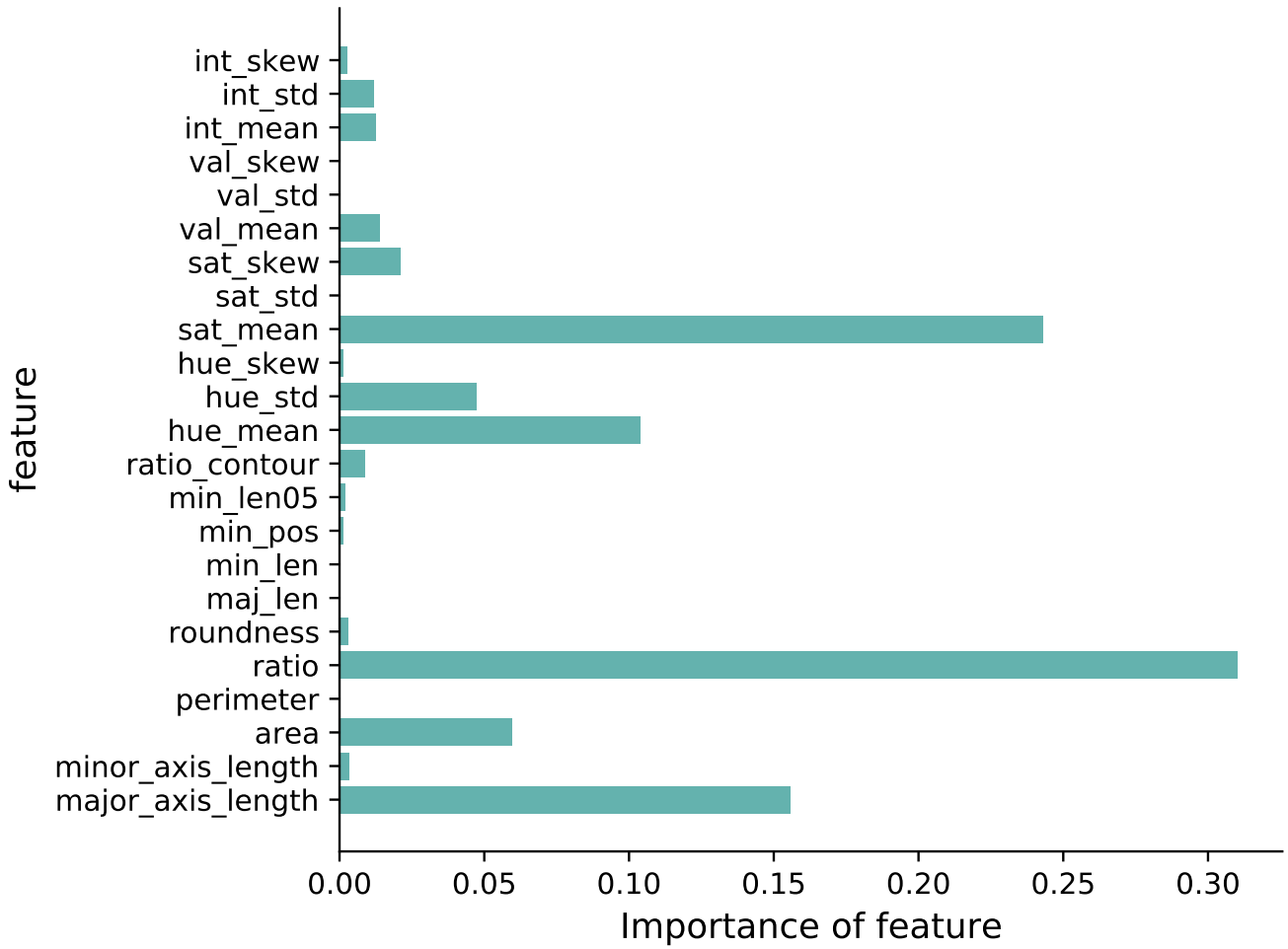


Figure: Importance of image descriptors for the decision tree

Discussion

- Segmentation of buds was done on a white paper

Discussion

- Segmentation of buds was done on a white paper
- Test performance of prediction with higher number of species, buds from different localities

Discussion

- Segmentation of buds was done on a white paper
- Test performance of prediction with higher number of species, buds from different localities
- Do the colour values show real differences?

Discussion

- Segmentation of buds was done on a white paper
- Test performance of prediction with higher number of species, buds from different localities
- Do the colour values show real differences?
- Lengths from estimated ellipse <> lengths from shape

Conclusion

- Used setup worked overall well, prediction showed high accuracy

Conclusion

- Used setup worked overall well, prediction showed high accuracy
- next steps:
  - analyse images taken directly on the tree outside
  - add more species
  - add pictures from more locations
  - implement convolutional network / deep learning approach

Conclusion

- Used setup worked overall well, prediction showed high accuracy
- next steps:
  - analyse images taken directly on the tree outside
  - add more species
  - add pictures from more locations
  - implement convolutional network / deep learning approach
- most important features: overall size (e.g. major axis length), colour values (e.g. mean saturation) and the ratio between the minor and the major axis length

References

Used packages in Python 3.8:

Blidh, H. PyEFD. https://github.com/hbldh/pyefd (2020).

Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585,** 357–362. https://doi.org/10.1038/s41586-020-2649-2 (2020).

Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **9,** 90–95 (2007).

Pandas development team, T. *pandas-dev/pandas: Pandas.* Version latest. 2020. https://doi.org/10.5281/zenodo.3509134.

Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12,** 2825–2830. https://jmlr.org/papers/v12/pedregosa11a.html (2011).

Terpilowski, M. A. scikit-posthocs: Pairwise multiple comparison tests in Python. *Journal of Open Source Software* **4,** 1169. https://doi.org/10.21105/joss.01169 (2019).

Van der Walt, S. *et al.* scikit-image: image processing in Python. *PeerJ* **2,** e453. https://doi.org/10.7717/peerj.453 (2014).

Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **17,** 261–272 (2020).

Thank you for your attention! :)

Do you have questions or comments?

All images and the Python code are available at: https://github.com/FelixNoessler/Buds-morphometrics