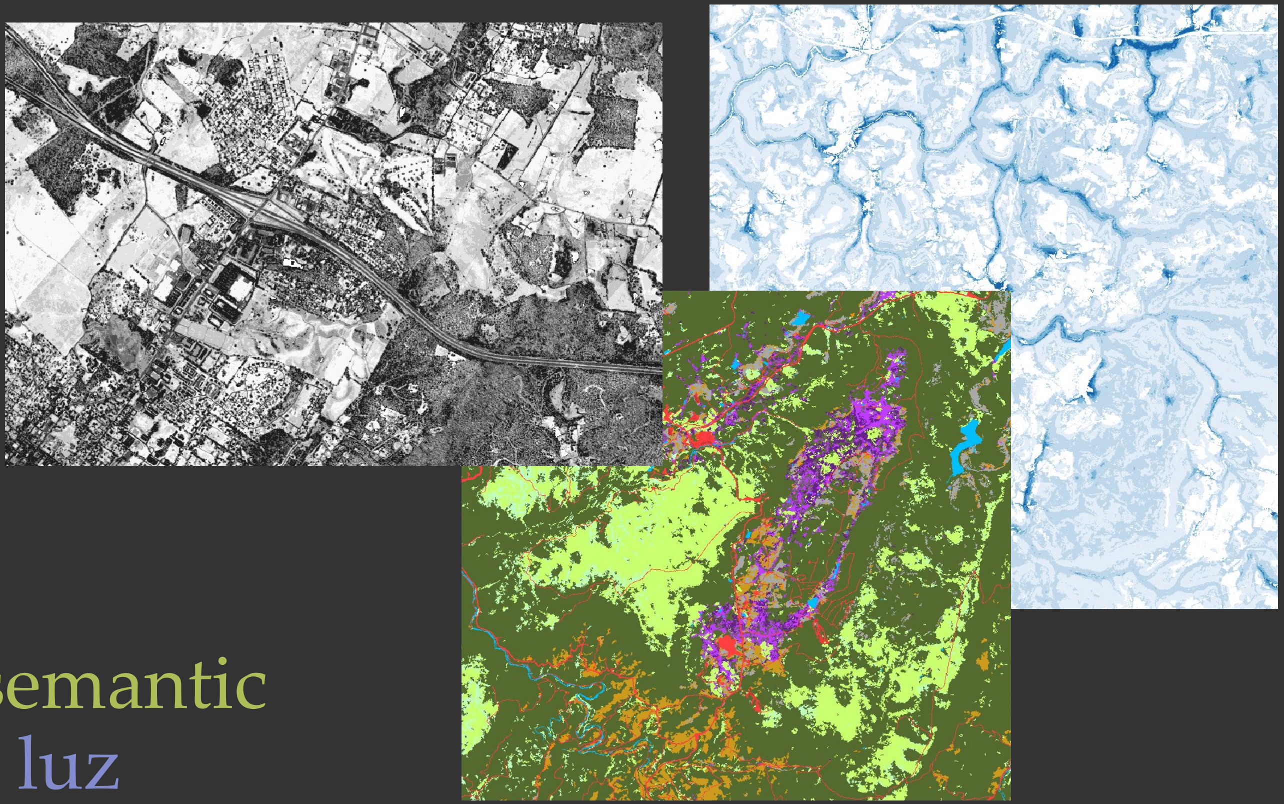
geodl: An R package for geospatial semantic segmentation using torch, terra, and luz



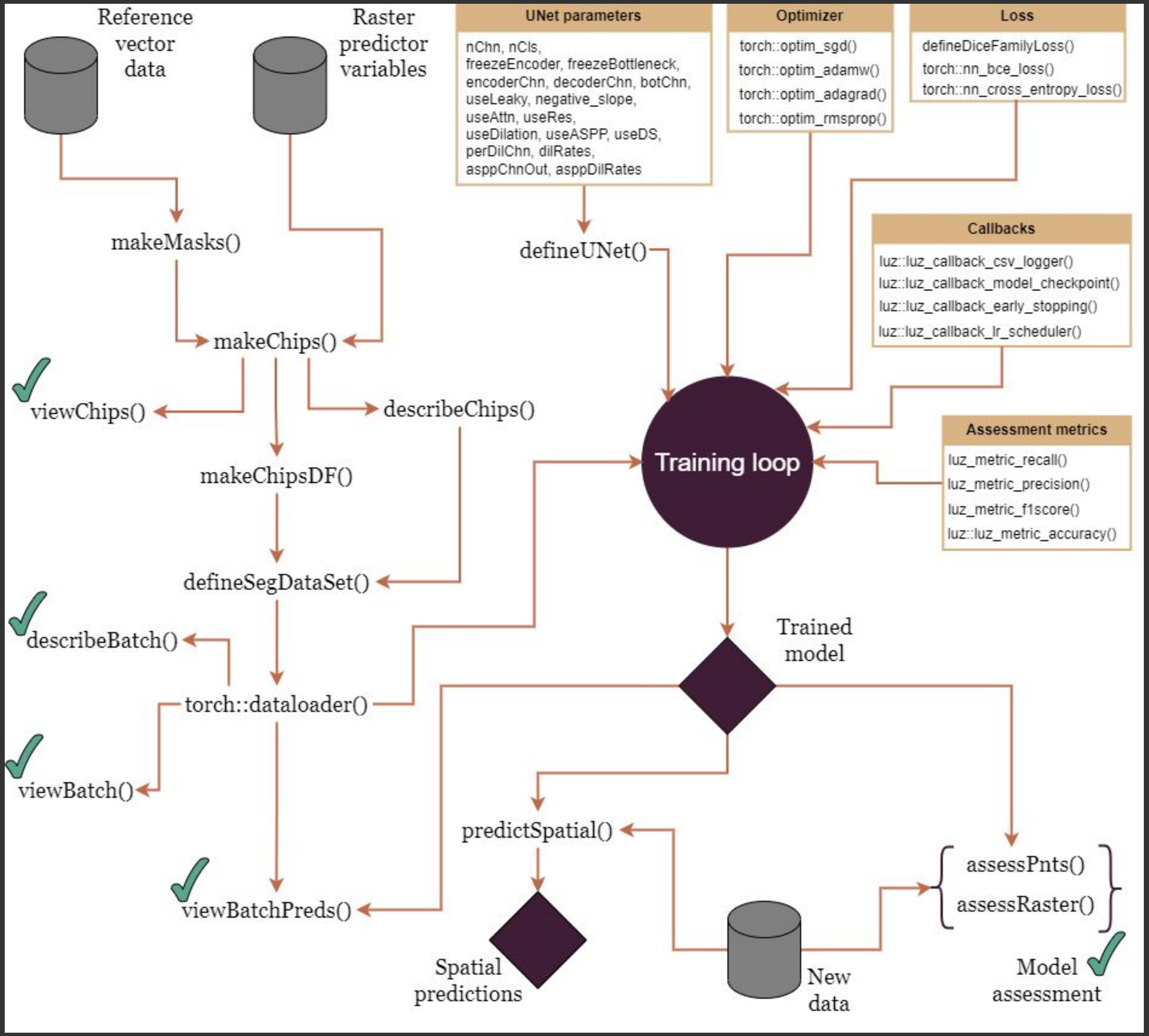
Aaron E. Maxwell
Associate Professor
West Virginia University
Department of Geology and Geography

About Myself

- Associate Professor WVU Department of Geology and Geography
- Faculty Director for WV GIS Tech Center
- PI for West Virginia View
- Interested in Geomorphology, GIS, Remote Sensing, and Spatial Modeling

http://www.wvview.org/Prof_Maxwell.html

Geospatial Deep Learning in R: geodl



https://github.com/maxwell-geospatial/geodl

https://cran.r-project.org/web/packages/geodl/index.html

Maxwell, A.E., Farhadpour, S., Das, S. and Yang, Y., 2024. geodl: An R package for geospatial deep learning semantic segmentation using torch and terra. *PLoS One*, *19*(12), p.e0315127.

Geospatial Packages

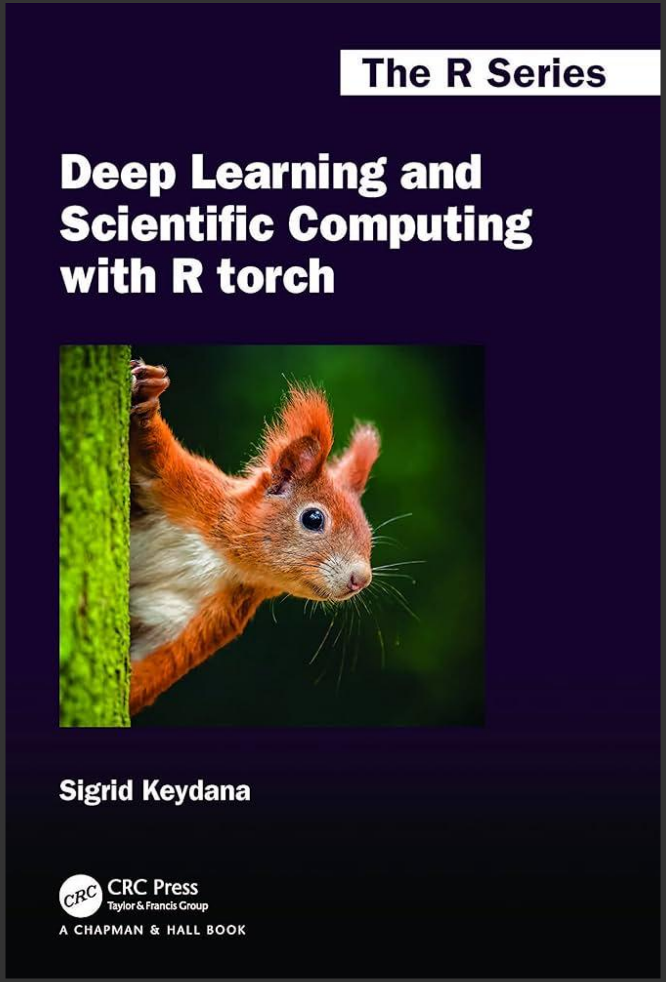
- sf (simple features for R) = read, write, and analyze vector geospatial data
- terra = work with raster geospatial data
- stars = work with spatiotemporal arrays and data cubes
- tmap = make thematic maps
- leaflet = interactive maps/interface to Leaflet JS
- lidR = work with lidar/point clouds in R
- spatstat = spatial stats/point pattern analysis

https://cran.r-project.org/web/views/Spatial.html

Deep Learning in R

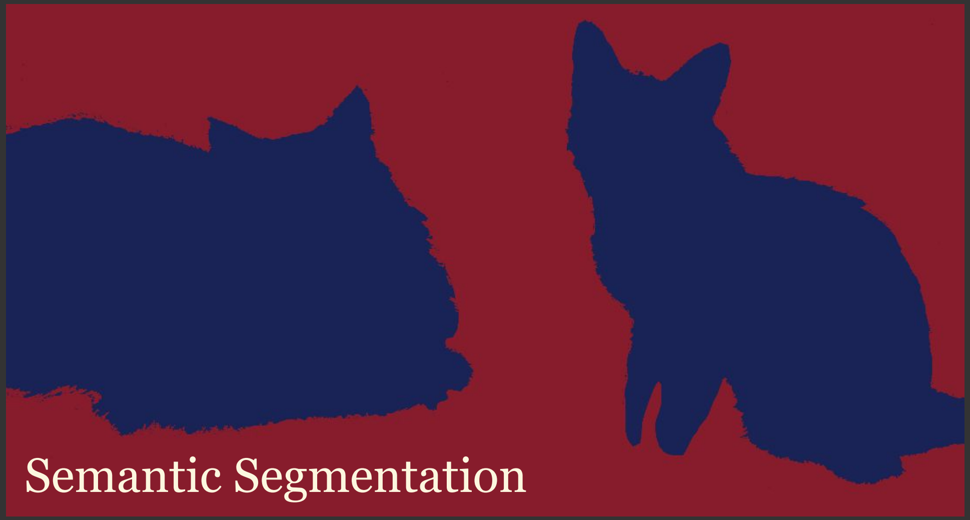
torch = PyTorch for R with C++ backend

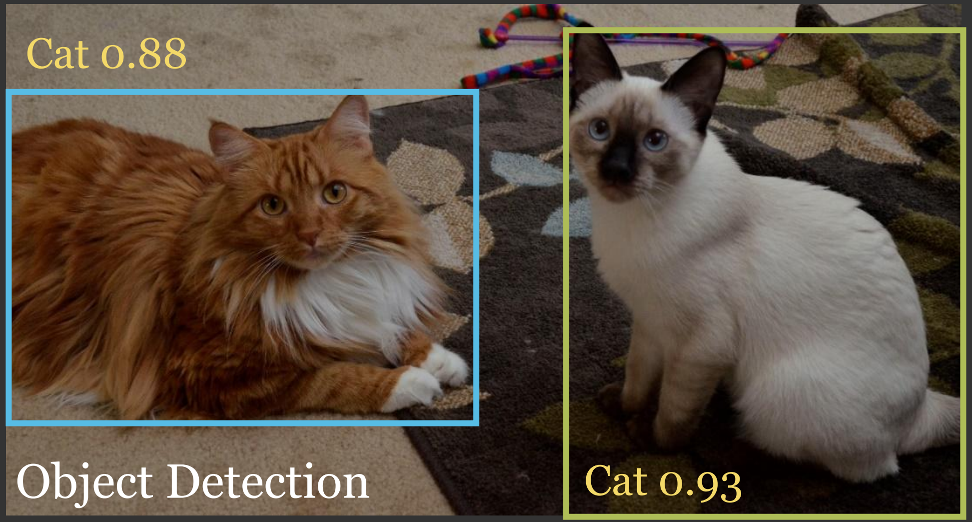
luz = simplifies training process

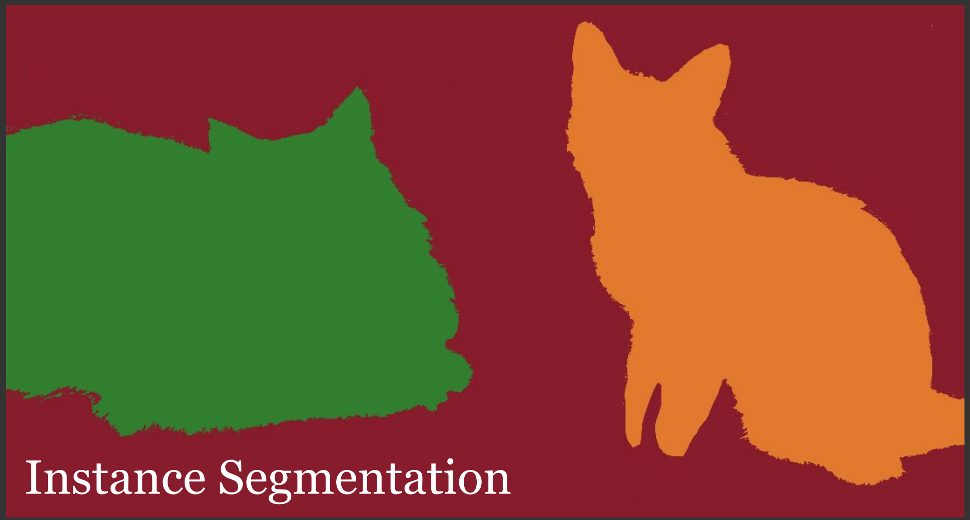


https://skeydan.github.io/Deep-Learning-and-Scientific-Computing-with-R-torch/

Semantic Segmentation







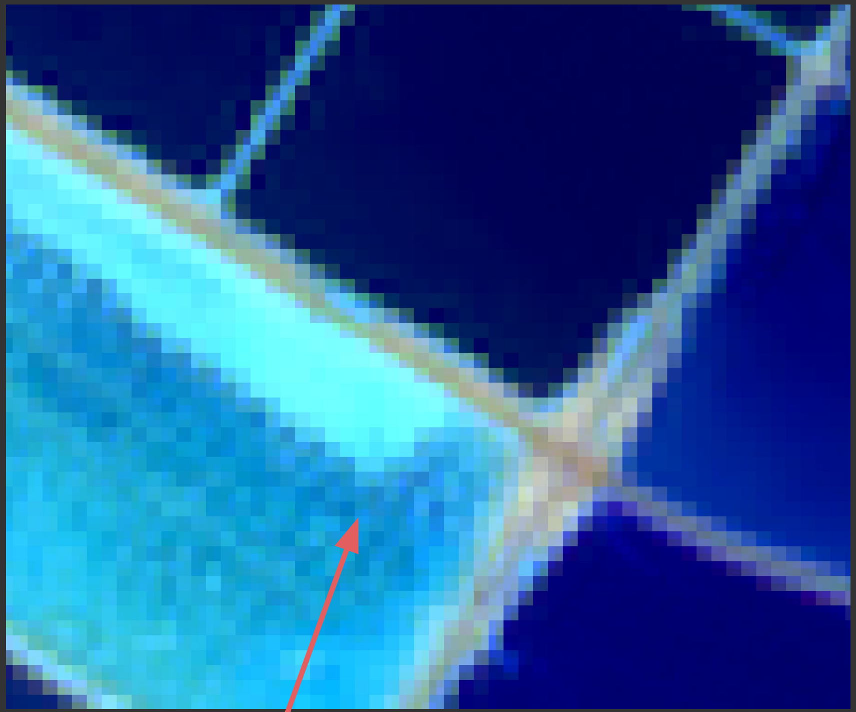
Why deep learning?

Predictor Variables

Label

Records or Samples

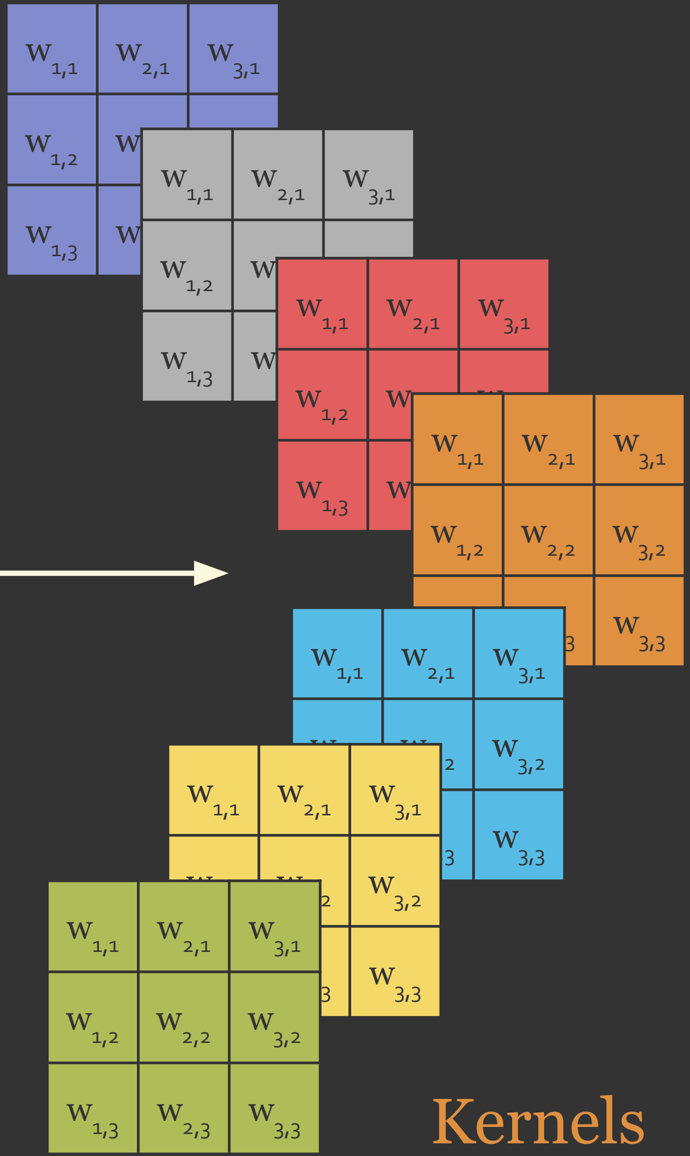
Class	Blue	Green	Red	NIR	SWIR1	SWIR2



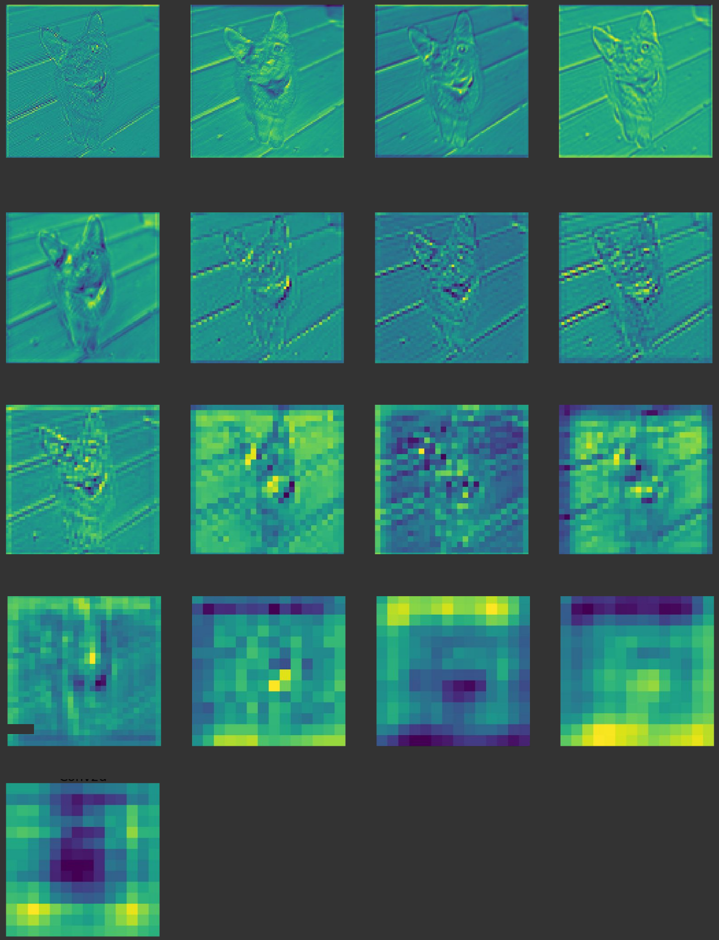
Unit of Analysis = Pixel

Why deep learning?

Iris

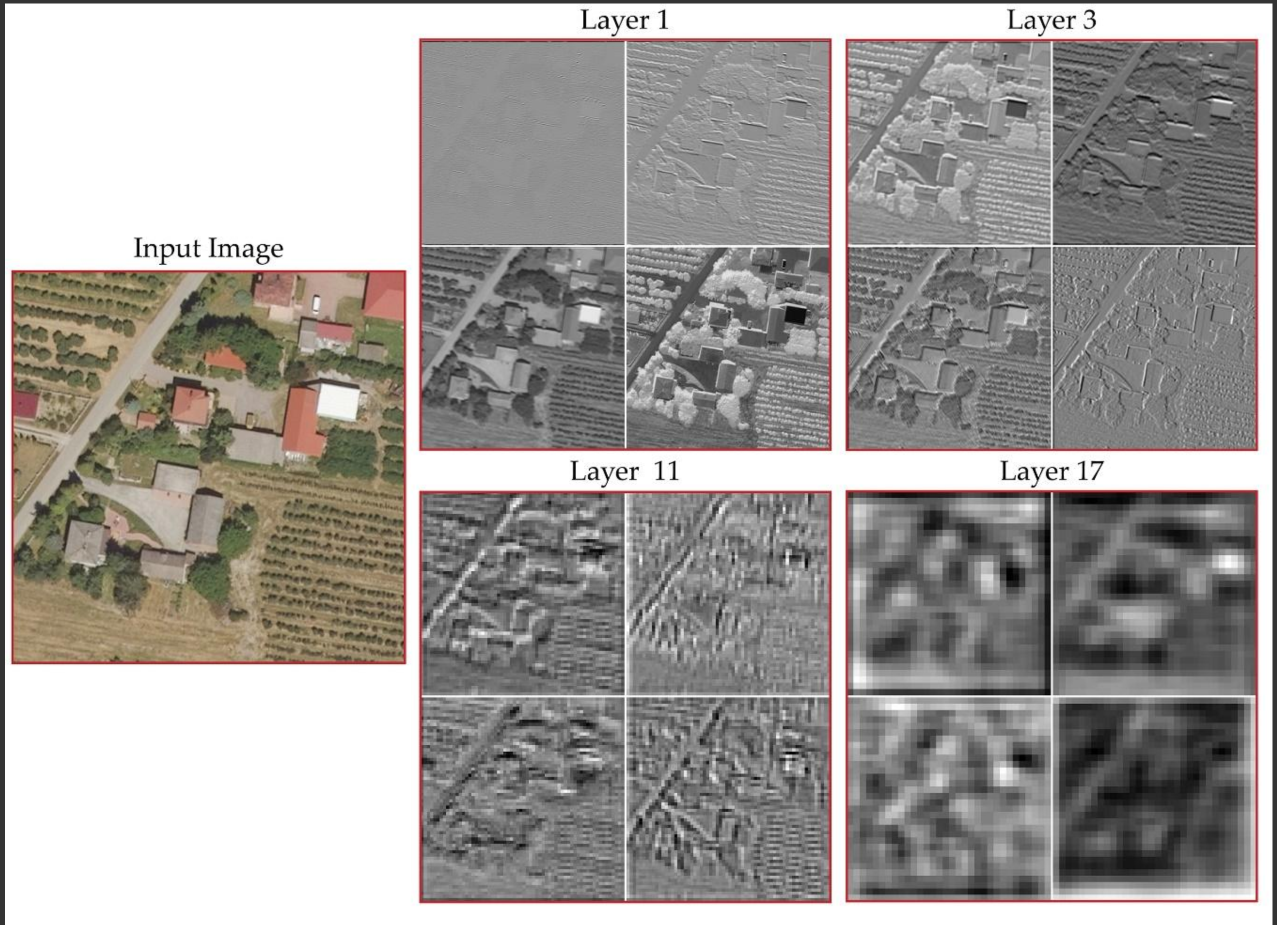


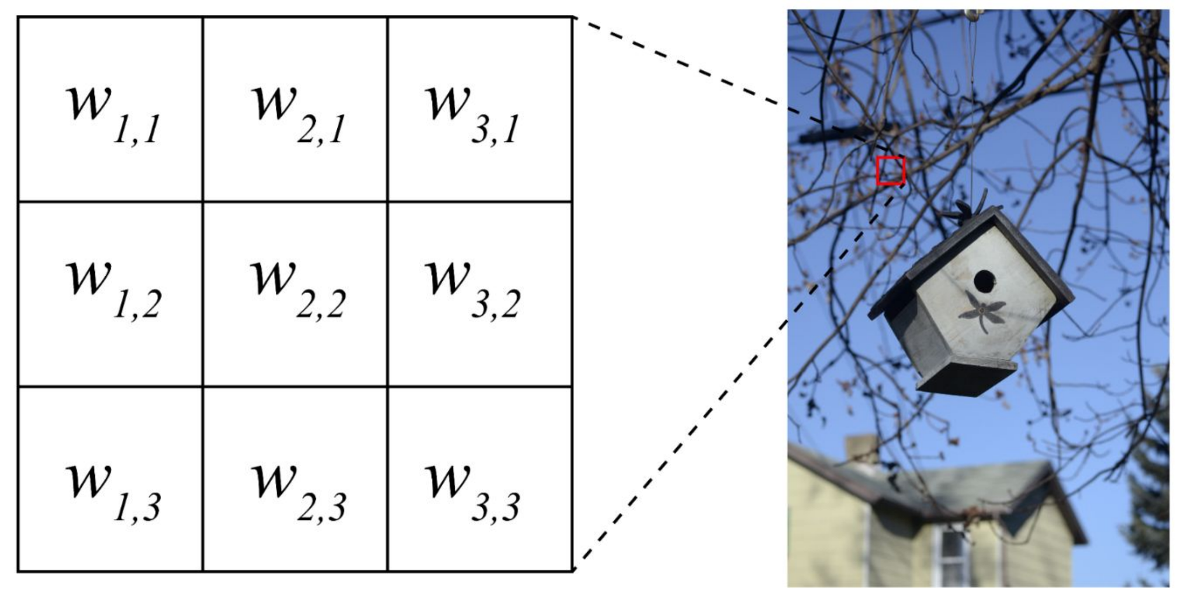
Kernels



Feature Maps

Why deep learning?

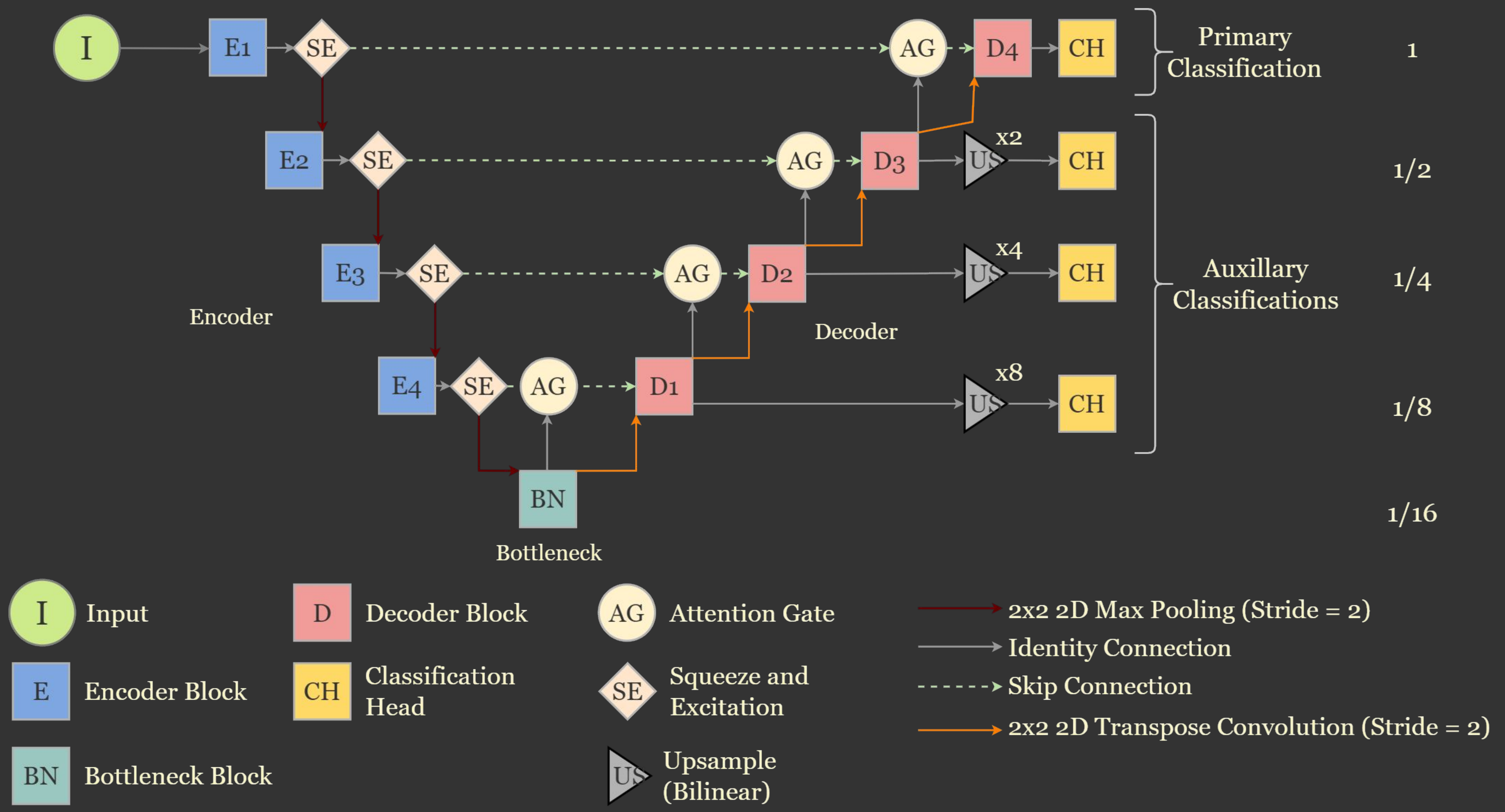




Architectures

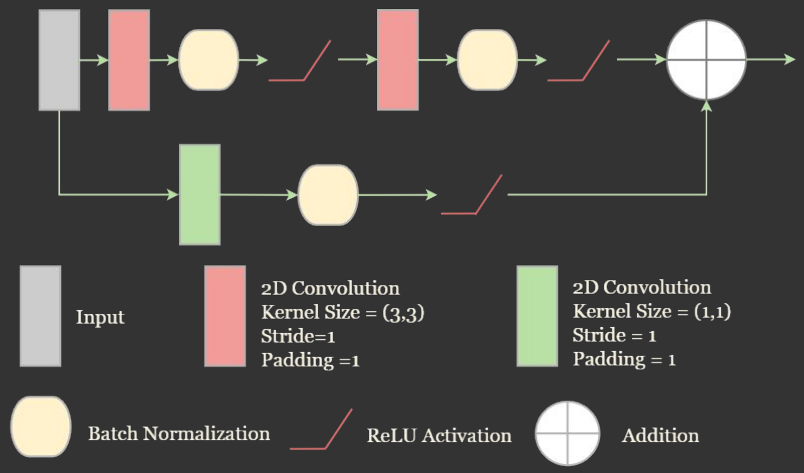
Configurable UNet

Ronneberger, O., Fischer, P. and Brox, T., 2015, October. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Cham: Springer international publishing.

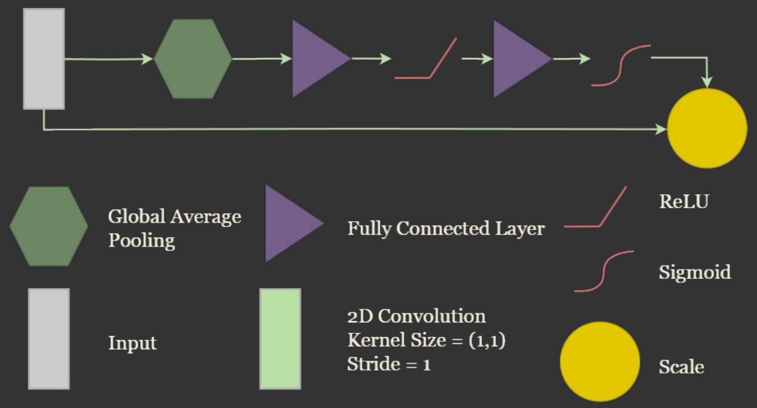


Configurations

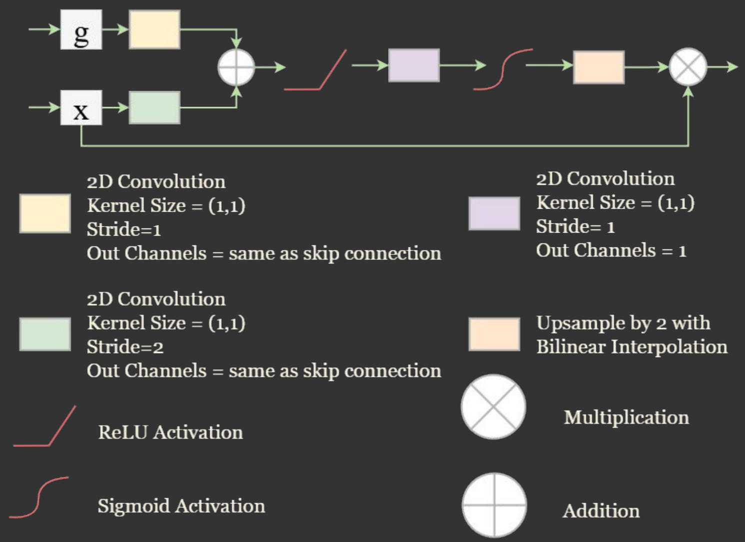
Residual Connections



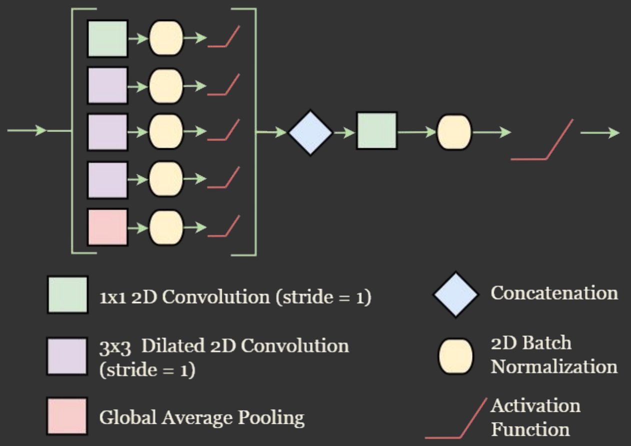
Squeeze and Excitation



Attention Gate

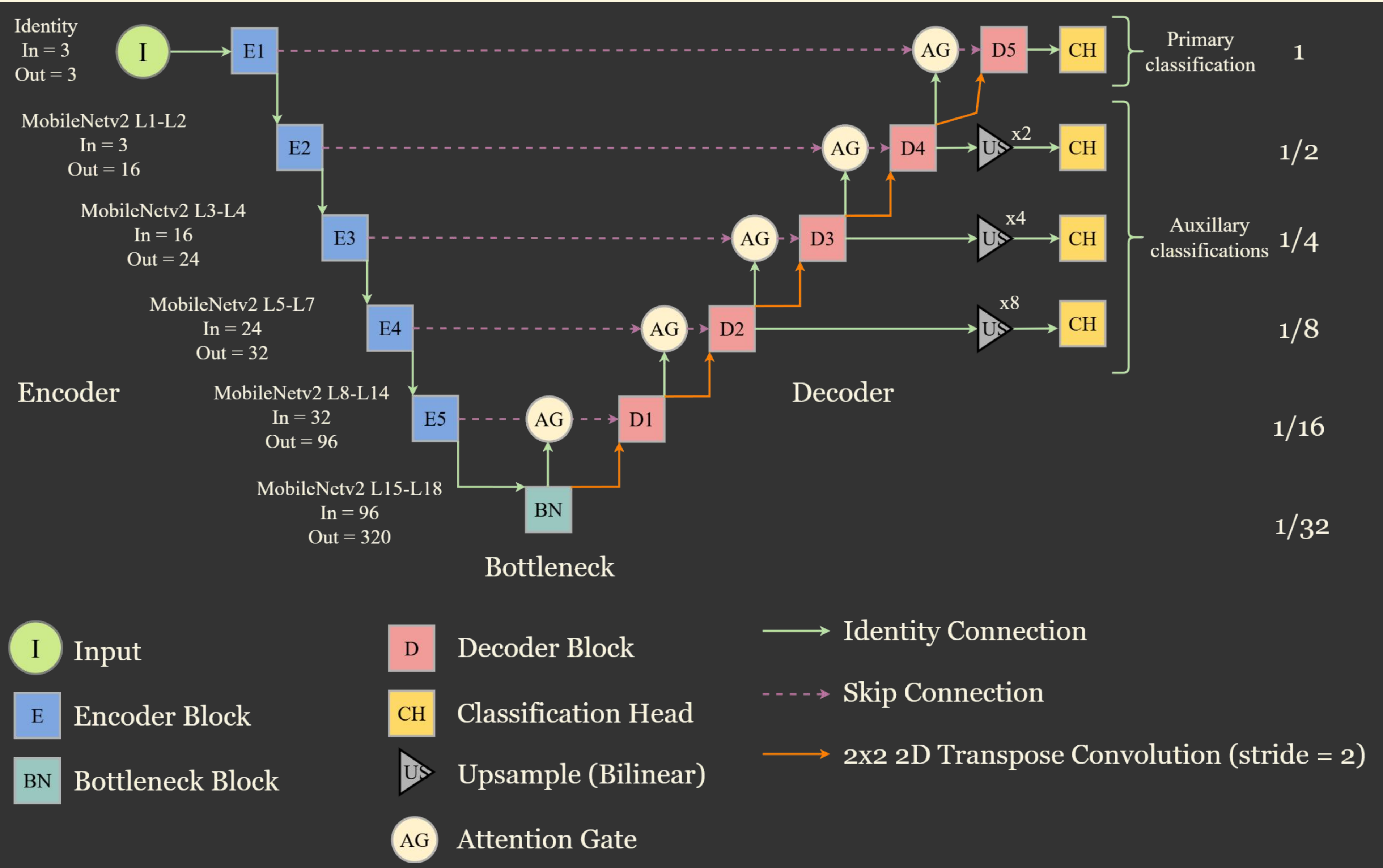


Atrous Spatial Pyramid Pooling (ASPP)



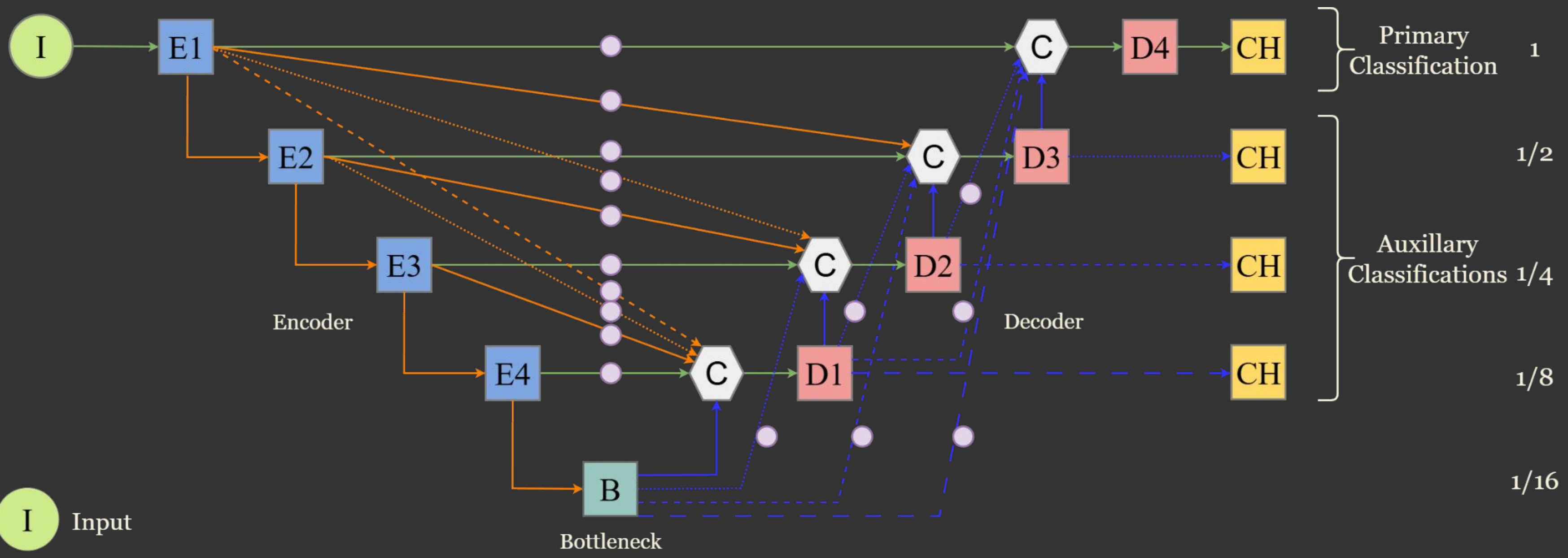
MobleNetv2 UNet

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L.C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).



Modified UNet3+

Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.W. and Wu, J., 2020, May. Unet 3+: A full-scale connected unet for medical image segmentation. In *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 1055-1059). Ieee.



I Input

E Encoder Block (3X3 2D Conv + BN + lReLU + 3x3 2D Conv + BN + lRelu

B Bottleneck (same as E or ASPP module)

D Decoder Block (3x3 2D Cov + BN + lReLU

C Concatenate

CH Classification Head (1x1 2D Convolution)

3x3 2D Convolution (Stride = 1, Padding = 1) + BN + lReLU

Identity Connection

2x2 2D Max Pooling (Stride = 2)

2x2 2D Max Pooling (Stride = 4)

2x2 2D Max Pooling (Stride = 8)

2x2 2D Transpose Convolution (Stride = 2) + BN + lReLU

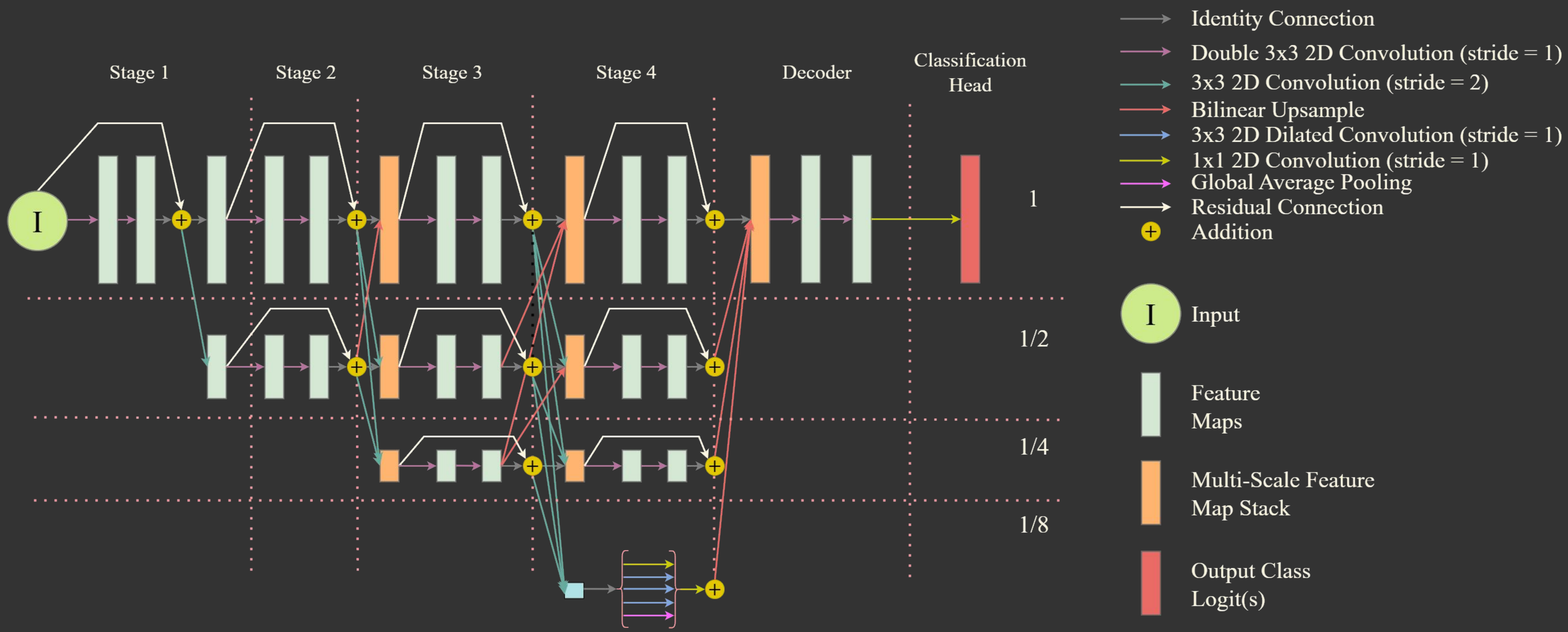
Bilinear Interpolation Upsample (Factor = 2)

Bilinear Interpolation Upsample (Factor = 4)

Bilinear Interpolation Upsample (Factor = 8)

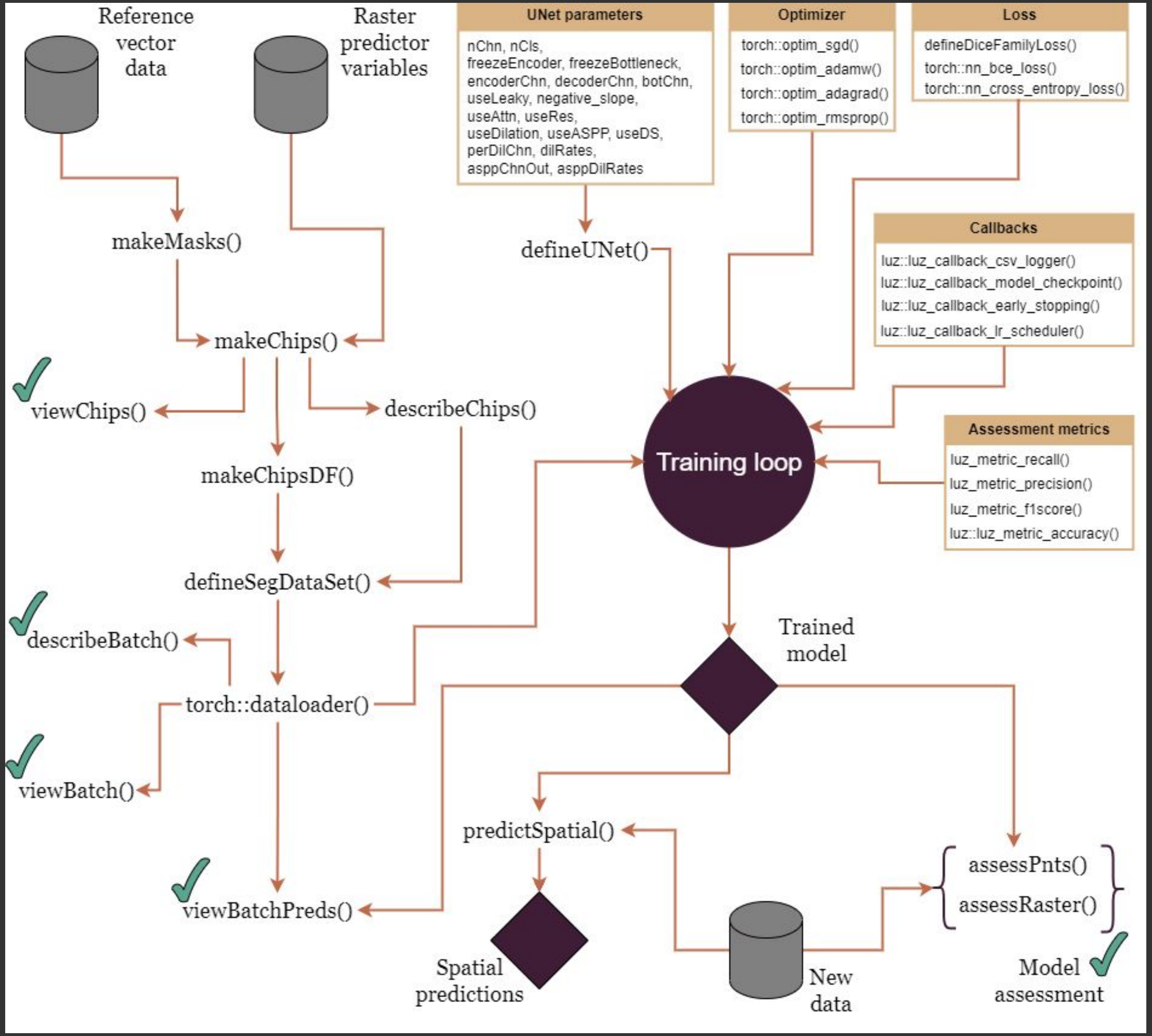
Modified HRNet

Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., Liu, D., Mu, Y., Tan, M., Wang, X. and Liu, W., 2020. Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, *43*(10), pp.3349-3364.



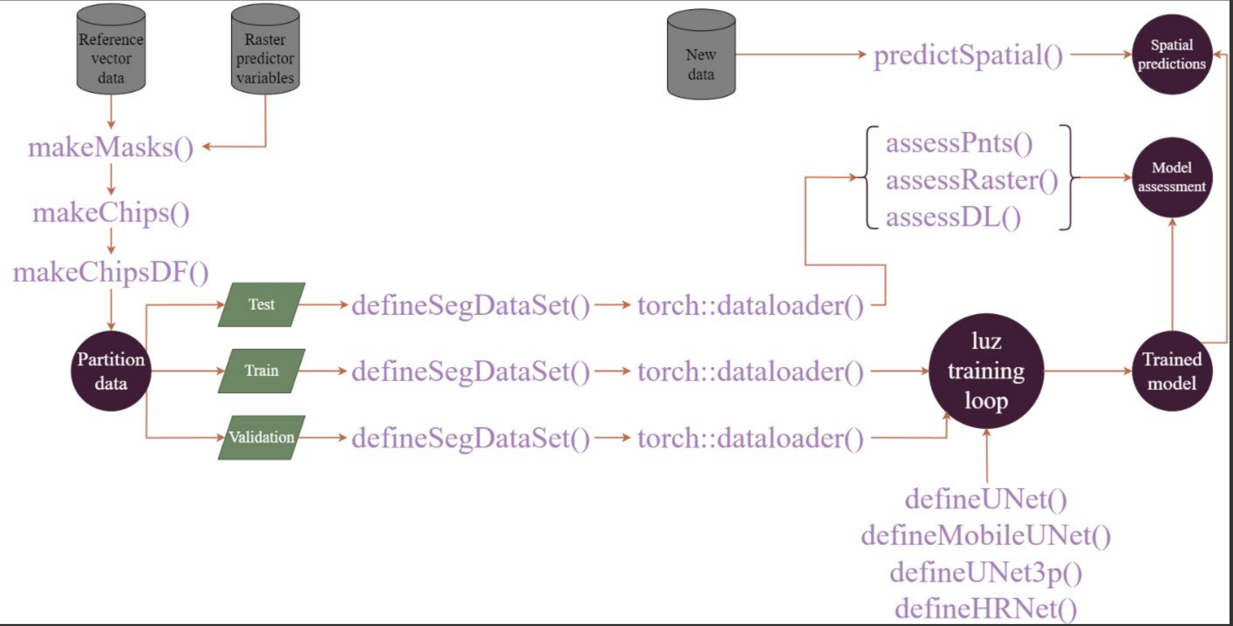
Workflows

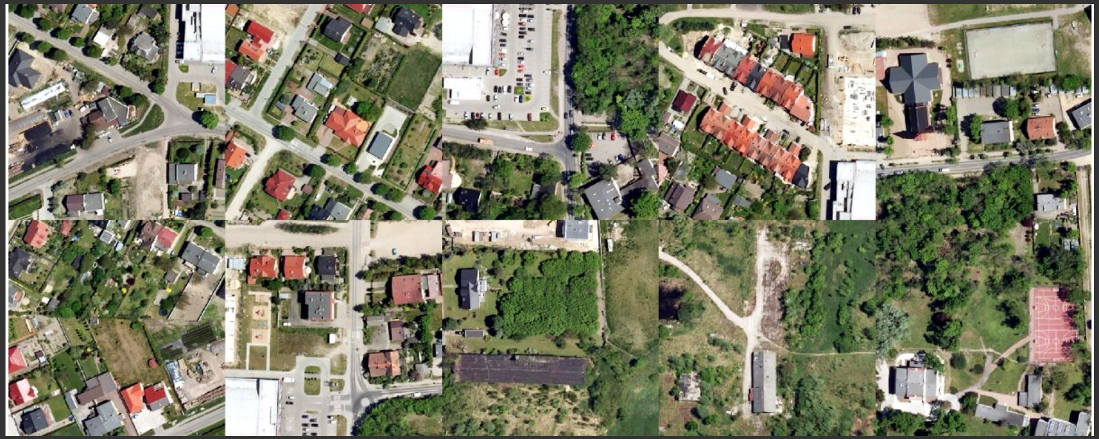
Data Prep Tasks

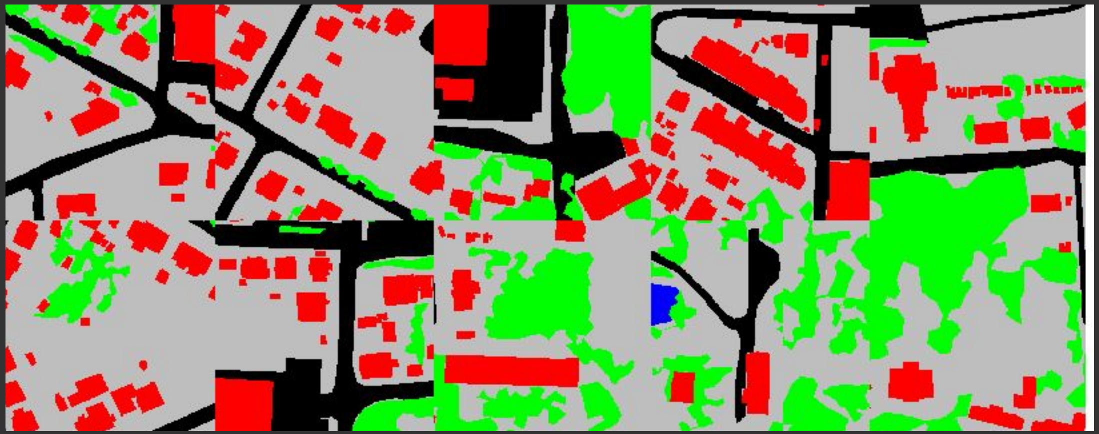


1. Generate raster masks
2. Chipping
3. Chip data frame
4. torch Dataset
5. torch DataLoader

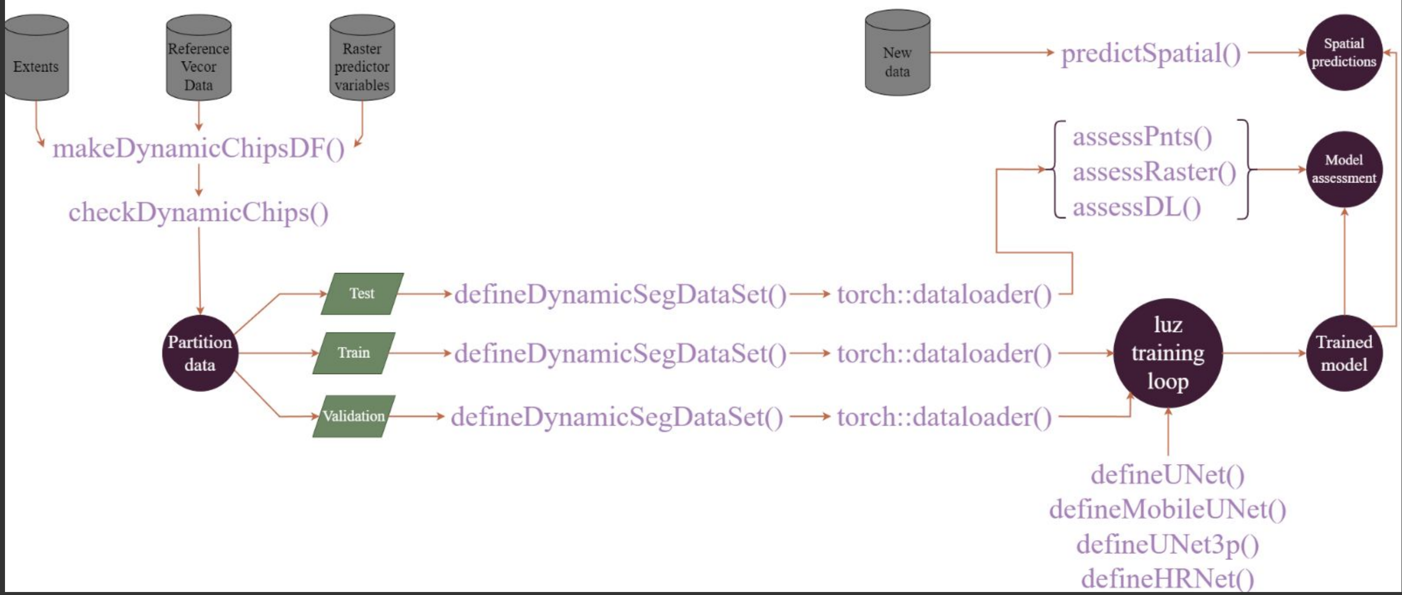
Option 1: Pre-Built Chips

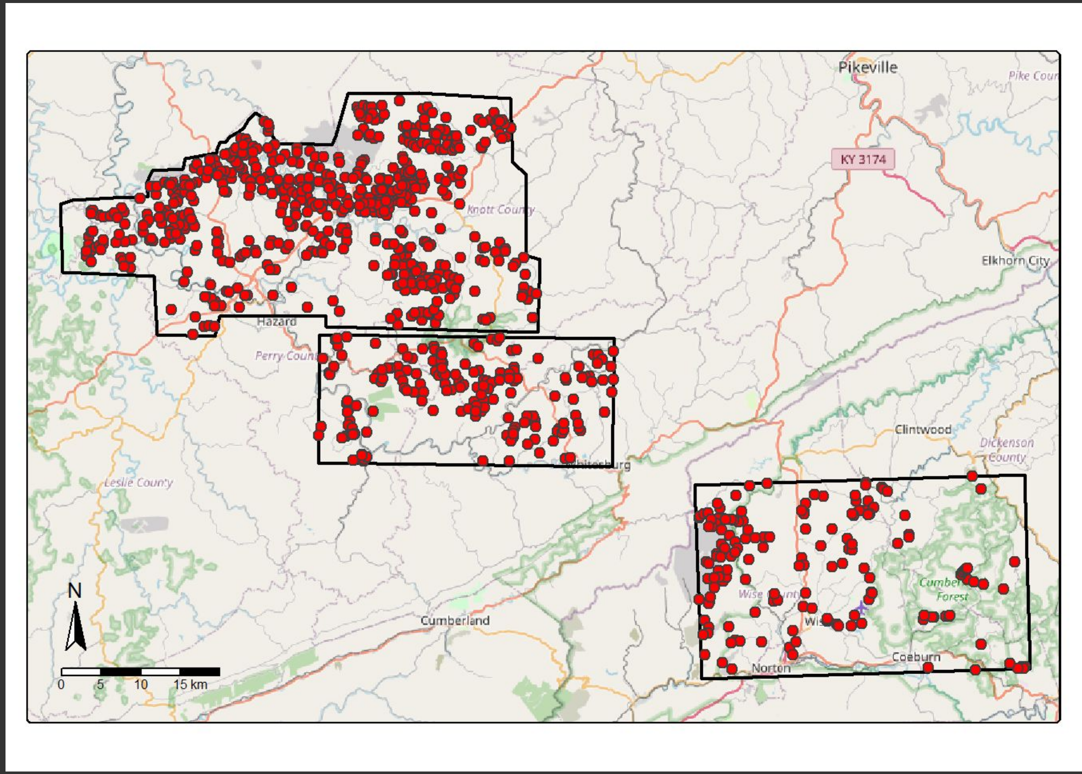






Option 2: Dynamically Generated Chips





Losses and Metrics

Unified Focal Loss Framework

	Distribution-Based	Compound	Region-Based
	$\lambda = 1$	$\lambda < 1$ & $\lambda > 0$	$\lambda = 0$
$\gamma > 0$ & $\gamma < 1$ $\delta \neq 0.5$	Focal CE Loss	Unified Focal Loss	Focal Tversky Loss
$\gamma = 1$ $\delta \neq 0.5$	CE Loss	Tversky + CE Loss	Tversky Loss
$\gamma = 1$ $\delta = 0.5$	CE Loss	CE + Dice Loss	Dice Loss

clsWghtsDist = relative weighting of classes in distribution-based loss (applied to each sample)
clsWghtsReg = relative weighting of classes in region-based loss (applied to each class when calculating a macro average)
useLogCosH = whether or not to apply a log cosh transformation to the region-based loss

```
myDiceLoss <- defineUnifiedFocalLoss(nCls=5,
                                  lambda=0, #Only use region-based loss
                                  gamma= 1,
                                  delta= 0.5, #Equal weights for FP and FN
                                  smooth = 1e-8,
                                  zeroStart=TRUE,
                                  clsWghtsDist=1,
                                  clsWghtsReg=1,
                                  useLogCosH =FALSE,
                                  device="cuda")
myDiceLoss(pred=pred,
           target=target)
```

Yeung, M., Sala, E., Schönlieb, C.B. and Rundo, L., 2022. Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized Medical Imaging and Graphics*, *95*, p.102026.

- `assessPnts()`
- `assessRaster()`
- `assessDL()`
- `luz_metric_overall_accuracy()`
- `luz_metric_f1score()`
- `luz_metric_recall()`
- `luz_metric_precision()`

$$\text{Overall Accuracy} = \frac{\text{Number of Features Correctly Classified}}{\text{Total Number of Features}}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

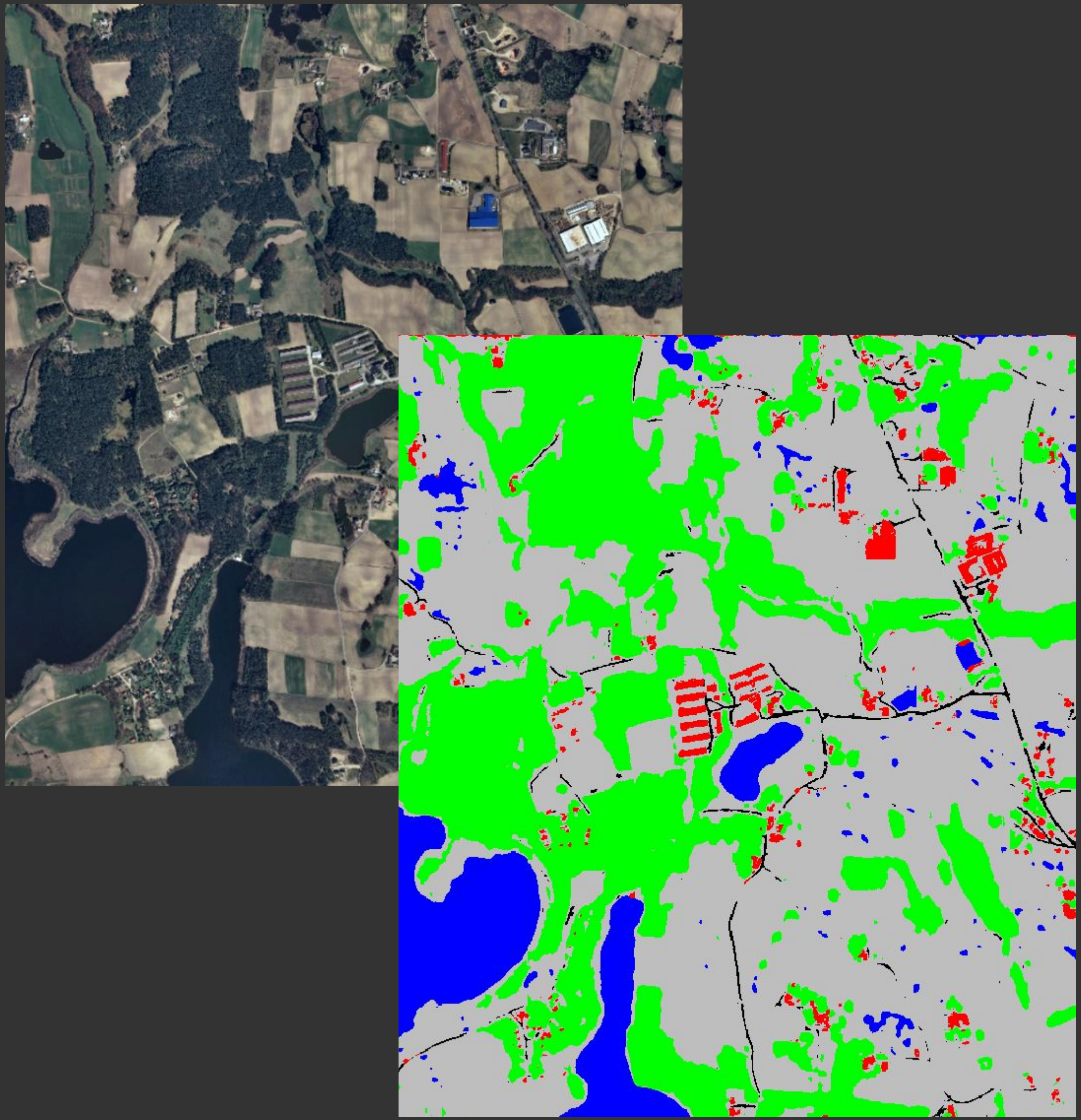
$$\text{F1-Score} = 2 \text{ X } \frac{\text{Precision X Recall}}{\text{Precision+Recall}}$$

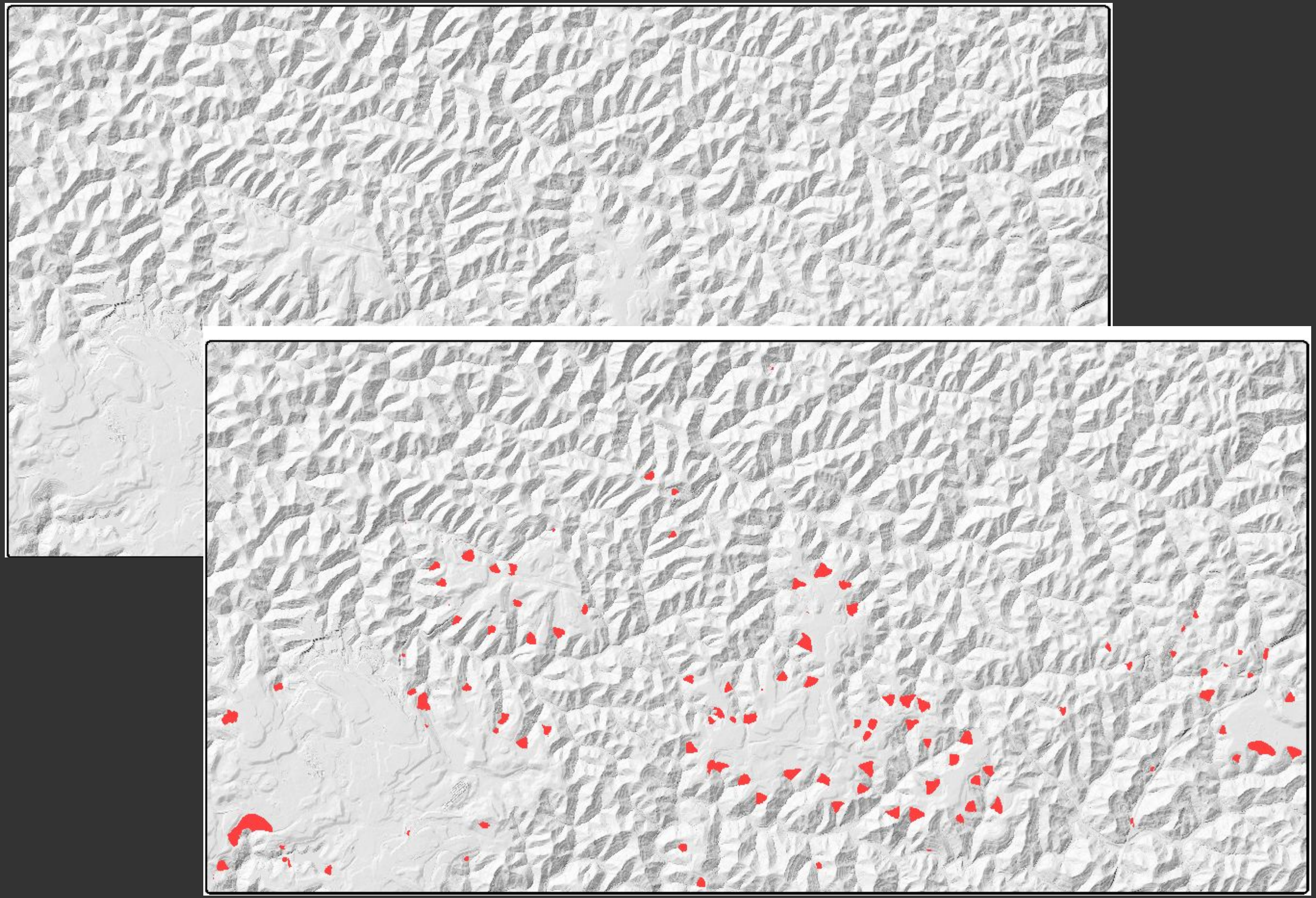
$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Negative Predictive Value} = \frac{TN}{TN + FN}$$

Spatial Predictions

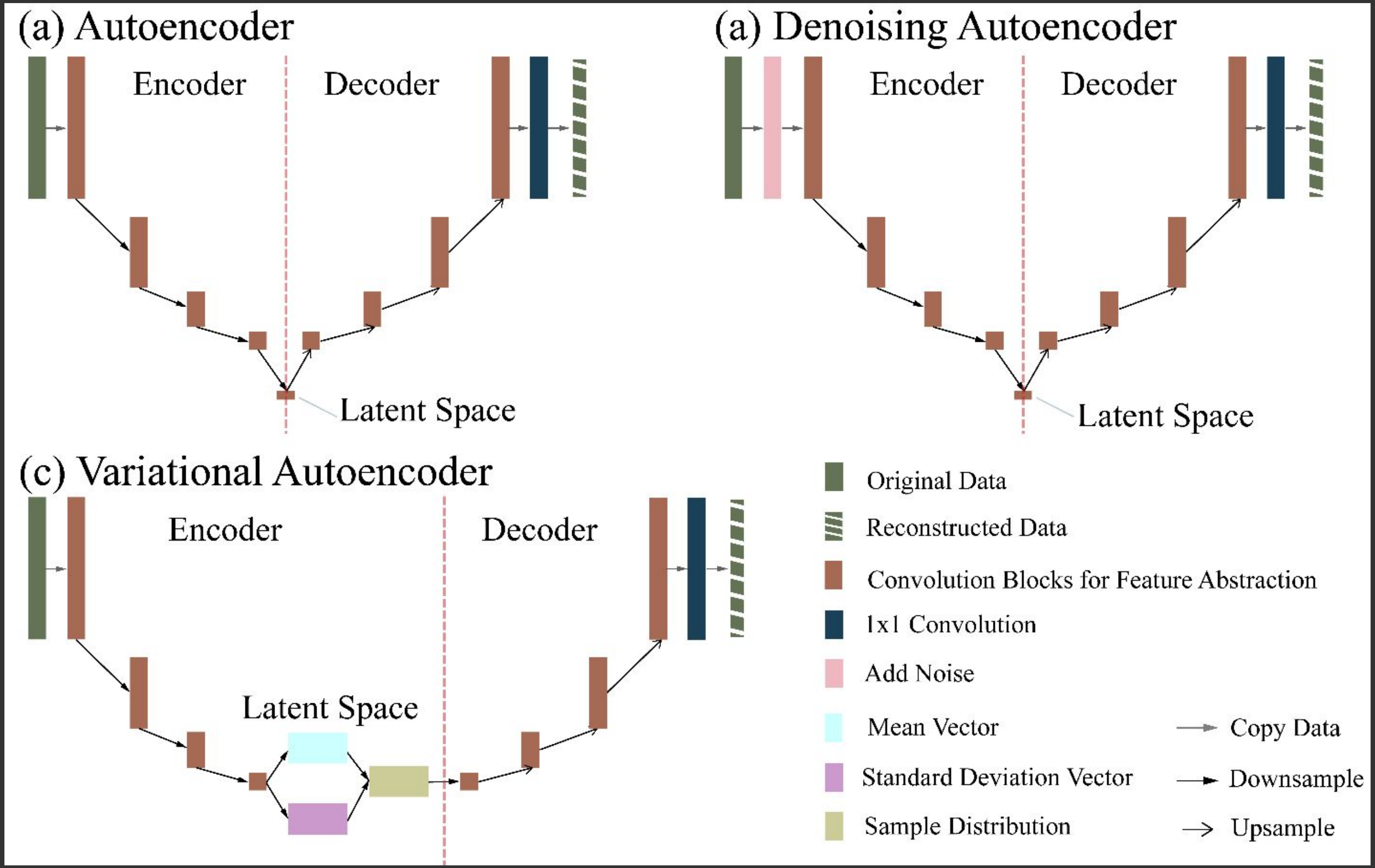
predictSpatial()





Further Development

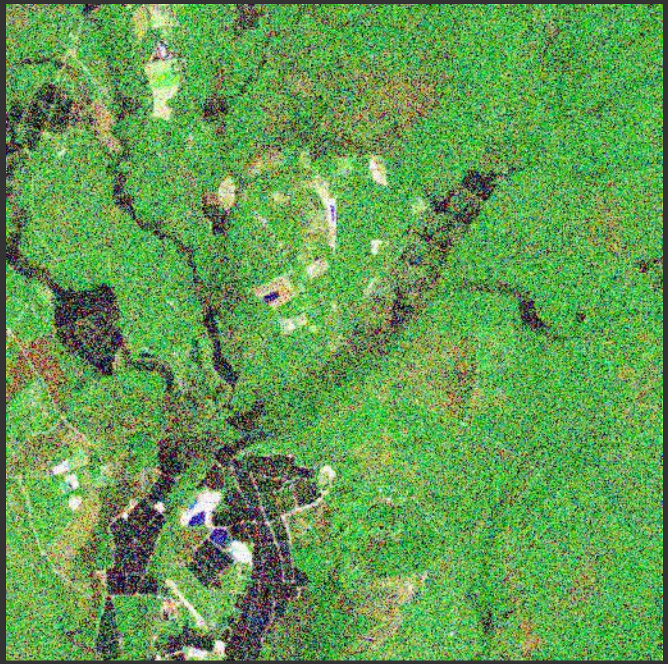
1. Autoencoders



1. Feature reduction
2. Anomaly detection

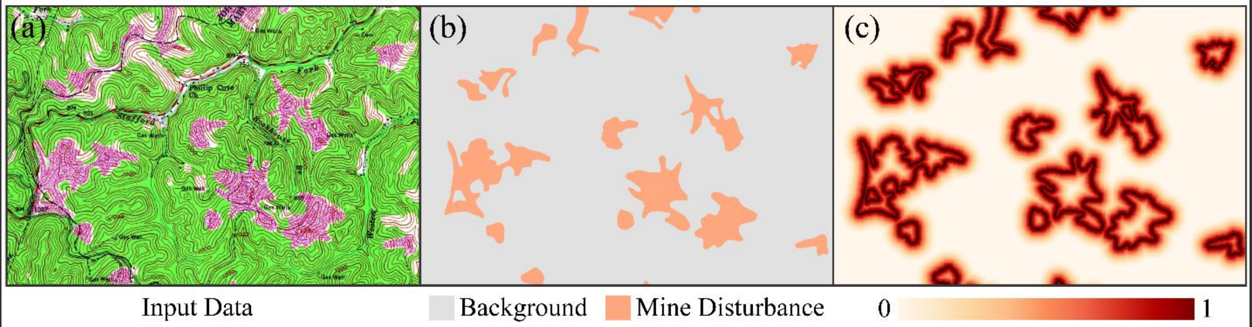
2. Pre-training

- Noise removal
- Gap filling



3. Edge-weighted loss functions

- Increase or decrease weight of cells near margins of features



4. Additional Architectures

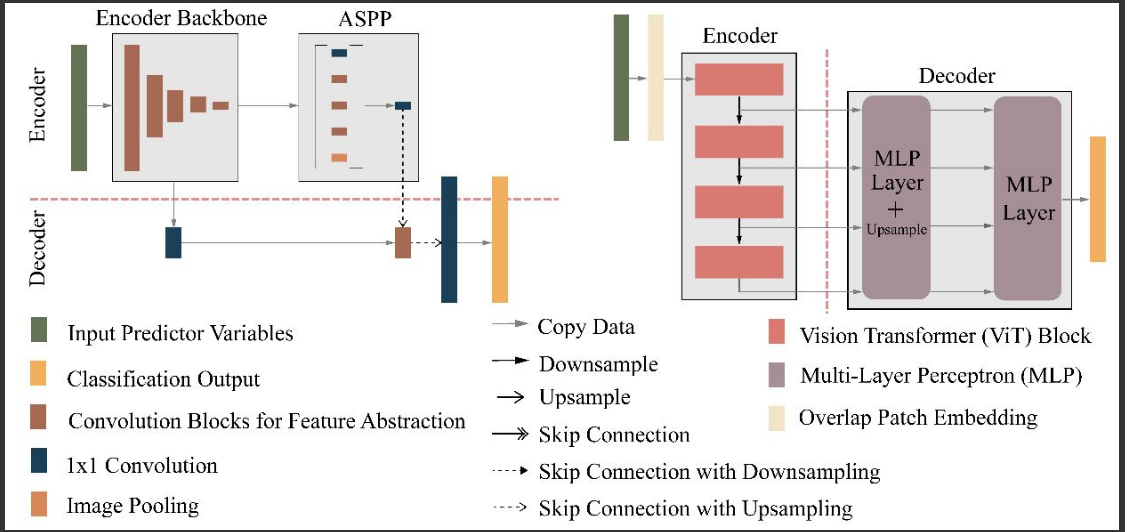
- DeepLabv3+

Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 801-818).

- SegFormer

Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J.M. and Luo, P., 2021. SegFormer: Simple and efficient design for semantic segmentation with transformers. *Advances in neural information processing systems, 34*, pp.12077-12090.

- Unet-SegFormer

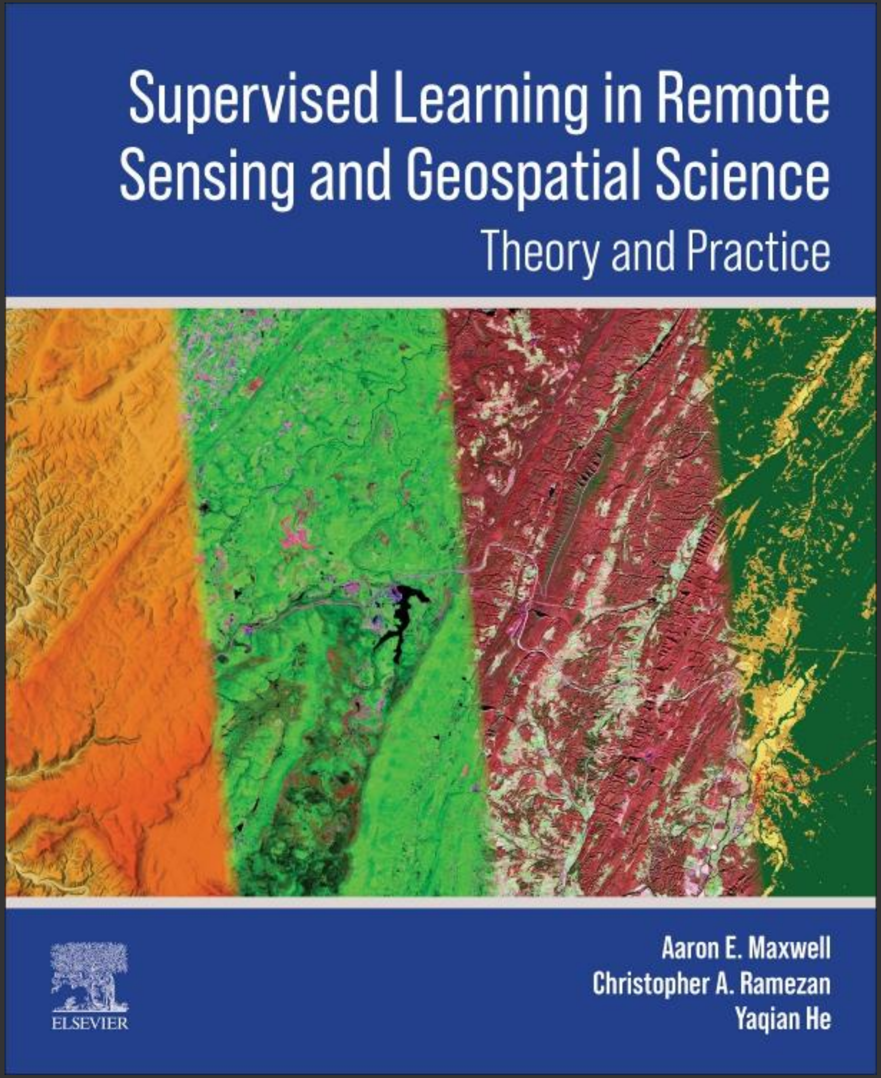


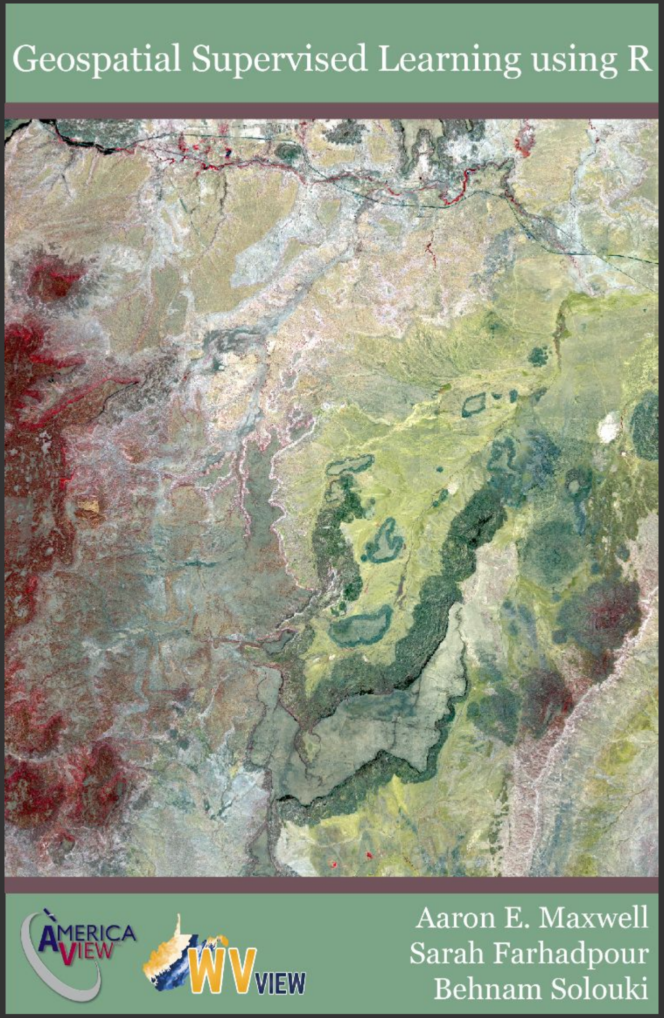
Others

1. Regression tasks
2. Variable learning rates
3. Ease loading, freezing, and unfreezing
4. Speed/efficiency

Conclusions

Geospatial Supervised Learning





See Chapter 15 and 16 of R text.

https://www.wvview.org/gslr.html

https://shop.elsevier.com/books/supervised-learning-in-remote-sensing-and-geospatial-science/e-maxwell/978-0-443-29306-1

Looking for collaborators/contributors!

Aaron.Maxwell@mail.wvu.edu

https://github.com/maxwell-geospatial/geodl

WV VIEW

https://www.wvview.org/

PLOS ONE

RESEARCH ARTICLE

geodl: An R package for geospatial deep learning semantic segmentation using torch and terra

Aaron E. Maxwell[1]*, Sarah Farhadpour[1], Srinjoy Das[2], Yalin Yang[1]

1 Department of Geology and Geography, West Virginia University, Morgantown, WV, United States of America, 2 School of Mathematical and Data Sciences, West Virginia University, Morgantown, WV, United States of America

Maxwell, A.E., Farhadpour, S., Das, S. and Yang, Y., 2024. geodl: An R package for geospatial deep learning semantic segmentation using torch and terra. *PLoS One*, *19*(12), p.e0315127.