Vesuvius Challenge

Deciphering the Vesuvius Scrolls Using AI

Andre Barle

Github: https://github.com/Cer001/papyrus_ink_detection

Agenda

1. EDA and data preparation
2. Considerations
3. Ink detector
4. Image processing
5. Letter classifier
6. Conclusion and key takeaways
7. Next steps

Background

Scrolls were carbonized from the eruption of Mt. Vesuvius in year 79 AD preserving the ink but making the scrolls impossible to open.

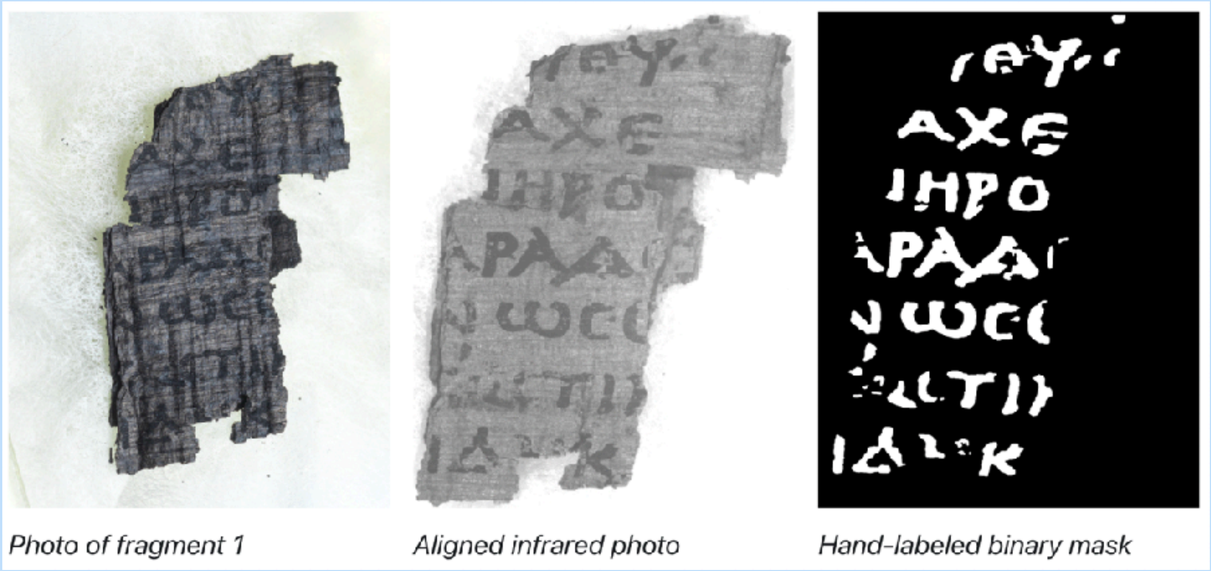


Photo of fragment 1 *Aligned infrared photo* *Hand-labeled binary mask*

Business Objective

Can we decipher the heavily damaged scrolls from the father-in-law of Julius Caesar?

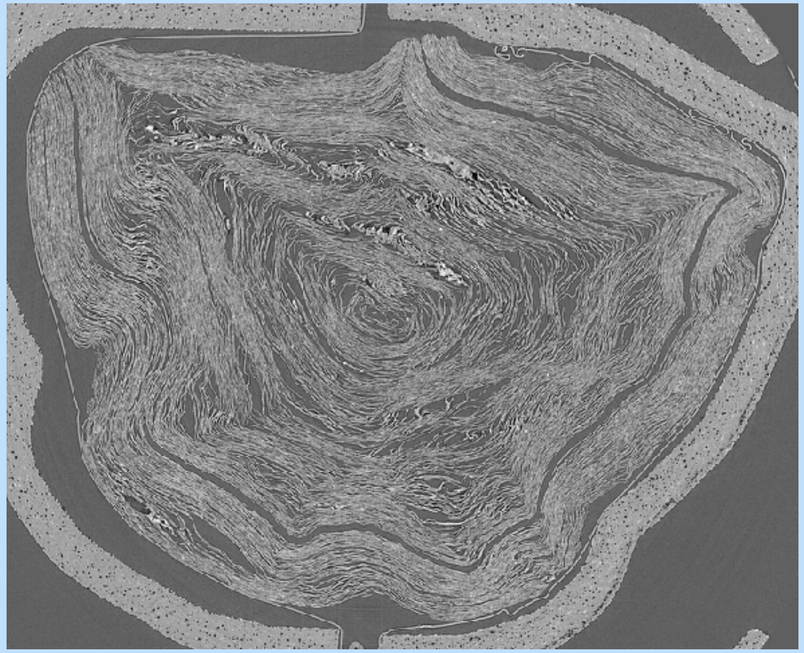
To improve our understanding of history by reading the library of heavily damaged papyrus scrolls by detecting ink letters. Grand prize for best model: $1,000,000

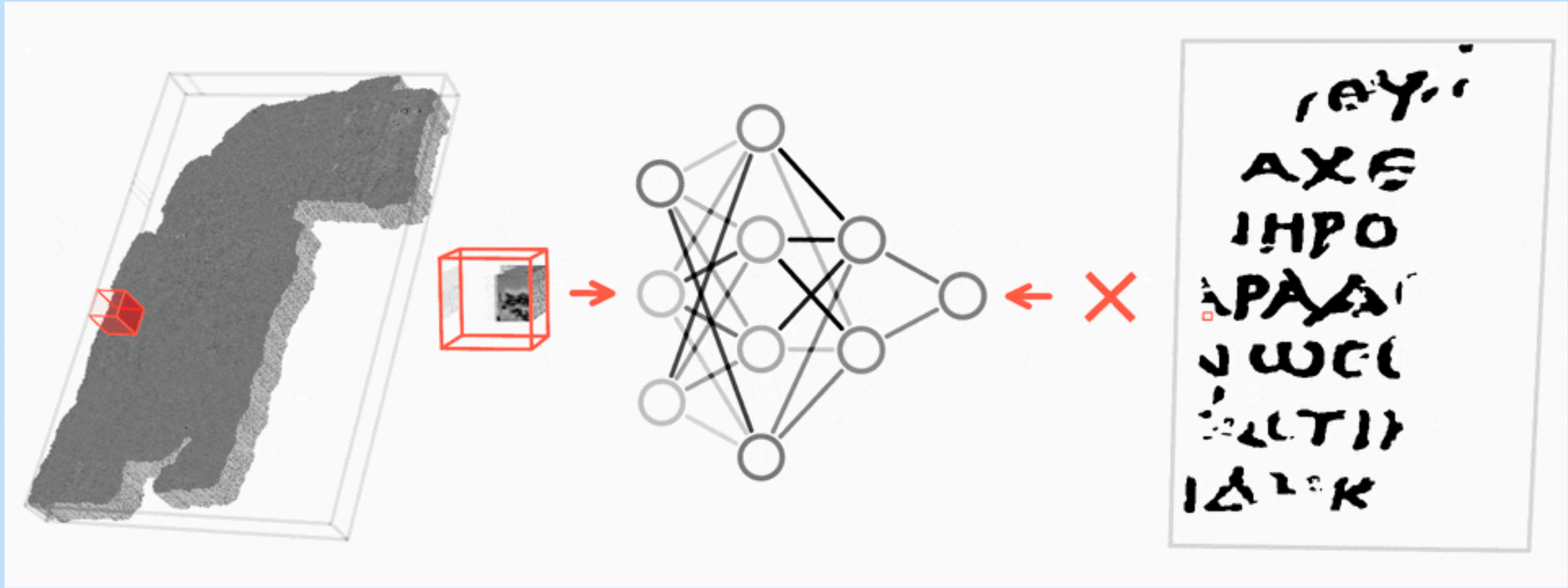
Questions we are hoping to answer:

- Can we build a neural network that can identify ink in multi-layered papyrus scrolls?
- Can we approximate the outputs from the neural network?
- If so, can we correct the text from the original output with a second AI system?

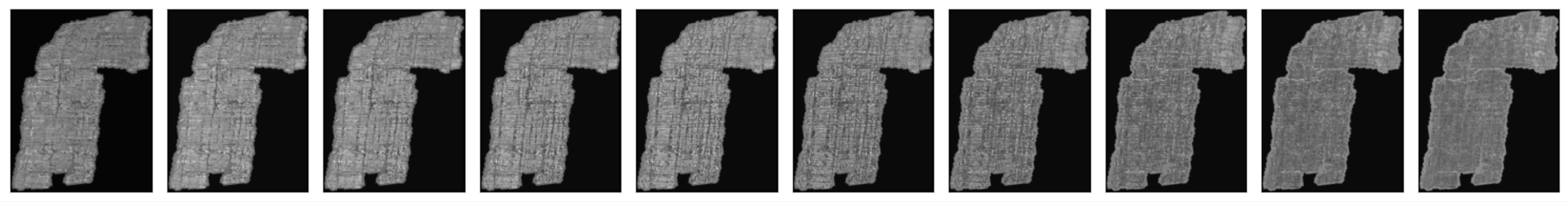
Rationale for a 3D Convolutional Neural Network

- Each slice is manually cropped and flattened
- Each slice contains about 50 z dimensions, only about 10 are relevant
- We have ground truth data, masks, labels, and perfectly prepped data
- Almost nothing is visible on the surface so we have to go deep

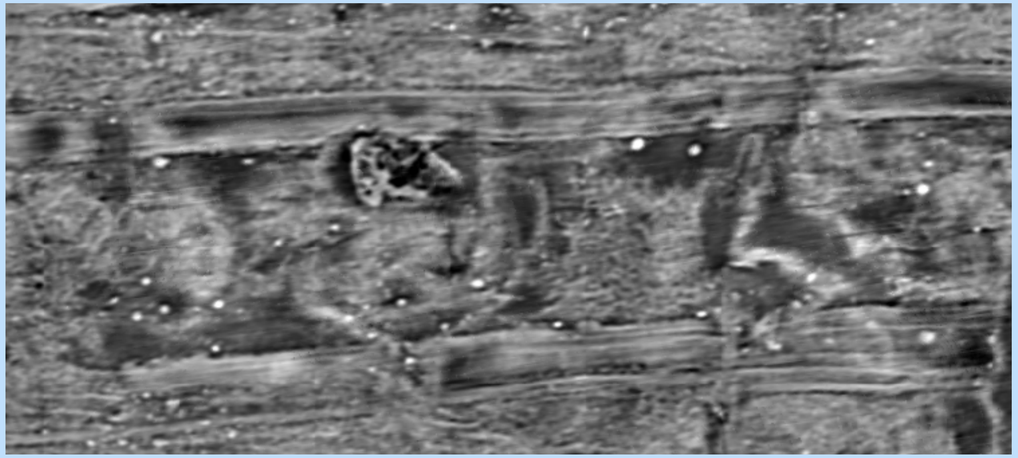




Ink Detection AI - EDA and data processing



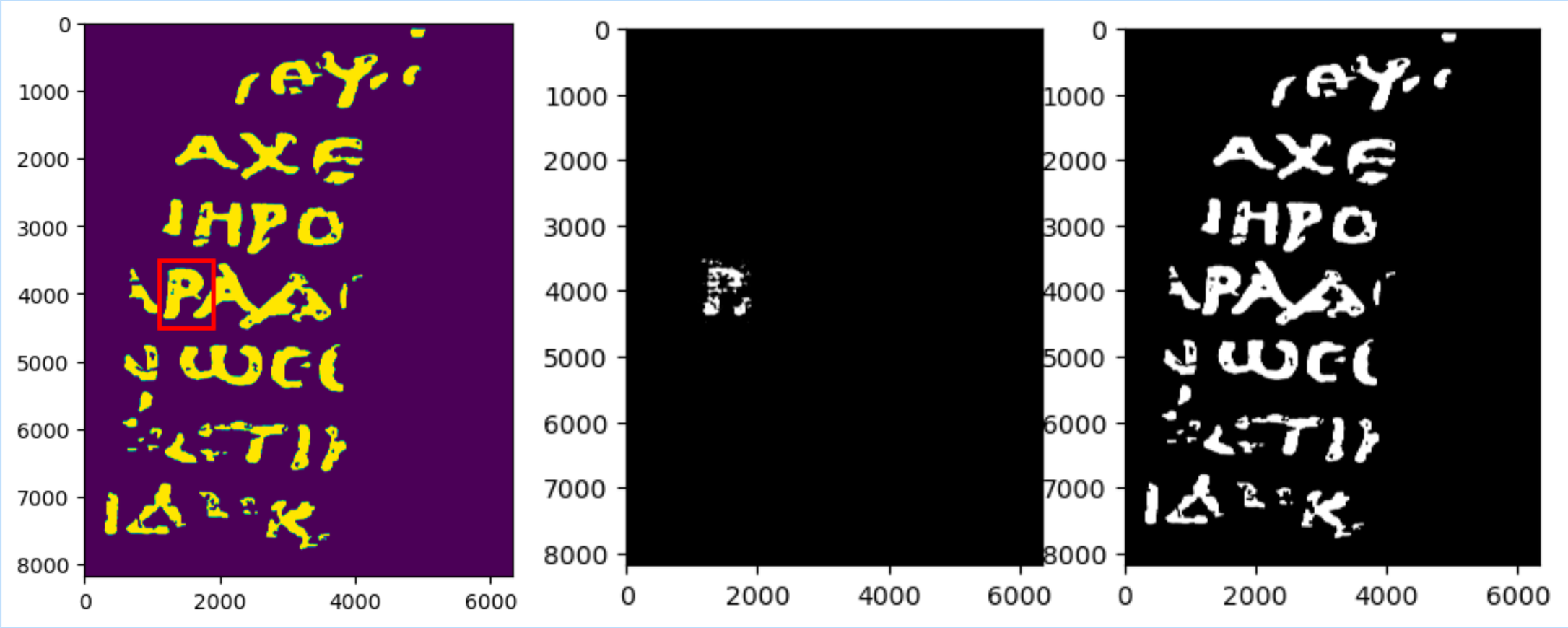
Raw Image



Kuwahara Filter

Final Model for Ink Detection

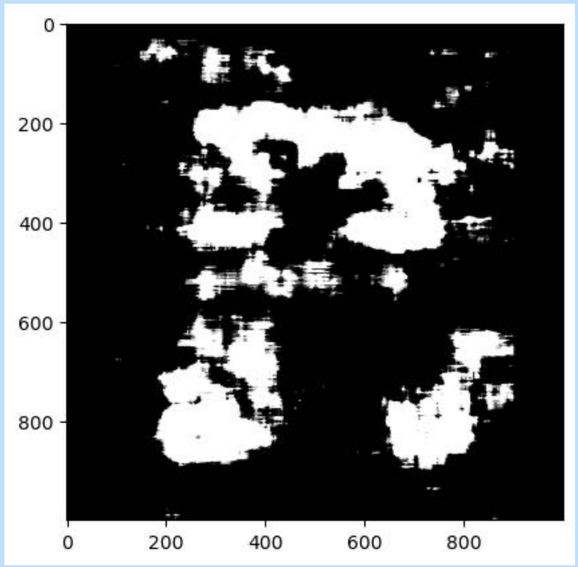
3D convolutional CNN with batch normalization and dropouts, SGD optimizer: focused on just the letter P for now, we can detect it but with noise



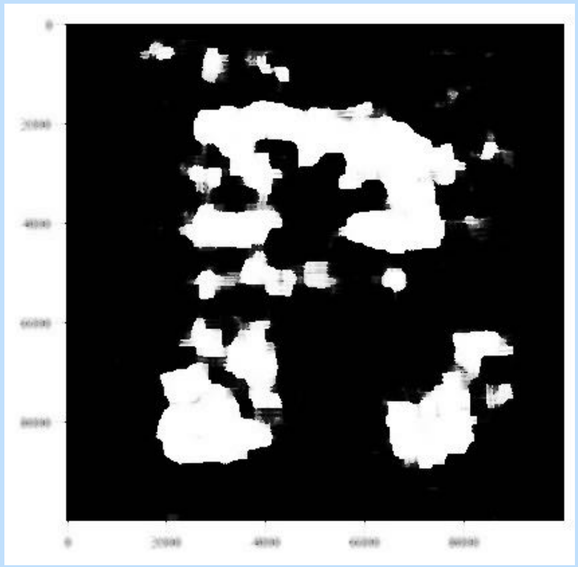
How can we make this better?

- Lets first enhance the letter as much as possible
- Then try to identify the letter using supervised deep learning
- We could correct the output by knowing the general shape it should be

Base output

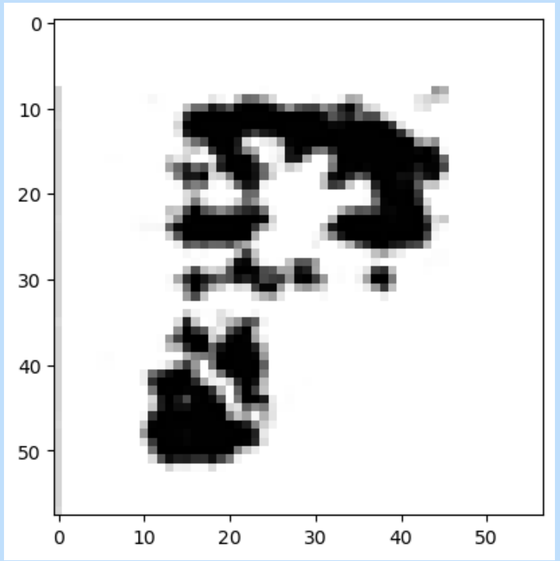


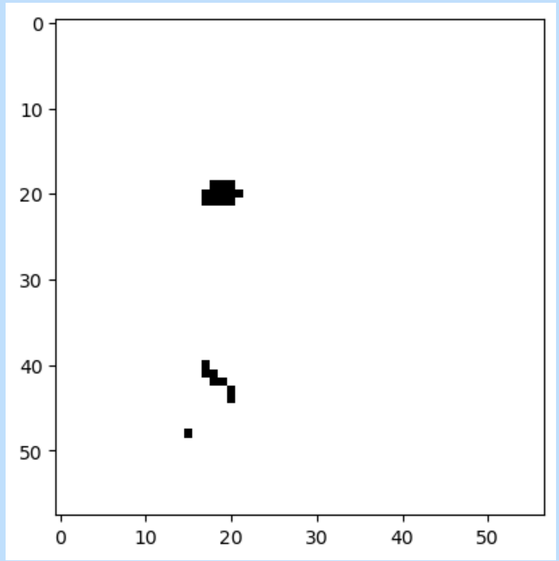
Mean Kuwahara

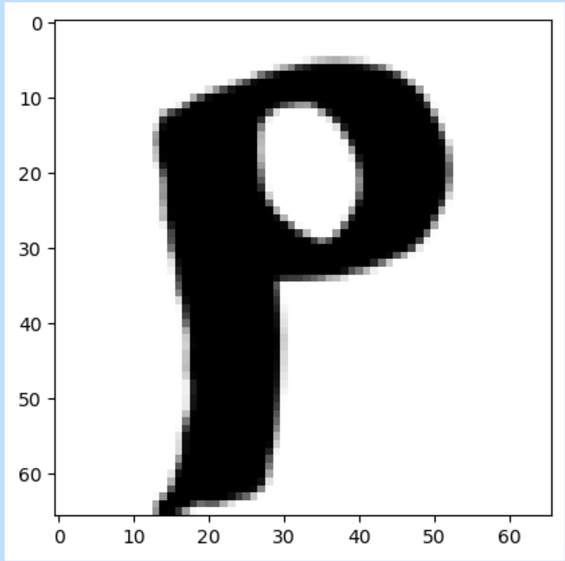


EDA/Data preparation for letter classifier AI

2D CNN + NN with batch normalization and Adam optimizer will take in a de-noised output image and approximate which letter it could be



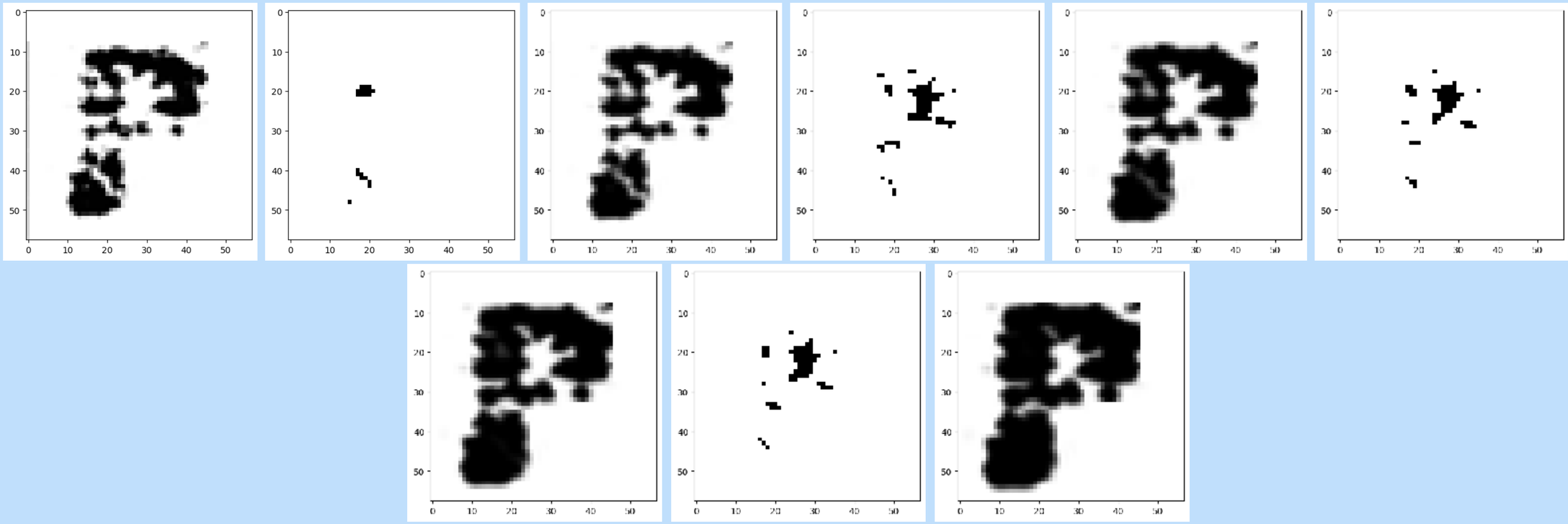




- Rescaled input images to 66*66
- Trained on distorted and normal greek letters using image datagen
- Image reconstruction and inpaint biharmonic restore the image from the other model enough to approximate the correct letter

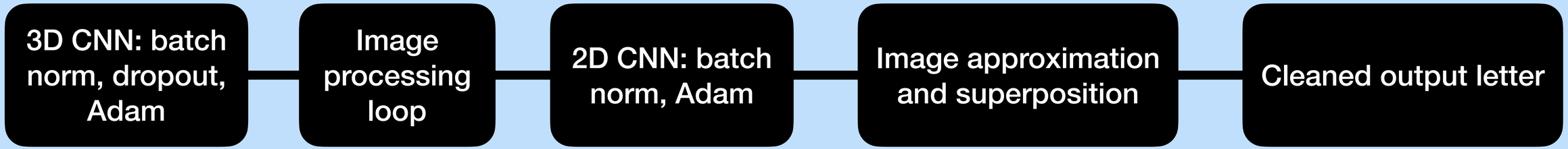
Strengthening ink detection using letter recognition

- We can further correct the image by superimposing an example of a letter P over it.
- Even something other than the exact letter will make the ink more legible
- With reconstruction and inpaint biharmonic we can reforge the original letters by inferring their shape



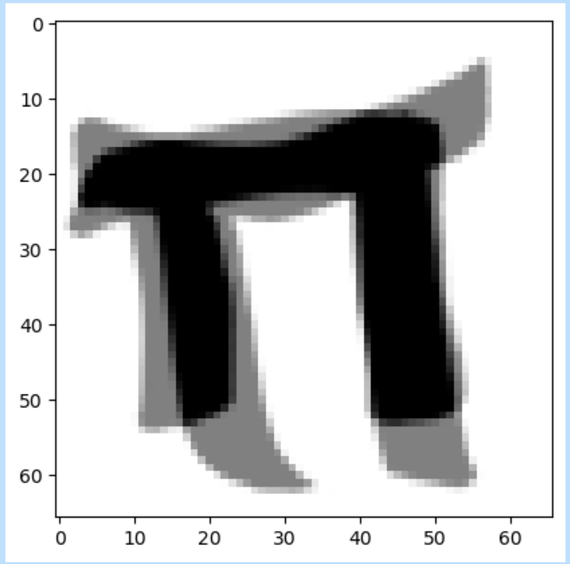
Recap

- We use a 3D CNN to extract the ink from the papyrus layers
- The output is then fed into an image processing loop
- Then the processed image is fed into a 2D CNN where it is approximated as a letter
- The approximated letter is then superimposed onto the letter it identified to increase its readability in theory



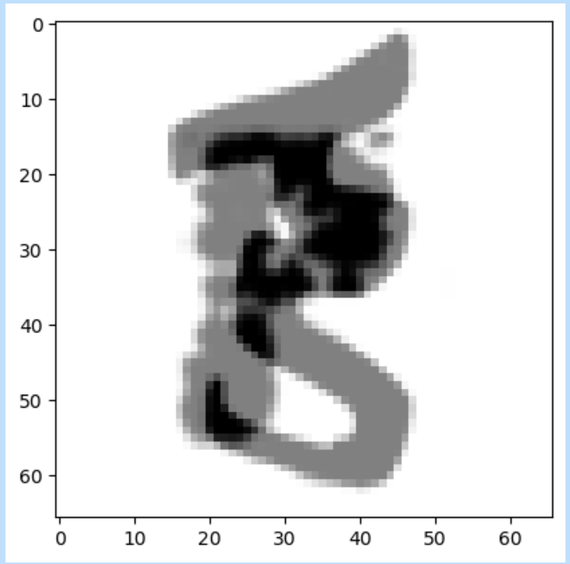
Conclusion

- The image processing loop enhances the image ink detection vastly and is more reliable than using the neural network
- If the neural network has more time to train and more data to train on, it can improve and start to recognize the noisy letters
- Even with its current predictions, the overlapping set allows us to increase the weights of the pixels where they intersect, in the future boosting accuracy if cross-combined with the image processing pipeline.
- We have a working proof of concept for a pipeline that can greatly enhance the outputs from the 3D CNN



Proper letter prediction

loss: 1.2186
accuracy: 0.6354



Live prediction

Key Takeaways and next steps

- The 3D CNN with my modifications improves from the base model presented by the Vesuvius Challenge team by reaching similar results with 1/3 of the training time
- The image restoration technique further increases correct pixel values for my submission
- My neural network is a stepping stone in the right direction for how these scrolls can be interpreted
- Next, I will train the network on heavily damaged images and with more data
  - Then I will cross-reference the output and amplify the intersection
  - Then I will feed that into the image processor

Visible improvement

Vesuvius Team Base Model

My Model

Ground Truth