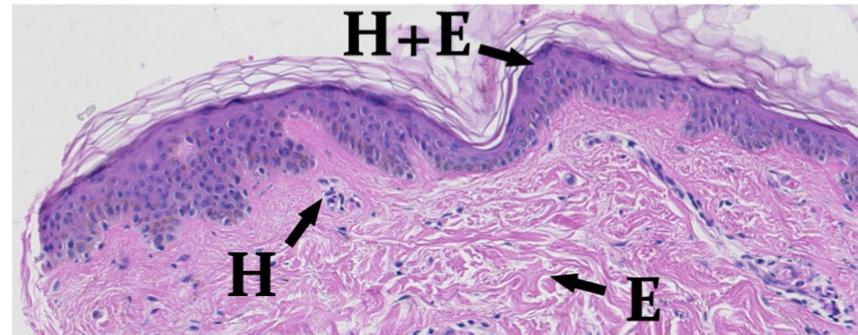


# TILseg

Ryan B., Braden C., Braden G., Cyrus H., Abishek S.

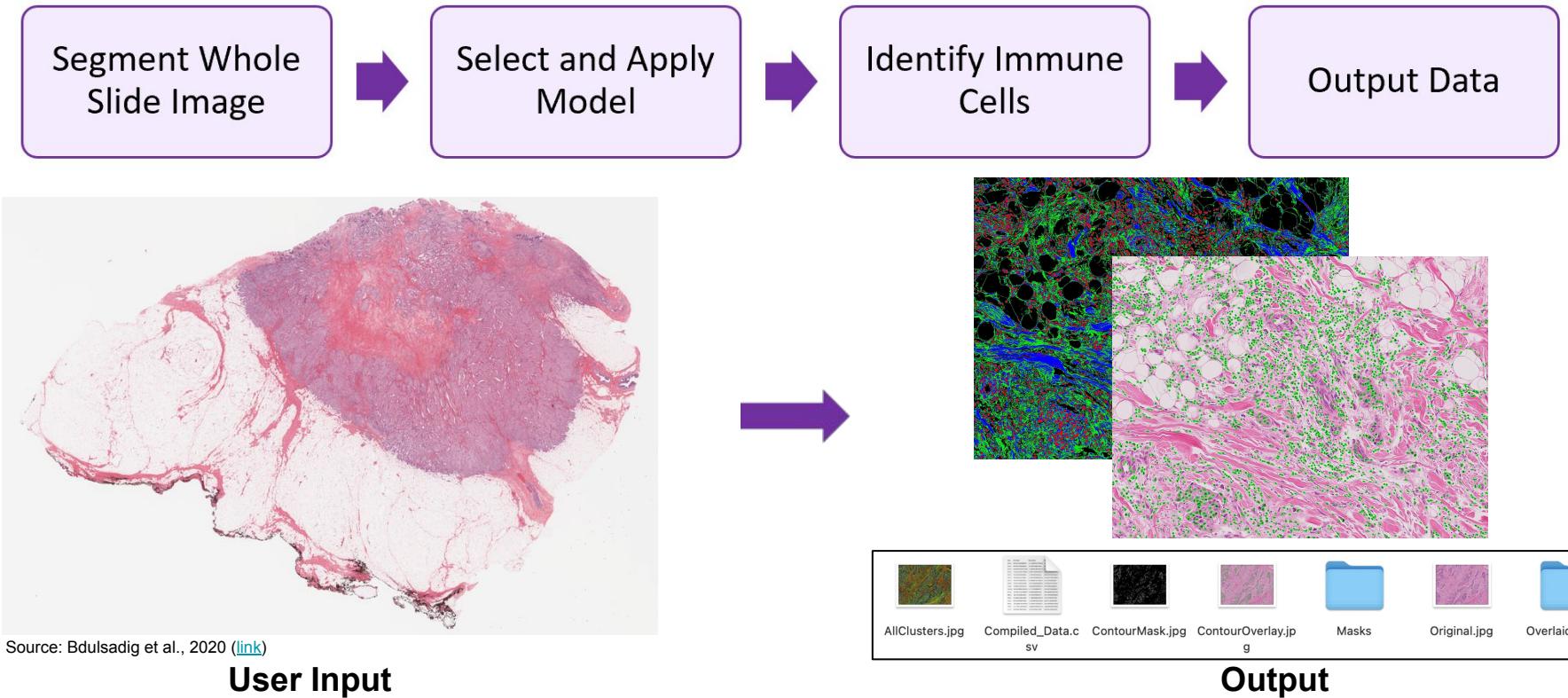
# Background

- Breast cancer diagnosis performed using hematoxylin & eosin (H&E) stained tumor biopsies
- Digital pathology can be integrated into diagnosis pipeline to mitigate inter-observer variability and for clinical diagnosis validation
- Tumor-infiltrating lymphocytes (TILs) are clinically relevant immune cells in the tumor microenvironment
- Stromal TIL density and architecture has statistical significance with clinical outcome
- This project aims at accurately and reliably segment TILs from digitized H&E images

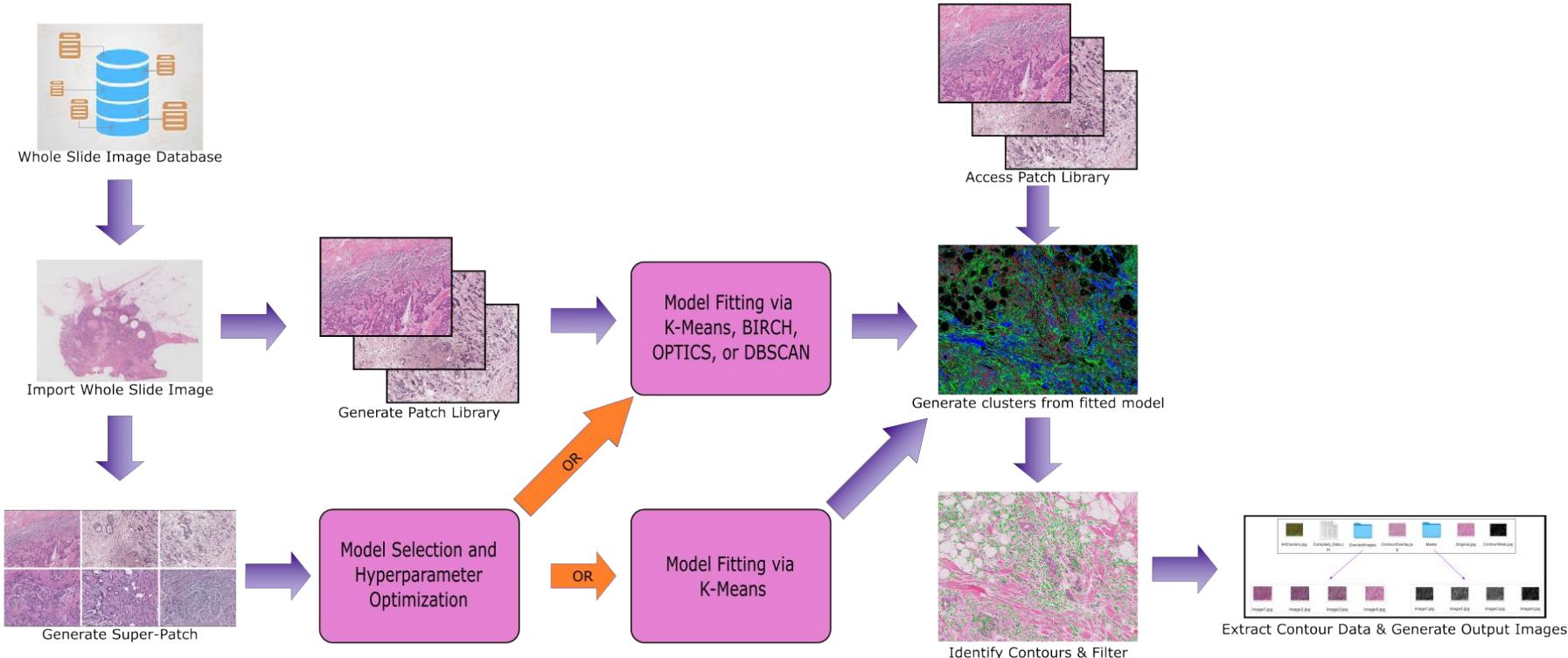


<https://www.sciencedirect.com/science/article/pii/S0895611119301016>

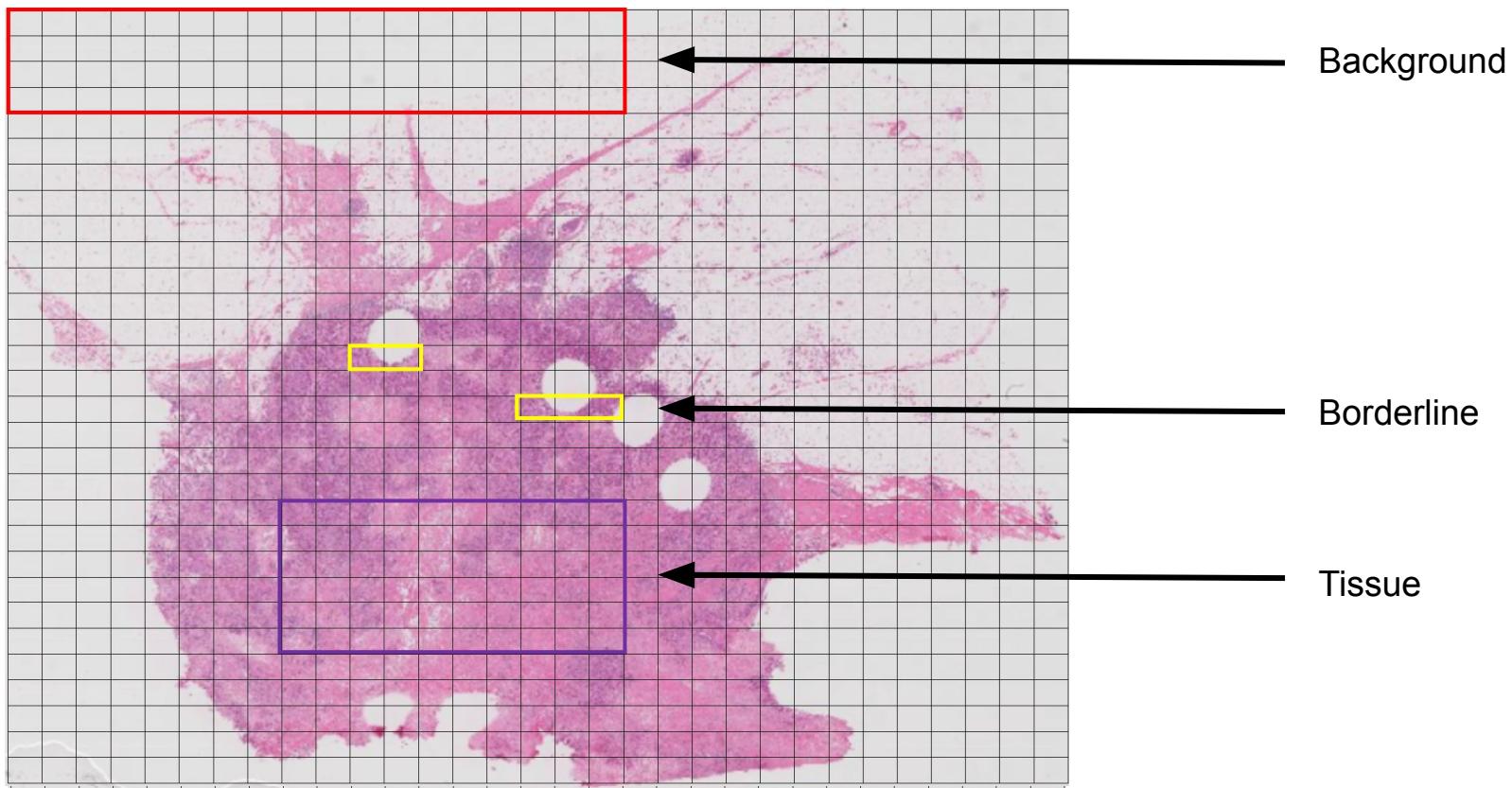
# Workflow Overview



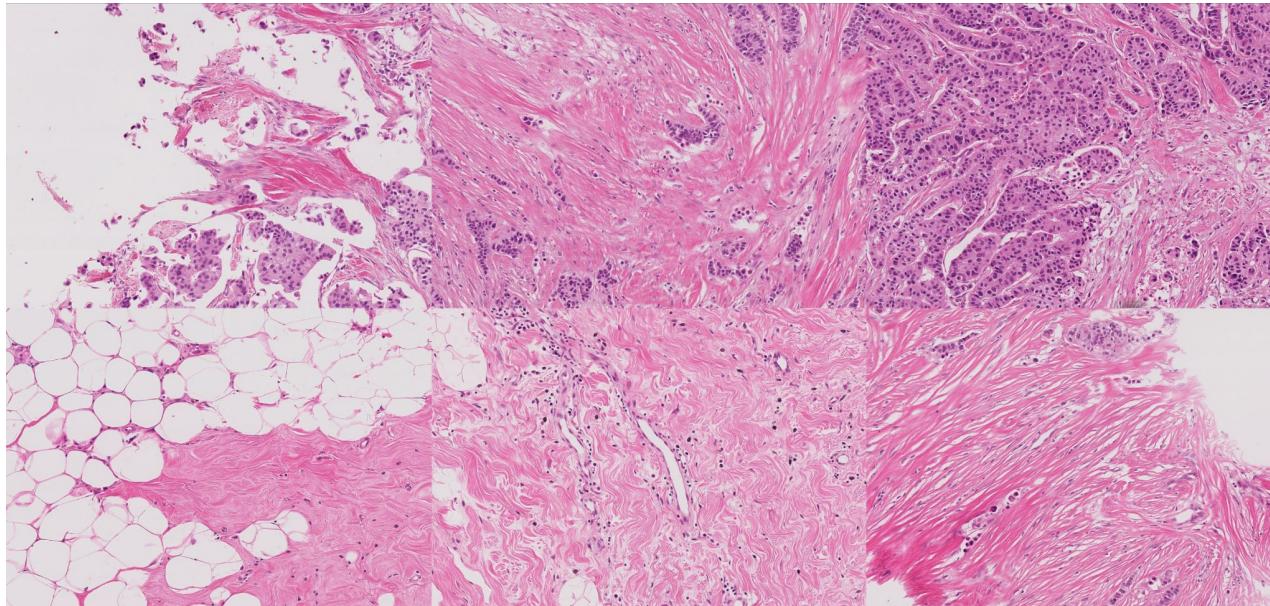
# Methodology



# Whole Slide Segmentation

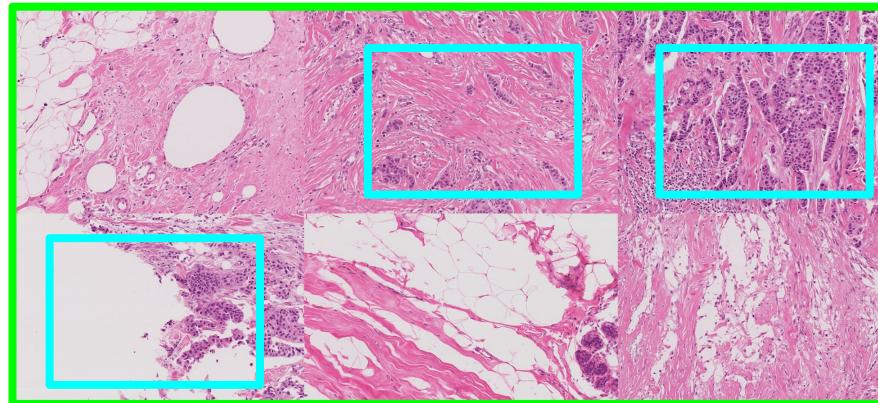


# Superpatch Creation



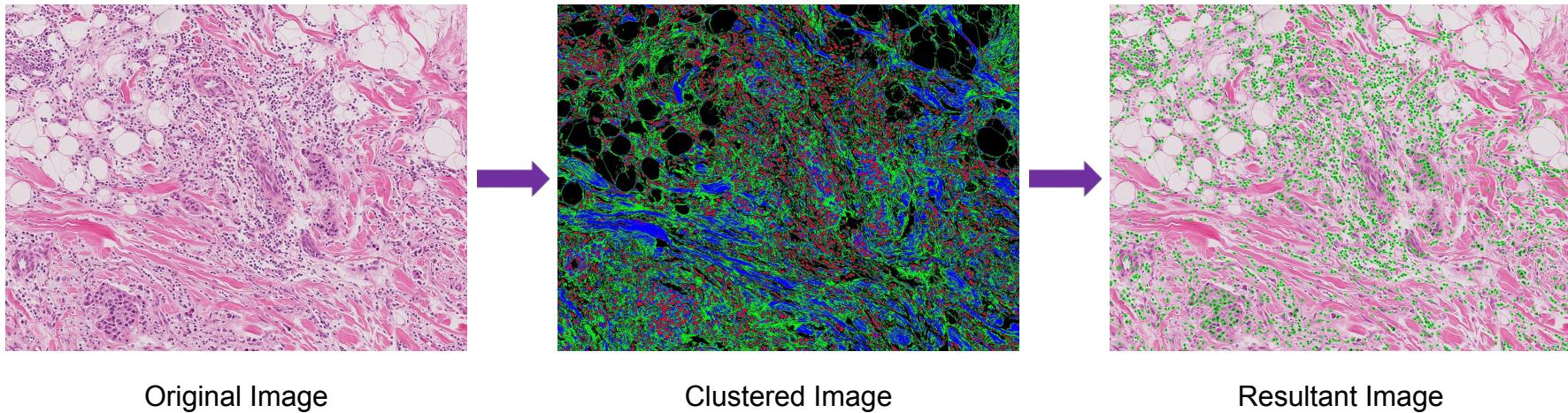
- Filter out background patches
- Stitch together tissue patches of varying color in user defined arrangement

# Model Fitting and Scoring



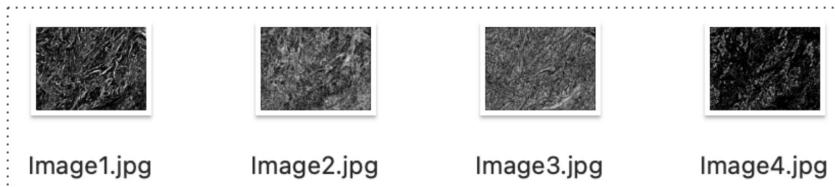
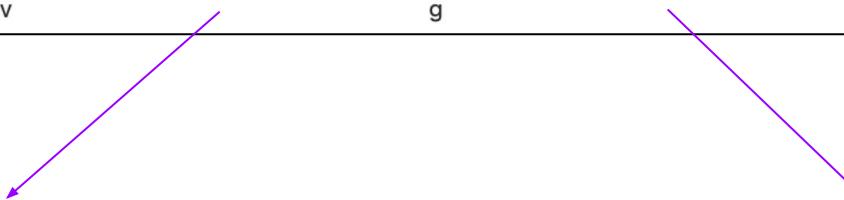
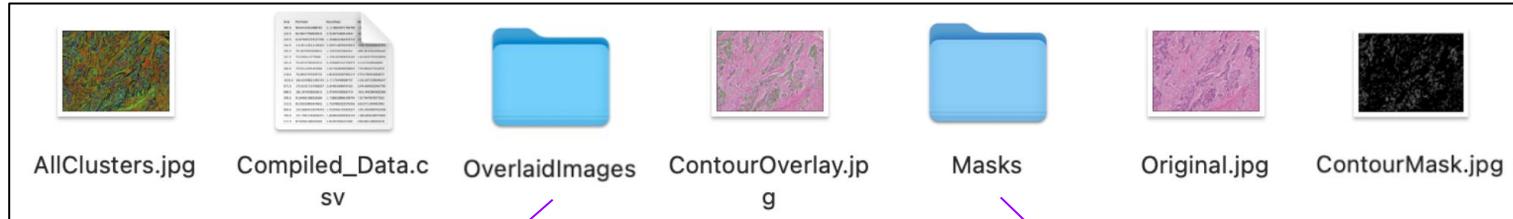
- Optimizes hyperparameters using the superpatch
- KMeans and BIRCH run on entire superpatch (green), but DBSCAN and OPTICS require smaller subpatches of superpatch or independent patches (blue)
- Scoring is performed on patches to determine clustering “goodness” using Silhouette score, Calinski-Harabasz index, and Davies-Bouldin score

# Immune Cell Identification



- Clustering is first performed using:
  - Fitted Model
  - Algorithm (K-Means, DBSCAN, OPTICS, BIRCH)
  - Optimized hyperparameters
- Contours derived from clusters
  - OpenCV functionality
  - Identifies region boundaries
  - Applied filters for roundness & area

# Data Output



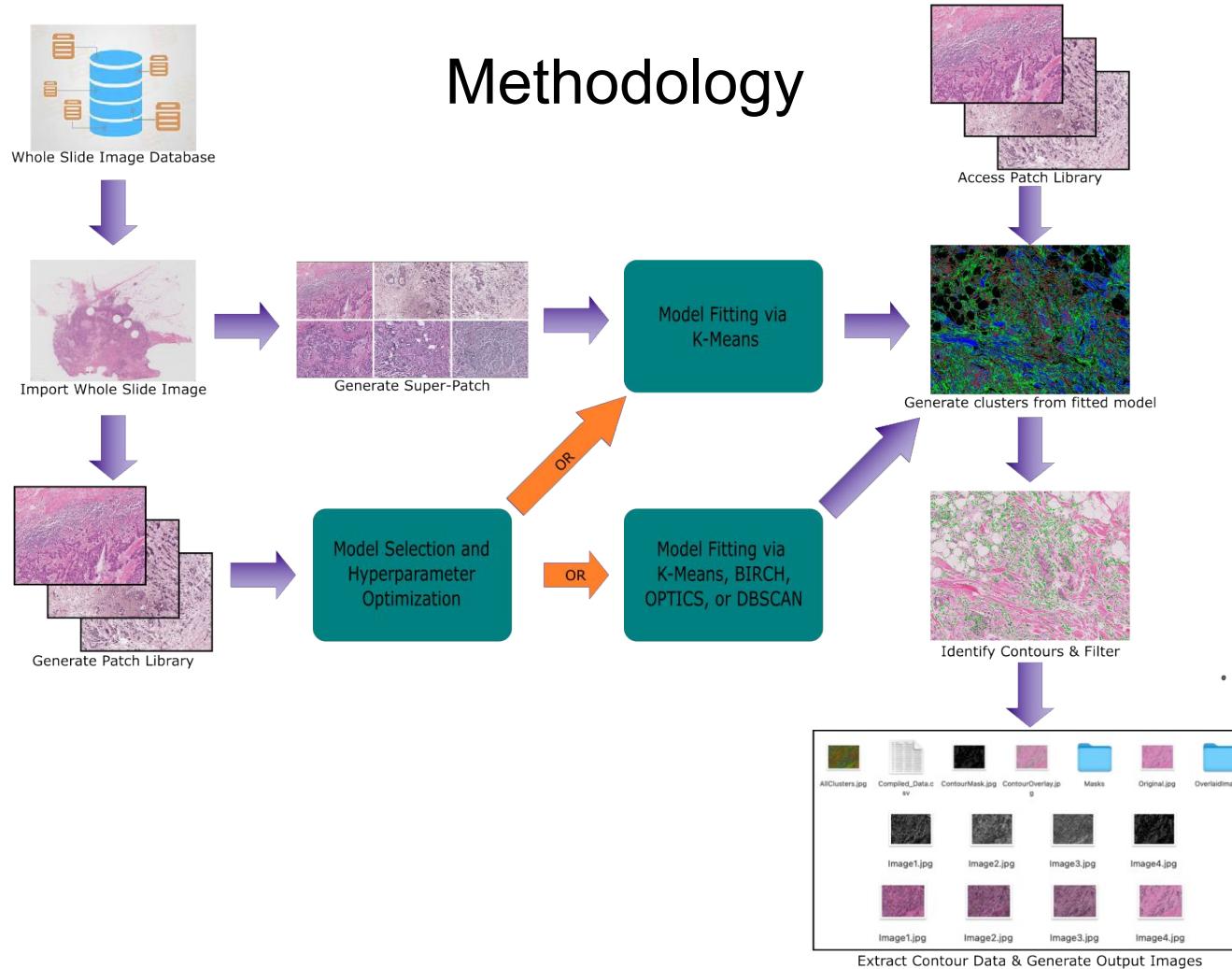
- Default output schema for each patch
- Users can specify desired file outputs

# Future Work

- Modify hyperparameter optimization to be algorithm specific
- Use supervised model to apply DBSCAN, and OPTICS, to other patches
- Find external computing source that can run software on entire superpatch
- Iteratively optimizing contour filters for TIL isolation from their cluster
- Performing clustering on contour features to identify TILs
- Use a supervised classification model to sort patches based on location (stromal or epithelial). Integrate tissue mask into tilseg pipeline to specifically segment stromal TILs.

# Questions?

# Methodology



# Workflow Overview

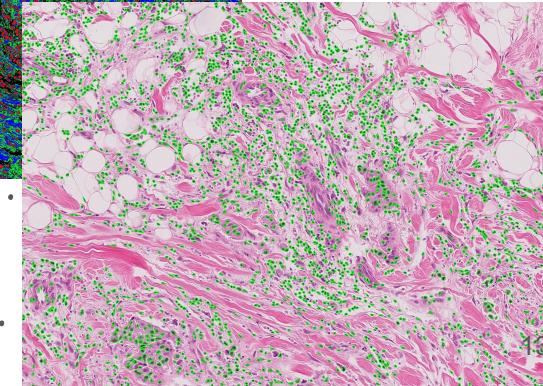
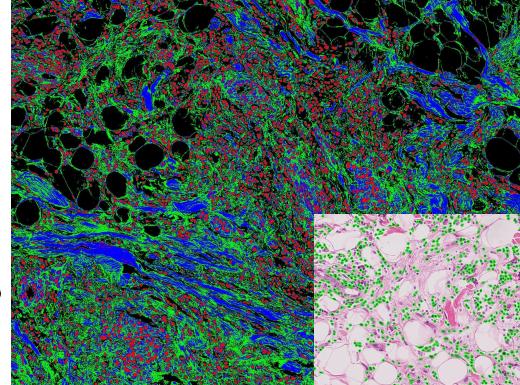
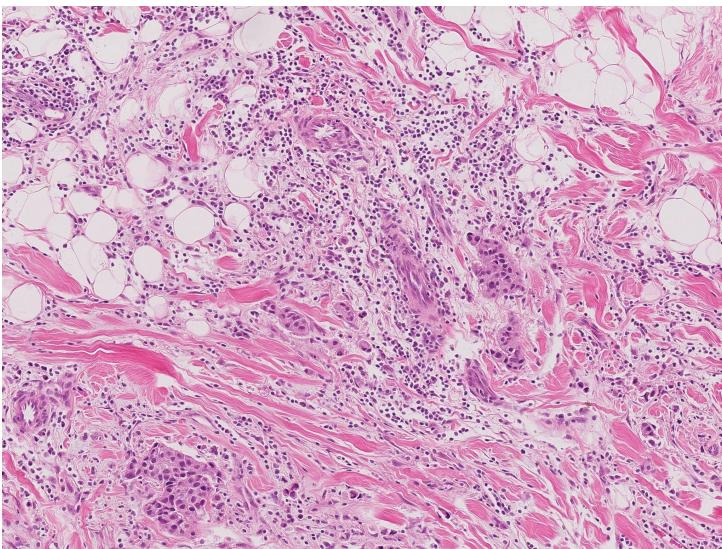
Patch & Super-Patch Segmentation

Model Identification & Training

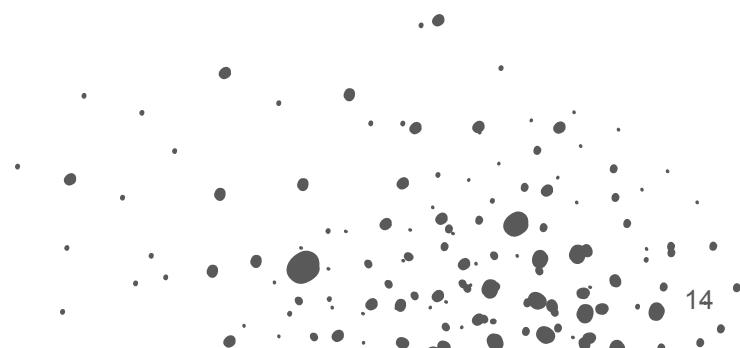
Model Fitting & Scoring

Immune Cell Identification

Image & Data Outputs

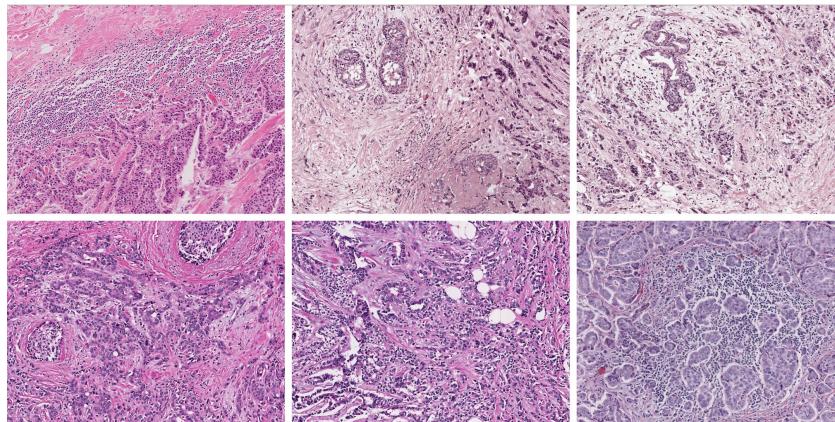


# Patch & Super Patch Creation



# Model Fitting & Scoring

- There are two ways of predicting clusters
  - Fitting a KMeans model to a super patch and then predicting clusters on a patch of interest

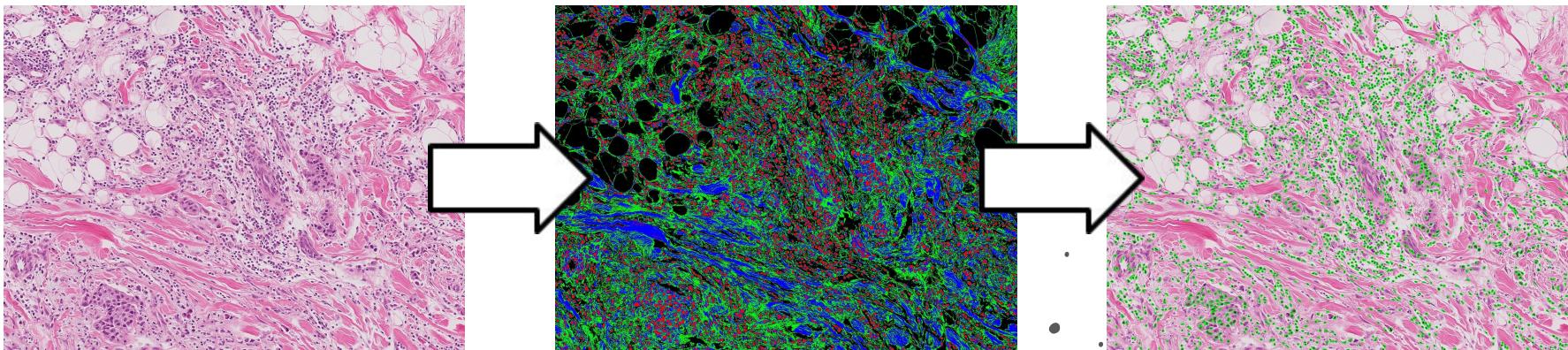


..

- Fitting a KMeans, DBSCAN, BIRCH, or OPTICS model to the same patch before predicting
- Scoring is performed on patches to determine clustering “goodness” using Silhouette score, Calinski-Harabasz index, and Davies-Bouldin score

# Immune Cell Identification

- Clustering is first performed using:
  - Fitted Model
  - algorithm (K-Means, DBSCAN, OPTICS, BIRCH)
  - Optimized hyperparameters
- Contours derived from clusters
  - OpenCV functionality
  - Identifies region boundaries
  - Applied filters for roundness & area



# Image & Data Outputs

- TILs in each patch are counted and outputted (to be used in further downstream pathological analyses)
- File outputs for each patch having TILs segmented:



AllClusters.jpg



sv



ContourMask.jpg



g



Masks



Original.jpg



OverlaidImages

- Masks directory contains binary masks for each cluster predicted:



Image1.jpg



Image2.jpg

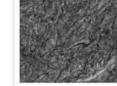


Image3.jpg



Image4.jpg

- Masks directory contains individual clusters overlaid on original H&E patch:

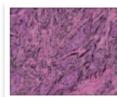


Image1.jpg

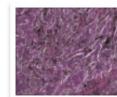


Image2.jpg

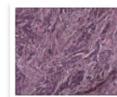


Image3.jpg

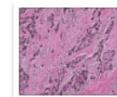


Image4.jpg

- Functionality for users can specify what file outputs they require

# Future Work

- Continue to optimize hyperparameters for DBSCAN, BIRCH, and OPTICS algorithms using `model_selection`
- Using computing resources to explore performing DBSCAN, BIRCH, and OPTICS on whole patches
- Iteratively optimizing contour filters for isolating TILs from their cluster with validation from a clinical pathologist
- Use output CSV file to apply clustering algorithms to contour features to better filter TILs
- Stromal TILs are statistically more significant than epithelial TILs. A supervised classification model can be used to stratify patches into stromal and epithelial regions. The tissue mask can then be integrated into the tilseg pipeline to segment stromal TILs in particular.