Semantic Segmentation

CP8307 - Group #3

Christopher Kolios (ckolios@ryerson.ca)
Yiannis Varnava (yvarnava@ryerson.ca)

Semantic Segmentation

- Process of sectioning an image into different classes (pixel-by-pixel) [1]

- Applications
  - Medical imaging
  - Self-driving cars
  - Augmented reality

Fig 1. Semantic segmentation. [7]

- Deep learning-based models
  - FCN [2]
  - **U-Net** [3]
  - RCNN [4]

- Traditional models
  - Thresholding
  - Chan-Vese [5]
  - **Region Adjacency Graph (RAG)** [6]

Convolutional Neural Networks (CNNs)

- Neural network architecture for image data [8]
- Why are CNNs used for images?
  - 2D input allows network to learn local and global image features
- Layers
  - Convolution
  - Pooling
  - Fully Connected

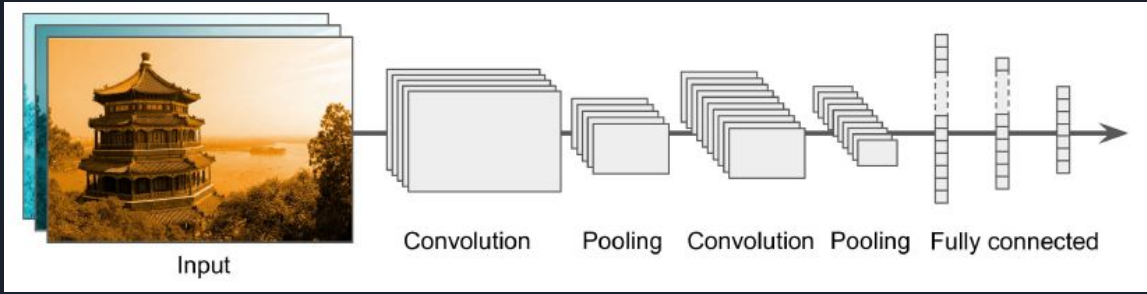


Fig 2. Basic CNN structure [8]

U-Net

- Published by Ronneberger et al. in 2015 with ~34000 citations
- Encoder-decoder based segmentation model for biomedical images
- Encoder-bottleneck-decoder stages with skip connections
- Outputs a segmentation map
    - Each pixel corresponds to a class

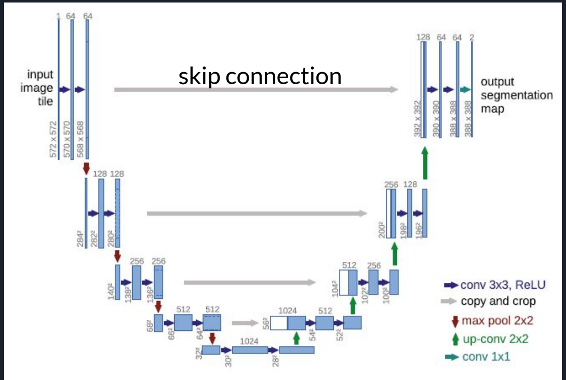


Fig 3. U-Net architecture [3]

0. Background
1. Person
2. Purse
3. Plants
4. Sidewalk
5. Building

Fig 4. Segmentation map [9]

Traditional Segmentation Methods

- There are a variety of segmentation techniques that can be used for binary segmentation and classification
- One of the simplest approaches is thresholding
- The threshold can be automatically determined

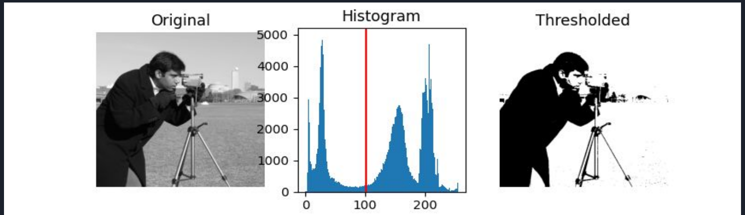


Fig 5. Thresholding [10]

- Another approach is Chan-Vese segmentation (first outlined in 1999 by Tony Chan and Luminita Vese) [5]
- Solves an energy minimization problem on intensities on the inside and outside of contours, providing good segmentation

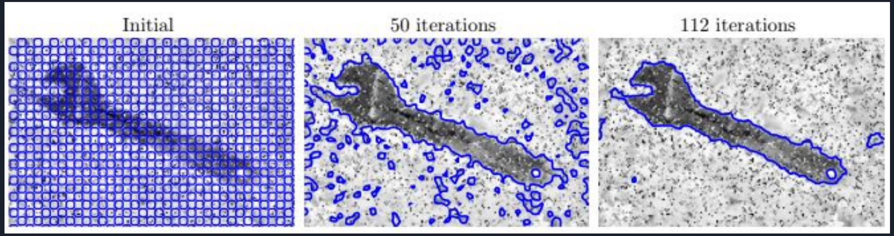


Fig 6. Chan-Vese Segmentation [11]

Traditional Segmentation Methods Cont.

- Finally, there is segmentation using Region Adjacency Graphs (RAGs) [6]
- Key idea is to transform images into superpixel regions based on colour similarity, then merge similar regions up to some threshold
- Algorithm procedure slightly modified for our dataset

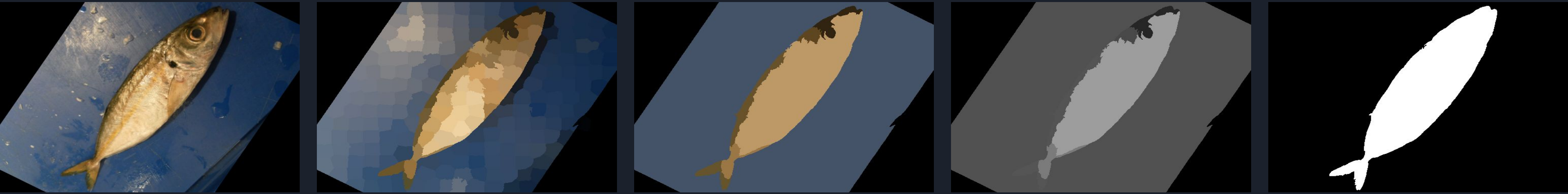


Fig 7-11. RAG Segmentation Procedure on Horse Mackerel

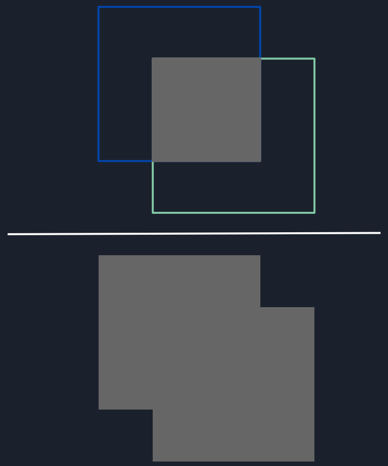
- There are many more classical segmentation methods, but those presented are relatively simple and have efficient implementations

Evaluation Metrics

Intersection over Union (IoU) [1]

$$\mathrm{IoU} = \frac{|\, A \cap B \,|}{|\, A \cup B \,|}$$

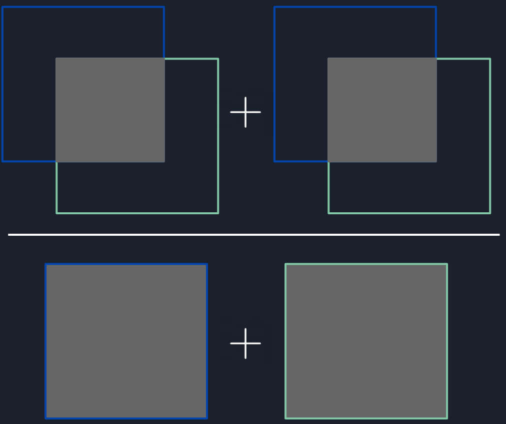
IoU =



Dice Score [1]

$$\mathrm{Dice} = \frac{2\,|\, A \cap B \,|}{|\, A \,| + |\, B \,|}$$

Dice =



Dataset

- “A Large Scale Fish Dataset” [12]
- 9000 image-mask pairs of fish
- 80%/20% training/test split
  - 7200 samples for training
  - 1800 samples for testing
- Includes data augmentation (rotation, flipping)
  - Method used to generate more training samples.

Fig 12. Sample of an image-mask pair [12]

Method

- Compared U-Net (TensorFlow implementation) with Region Adjacency Graph, Chan-Vese and thresholding segmentation methods

- Test set
  - 1800 (20%) image-mask pairs

- Calculated IoU and Dice score

Results - Foreground Class

Test Set - 1800 Samples

Table 1. IoU and Dice scores on test set

	U-Net	RAG	Chan-Vese	Threshold
IoU	0.89	0.72	N.C.	N.C.
Dice	0.94	0.81	N.C.	N.C.

Results Histograms

Test Set - 1800 Samples

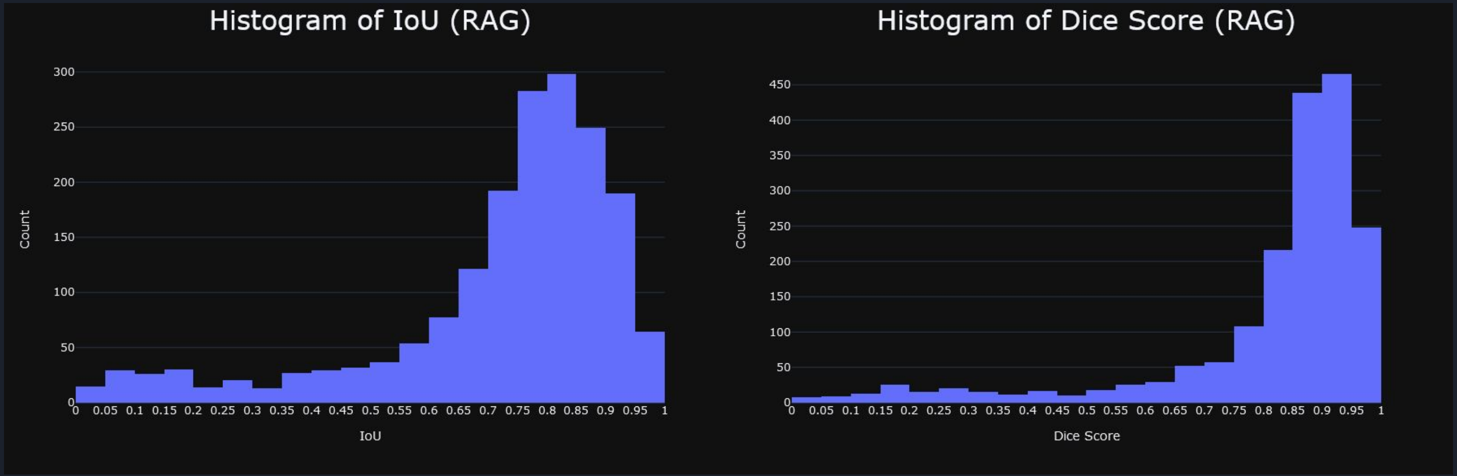


Fig 13. RAG Result Histograms

Experimental Results - Visual

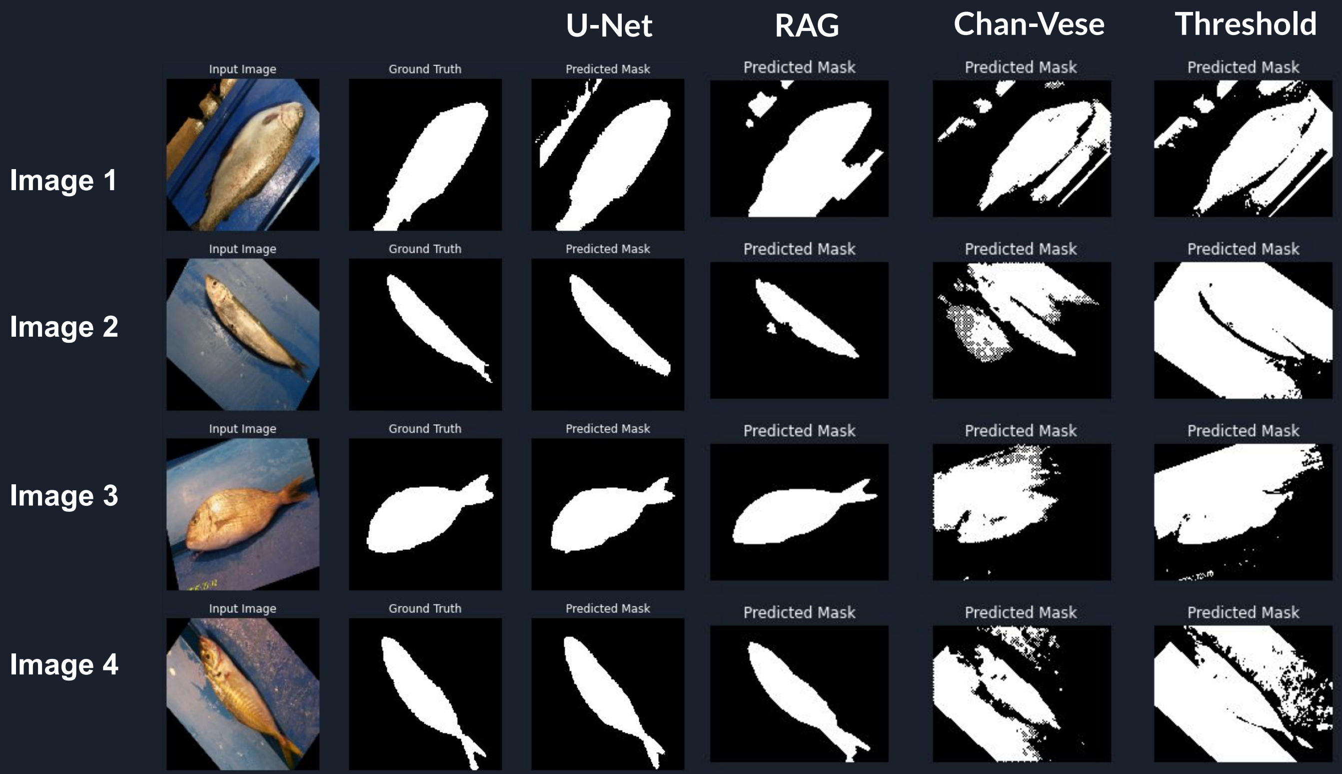


Fig 14. Experimental results

Experimental Results - Metrics

Table 2. IoU results for four images

	U-Net	RAG	Chan-Vese	Threshold
Image 1	0.80	0.64	0.53	0.21
Image 2	0.92	0.84	0.26	0.49
Image 3	0.92	0.90	0.35	0.14
Image 4	0.94	0.87	0.26	0.32

Table 3. Dice score results for four images

	U-Net	RAG	Chan-Vese	Threshold
Image 1	0.89	0.78	0.69	0.34
Image 2	0.96	0.91	0.42	0.66
Image 3	0.96	0.94	0.51	0.25
Image 4	0.97	0.93	0.41	0.49

Conclusions

- U-Net offers superior performance to classical methods, but depends on large training set
- Of the classical segmentation methods, only RAG had good performance
  - Thresholding is too simple
  - Chan-Vese is not well-suited to Fish dataset
  - Colour differences between fish and background cause RAG thresholding to perform well
- Semantic segmentation is much harder than just contour detection

- Future Work
  - More segmentation techniques
  - U-Net improvements
    - Dropout layers
    - More training data
  - Adding ML to tune RAG parameters

Thank you!

References

1. S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2021, doi: 10.1109/TPAMI.2021.3059968.
2. J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," 2015, pp. 3431–3440. Accessed: Dec. 04, 2021. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2015/html/Long_Fully_Convolutional_Networks_2015_CVPR_paper.html
3. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Cham, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
4. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," 2014, pp. 580–587. Accessed: Dec. 04, 2021. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2014/html/Girshick_Rich_Feature_Hierarchies_2014_CVPR_paper.html
5. T. Chan and L. Vese, "An Active Contour Model without Edges," in *Scale-Space Theories in Computer Vision*, Berlin, Heidelberg, 1999, pp. 141–151
6. A. Tremeau and P. Colantoni, "Regions adjacency graph applied to color image segmentation," *IEEE Transactions on Image Processing*, vol. 9, no. 4, pp. 735–744, Apr. 2000, doi: 10.1109/83.841950.
7. "A 2021 guide to Semantic Segmentation," *AI & Machine Learning Blog*, May 19, 2021. https://nanonets.com/blog/semantic-image-segmentation-2020/ (accessed Dec. 04, 2021). [Online].
8. G. Aurélien, *Hands-on machine learning with Scikit-Learn and TensorFlow concepts, tools, and techniques to build intelligent systems*. OReilly, 2019.
9. "An overview of semantic image segmentation.," *Jeremy Jordan*, May 22, 2018. https://www.jeremyjordan.me/semantic-segmentation/ (accessed Dec. 04, 2021). [Online].
10. "Thresholding — skimage v0.20.0.dev0 docs." https://scikit-image.org/docs/dev/auto_examples/applications/plot_thresholding.html (accessed Dec. 04, 2021). [Online].
11. P. Getreuer, "Chan-Vese Segmentation," *Image Processing On Line*, vol. 2, pp. 214–224, 2012.
12. O. Ulucan, D. Karakaya, and M. Turkan, "A Large-Scale Dataset for Fish Segmentation and Classification," in *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, 2020, pp. 1–5.