THE BOUNDING SHAPE GENERATOR

Charles Garrett Eason

Binbin Wu

Project Goals

- The basic idea: If we have a well trained classifier we get can get location and shape information quickly and for free.
- No need for:
  - *Additional training*
  - *Network augmentation*
  - *Additional human annotation*
- Process should be:
  - *Quick*
  - *Not computationally intensive*
  - *Simple*

Prior Studies

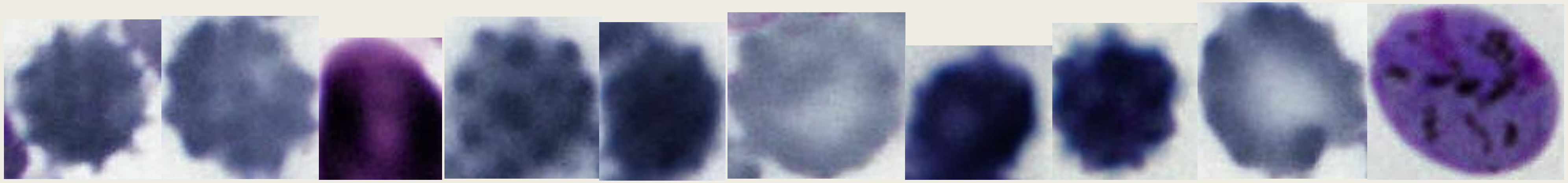
- Some similar prior studies focused on:
  - *Augmenting typical CNN architecture (Weakly-supervised object localization)*
    - Alternative Loss functions (Oquab, Bottou, Laptev, & Sivic, 2015; Ribera, Guera, Chen, & Delp, 2019)
    - Replacing fully connected layer with convolution layer (Oquab et. al., 2015)
    - Using global max pooling to “highlight” learned patterns (Oquab et. al., 2015)
  - *Using human verification to assist in network training.*
    - Iterative training with subset annotations (Papadopoulos, Uijlings, Keller, & Ferrari, 2016)
    - Iterative training, bounding box re-localizing / generation, and human verification (Papadopoulos et. al., 2016)

Data Set – Fruits 360

- Total Images: 82213
- Number of Classes: 120
- Image Size: 100x100x3

Data Set – Exam 1 and Exam2

- Total Images: 1364 (~80,000 cells)
- Number of Classes: 7 (we only classified 4 types)
- Image Size: 1200x1600x3 (median size)
- For training the MLP we used the data extrapolated from this data set for Exam 1.



Cropping - Predicting - Mapping

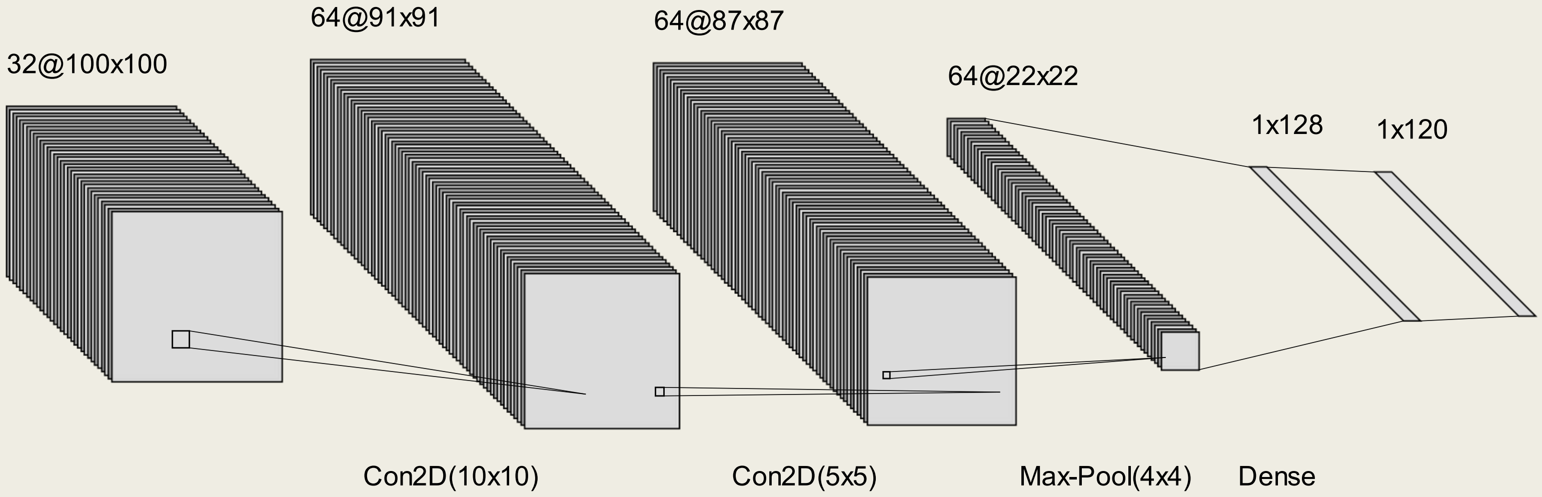
- The Cropping - Predicting - Mapping Algorithm is comprised of 4 steps:
  1. *Crop sub-images using a sliding window kernel.*
  2. *Feed sub-images into neural network classifier and obtain predictions.*
  3. *Map the original locations of the sub-images back onto the original image with binary mapping (0 or 1 for the entire sub-image area).*
  4. *Sum the binary maps across the original image to obtain the full map of identified object shapes and locations.*

Cropping - Predicting - Mapping

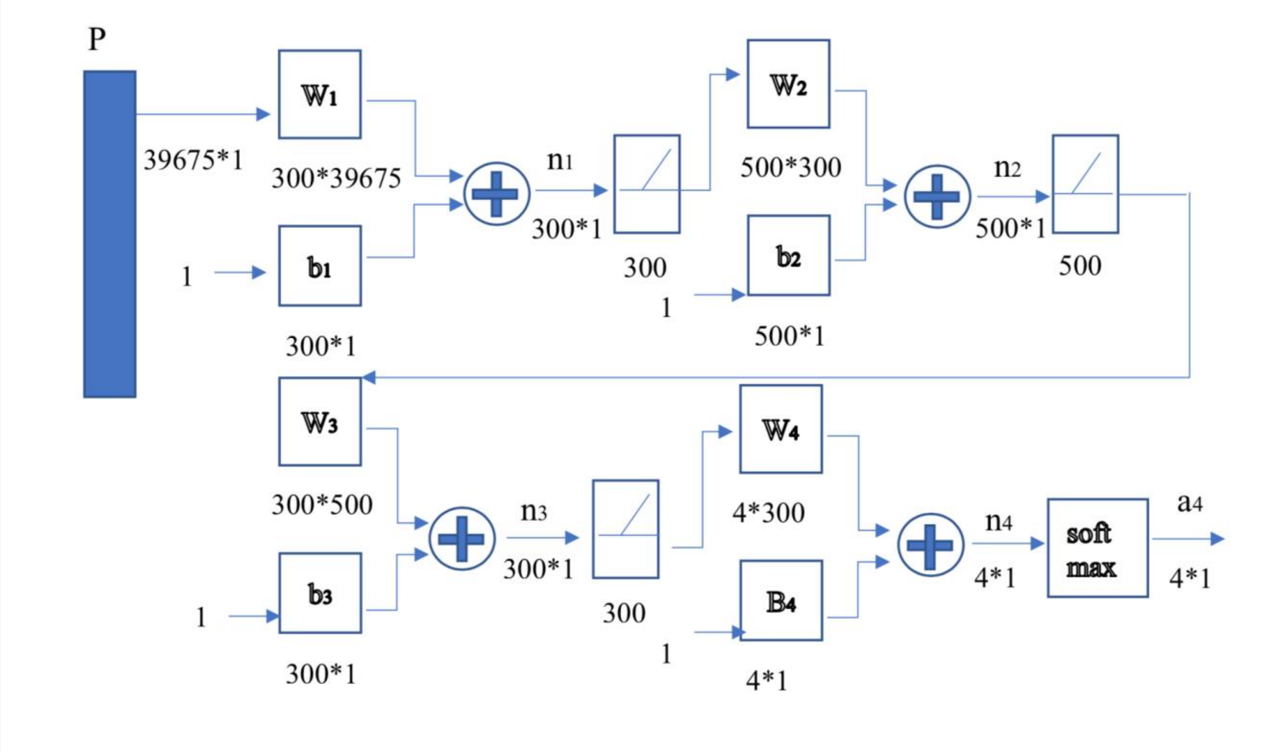
- Hyperparameters:
  1. *Kernel Size*
     - *The size of the cropping window.*
  2. *Stride*
     - *The step of the cropping window.*
  3. *Resize*
     - *The size the image needs to be as an input into the classifier.*
  4. *Threshold*
     - *The value of the softmax probability output to be ignored if it is too low.*

Explanation Example

Fruit-360 Convolution Network Design



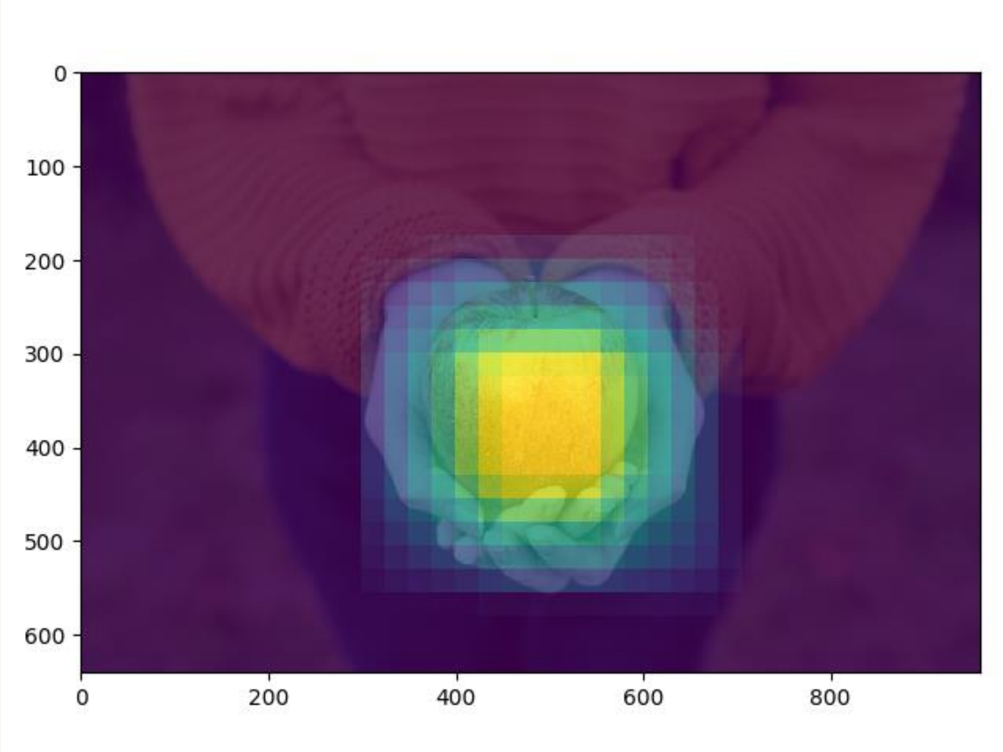
Cells Multi-Layered Perceptron Network



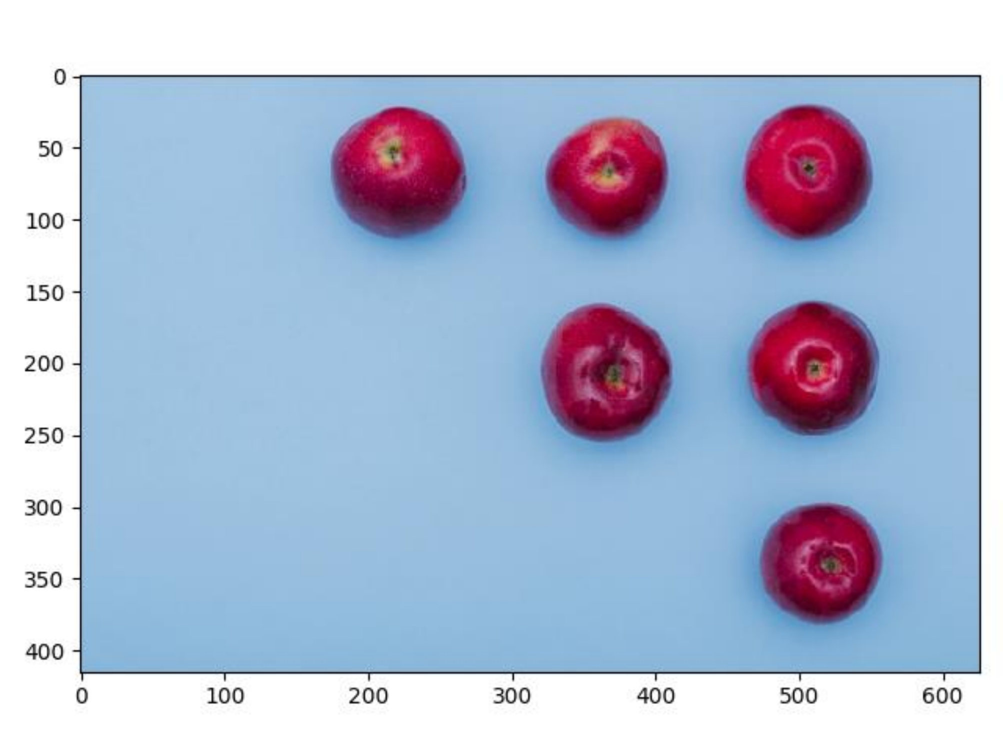
Other Attempted Avenues

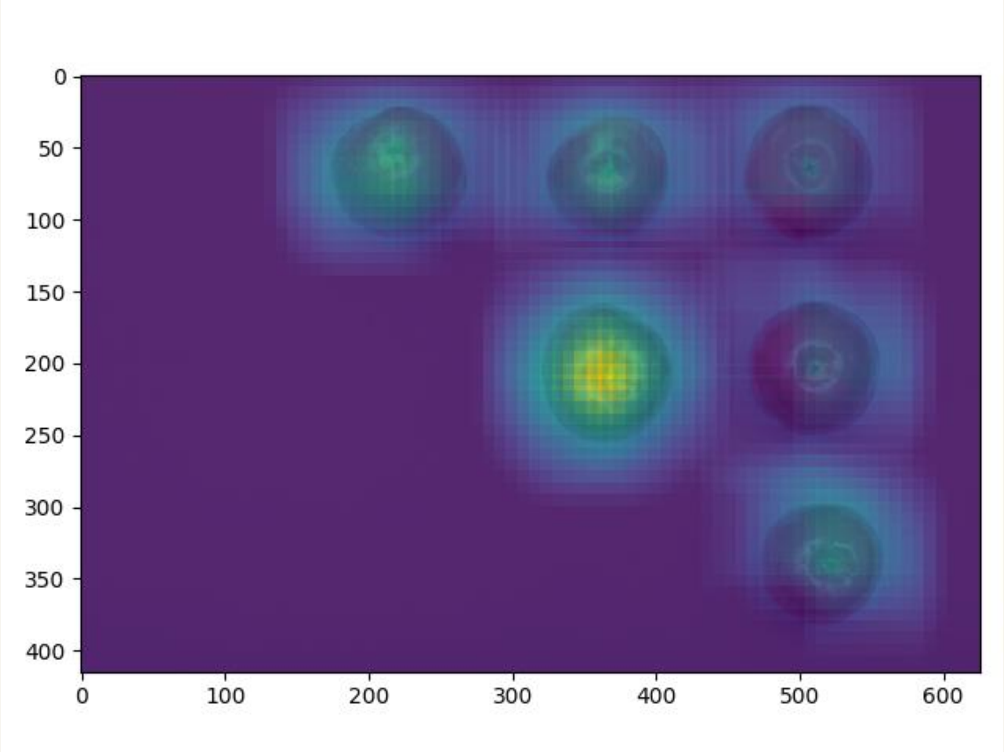
- Pre-trained Densenet
  - *Took too long to load framework and performed equally or worse than our non-pretrained networks.*
- Image Augmentations
  - *Shift, rotation, mirror, brightness, channel-shift, zoom, rescale*
- Manipulated network architecture
  - *Layers, kernel sizes, learning rates, batch sizes (need smaller batch sizes due to hardware constraints), dropout, pooling, multilabel vs. multiclass, etc.*

Fruit-360 Results – Bounding Shape

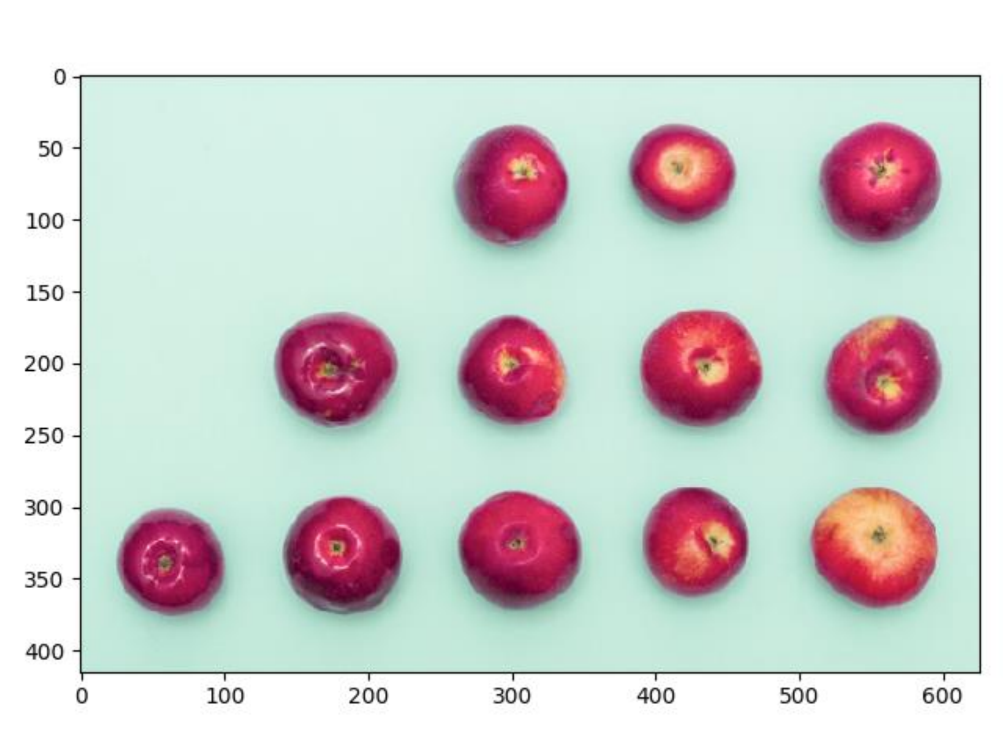


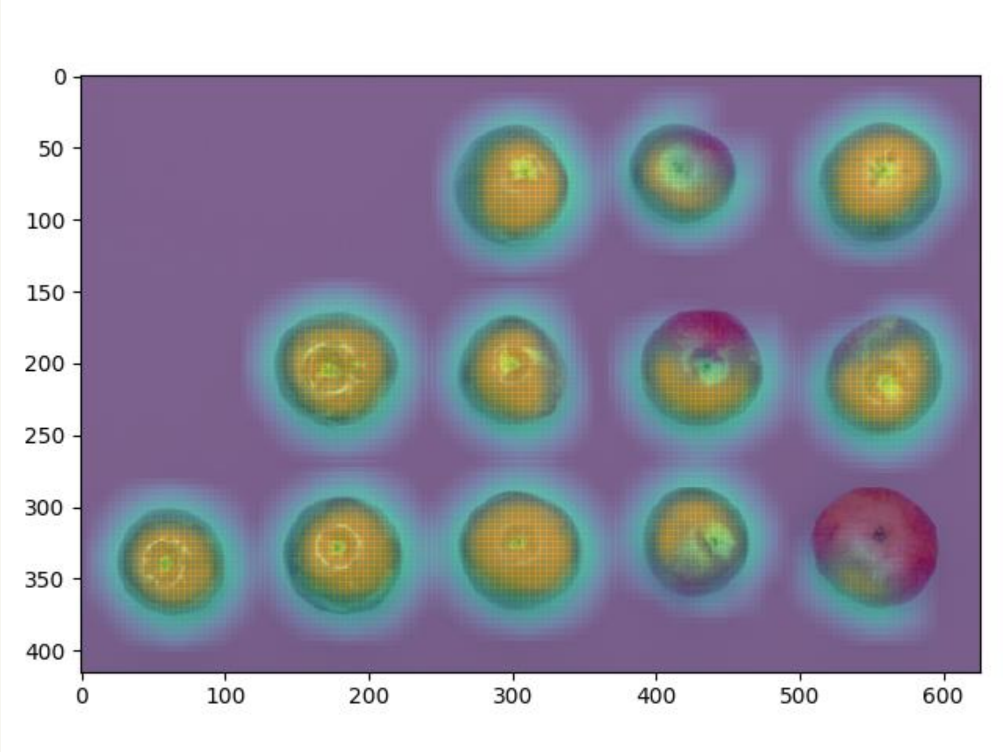
Fruit-360 Results – Bounding Shape



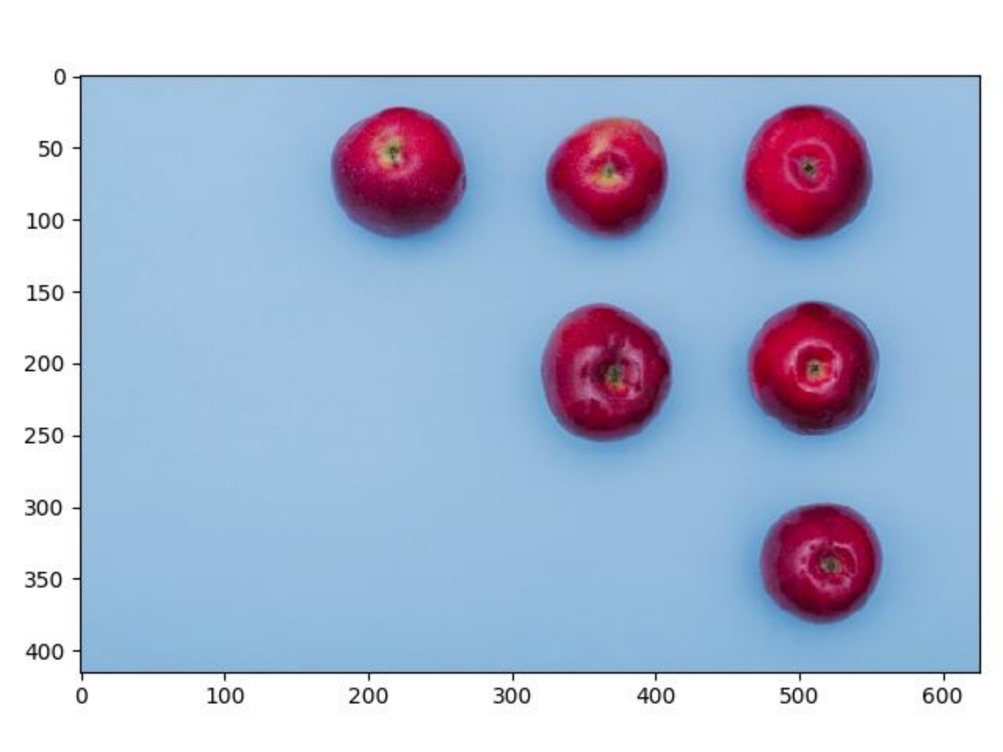


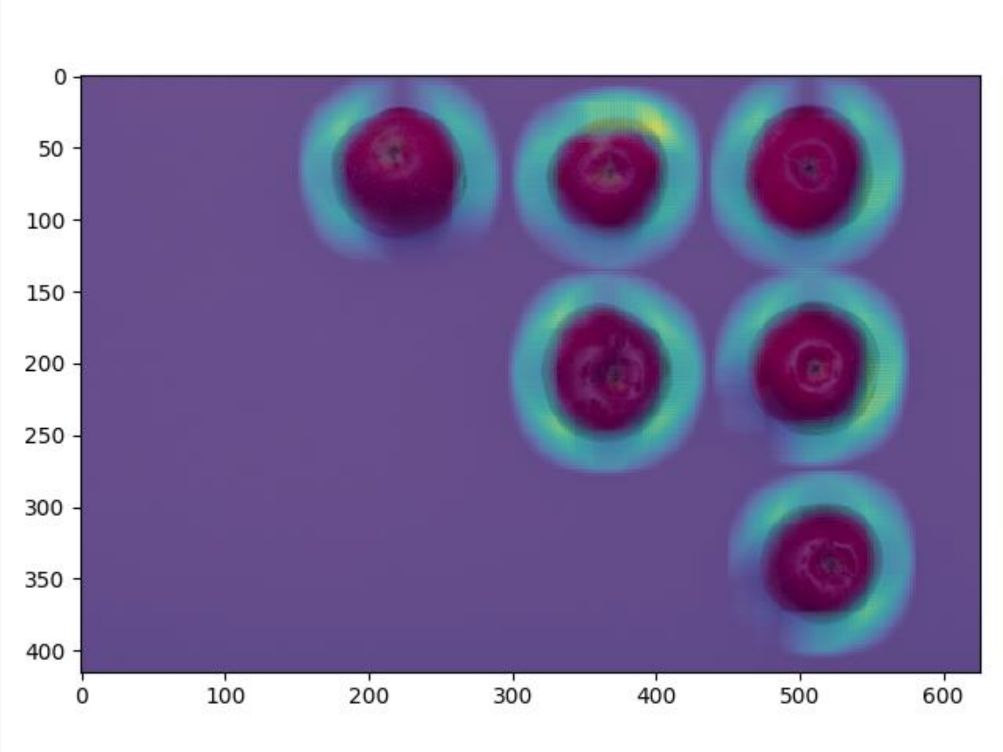
Fruit-360 Results – Bounding Shape



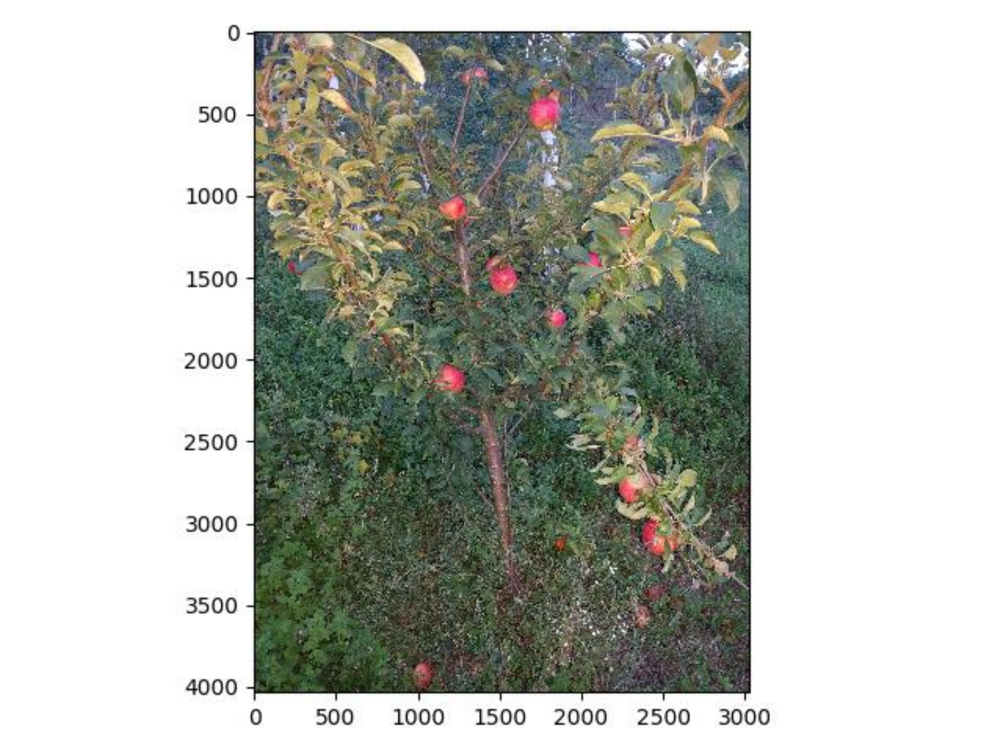


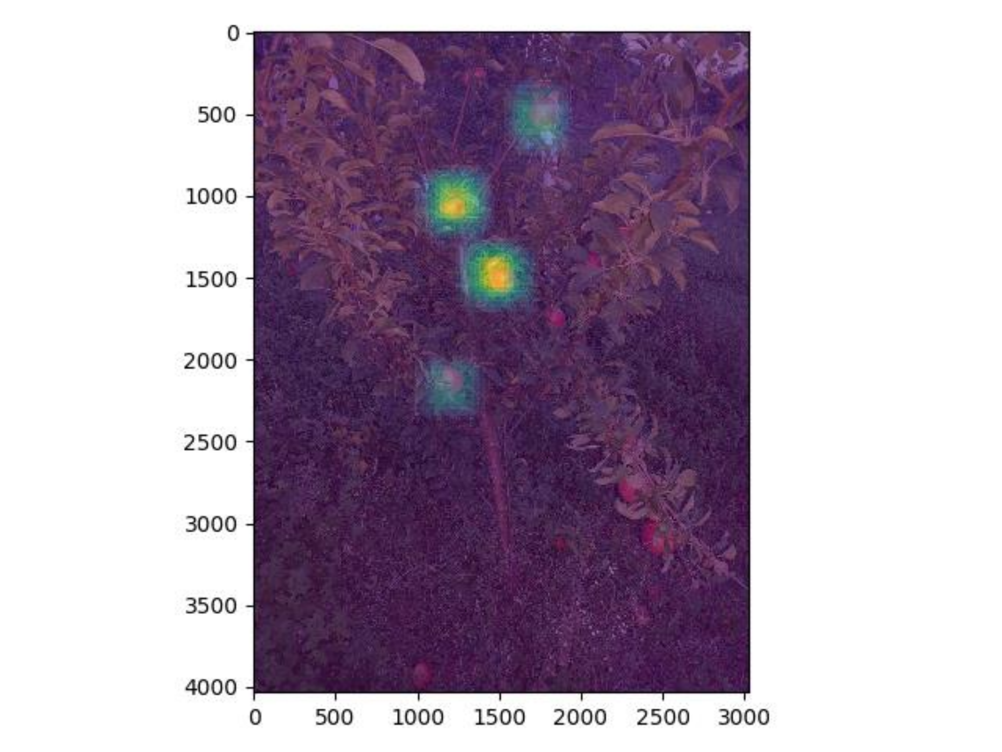
Fruit-360 Results – Smaller Kernels



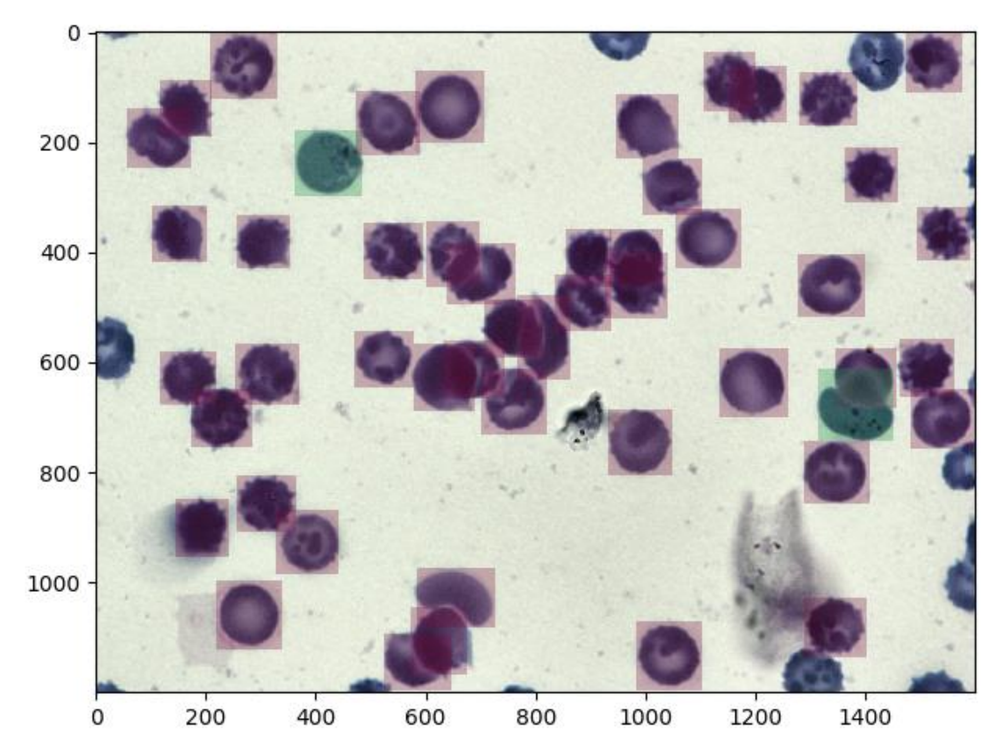


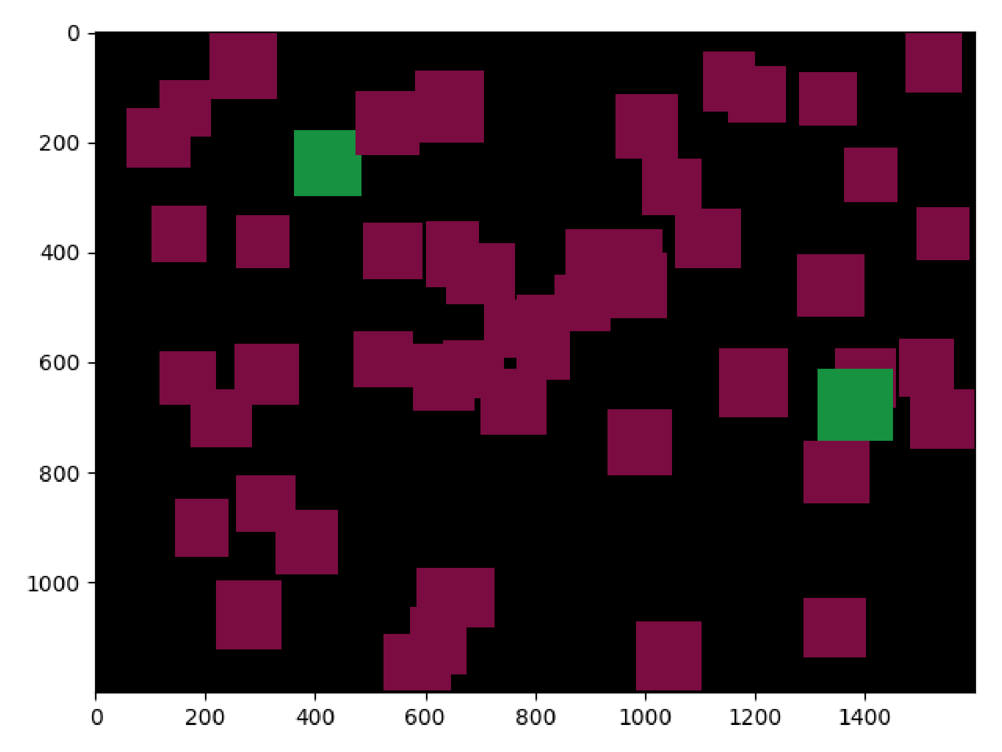
Fruit-360 Results – Real World Images



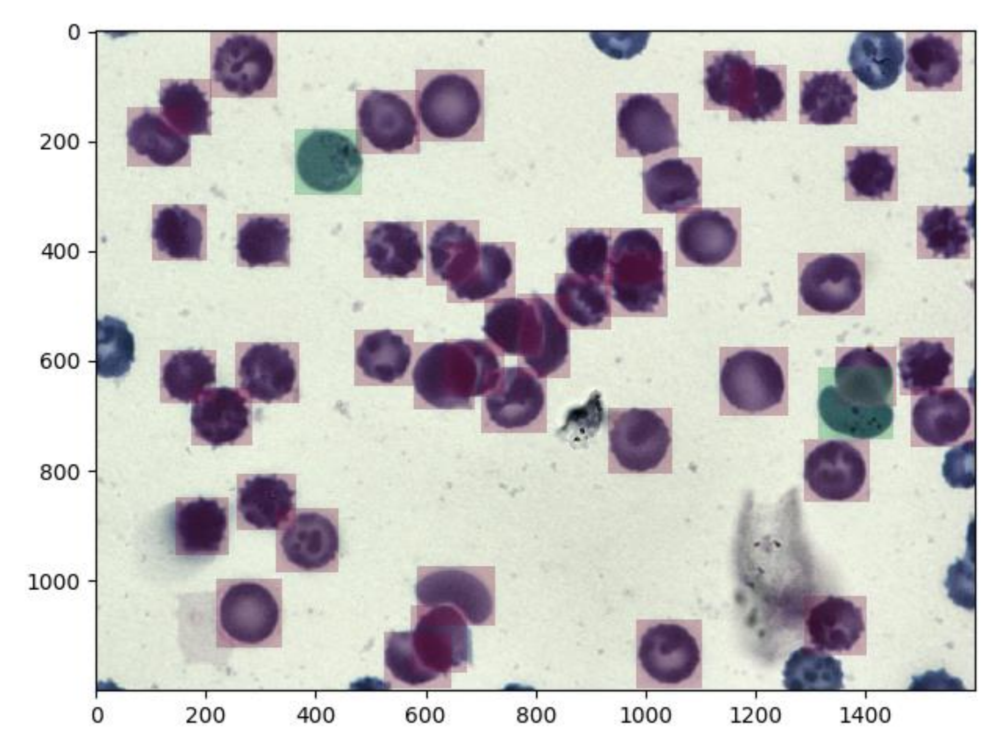


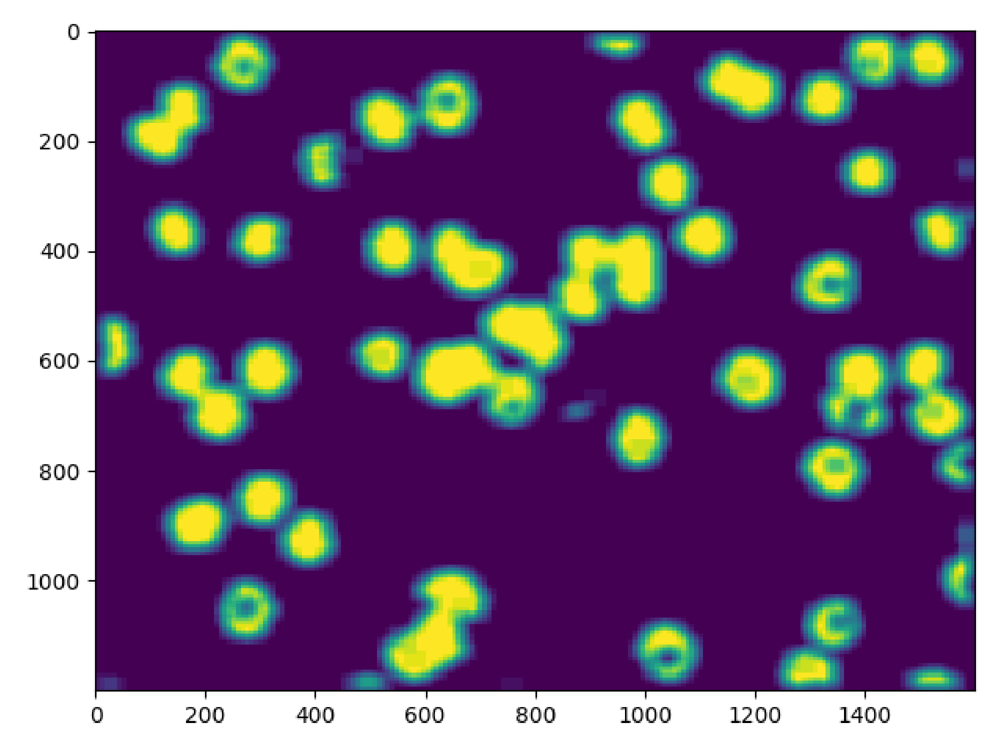
Cell Results



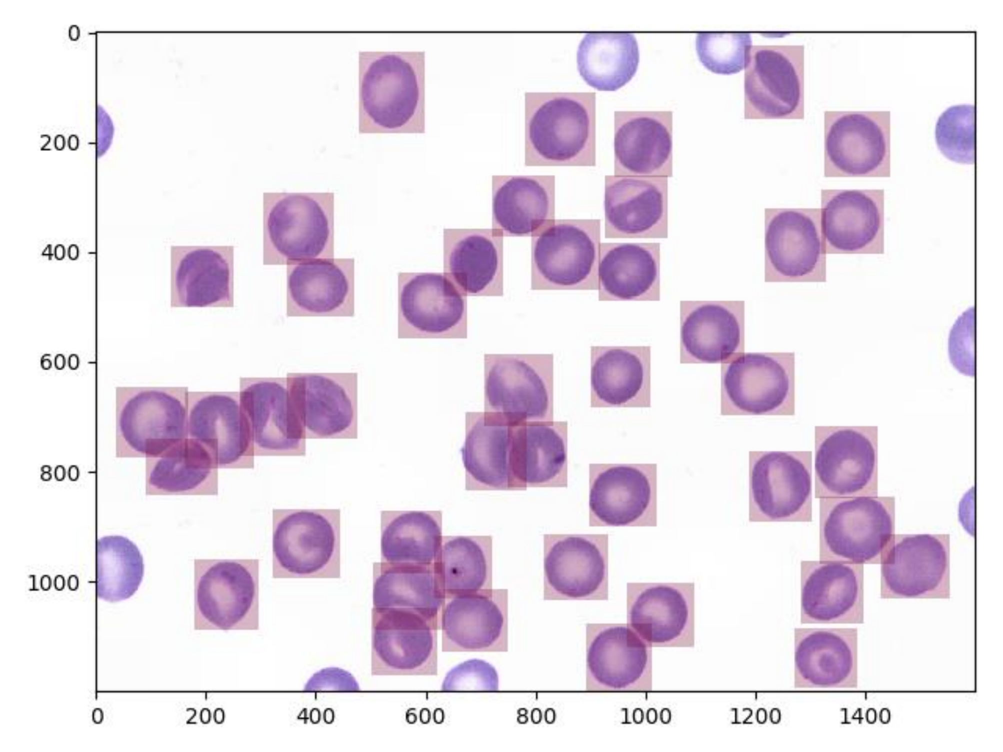


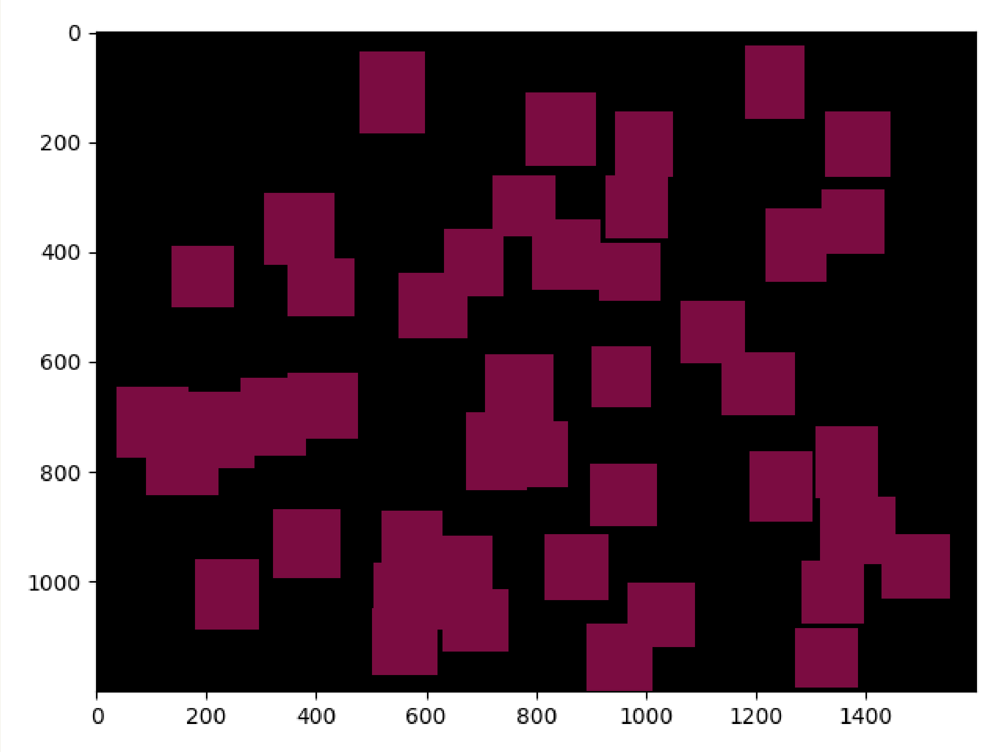
Cell Results



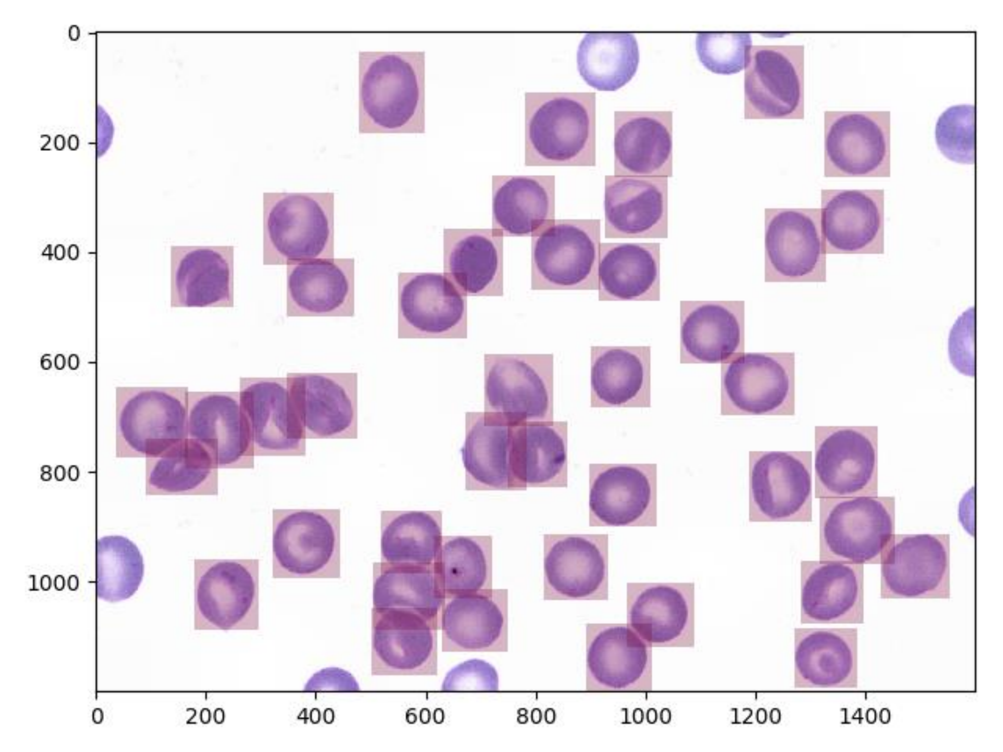


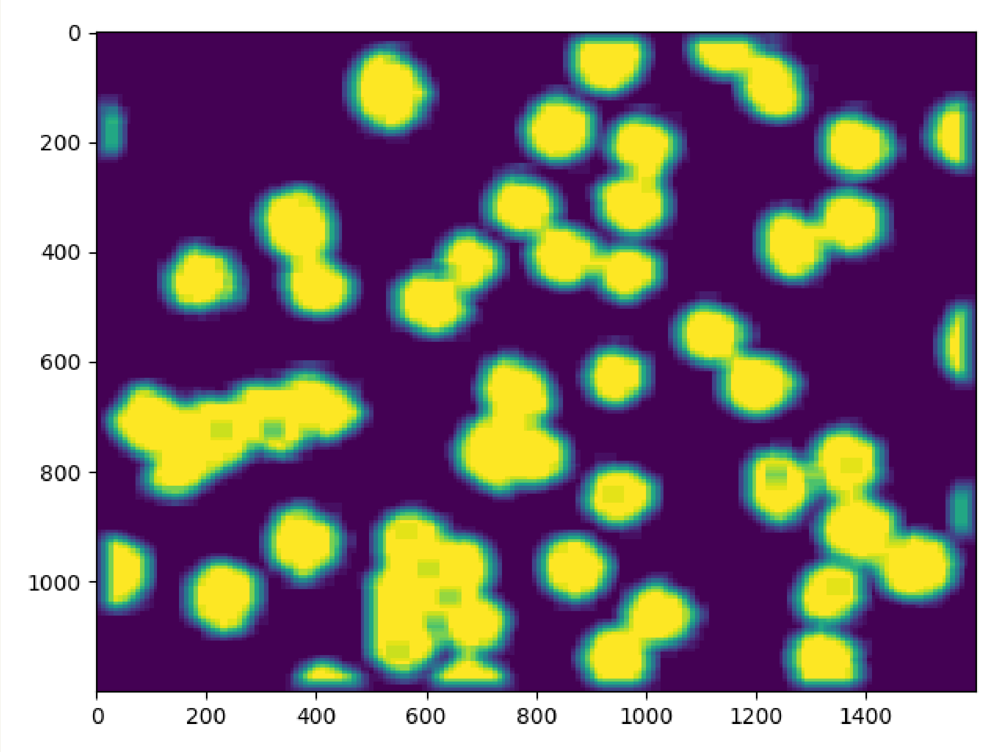
Cell Results





Cell Results

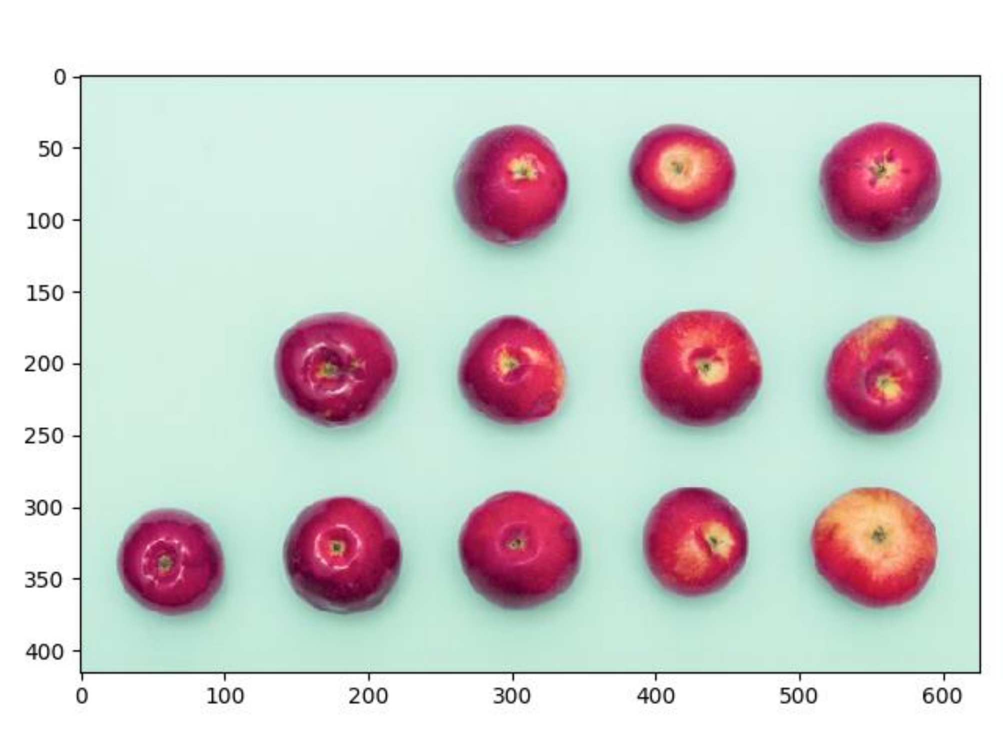


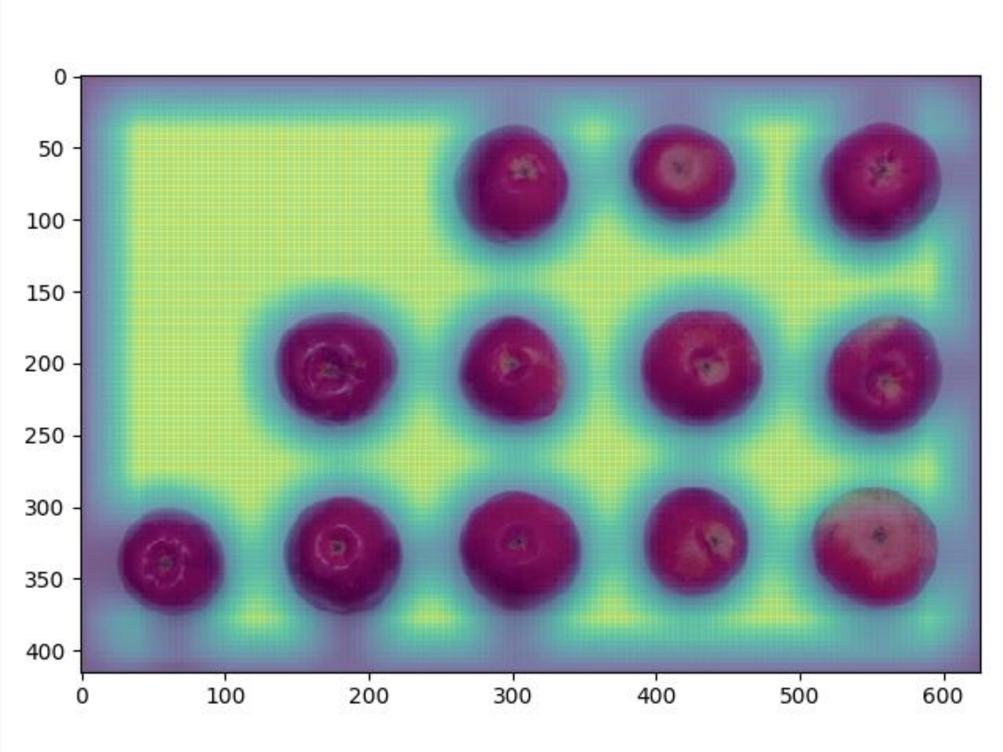


Limitations

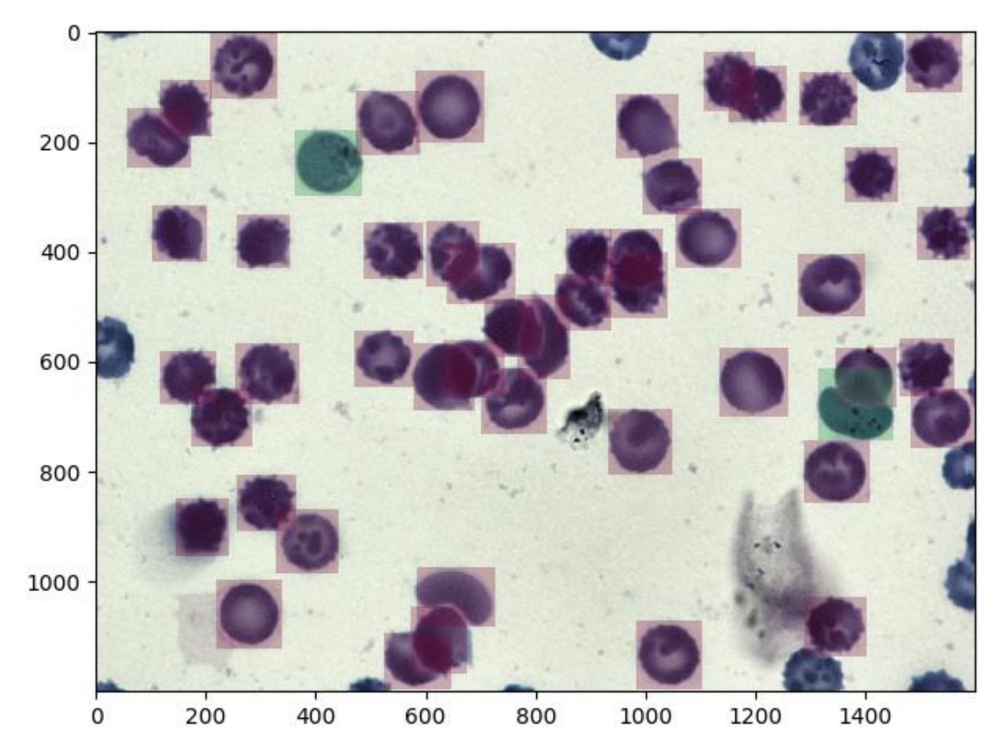
- Results highly depend on classifier quality.
  - *In our results the classifiers were poor and trained on isolated data, impacting the ability of the mapping algorithm.*
- Results can very with significantly large or small kernel size.
  - *Ideally smaller kernel sizes work better, but a good classifier is required.*
- A null class is required or one class will be associated with the null class.
- Image size can be an issue if storage space is limited.
  - *Each image will produce many cropped images to be stored.*

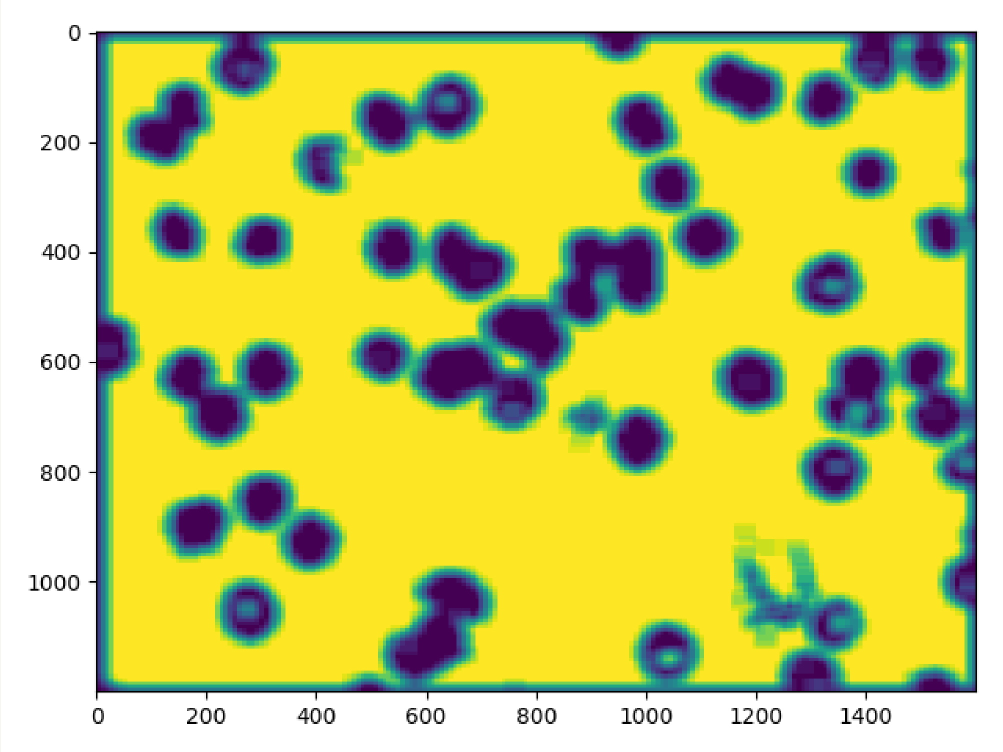
Limitations– Need a Null Class



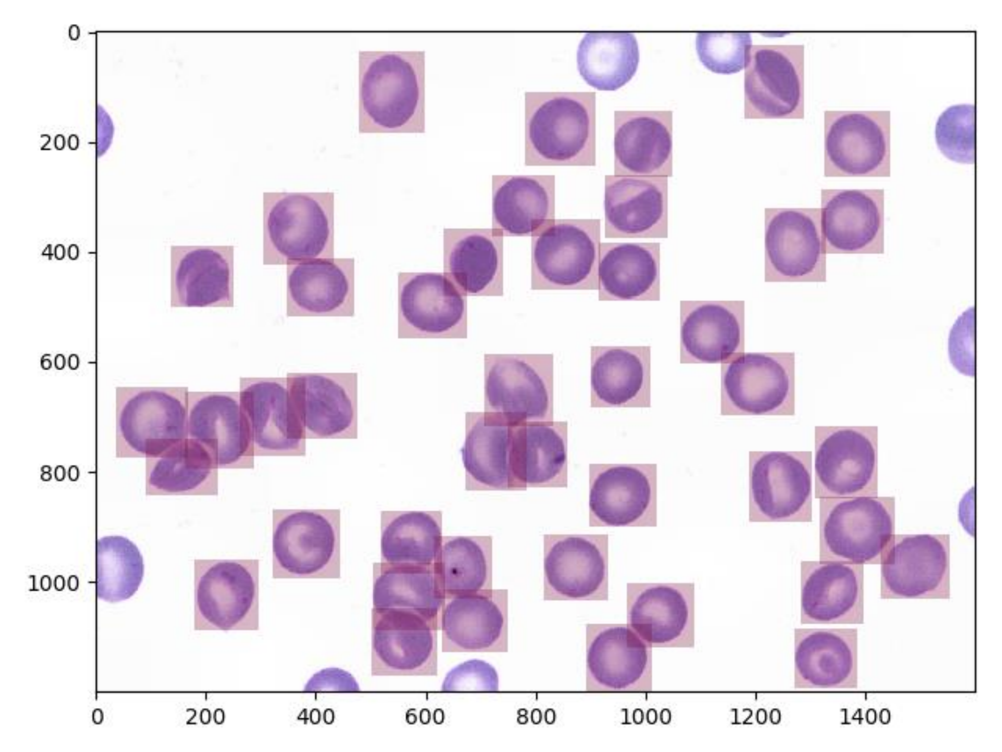


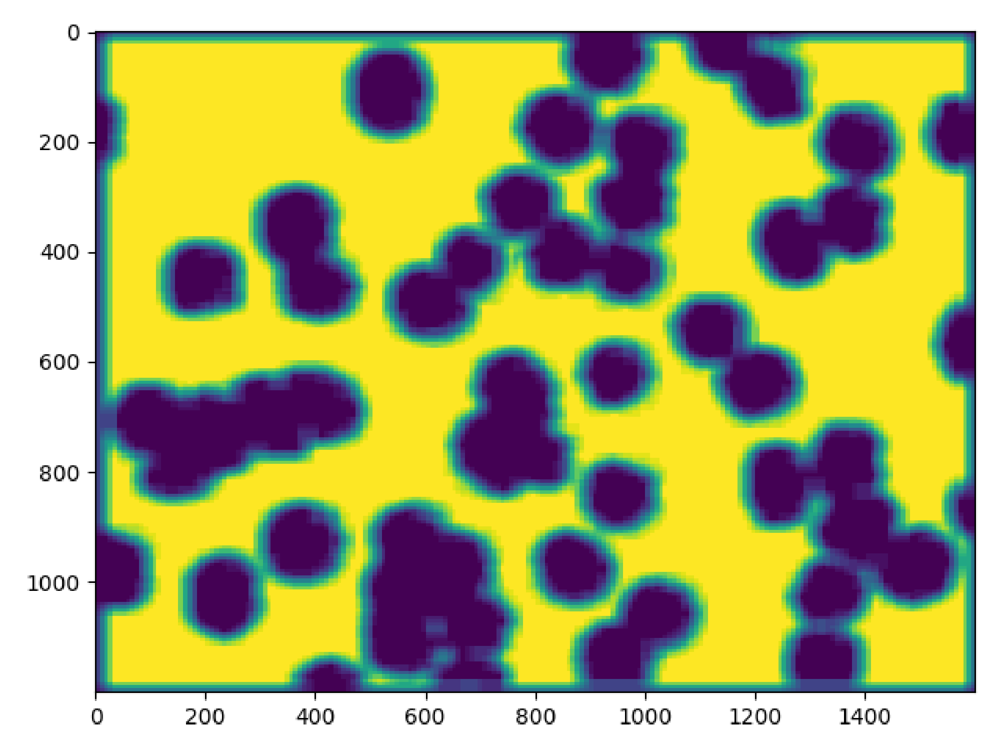
Limitations – Need a Null Class





Limitations – Need a Null Class





Future Directions

- Make a better classifier using relevant data that includes a null class.
- Use the Cropping Algorithm with CUDA to improve image processing speed.
- Use multiple kernel sizes (or random cropping) to better establish an image shape and location.
- Train on a new partial object dataset to better identify images
- explore using kernels to trace boundaries directly.
- Exploring using RNN to take into account the relationships between cropped images.
- Explore combining edge detection techniques with our approach.
- Need more computational power for classifiers.

Conclusions

- Simple trained neural network classifiers can be used to identify both the location and shape of objects, without the need of bounding boxes.
  - *Assuming the classifier is powerful enough.*
  - *Assuming the kernel size is appropriate.*
- Once training has occurred, the cropping and mapping sequence defined here is computationally efficient.
- A null class is necessary for proper classification.

References

- Horea Muresan, Mihai Oltean, Fruit recognition from images using deep learning, Acta Univ. Sapientiae, Informatica Vol. 10, Issue 1, pp. 26-42, 2018.
- Oquab, M., Bottou, L., Laptev, I., & Sivic, J. (2015). Is object localization for free? - Weakly-supervised learning with convolutional neural networks. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. doi: 10.1109/cvpr.2015.7298668
- Papadopoulos, D. P., Uijlings, J. R. R., Keller, F., & Ferrari, V. (2016). We Don't Need No Bounding-Boxes: Training Object Class Detectors Using Only Human Verification. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. doi: 10.1109/cvpr.2016.99
- Ribera, J., Guera, D., Chen, Y., & Delp, E. J. (2019). Locating Objects Without Bounding Boxes. *Computer Vision and Pattern Recognition*. Retrieved from https://arxiv.org/abs/1806.07564
- We used image set BBBC041v1, available from the Broad Bioimage Benchmark Collection [Ljosa et al., *Nature Methods*, 2012].