Traditional Machine Learning for Flower Classification: Improvements on Oxford Flowers 102

Anonymous ACL submission

Abstract

This paper explores flower classification on the Oxford Flowers 102 dataset using traditional machine learning techniques, revisiting the seminal work of Nilsback and Zisserman (2008). The original approach employed handcrafted features—HSV for color, SIFT for texture and boundary, and HOG for structure—combined with Multi-Kernel SVM (MKL), achieving 72.8% accuracy across 102 classes. We replicate this pipeline, including image segmentation, feature extraction, Bag-of-Words representation, and MKL-based classification. To address the computational inefficiency of the original segmentation method, we propose a hybrid approach combining K-Means clustering and contour detection, which is 5-10x faster while preserving quality for feature extraction. We further enhance the feature set by incorporating Local Binary Patterns (LBP) to capture texture details and optimize MKL parameters using the EasyMKL algorithm. Our replicated pipeline achieves an accuracy of [TBD]%, closely matching the original performance. With our improvements, we attain an accuracy of [TBD]%, surpassing the baseline. This work underscores the viability of traditional computer vision methods in modern contexts, offering insights into their practical implementation and optimization for large-scale classification tasks.

1 Introduction

This paper explores flower classification on the Oxford Flowers 102 dataset love using traditional machine learning techniques, revisiting the seminal work of Nilsback and Zisserman (2008). The original approach employed handcrafted features—HSV for color, SIFT for texture and boundary, and HOG for structure—combined with Multi-Kernel SVM (MKL), achieving 72.8% accuracy across 102 classes. We replicate this pipeline, including image segmentation, feature extraction, Bag-of-Words representation, and MKL-based classification. To address the computational inefficiency of the original segmentation method, we propose a hybrid approach combining K-Means clustering and contour detection, which is 5-10x faster while preserving quality for feature extraction. We further enhance the feature set by incorporating Local Binary Patterns (LBP) to capture texture details and optimize MKL parameters using the EasyMKL algorithm. Our replicated pipeline achieves an accuracy of [TBD]%, closely matching the original performance. With our improvements, we attain an accuracy of [TBD]%, surpassing the baseline. This work underscores the viability of traditional computer vision methods in modern contexts, offering insights into their practical implementation and optimization for large-scale classification tasks.

2 Related Work

3 Methodology

3.1 Problem Statement

4 Experimental Settings

4.1 Datasets

4.2 Baselines

4.3 Settings

4.4 Evaluation metrics

5 Result and Discussion

5.1 Performance Comparison

Limitations

ACL 2023 requires that all submissions have a section titled “Limitations”, to discuss the limitations of the paper as a complement to the discussion of strengths in the main text. This section should occur after the conclusion, but before the references. It will not count towards the page limit. The discussion of limitations is mandatory. Papers without a limitation section will be desk-rejected without review. While we are open to different types of limitations, just mentioning that a set of results have been shown for English only probably does not reflect what we expect. Mentioning that the method works mostly for languages with limited morphology, like English, is a much better alternative. In addition, limitations such as low scalability to long text, the requirement of large GPU resources, or other things that inspire crucial further investigation are welcome.

Ethics Statement

Scientific work published at ACL 2023 must comply with the ACL Ethics Policy.[1] We encourage all authors to include an explicit ethics statement on the broader impact of the work, or other ethical considerations after the conclusion but before the references. The ethics statement will not count toward the page limit (8 pages for long, 4 pages for short papers).

Acknowledgements

This document has been adapted by Jordan Boyd-Graber, Naoaki Okazaki, Anna Rogers from the style files used for earlier ACL, EMNLP and NAACL proceedings, including those for EACL 2023 by Isabelle Augenstein and Andreas Vlachos, EMNLP 2022 by Yue Zhang, Ryan Cotterell and Lea Frermann, ACL 2020 by Steven Bethard, Ryan Cotterell and Rui Yan, ACL 2019 by Douwe Kiela and Ivan Vulić, NAACL 2019 by Stephanie Lukin and Alla Roskovskaya, ACL 2018 by Shay Cohen, Kevin Gimpel, and Wei Lu, NAACL 2018 by Margaret Mitchell and Stephanie Lukin, BibTEX suggestions for (NA)ACL 2017/2018 from Jason Eisner, ACL 2017 by Dan Gildea and Min-Yen Kan, NAACL 2017 by Margaret Mitchell, ACL 2012 by Maggie Li and Michael White, ACL 2010 by Jing-Shin Chang and Philipp Koehn, ACL 2008 by Johanna D. Moore, Simone Teufel, James Allan, and Sadaoki Furui, ACL 2005 by Hwee Tou Ng and Kemal Oflazer, ACL 2002 by Eugene Charniak and Dekang Lin, and earlier ACL and EACL formats written by several people, including John Chen, Henry S. Thompson and Donald Walker. Additional elements were taken from the formatting instructions of the *International Joint Conference on Artificial Intelligence* and the *Conference on Computer Vision and Pattern Recognition*.

[1] https://www.aclweb.org/portal/content/acl-code-ethics

References

Alfred V. Aho and Jeffrey D. Ullman. 1972. *The Theory of Parsing, Translation and Compiling*, volume 1. Prentice-Hall, Englewood Cliffs, NJ.

American Psychological Association. 1983. *Publications Manual*. American Psychological Association, Washington, DC.

Rie Kubota Ando and Tong Zhang. 2005. A framework for learning predictive structures from multiple tasks and unlabeled data. *Journal of Machine Learning Research*, 6:1817–1853.

Galen Andrew and Jianfeng Gao. 2007. Scalable training of L_1-regularized log-linear models. In *Proceedings of the 24th International Conference on Machine Learning*, pages 33–40.

Isabelle Augenstein, Tim Rocktäschel, Andreas Vlachos, and Kalina Bontcheva. 2016. Stance detection with bidirectional conditional encoding. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 876–885, Austin, Texas. Association for Computational Linguistics.

Ashok K. Chandra, Dexter C. Kozen, and Larry J. Stockmeyer. 1981. Alternation. *Journal of the Association for Computing Machinery*, 28(1):114–133.

Marc-Alexandre Côté, Xingdi Yuan, and Prithviraj Ammanabrolu, editors. 2022. *Proceedings of the 3rd Wordplay: When Language Meets Games Workshop (Wordplay 2022)*. Association for Computational Linguistics, Seattle, United States.

James Goodman, Andreas Vlachos, and Jason Naradowsky. 2016. Noise reduction and targeted exploration in imitation learning for Abstract Meaning Representation parsing. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1–11, Berlin, Germany. Association for Computational Linguistics.

Dan Gusfield. 1997. *Algorithms on Strings, Trees and Sequences*. Cambridge University Press, Cambridge, UK.

Mary Harper. 2014. Learning from 26 languages: Program management and science in the babel program. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, page 1, Dublin, Ireland. Dublin City University and Association for Computational Linguistics.

Mohammad Sadegh Rasooli and Joel R. Tetreault. 2015. Yara parser: A fast and accurate dependency parser. *Computing Research Repository*, arXiv:1503.06733. Version 2.

A Example Appendix

This is a section in the appendix.