

Using Traditional Machine Learning Methods to extract information from GBM and LGG pathology images

Matthew Millett

Stanford University

450 Serra Mall, Stanford, CA 94305

https://github.com/millett/pathology_learning

millett@stanford.edu

Abstract

Pathology slides have long been useful in various diagnoses, including cancer. Pathologists look at many large pathology slides every day for signs of cancer diagnosis, and stand to gain from computer-assisted diagnosis. This report combines pathology slides and semantic data from The Cancer Genome Atlas to predict tumor type, survival, and tumor grade from Glioblastoma and Lower Grade Glioma pathology slides.

1. Introduction

One in four deaths in the United states is due to cancer [10]. Individual cancers have unique characteristics that affect their diagnosis and treatment. Additionally, different treatments- types of radiation therapy, pharmaceutical interventions- for a given cancer will be more and less effective at different stages of a particular cancers evolution. One tool commonly used by pathologists is the hematoxylin and eosin (H&E) stain, which gives valuable information about cell morphology and spatial arrangement. A slide consists of a slice of tissue that has been fixed and stained with certain dyes to make morphological features, e.g. nuclei, more visible. Currently H&E slides are usually interpreted by unaided pathologists. Inter-reader variability has been shown to be quite high between pathologists [6], who stand to gain significantly from computer assistance. Digital image processing methods have already been shown to improve consistency, efficiency, and accuracy in evaluating histopathology stains, and can be used for decision support [13].

Glioblastoma (GBM) and Lower Grade Glioma (LGG) account for 28% of brain tumors in the United States and constitute 80% of the malignant brain tumors. GBM has a 5-year survival rate of 5%, whereas LGG has a 51.4% 5-year survival rate, so differentiating between these tumors is important for patient treatment [2].

2. Problem Statement

This project attempts to classify 2 different types of cancer, Glioblastoma (GBM) and Low Grade Glioma (LGG) based on morphological and color features from their pathology slide images along with the patient's age and gender. This project also attempts to predict tumor grade and patient mortality from these features.

3. Background

Currently, most pathologists interpret images without much computer assistance [2]. Even many machine learning approaches have had a pathologist manually designate areas of interest on the pathology slide [1] [3] [8]. Mousavi et al specifically had a pathologist highlight several areas of interest on potential LGG and high-grade glioma (GBM and clinically more dire LGG) slides to separate them by pathology [16], achieving 95% detection accuracy of high-grade glioma, but of course the model used the pathologist's time, attention and expertise. Manual selection is also so common because these slides can contain around 10^{10} pixels [2], making most classical feature extraction from a whole image computationally intractable.

Barker et al calculated manually-selected features of pathology images to attain 90% accuracy on the same dataset, selecting which images from a slide to use with a voting scheme [2]. Deep learning for tumor grade classification has also been done using an ensemble method with gliomas [9]. Xu et al [21] classified tumors into different subtypes using transfer learning with an AlexNet model. They first trained with ImageNet [7] and fine-tuned on the MICCAI 2014 Brain Tumor Digital Pathology Challenge dataset to account for small sample sizes and large images, and achieved 97.5% classification accuracy on the dataset. Their CNN goes across the entire pathology image, which can go up to 100,000 by 100,000 pixels, but is the highest reported accuracy for this particular problem. Deep learn-

ing for tumor grade classification has been done using an ensemble method with gliomas [9]. Xu et al [21] classified tumors into different subtypes using transfer learning with a Convolutional Neural Network AlexNet model, training first with ImageNet and fine-tuning on the MICCAI 2014 Brain Tumor Digital Pathology Challenge dataset to account for small sample sizes and large images.

4. Data

The Cancer Genome Atlas (TCGA) contains large amounts of H&E pathology stain images from cancer patients and has many opportunities for learning [20]. Here I use traditional machine learning and image processing techniques to predict if a tumor is GBM or Lgg, the tumor grade, and the survival rate. grade of a tumor and other relevant statistics just from pathology slides. A tumor grade tells us how abnormal tumor cells look under a microscope compared to their environment (Cancer.gov). TCGA was chosen because the data is publicly accessible and well-labeled, has large amounts of pre-existing published literature, and contains large amounts of different semantic data that could potentially be useful in future studies, such as genetic mutations. Predicting genetic mutations from pathology stains could be useful for genome-appropriate chemotherapeutic treatments down the line without having to wait on genetic test results.

4.1. Images

The input images are a scanned Hematoxylin and Eosin (H&E) pathology slide, linked with semantic information about each patient that a previous study extracted from TCGA. [4]. Histology stains pose unique challenges as compared to other modalities such as CT or MRI scans. The images themselves are massive, are not uniformly sized, and contain large amounts of detail. For example, one randomly selected slide was 21973x17532 pixels, and some pathology slides can contain more than 10^{10} pixels. Some pathology slides are from the same patient, and other pathology slides may not have semantic data associated with them. In this case I will fill in that data with averages from the training set. The slides can also contain one or multiple globs of cells. I ran parallel analysis using convolutional neural networks (my CS231N final project) to see how the respective models perform. Currently, the neural network uses both padded versions of the entire image and the preprocessing scheme described below.

The tumors in question are Lower Grade Gliomas (LGG) or Glioblastomas (GBM), which were chosen because there has been limited work on GBM and LGG pathology slide processing that does not use neural networks. The images were hosted in .svs format on a Google Cloud instance, so it was fairly simple to download them.

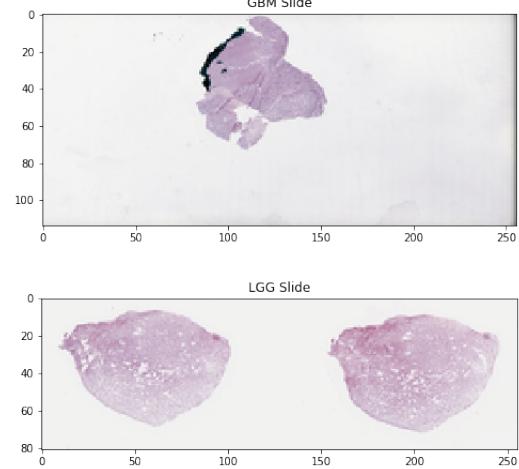


Figure 1: Examples of the input data, significantly scaled down. A primary challenge in working with pathology image data is their size and heterogeneity. The full-size images contain on the order of 10^{10} pixels. Note the morphological differences in numbers of globs, different locations, different shapes, and different colors. Images courtesy TCGA.

4.2. Semantic Information

The semantic data includes information such as gender, age, tumor grade labeled by a radiologist, survival, and various genetic information such as mutation count and methylation at the MGMT promoter site. This project attempts to classify tumor type, grade and mortality, some of the most impactful and accessible features. Tumor grade and type are useful sanity checks to see early on how well the manually chosen features represent the data, and how well the model interprets the data. Well-chosen features should be able to represent key factors of the image.

The semantic information from each patient was on the TCGA website and posted as supplementary data to a publication [5], and was a simple way to acquire useful and interesting semantic data without extensive mining of TCGA.

TCGA gives each of its patients a unique identifier, which allows for linking together semantic data with the relevant pathology images. All these data were aggregated and put into a Pandas dataframe [15], which is amenable to learning via various scikit-learn modules.

5. Methods

5.1. Data Preprocessing

Because slides are heterogeneous and include significant portions of blank space, this project adapts an approach by Yu et al [22] to select the most cellularly dense portions of each slide for training. For each image, the top 10 most

dense 1000x1000 tiles were chosen, where image density was calculated as the percentage of nonwhite (all RGB values under 200) pixels in the tile. These images were selected using a min-heap. While choosing the most dense images only might miss out on important data needed to make a diagnosis, it is one of the simple ways to make studying such large images possible and on the whole seems to perform adequately.

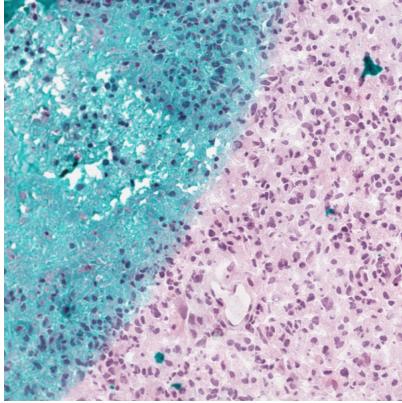


Figure 2: An example dense image selected by the pipeline.

5.2. Feature Selection

While Yu et al [22] computed many features in the lung dataset, relating to the cytoplasm and the entire image, for a preliminary analysis this study only focuses on morphology features of the nucleus. Cancer nuclei have long been known to be abnormal, and the particular abnormalities might help distinguish types of tumors [23]. As such, this project used HistomicsTK [12], a library provided by the Digital Slide Archive [11], to segment nuclei by adapting their provided example pipeline for cell segmentation.

Petushi et al [17] calculated various nuclei morphological features to differentiate grades of breast cancer; this uses a similar approach.

The pipeline first normalizes the 1000x1000 subimage to the most upper-left selected dense tile also taken from the same slide. It then uses color deconvolution to separate the hematoxylin and eosin stains, and uses the hematoxylin channel with local maximum clustering to label nuclei and find nucleus features. It then uses the HistomicsTK package to compute various morphological statistics on the nuclei, such as area, eccentricity, and major and minor axes. Statistics were calculated for each nucleus, and aggregated for each image as simply the mean morphological statistics of all the nuclei in the slide. Nuclei were not labeled, so segmentation quality was evaluated visually and deemed adequate, as shown in figure 3. As this is a step in the pipeline and the final result- largely based on this step- proved impressive, nucleus segmentation quality

was not further investigated. Color features were calculated by using the positive pixel count algorithm, which allows pathologists to quantify how much their stains have been absorbed and have been used as immunohistochemical indicators [18]. Essentially it counts how many pixels are within particular intensity thresholds [12]. These hyperparameters were also tweaked manually from default settings, where I examined several slides passing through this threshold using masked images as seen (see Appendix for an example).

Ultimately, the morphometry features chosen from cell nuclei were average values for area, major axis length, minor axis length, perimeter, circularity, eccentricity, equivalent diameter, extent, minor to major axis ratio, solidity. The color features chosen were proportions and ratios of how many pixels were in each threshold zone. The semantic features chosen as prediction inputs were age at initial diagnosis and gender. The models attempted to predict tumor grade and survival.

All features were normalized to mean 0 and unit variance so variable scale and magnitude would not disproportionately affect coefficients in classification and regression, which were used as proxies for feature importance. Some color features were not able to be calculated due to division by zero, and were imputed to be the mean color values of the entire dataset before splitting into training and test groups.

The pipeline benefited significantly from using the python multiprocessing library and achieved a 3x speedup once multiprocessing was added. Processing time went from 8 hours for 1000 images to under 3 hours.

5.3. Prediction Methods

All models used a train-test split of 90%-10%. The model chosen to differentiate tumors and to predict tumor grade was logistic regression, mostly because it is simple technique. For the final project, other classifiers from the scikit-learn toolkit will be tried and compared. Tumor grade is progressive, so a linear regression could have been used, but evaluation is more clear with a categorical classifier. The model chosen to predict survival was linear regression, mostly again because it is a simple technique to implement and interpret. For the final project, other regression models will be used.

6. Results

6.1. Tumor type classification: LGG vs GBM

The most basic task to do with this data is to classify GBM tumors as separate from LGG tumors. To that extent, several models were tried on the extracted features.

The data contain a slight class imbalance in the number of images from each type of tumor. The dataset contains 1194 GBM images and 730 LGG images, which were then extracted into densest 1000x1000 images of counts 11458

for GBM images and 7211 for LGG. A majority classifier would achieve an accuracy of 61%.

Several different classifiers were tried with various characteristics. Ultimately the best model was selected by the maximum area under the ROC curve (AUROC) out of all those tested. An unbounded decision tree was able to achieve an AUROC of 1.0, but upon further examination the tree was over 6 levels deep. To reduce overfitting, I set the maximum depth of the tree to 3 levels. The ROC curves for all models are shown in Figure 3. Ultimately,

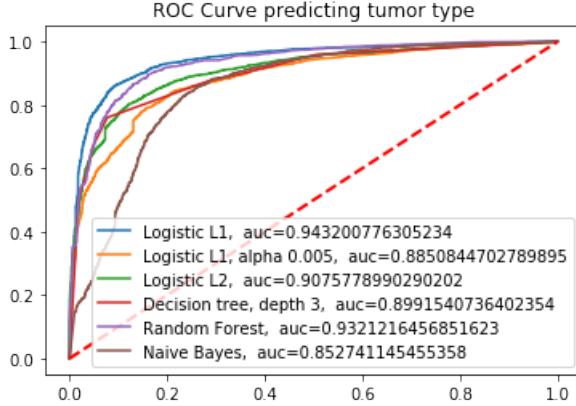


Figure 3: The Tumor Classification ROC curves for all models attempted. AUC listed in legend.

the most accurate and highest AUROC classifier was an L1-penalized logistic regression, which cuts out unimportant feature weights to 0 and leaves important weights. The tumor type classifier using the nuclei and color features selected, age, and gender, achieves an accuracy on the test set of 88.1%. A confusion matrix for the logistic regression is also shown in Figure 4. The most important features var-

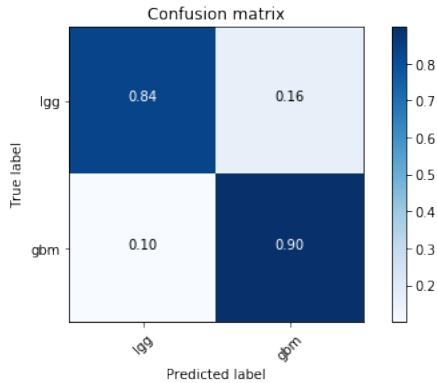


Figure 4: Tumor type confusion matrix for the L1-penalty logistic regression classifier.

ied slightly from classifier to classifier (see coefficients in

Appendix), but the highest-weighted selected features for the best classifier were average equivalent diameter of nuclei, average minor axis length of nuclei, average nucleus perimeter, sum of positive-intensity pixels, and average major axis length of nuclei. Interestingly, for L1 logistic regression with an α set to only allow 5 variables to be chosen, those variables are average nuclei solidity, age (years at diagnosis), average nuclei shape extent, and gender. The diversity of features selected by different models suggests the utility of each on its own.

6.1.1 Convolutional Neural Network Models

For CS231N, I built several different neural network models to try to solve this same problem. The best model came out with 95% accuracy. Sensitivity was 90%, but specificity was only 62% for the best model. The best model architecture used an 11-layer VGG net pretrained with ImageNet, then fine-tuned on the dense 1000x1000 images.

Interestingly, an 18-layer ResNet model pretrained on ImageNet and fine-tuned on 256x256 scaled-down and padded whole-slide images achieved 83% accuracy which is actually also better than the depth-3 decision tree, Naive Bayes, and the 5-feature logistic regression. See Figure 5 for example predictions from the ResNet.

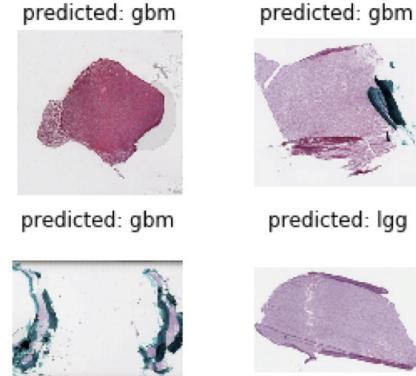


Figure 5: ResNet18 predictions based on whole slide images. This model achieved an 83% accuracy on its test set.

6.2 Predicting Survival

It is notoriously difficult to predict survival in cancer patients [14]. Thus it is unsurprising that initial regression models had R^2 values of approximately 0.06-0.11. An attempt was made to use quadratic features, which brought the R^2 value up to approximately 0.2, where the largest coefficients are all related to pixel intensity levels and counts of intense pixels.

6.3. Tumor Grade Classification

The best tumor grade classifier achieves an accuracy score of approximately 68% on the test set. As with tumor type classification, the best model chosen was an L1-penalized logistic regression. The confusion matrix is shown in Figure 6. Curiously, despite plenty of grade 4 tumors appearing in the test set, the model almost never classifies any slides as grade 4. (The confusion matrix hides 2 cases where the tumor was grade 4 and correctly predicted to be grade 4). This is especially intriguing because grade 4 tumors are GBM and it seems the model is reluctant to make that classification. Despite this oddity, the model has a strong tendency to classify on the diagonal of the confusion matrix in Figure 6, which suggests that for the most part, the classifier is robust. As mentioned in Methods, tumor grade classification is arguably a regression problem, and the diagonals of the confusion matrix reflect that, especially for the tumors predicted as grade 3 but are actually grade 4.

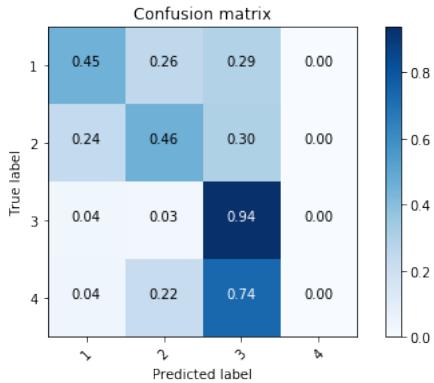


Figure 6: Tumor grade confusion matrix for the L1-penalty logistic regression classifier.

7. Discussion

One important note is that these results are all models are trained on features extracted from the 1000x1000 chunks from each image. It is somewhat difficult to directly compare the results of this paper to others because classification accuracy is also rated on chunks instead of whole images. Future work could include a voting scheme where each chunk has an equally weighted vote in determining whether a slide is GBM or LGG. Despite this, it seems fair to make these comparisons given the high accuracy would mean most of the tiles probably would be in accordance with each other. A pathologist could even just take one of the random top 10 most dense chunks of a pathology slide and use the classifier solely on that chunk.

7.1. Tumor Type Classification

It appears that Yu’s method [22] to select the most cellularly dense images from a subset works reasonably well to select chunks of the slide that actually have cells in them (see Figure 2), despite the heterogeneity and size of pathology slides (see Figure 1), which can be on the order of 10^{10} pixels in size. A prime difficulty for machine learning algorithms with pathology slides is deciding what to look at, and this method takes care of it. (See Background for a the different ways other researchers have dealt with this problem.)

Many papers use the MICCAI 2014 dataset as training data, which is more carefully curated than TCGA. It may be less difficult to get better accuracy on that dataset than TCGA, as reported in papers such as Xu et al and Barker et al. [21], [2]. Barker et al [2] were still able to achieve an impressive 93.1% classification accuracy on 604 images from TCGA and using over 100 features to create their model. Xu et al [21] also uses a complicated model, using a convolutional neural network (CNN) to segment nuclei and extract features to achieve 97.5% accuracy on their MICCAI test set (without attempting TCGA classification).

This project uses over 1800 original slides from TCGA and achieves a classification accuracy of 96% on a more heterogeneous dataset. The pipeline is relatively simple and there are fewer than 25 features total. This makes the model more interpretable by clinicians as well as less likely to overfit one particular dataset. Also, the fact that the different logistic regression models were close in accuracy but used different features could explain the quality of each feature instead of finding accuracy through quantity of features.

I also built several of my own CNNs for this problem (as part of CS231N). They were able to achieve a 95% accuracy with a VGG Net pretrained on ImageNet and fine-tuned on the dense images. Fascinatingly, an 18-layer ResNet pretrained on ImageNet can achieve 83% accuracy by looking at 256x256 versions of whole-slide images.

7.2. Tumor Grade Classification

Tumor grade classification was chosen as a subproblem because tumor grade is important in determining how serious a tumor is and is usually also determined by a pathologist directly from the pathology slide [9]. Because there are more classes to choose from, it makes classification accuracy lower. Ertosun et al [9] achieved approximately 70% classification accuracy between grade II and III tumors on an independent test set. The model from this paper does not use an ensemble classifier, but still achieves 68% accuracy on a TCGA test set. All GBM is a grade IV tumor, so an ensemble classifier may have made sense in this context and definitely warrants another look with these new classifiers.

7.3. Features

Interestingly, useful features seemed to vary based on the classifier and task, although most of the useful features were nucleus characteristics. Somewhat surprisingly, age and gender rarely showed up as important factors in tumor grade and type, despite age and gender being strong factors in determining tumor type [19].

8. Conclusion

In this work, a pipeline was developed to take a pathology image and decide if it was LGG or GBM, predict survival, and to predict tumor grade. While survival prediction has always been difficult for cancers, the more concrete morphological predictions of tumor grade and type proved quite successful from a logistic regression over approximately 20 features only. The simplicity of the model makes it interpretable and keeps it from overfitting, and achieves an accuracy of 88.1% on a test set. This opens up exciting new territory and shows that simpler methods than neural networks work reasonably well if the correct features are chosen. However, neural networks do also show significant promise as the 11-layer VGGNet achieved 95% accuracy on the test set for tumor classification, and an 18-layer ResNet achieved similar accuracy to the best traditional classifier developed in this paper on 256x256 pixel whole-slide images. All of these areas warrant further exploration.

9. Future Work

As mentioned in the discussion, classification accuracy is currently rated on a per-dense-image basis. It may be useful to turn this into a voting scheme where each chosen dense tile of an image would vote to determine whether a slide is LGG or GBM. TCGA has comprehensive genetic information on many of the patients that have submitted these slides. It is possible that simply looking at pathology slides could give insights into genetic information without having to do traditional genotyping. The features extracted here have proven to be useful in predicting things about the tumor that a pathologist can see. Machine Learning can help clinicians see things they couldn't before, and methods like the ones in this paper provide stepping stones into that world.

10. Differences from CS231N work

While I am submitting my CS231N final project (labeled as such) along with this report, hopefully it is clear that the processes used in each were different. The convolutional neural network used transfer learning from a VGGNet with PyTorch on rescaled whole-slide images to classify tumor type.

The files padder.py and PyTorch.py and train_nets.py and

train_nets_256.py were used for CS231N, and most of the rest of the files were created and used for BMI260. Basically anything committed since June 7 at 11:59 was for BMI260.

References

- [1] O. S. Al-Kadi. Texture measures combination for improved meningioma classification of histopathological images. *Pattern recognition*, 43(6):2043–2053, 2010.
- [2] J. Barker, A. Hoogi, A. Depeursinge, and D. L. Rubin. Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles. *Medical image analysis*, 30:60–71, 2016.
- [3] A. N. Basavanhally, S. Ganesan, S. Agner, J. P. Monaco, M. D. Feldman, J. E. Tomaszewski, G. Bhanot, and A. Madabhushi. Computerized image-based detection and grading of lymphocytic infiltration in her2+ breast cancer histopathology. *IEEE Transactions on biomedical engineering*, 57(3):642–653, 2010.
- [4] M. Ceccarelli, F. P. Barthel, T. M. Malta, T. S. Sabedot, S. R. Salama, B. A. Murray, O. Morozova, Y. Newton, A. Radenbaugh, S. M. Pagnotta, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell*, 164(3):550–563, 2016.
- [5] M. Ceccarelli, F. P. Barthel, T. M. Malta, T. S. Sabedot, S. R. Salama, B. A. Murray, O. Morozova, Y. Newton, A. Radenbaugh, S. M. Pagnotta, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell*, 164(3):550–563, 2016.
- [6] S. W. Coons, P. C. Johnson, B. W. Scheithauer, A. J. Yates, and D. K. Pearl. Improving diagnostic accuracy and interobserver concordance in the classification and grading of primary gliomas. *Cancer*, 79(7):1381–1393, 1997.
- [7] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 248–255. IEEE, 2009.
- [8] S. Doyle, M. D. Feldman, N. Shih, J. Tomaszewski, and A. Madabhushi. Cascaded discrimination of normal, abnormal, and confounder classes in histopathology: Gleason grading of prostate cancer. *BMC bioinformatics*, 13(1):282, 2012.
- [9] M. G. Ertosun and D. L. Rubin. Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks. In *AMIA Annual Symposium Proceedings*, volume 2015, page 1899. American Medical Informatics Association, 2015.
- [10] U. C. S. W. Group et al. United states cancer statistics: 1999–2014 incidence and mortality web-based report. Atlanta: US Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute, 2017.
- [11] D. A. Gutman, M. Khalilia, S. Lee, M. Nalisnik, Z. Mullen, J. Beezley, D. R. Chittajallu, D. Manthey, and L. A. Cooper. The digital slide archive: A software platform for manage-

- ment, integration, and analysis of histology for cancer research. *Cancer research*, 77(21):e75–e78, 2017.
- [12] G. D. Helba, C. D. Manthey D, C. L. Beezeley J, and M. Lee S. Histomicstk. <https://github.com/DigitalSlideArchive/HistomicsTK>, 2018.
 - [13] J. Hipp, T. Flotte, J. Monaco, J. Cheng, A. Madabhushi, Y. Yagi, J. Rodriguez-Canales, M. Emmert-Buck, M. C. Dugan, S. Hewitt, et al. Computer aided diagnostic tools aim to empower rather than replace pathologists: Lessons learned from computational chess. *Journal of pathology informatics*, 2, 2011.
 - [14] D. Kondziolka, P. V. Parry, L. D. Lunsford, H. Kano, J. C. Flickinger, S. Rakfal, Y. Arai, J. S. Loeffler, S. Rush, J. P. Knisely, et al. The accuracy of predicting survival in individual patients with cancer. *Journal of neurosurgery*, 120(1):24–30, 2014.
 - [15] W. McKinney et al. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, volume 445, pages 51–56. Austin, TX, 2010.
 - [16] H. S. Mousavi, V. Monga, G. Rao, and A. U. Rao. Automated discrimination of lower and higher grade gliomas based on histopathological image analysis. *Journal of pathology informatics*, 6, 2015.
 - [17] S. Petushi, F. U. Garcia, M. M. Haber, C. Katsinis, and A. Tozeren. Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer. *BMC medical imaging*, 6(1):14, 2006.
 - [18] M. G. Rojo, G. Bueno, and J. Slodkowska. Review of imaging solutions for integrated quantitative immunohistochemistry in the pathology daily practice. *Folia histochemica et cytopathologica*, 47(3):349–354, 2009.
 - [19] T. Sun, A. Plutynski, S. Ward, and J. B. Rubin. An integrative view on sex differences in brain tumors. *Cellular and molecular life sciences*, 72(17):3323–3342, 2015.
 - [20] J. N. Weinstein, E. A. Collisson, G. B. Mills, K. R. M. Shaw, B. A. Ozenberger, K. Ellrott, I. Shmulevich, C. Sander, J. M. Stuart, C. G. A. R. Network, et al. The cancer genome atlas pan-cancer analysis project. *Nature genetics*, 45(10):1113, 2013.
 - [21] Y. Xu, Z. Jia, L.-B. Wang, Y. Ai, F. Zhang, M. Lai, I. Eric, and C. Chang. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC bioinformatics*, 18(1):281, 2017.
 - [22] K.-H. Yu, C. Zhang, G. J. Berry, R. B. Altman, C. Ré, D. L. Rubin, and M. Snyder. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nature communications*, 7:12474, 2016.
 - [23] D. Zink, A. H. Fischer, and J. A. Nickerson. Nuclear structure in cancer cells. *Nature reviews cancer*, 4(9):677, 2004.

11. Appendix

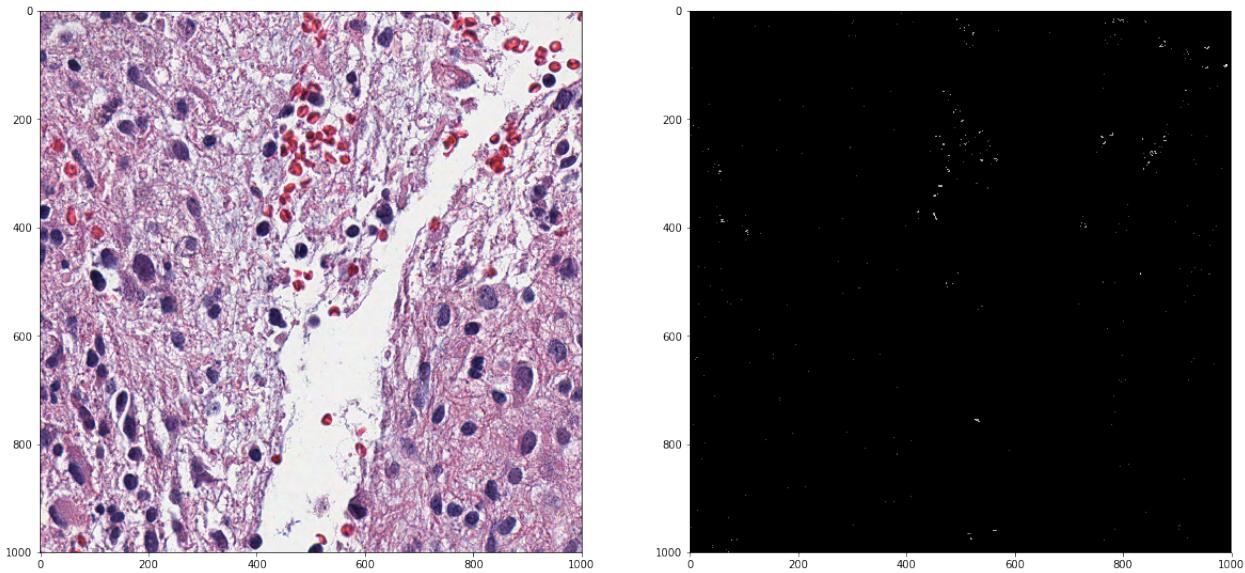


Figure 7: An example of a bad segmentation, where the masking image appears to represent little of the original image. (Masking on right)

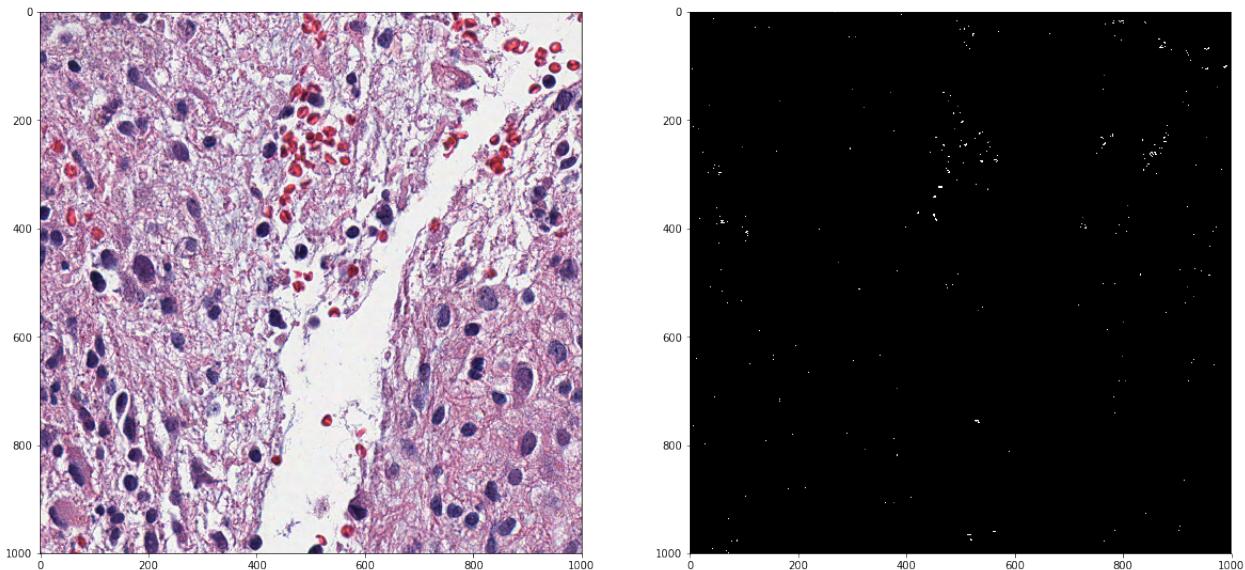


Figure 8: An example of a better segmentation, where the masking image appears to represent more of the original image. (Masking on right)