

# Contents

<b>1</b>	<b>Make dataset</b>	<b>2</b>
<b>2</b>	<b>Build features</b>	<b>2</b>
<b>3</b>	<b>Train model</b>	<b>2</b>
<b>4</b>	<b>Predict model</b>	<b>2</b>
<b>5</b>	<b>Evaluate model</b>	<b>2</b>
<b>6</b>	<b>Visualize</b>	<b>2</b>
<b>7</b>	<b>Conclusion</b>	<b>2</b>
	<b>References</b>	<b>3</b>

## 1 Make dataset

```
1 import pandas as pd
```

## 2 Build features

## 3 Train model

## 4 Predict model

## 5 Evaluate model

## 6 Visualize

## 7 Conclusion

## References

- [1] Ding-Jie Chen et al. “See-Through-Text Grouping for Referring Image Segmentation.” In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, Oct. 2019, pp. 7453–7462. ISBN: 978-1-72814-803-8. DOI: [10.1109/ICCV.2019.00755](https://doi.org/10.1109/ICCV.2019.00755). URL: <https://ieeexplore.ieee.org/document/9009843/> (visited on 12/17/2023).
- [2] Shih-Cheng Huang et al. “GLoRIA: A Multimodal Global-Local Representation Learning Framework for Label-efficient Medical Image Recognition.” In: *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada: IEEE, Oct. 2021, pp. 3922–3931. ISBN: 978-1-66542-812-5. DOI: [10.1109/ICCV48922.2021.00391](https://doi.org/10.1109/ICCV48922.2021.00391). URL: <https://ieeexplore.ieee.org/document/9710099/> (visited on 12/17/2023).
- [3] Fabian Isensee et al. “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation.” In: *Nature Methods* 18.2 (Feb. 2021), pp. 203–211. ISSN: 1548-7091, 1548-7105. DOI: [10.1038/s41592-020-01008-z](https://doi.org/10.1038/s41592-020-01008-z). URL: <https://www.nature.com/articles/s41592-020-01008-z> (visited on 12/17/2023).
- [4] Wonjae Kim, Bokyung Son, and Ildoo Kim. “ViLT: Vision-and-Language Transformer Without Convolution or Region Supervision.” In: ().
- [5] Zihan Li et al. *LViT: Language meets Vision Transformer in Medical Image Segmentation*. June 26, 2023. arXiv: [2206.14718\[cs\]](https://arxiv.org/abs/2206.14718). URL: <http://arxiv.org/abs/2206.14718> (visited on 12/17/2023).
- [6] John Schulman et al. *Proximal Policy Optimization Algorithms*. Aug. 28, 2017. DOI: [10.48550/arXiv.1707.06347](https://doi.org/10.48550/arXiv.1707.06347). arXiv: [1707.06347\[cs\]](https://arxiv.org/abs/1707.06347). URL: <http://arxiv.org/abs/1707.06347> (visited on 08/17/2023).
- [7] Zhao Yang et al. *LAVT: Language-Aware Vision Transformer for Referring Image Segmentation*. Apr. 5, 2022. arXiv: [2112.02244\[cs\]](https://arxiv.org/abs/2112.02244). URL: <http://arxiv.org/abs/2112.02244> (visited on 12/17/2023).