Interactive Attribution-based Explanations for Image Segmentation

C. Humer, M. Elharty, A. Hinterreiter, and M. Streit
Johannes Kepler University Linz, Austria

JOHANNES KEPLER UNIVERSITÄT LINZ

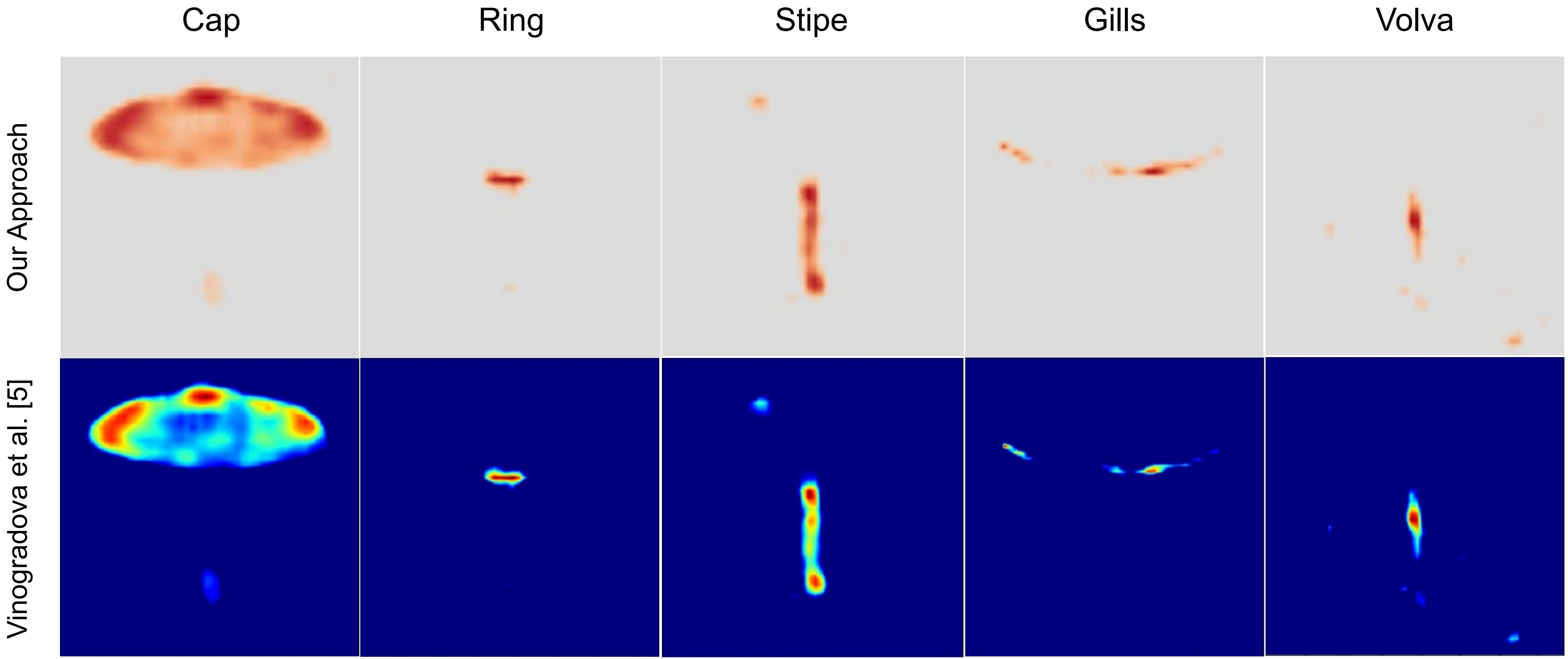
Visual Data Science Lab

LINZ INSTITUTE OF TECHNOLOGY

Der Wissenschaftsfonds.

LAND OBERÖSTERREICH

Comparison of Approaches

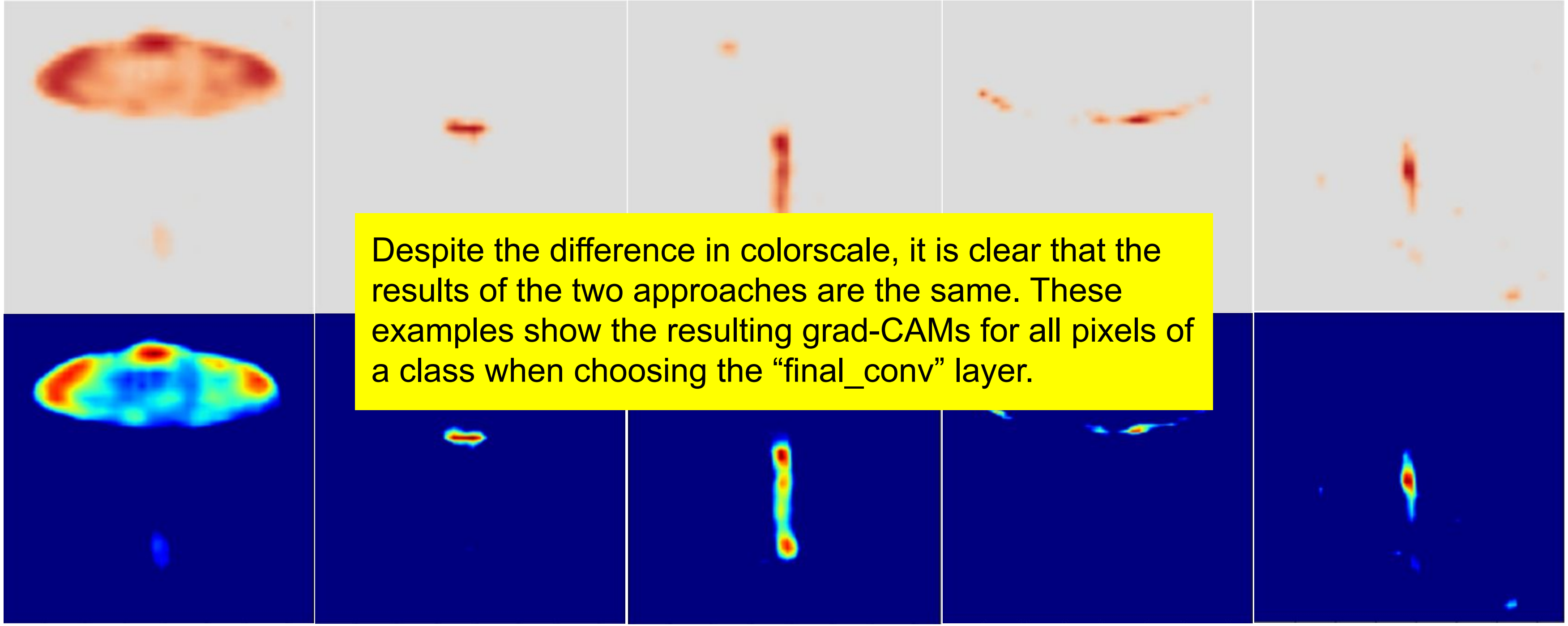


Comparison of Approaches

Cap | Ring | Stipe | Gills | Volva

Our Approach

Vinogradova et al. [5]



Despite the difference in colorscale, it is clear that the results of the two approaches are the same. These examples show the resulting grad-CAMs for all pixels of a class when choosing the “final_conv” layer.

Example Use Case: Mushroom Segmentation

Input Image

Predicted Mask

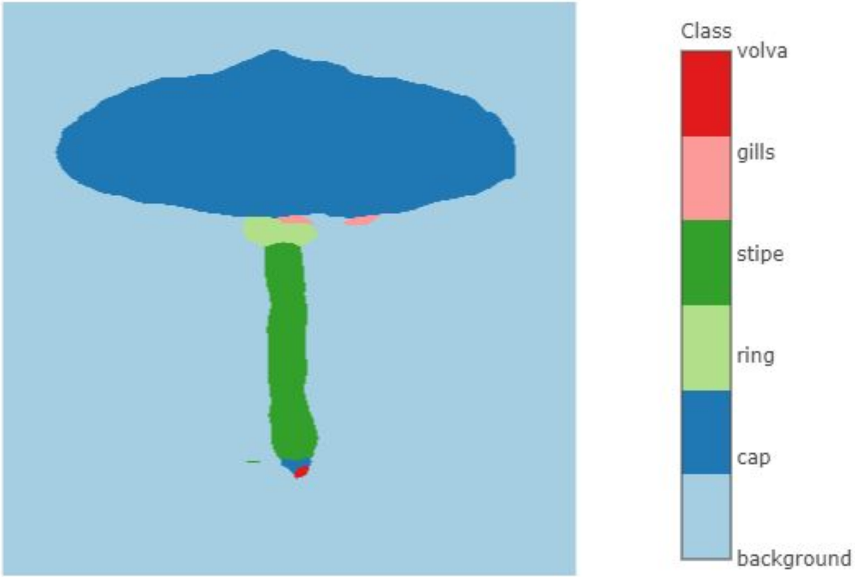
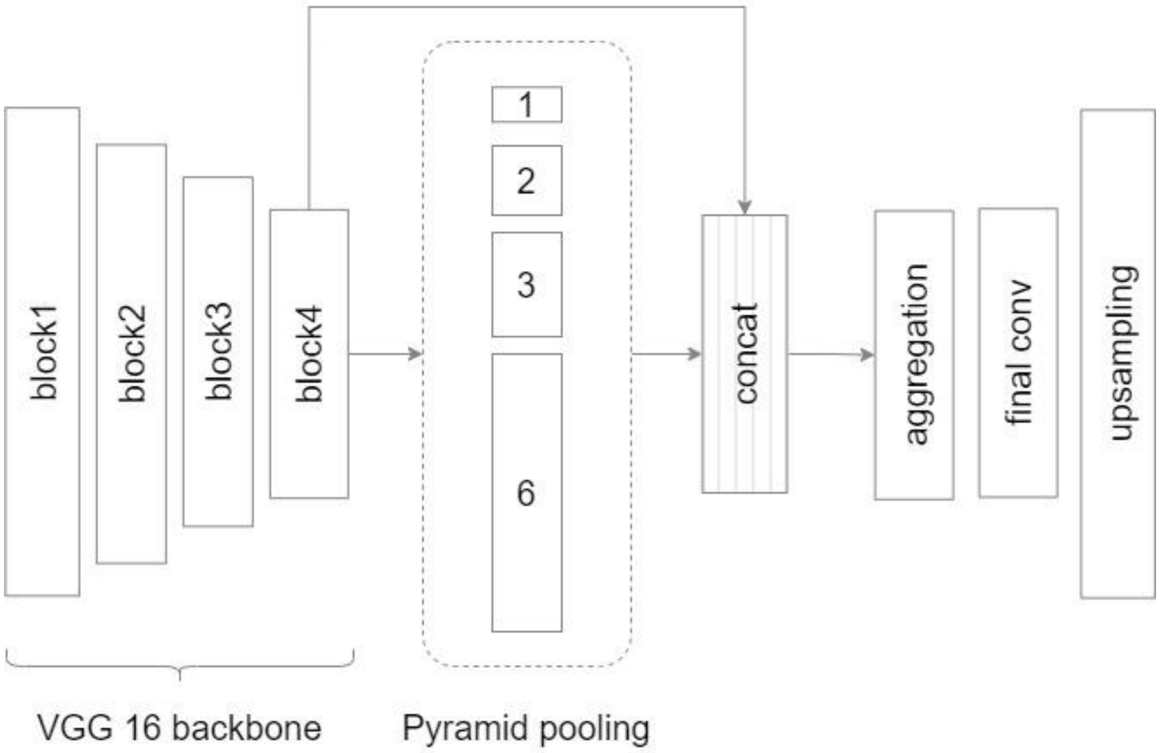


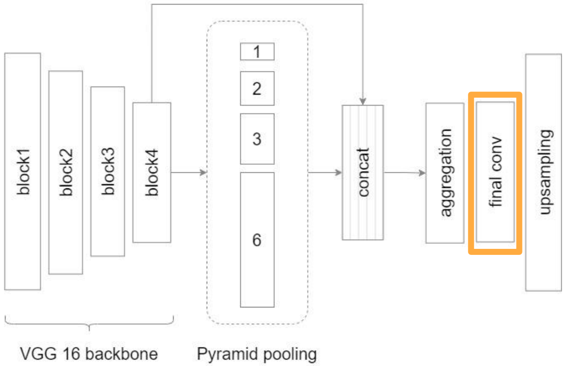
Image source:
[1] Statens Naturhistoriske Museum et al. (2022). Danmarks svampeatlas [Accessed March 29, 2022]. https://svampe.databasen.org
[2] 2018 FGCVx fungi classification challenge [Accessed March 29, 2022]. https://www.kaggle.com/competitions/fungi-challenge-fgvc-2018

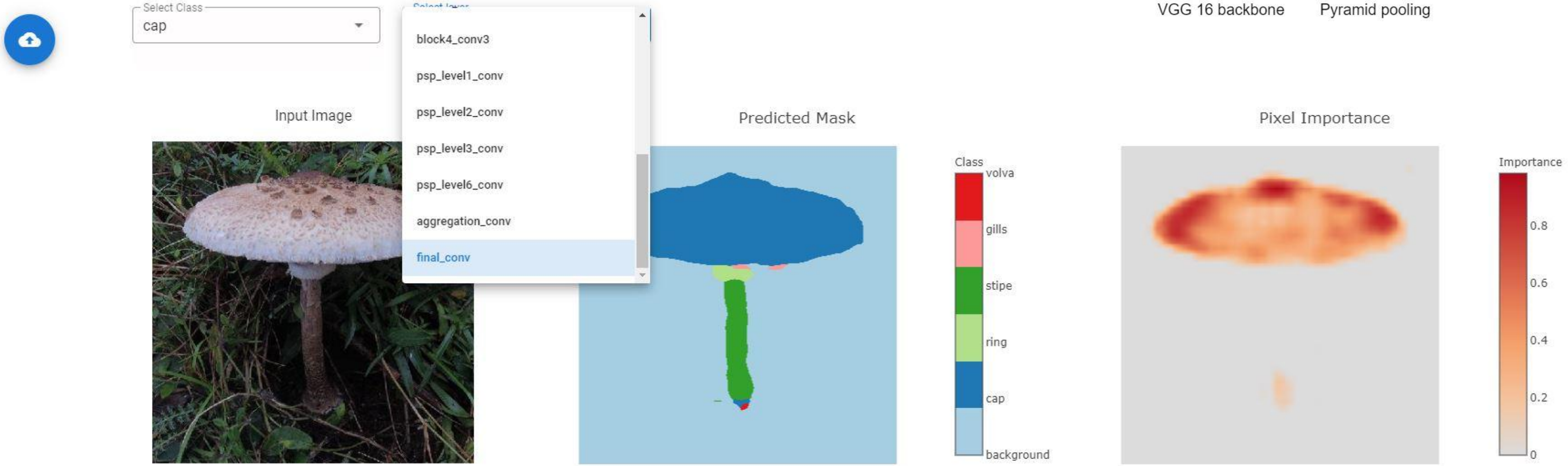
Example Network: PSPNet with VGG16 backbone



[3] H. Zhao, J. Shi, X. Qi, X. Wang and J. Jia, "Pyramid Scene Parsing Network," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6230-6239, doi: 10.1109/CVPR.2017.660.
[4] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” 2015 arXiv:1409.1556, 2015, doi: 10.48550/arXiv.1409.1556.

Example Task: Explore Network Architecture

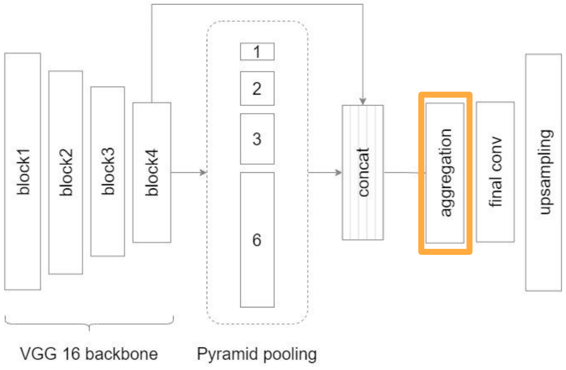


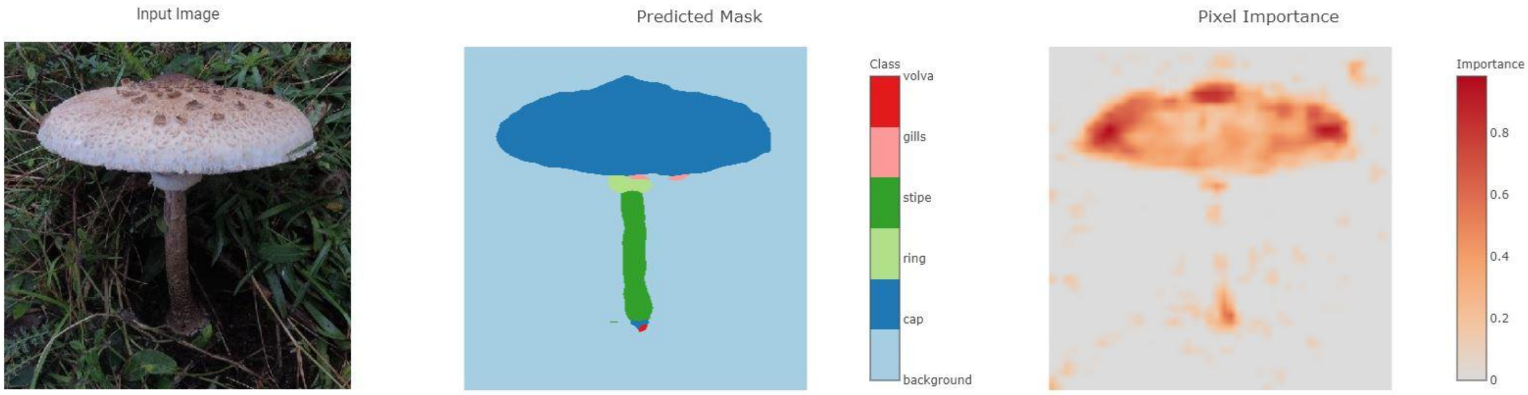


Smooth heatmap that mainly focuses on the cap, but also a bit on the stem.

[1], [2]

Example Task: Explore Network Architecture

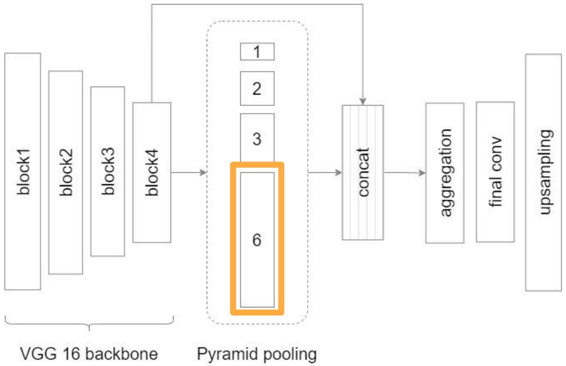


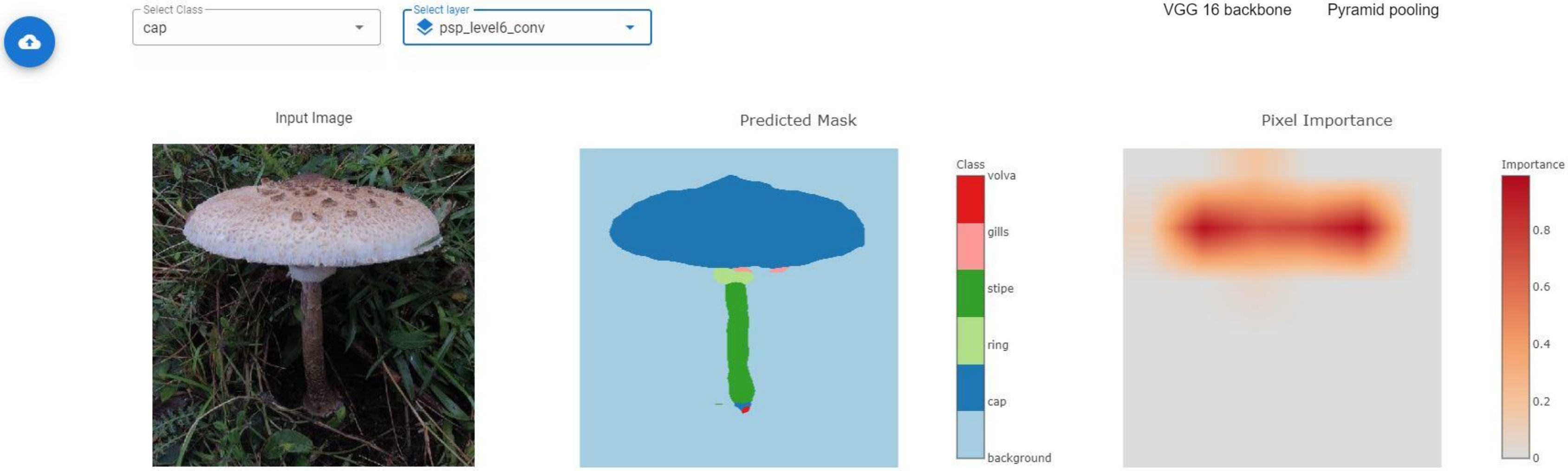


Heatmap is less smooth. This layer seems to focus a bit more on the context. Receptive field might play a role.

[1], [2]

Example Task: Explore Network Architecture

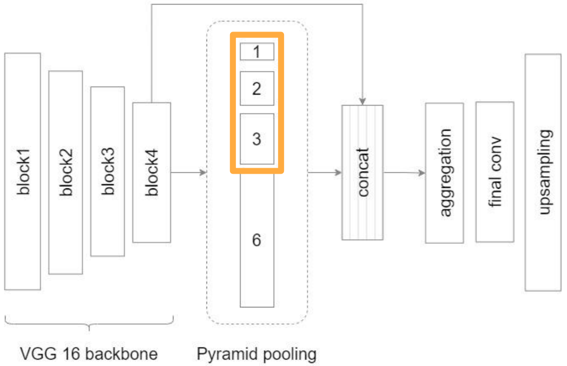


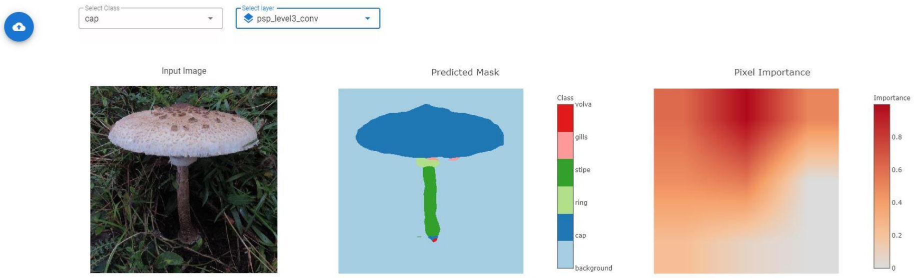


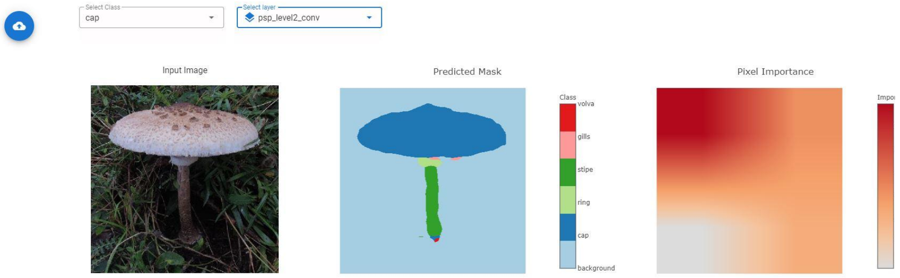
This layer seems to find the global location of the class.

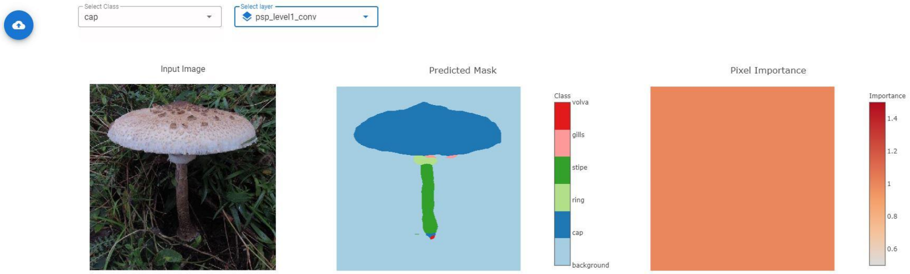
[1], [2]

Example Task: Explore Network Architecture





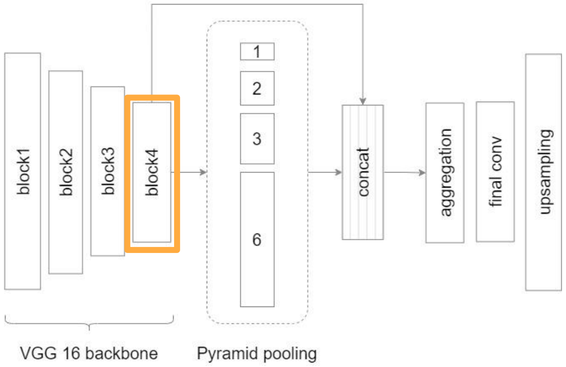


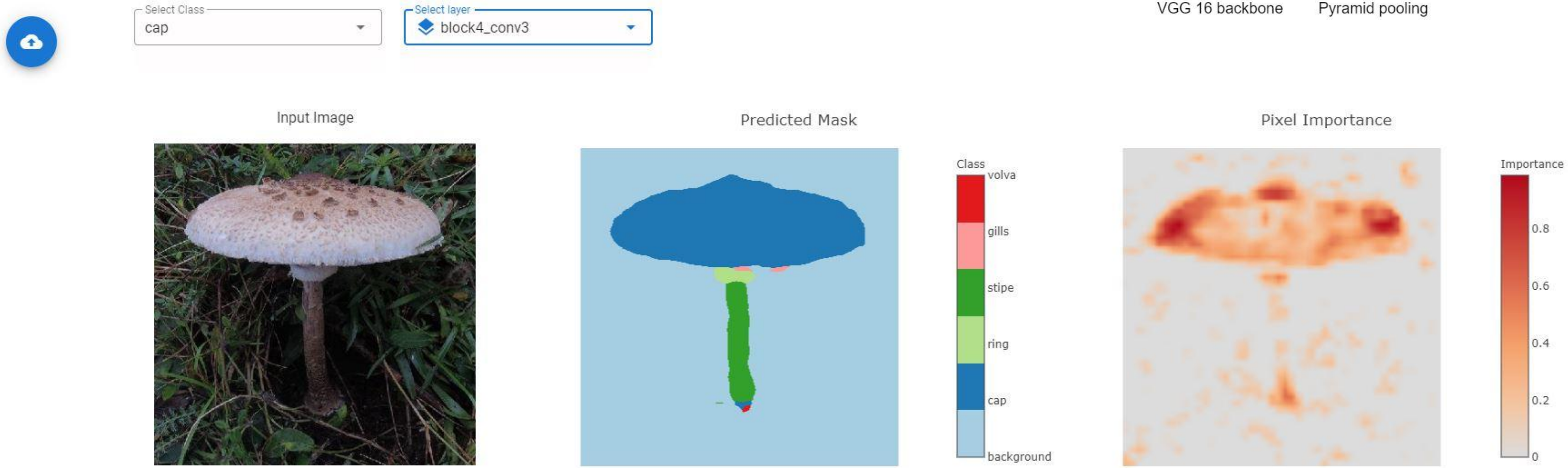


3x3, 2x2, and 1x1 Pyramid Pooling layers only seem to have approximate global location of the class.

[1], [2]

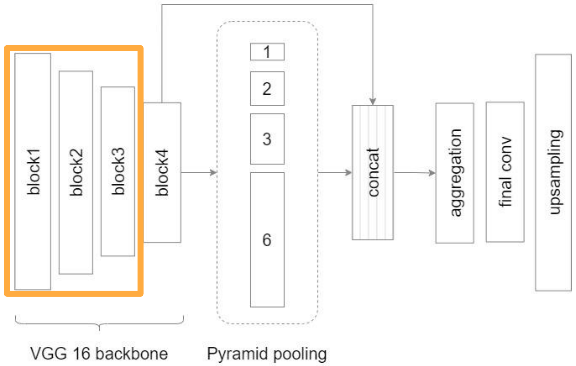
Example Task: Explore Network Architecture

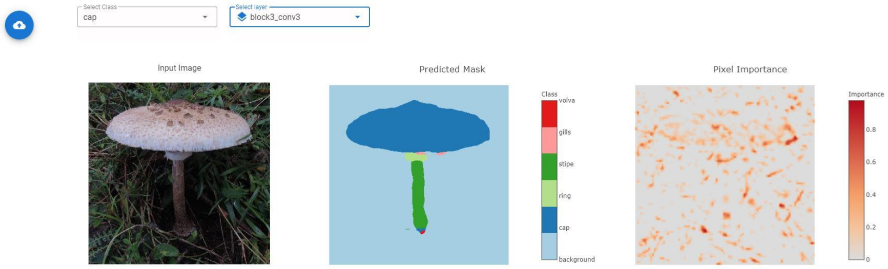


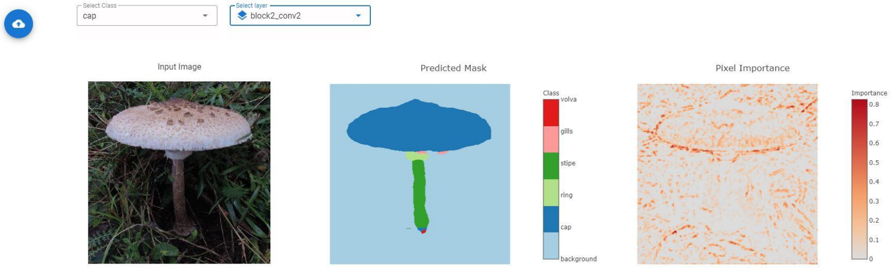


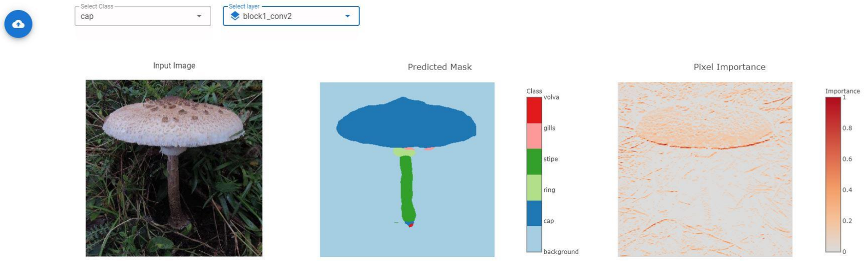
This layer is similar to the aggregation layer, but seems to have more local information about the class. E.g., There is more focus on the ring, probably because it has a similar color as the cap. It seems that combining this local information and the global information from the pyramid pooling results in a more accurate representation. [1], [2]

Example Task: Explore Network Architecture









The closer layers are to the input, the sparser their representations get due to a lower receptive field. Those layers seem to have more local information, but almost no global information.

[1], [2]

Tool Enables Exploration of:

- Network Architecture
  - shown in previous example
- Classes
  - verify that predictions are made for the “right” reason
- Images
  - explore images of interest
  - investigate for which images a segmentation works or does not work
- Pixels and Areas
  - investigate pixels or areas for which the segmentation did not work, or behaved in an (un)expected way

Resources

[1] Statens Naturhistoriske Museum et al. (2022). Danmarks svampeatlas [Accessed March 29, 2022]. https://svampe.databasen.org

[2] 2018 FGCVx fungi classification challenge [Accessed March 29, 2022]. https://www.kaggle.com/competitions/fungi-challenge-fgvc-2018

[3] H. Zhao, J. Shi, X. Qi, X. Wang and J. Jia, "Pyramid Scene Parsing Network," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6230-6239, doi: 10.1109/CVPR.2017.660.

[4] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” 2015 arXiv:1409.1556, 2015, doi: 10.48550/arXiv.1409.1556.

[5] Vinogradova, K., Dibrov, A., & Myers, G. (2020). Towards Interpretable Semantic Segmentation via Gradient-Weighted Class Activation Mapping (Student Abstract). Proceedings of the AAAI Conference on Artificial Intelligence, 34(10), 13943-13944. https://doi.org/10.1609/aaai.v34i10.7244