Level 4 Project

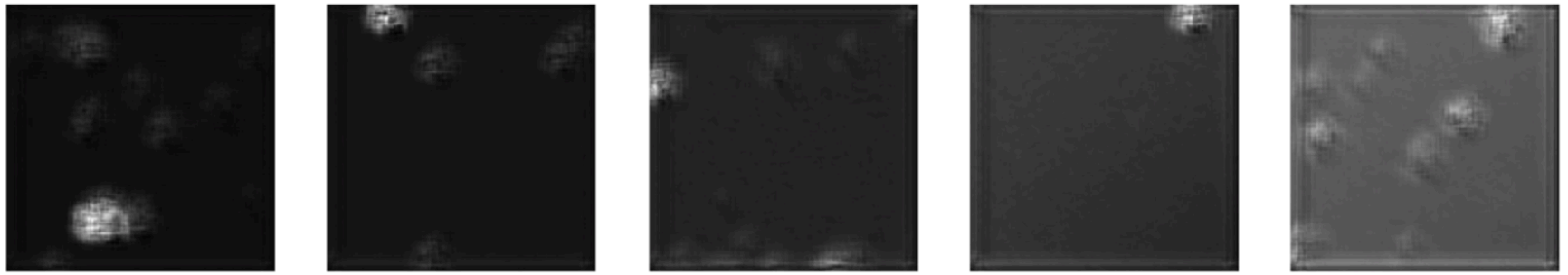
Week 6 Meeting

(Week 5 Recap)

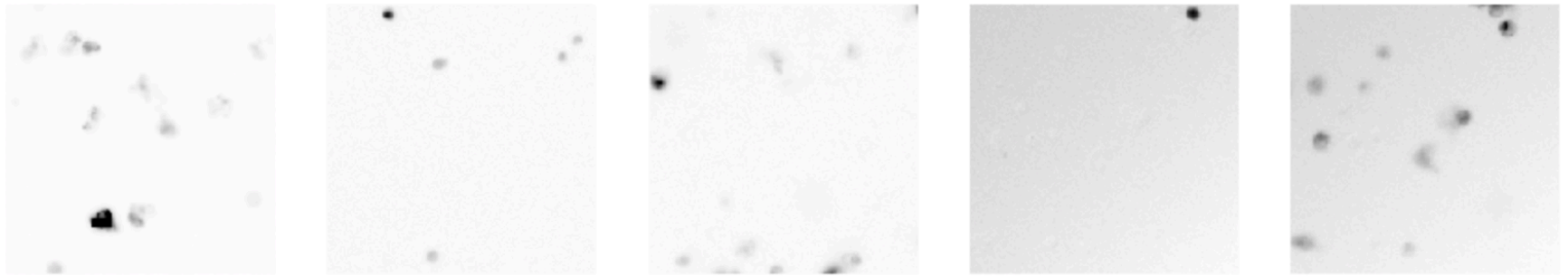
Completed work

- Cropped images to 200x200 as suggested
  - Got better results
- Tuned autoencoder
  - Got better results with the 200x200 cropped
- Made function to calculate overlap
  - Absolute difference
  - Is this the right way?
- Made function to calculate labels
  - Hardcoded to one Excel sheet
  - Are the Excel sheets always the same format?

Reconstructed



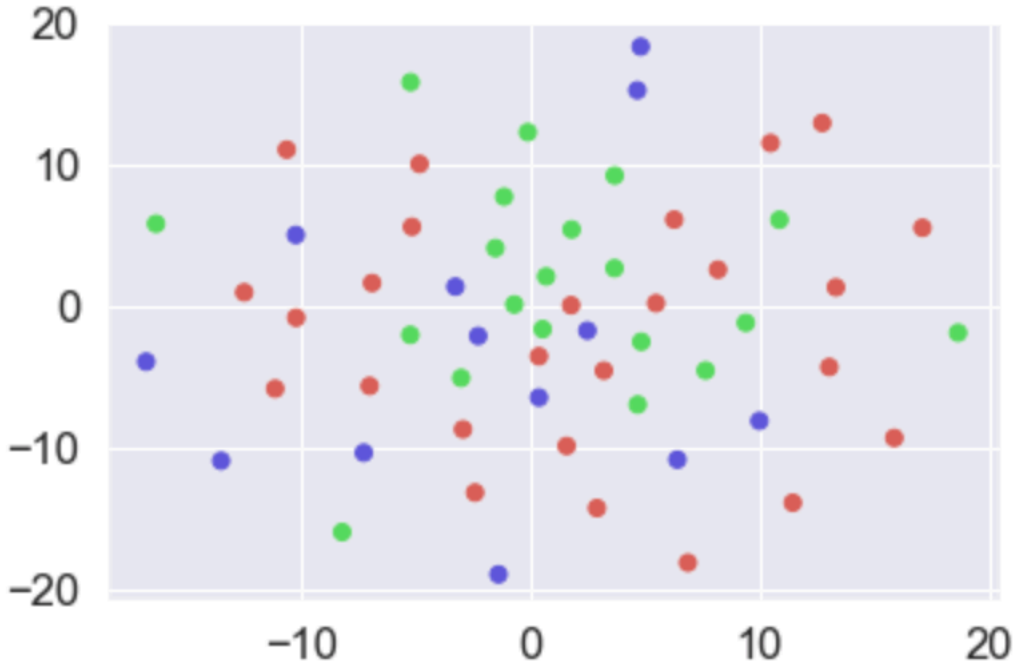
Original



Compressed

Completed work (2/2)

- Read about PCA and tsne
- Tried to cluster the compressed and overlapped images...
- Not much luck:



Work to come

- Script to read Excel sheet into labels
- Spending time on getting resized images read back to disk so I can transfer them easier to GPU cluster or Colab
  - Running into a lot of issues for this though
- Research on how PCA + tsne can work better for this.
  - Right now, even with different parameters not giving much results.
  - Increase dataset?
  - Then linked to resizing data work.

Rough plan for semester, reworked

- ***Week 5***
  - *Tuned autoencoder*
  - *Calculated overlap of images*
  - *Start working on clustering algorithm for image overlaps*
- **Week 6**
  - HPC training day (hopefully will help with running some models)
  - Find way to write images back to disk without losing data
- **Week 7**
  - Read on PCA/tsne
  - See if dataset can be expanded
  - Fit model, save weights
- **Week 8**
  - Tune clustering algorithms to improve performance
- **Week 9**
  - Tune clustering algorithms to improve performance
  - Evaluate model performance