

Aging Cell Morphology: Identifying cellular transformations associated with aging using image processing

SUMI SINGH*, MAHDI MOQRI, KEN XIONG, and KAVITA KULKARNI

Abstract

Aging is a profound risk factor for the most common fatal diseases in the developed world. It has been shown in literature that there are visually detectable changes in liver cell morphology (e.g. cell volume, nuclear content) that are correlated with aging. Predictors of cellular aging could help identify important age-induced transformations in cells.

In this project, we used digital microscopic images of liver tissue sections harvested from mice at different ages to identify subtle age-related morphological features. Two approaches to segmentation were explored: (1) simple blob detector and (2) fuzzy c-means clustering. Standard geometric and textural features were extracted from the segments and fed into six classification models. Our evaluations suggest that a decision tree model can be used to predict the age of a mouse based on an image of a hematoxylin and eosin stained liver section with an accuracy of 77 percent.

Additional Key Words and Phrases: aging and image processing, image segmentation, convolutional neural networks, liver histopathology

1 INTRODUCTION

For most of the last century, social policy focused on increasing life expectancy in the population; but in recent decades, policy and research are increasingly focused on the potential of increasing healthy life or healthspan.

The process of aging or growing older is an inevitable process. This depends on genetics, physiology, and environmental factors. Aging itself is not a disease but the process represents increased susceptibility to diseases, frailty, and lack of resilience resulting in one or more chronic illness. The understanding of cellular mechanism, molecular changes and genetic basis associated with aging can help in finding the underlying relationship [7].

Aging-related non-infectious conditions in elderly are the leading cause of premature mortality. It is the most profound risk factor for the most common fatal diseases in the developed world [8]. As the access to health care improves, there will be an increase in the global elderly population of age 65 and greater from 8% to 16%. In the year 2030, 16% of world's population will equate to 1.5 billion [13].

It is estimated that by 2030, the elderly population in the United States will be nearly 20% of the total population of the country. With the shifting population demographics, there is a need for collaborative interdisciplinary efforts to investigate different aspects of aging starting from basic science to computational sciences [14]. This also plays well with the shifting focus on increasing healthy lifespan from just increasing lifespan of individuals or population [2].

Biomarkers of cellular aging could help better measure aging at a cellular level and potentially identify important age-induced cellular transformations in cells. In this project, we use microscopic images of liver cells of mice at different ages to identify cellular transformations associated with aging. It has been seen in literature that aging-related changes in liver cells include volume changes and nucleus ploidy [5]. We hope to extract features of geometric and textural type for identifying biomarkers of aging in liver cells.

*All authors contributed equally to this research.

Authors' address: Sumi Singh, forsumi2@gmail.com; Mahdi Moqri, moqri@iastate.edu; Ken Xiong, kfxyooj@stanford.edu; Kavita Kulkarni, kavita_kulkarni@stanford.edu.

2 BACKGROUND

Cell area has been shown to be statistically significant in describing age related morphological changes in bone marrow derived mesenchymal stromal cells (MSCs) [9]. In Oja et al's 2018 paper, MSC images were used to extract nine morphological parameters describing cell age. These parameters include two groups. The first group described the cell size, which included length, width, area and perimeter. The second group of parameters included descriptions of cell shape, and consisted of perimeter-to-area ratio, length-to-width ratio, boxed frame ratio, convex hull area, and convex-hull-perimeter ratio. The images were taken from six tissue culture plates, each one containing 3000 cells per square centimeter that were fixed and stained using Mask Deep Red Stain for cytoplasm and DAPI for nuclei. Several images were used for calibrating for object separation, object-edge thresholding, signal intensity and minimum object size. An object was defined as a cell with only one nucleus and a size of area greater than 765 square micrometer. Each image was acquired using a high-content screening microscope with a 10x objective lens. A 630 nm optical filter and 386 nm optical filter were used to, respectively, detect Cell Mask Red and DAPI. Images were acquired from the inside of the well towards the outside in a spiral path to minimize convexity-related optical distortion.

Another outstanding piece of literature described associations between aging and changes in hepatic macro- and micro-cellular structure and function. Kim et al stated those changes occurred in nuclear ploidy, lipofusion and cell volume [5].

In addition to leveraging existing literature for domain familiarization, we explored some several technical resources for segmentation, feature extraction and model development [15][1][6][3]. Gray-level intensity based nuclei segmentation using maximum entropy thresholding has been shown to give very promising results [15]. Win et al demonstrated in their 2018 paper that maximum entropy thresholding can find an optimum threshold using Shanon's entropy theory without prior knowledge of image levels. They applied the method to extract 16 geometric and 10 textural features for classification of single and polypoid nuclei cells with a random forest classifier.

3 MATERIAL

For this project, we analyzed images of cells in mice liver tissue. The data is sourced from the Atlas of Gene Expression in Mouse Aging Project (AGEMAP)[17]. This dataset is comprised of images from the livers of 30 male and female mice, ranging in age from 1, 6, 16, and 24 months. The images were acquired using a bright-field Carl Zeiss Axiovert 200 microscope. There are 1500 images, each at 40x magnification and 1388x1040 resolution in RGB TIFF format with 12 bits of quantization per channel, totaling 12 Gigabytes of compressed data.

The liver tissue sections are stained with hematoxylin and eosin (HE), two common contrast stains in tissue histology. HE stained tissue sections have purple nuclei, pink cytoplasm and red erythrocytes. The background is white.

4 METHODS

Image segmentation is necessary for well-performing classification. It isolates interesting objects from irrelevant data for downstream processing. Though often considered a pre-processing step, it can also be the most important and computationally challenging part of an image analysis pipeline. In our own pipeline, we segment the nuclei in each image and assign it the averaged cell features. We explored two segmentation methods and several classification models.

Table 1. FCM Parameter Description

Parameter Name	Parameter Value	Parameter Description and Significance
Cluster	3	Number of clusters we expect to see in a given image. We set this to 3 for the best segmentation result.
Epsilon	0.05	Sets the stopping condition for the fuzzy c-mean iteration. As long as the difference between the current and the previous iteration is greater than Epsilon, the iteration will continue.
Iterations	25	Every iteration results in reduction of “cost” associated with the result. Ideally, the cost should be as close as possible to the defined epsilon

Semantic features describe the visual content of an image by correlating low level features such as colour, gradient orientation with the content of an image scene. We used semantic information to describe the desirable properties of segments.

4.1 Segmentation

4.1.1 Simple Blob-Detection based Image Segmentation (SBD). There were two obvious candidates to be segmented for downstream processing - cell nuclei and cell cytoplasm. Segmenting the former proved to be a much more straightforward task than so with the latter, which made it the more productive endeavor. The circular trend in nuclei morphology inspired us to utilize a simple blob-detection approach. Our method for quick and efficient segmentation of cell nuclei in microscopic images of hematoxylin- and eosin-stained histology tissue sections is as follows. First, each RGB image was deconvolved into channels representative of hematoxylin (hema), eosin (eos), and 3,3'-Diaminobenzidine (DAB) [12]. The chemistry of the stains is such that hematoxylin is concentrated in cell nuclei and eosin gathers in the cytoplasm. Accordingly, the hematoxylin channel was used to segment nuclei. Proceeding with just the hematoxylin channels, images were denoised with a Gaussian filter, and then binarized with an Otsu threshold to create a mask of the regions of interest. The produced binary regions were tidied: Edges were smoothed, small holes were filled, and objects too small to be considered nuclei were deleted. A Laplacian-of-Gaussian-based blob-detection algorithm was subsequently applied to the resulting mask, which produced coordinates to the center of each identified blob. These coordinates were mapped to labelled regions in the binary mask, yielding the nuclei segments from which geometric and texture features were extracted and averaged.

4.1.2 Fuzzy C-Mean Clustering based Image Segmentation (FCM). FCM method is similar to a KNN clustering technique except that a given pixel in an image can belong to multiple clusters rather than just one cluster. This ‘fuzziness’ is bounded by a parameter called membership value which provides a distance measure between the data point of interest and the centroid of clusters. The farther away a given data point is from a cluster centroid, the lower its membership value. The number of clusters is user defined and often requires extensive trial-and-error experiments or expert inputs to ensure the right number of clusters are defined for a given set of images. Large variance results with changes to cluster size. Setting the cluster size at 2 gave most optimal results. As the number of iterations exceeded 25, the algorithm’s performance plateaued. The table below gives the parameters and their descriptions for running FCM on mice image slices.

4.2 Feature Extraction

Creating scale space: We created a scale space of 100 to create 100x tiers of the input image for analysis. Setting the scale to 100 gave us the highest number of blobs(nuclei) detected in each slice.

Laplacian filtering: The images were convolved with a Laplacian of Gaussian filter to obtain a scale-space where high responses denote the location of blobs with a specific size.

Loop through scale space and detect blobs: We set a threshold of 20 across all scales as we were only interested in similar sized blobs (nuclei) rather than a range of blob sizes. For every scale image, peaks were detected using the maximum filter technique. All pixels in the neighborhood are set to the maximal value. Peaks were defined as where the image equals the value of maximum filter of that image.

Prune blobs using local maxima: We looped over all pair-wise blob combinations and check for overlapping blobs. If the blobs overlapped beyond a defined degree, the largest blob was kept and the other discarded.

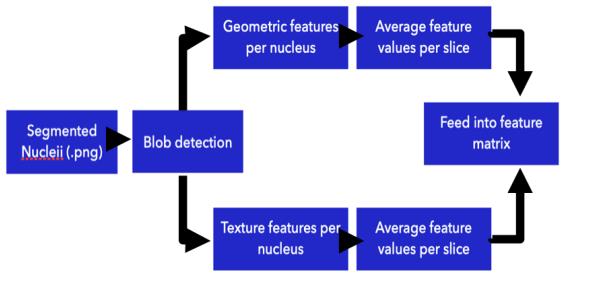


Fig. 1. Feature extraction

4.3 Data Preparation

There were 500 tissue sections per mouse. The entire set was split into training, validation and testing sets, so that no mouse was a part of more than one set. This was done to prevent the introduction of classification bias due to similarity in slices from the same organism.

The images are categorized into four classes based on the age of the represented mice: 1 month, 6 month, 16 month, and 24 month. The class labels can be considered ordinal, based on the fact that each label is associated with an age, or nominal, where the age information is discarded. For the purpose of this task, we consider the class membership to be nominal and assume that the order of the classes based on the age has no inherent effect on the prediction. Classification is performed in three different feature spaces. The first feature space includes feature extraction on all instances using only the SBD method.

Table 2. Geometric and Texture Features Extracted using SBD

Geometric Feature	Geometric Feature Description	Texture Feature	Texture Feature Description
Area	Area of the blob	Energy	$\sqrt{\sum_{i,j=0}^{levels-1} P_{i,j}^2}$
Bounding Box Area	Area of the bound box around the blob	Homogeneity	$\sum_{i,j=0}^{levels-1} \frac{P_{i,j}}{1+(i-j)^2}$
Convex Area	Area of convex hull	Correlation	$\sum_{i,j=0}^{levels-1} \frac{P_{i,j}}{1+(i-j)^2}$
Eccentricity	Ratio of the focal distance over the major axis length	Dissimilarity	$\sum_{i,j=0}^{levels-1} P_{i,j} i - j $
Equivalent Diameter	Diameter of a circle with the same area as the region		where P is the grey level co-occurrence matrix
Euler Number	Number of objects subtracted by number of holes		
Extent	Ratio of pixels in the region to pixels in the total bounding box		
Filled Area	Number of pixels of the region with all the holes filled in		
Major Axis Length	Length of the major axis of the ellipse that has the same normalized second central moments as the region		
Max Intensity	Value with the greatest intensity in the region		
Mean Intensity	Value with the mean intensity in the region		
Minimum Intensity	Value with the least intensity in the region		
Minor Axis Length	Length of the minor axis of the ellipse that has the same normalized second central moments as the region		
Orientation	Angle between the 0th axis (rows) and the major axis of the ellipse that has the same second moments as the region		
Perimeter	Perimeter of object approximated as the length of the line through the centers of border pixels		
Solidity	Ratio of pixels in the region to pixels of the convex hull image		

The second feature space includes features extracted only using the FCM method. Here, it was observed that the segmented images taken from subjects 1 month of age seemed to contain some spurious values. Only for the purpose of exploration, we removed all the instances for age 1 month from FCM only classification/prediction on FCM segmented feature space.

Table 3. Geometric and Texture Features Extracted using FCM

Geometric Feature	Geometric Feature Description	Texture Feature	Texture Feature Description
Number of Nuclei	Computed by calculating the number of blobs found in a slice	Average Energy	GLCM feature average of energy across image patches created around each blob coordinate
Number of overlapping nuclei	Computed by determining overlap area	Average Homogeneity	GLCM feature average of homogeneity across image patches created around each blob coordinate
Average Area	Total area of all blobs/number of blobs	Average Correlation	GLCM feature average of correlation across image patches created around each blob coordinate
Average Compactness	Compactness = Area/(perimeter) ² . Averageof Compactnessacrossallblobs		
Circular blobs	Roundness = 4 * pi * area/(perimeter) ² . If Roundness<1, blob is circular		

The third feature space combines the features from simple SBD and FCM to derive a composite feature space for classification. For all three feature spaces, the image feature values were averaged over an image slices and have not been normalized over each category or the sample.

4.4 Prediction

The goal of this work is to be able to classify new liver histology images into categories based on the age of subject from which these liver tissue images are extracted. To categorize the unknown images based on their age, we perform classification. We used two different approaches for our classification: 1) using Weka on FCM, and 2) using Python prediction scripts on Composite dataset. In this section, we present both the methodology for each approach and the results of each will be presented in the following section.

4.4.1 WEKA for prediction. For FCM derived feature spaces, classification is run on Waikato Environment for Knowledge Analysis (WEKA, Version 3.8.3, for macOS Mojave, version 10.14, processor 1.4 GHz Intel Core i5, memory 4GB 1600 MHz DDR3). WEKA is a Java based machine learning environment popularly used for data mining tasks [16]. To predict class membership we use three different classifiers that are known to work well with nominal classes and numeric valued instances. Respectively, these classifiers are tree-based, function-based, and neighborhood-based lazy learner. For tree-based learning, we used a java-based j48 implementation of the C4.5 algorithm optimized for continuous valued attributes [11]. For functional machine learning, we used Sequential Minimal Optimization (SMO), which is

a faster implementation of support vector machines. SMO uses heuristics based partitioning of training into smaller problems and analytically combines the results. It is a strong algorithm for automatic hyper-parameter tuning for finding good separation plain [10]. The SMO classifier replaces all missing values, transforms nominal attributes into binary ones, and normalizes all attributes by default. A multi-class classification problems is decomposed into pairwise binary classification or 1-vs-1 classification problem, the solved using pairwise classification (aka 1-vs-1). The final class probabilities are predicted using the pairwise coupling method described by Hastie and Tibshirani in 1998 [4]. A linear kernel, implemented as a polynomial kernel of power 1 is used as the classifying kernel. Modified logistic regression with a ridge estimator is used as the calibrating function. The modification was made to utilize the instance weights. For instance-based learning, we used a Java implementation of K-nearest neighbor with cross validation, where the target outcome is determined by a majority vote. A brute force search algorithm was implemented for nearest neighbour search based on Euclidean distance, which built data models that computed the distances between objects in the data model. The value of K was set to 3.

4.4.2 Python Models Approach. In order to explore more prediction (classification models) and verify the result of our packaged software (WEKA) approach, we developed a set of six prediction models, using the scikit-learn Python module. The code is developed using Python 3.4. The feature space we used for our models were the third feature space, mentioned in the previous section, combining all the features from FCM and SBD (combined based on animal ID as well as case ID). We normalized all the features using a min-max scaling to take care of differences in the features distribution. We then trained the model on 70 percent of the cases (split solely based on animal ID) and tested the model on the remaining 30 percent of the data.

The six prediction models that were fitted to the data are as follows: 1) Logistic Regression, 2) Decision Tree, 3) K-Nearest Neighbors (KNN), 4) Latent Dirichlet allocation (LDA), 5) Gaussian Naive Base, and 6) Support Vector Machines (SVM). Together these models present a set of diverse classification methods that are able to group data based on different underlying similarities, ranging from linear relations to more complex relation in higher feature spaces.

5 RESULTS, EVALUATION AND CONCLUSIONS

5.1 Segmentation and Feature Extraction

Segmentation and feature extraction results for both SBD as well as FCM based method were visually inspected and verified. In some instances, inversion masks were applied to FCM when segmentation resulted in opposite segments from what was desired.

Figures 2 and 3 depict the two segmentation strategies used in our methods. The single tissue section in figure 2 shows that our blob detection implementation, in its simplicity, does not perfectly isolate and label nuclei. The deconvolution and binarization performed well. In the hematoxylin channel, nuclei were easily distinguishable bright spots, while cytoplasm faded into the background. This resulted in enough separation that the binary mask produced by Otsu's threshold included nearly all of the nuclei and very little of the cytoplasm. Unfortunately, the granularity of the nuclei resulted in lobular binary regions, reducing the accuracy of the blob detection algorithm. Over- and under-segmentation were both present in the final output. Nuclei with many lobes were segmented into multiple blobs, while multiple closely-packed nuclei were identified as single blobs. Figure 3 shows the FCM-based image segmentation results for a single image.

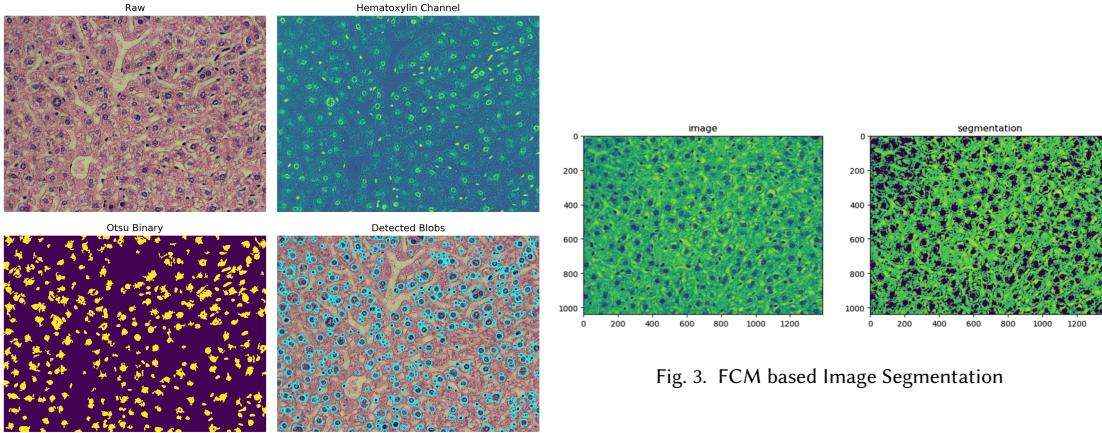


Fig. 2. SBD based Image Segmentation

Fig. 3. FCM based Image Segmentation

5.2 Classification Model Selection

5.2.1 Model selection on FCM only data using WEKA. In Figure 4 Decision Tree has an AUC=1 for all three classes, indicating that, for this training-test set, decision tree has good measure of separability of classes.

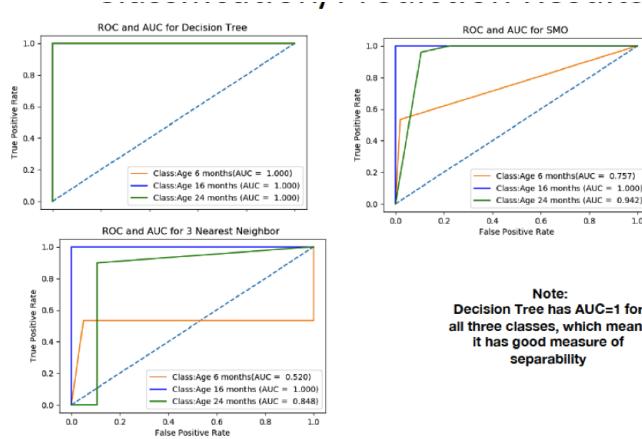


Fig. 4. FCM ROC and AUC for classifiers

5.2.2 Model results on composite data. The model results below is on data that has combined features from FCM and SBD. Each of the six prediction models (Logistic Regression, Decision Tree, KNN, LDA, Gaussian Naive Base, and SVM) were separately trained on the train set and the accuracy of each were calculation on the train as well as the test data sets. The table below presents the accuracy of each model on both data sets.

Table 4. Accuracy of the Decision Tree on the training and testing data

Model	Accuracy on Training	Accuracy on Testing
Logistic Regression	0.87	0.59
Decision Tree	1.00	0.77
KNN	0.97	0.61
LDA	0.98	0.83
Gaussian Naïve Base	0.87	0.52
SVM	0.77	0.67

Because it achieved the highest accuracy of all of the classification models and easily identified important features, we recommend the Decision Tree for such a classification task. To further investigate the Decision Tree model, we calculated the prediction accuracy in each age category. The table below presents the accuracy of the model in each age-group of the test set.

Table 5. Accuracy of the Decision Tree for different age categories on testing data

\	1 Month	6 Months	16 Months	24 Months
1 Month	0	0	0	0
6 Months	25	75	0	0
16 Month	0	0	50	0
24 Month	37	7	0	106

The accuracy of the model is highest (100 percent accurate) within the 16 months age group, and lowest within the 1 month age group. We will explore and verify this further in the next section, using a plot of the data in the feature space of our two most important features, based on the Decision Tree model.

6 DISCUSSION

Error Analysis: The small sample size simplifies our error analysis. From the confusion matrices, it can be seen that most mis-classifications occur for samples that were 6 months of age. Considering that the spurious nature of the FCM and Scale Space Data rendered features from 1 month old mice non-predictive, it may be that 6 month old mice, being closest temporally to the 1 month old mice, also contained features that were non-predictive.

Decision Nodes: The decision nodes of decision tree classifier can be considered as potential biomarkers for aging. Decision nodes include both texture and geometric features. We noticed that Average Cell Area is one of the decision nodes of decision tree classifier. This is in line with prior research, where cell area was shown to be statistically significant in describing age-related morphological changes in bone marrow derived mesenchymal stromal cells (MSCs) [9]. This agreement between our study and Oja et al's strengthens the argument that there may be some relationship between cell area and organism age.

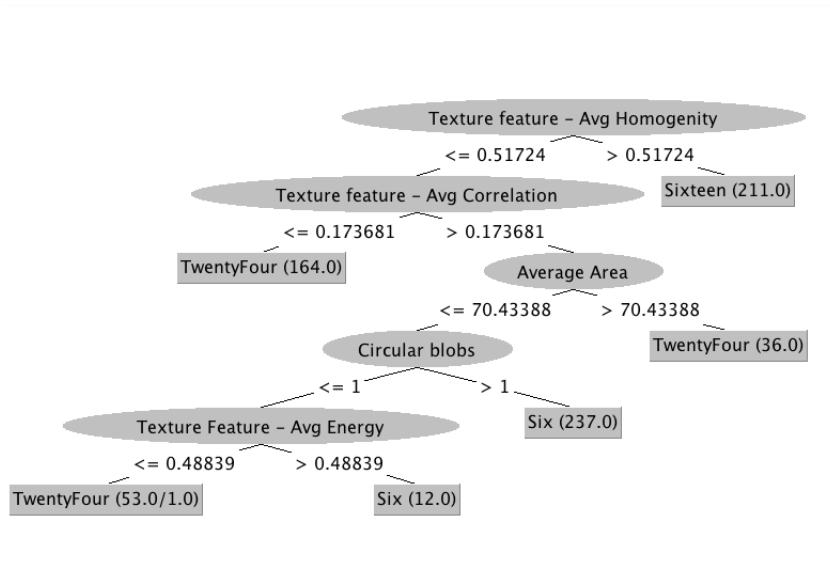


Fig. 5. Decision Nodes

In addition, our Python Decision Tree gave us some complementary information and insights. The most important features detected by our Python models are "Texture - Average Correlation," "Solidity," and "Texture - Average Homogeneity" with the relative weights of 0.50, 0.13, and 0.10, respectively. Figure 6 shows classification based on the four age groups in the space of the two former features. It visually confirms the separability of the data in this space.

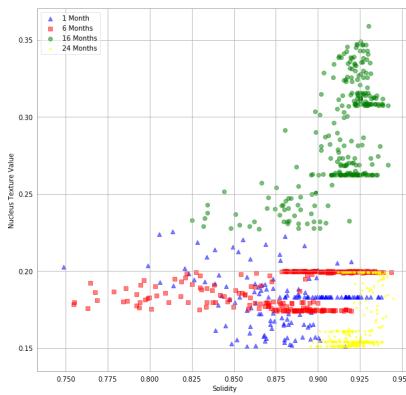


Fig. 6. Classification plot in the feature space

The results of our Python prediction models, particularly those of the decision tree, agree with the results of our WEKA models. Specifically, two of the most important features identified by our Python model - "Texture - Average Correlation" and "Texture - Average Homogeneity" - are at the top levels of our WEKA Decision Tree (signifying its importance).

All the scripts for our prediction models are publicly available, in the form of a Python Notebook, on Github at <https://github.com/moqri/Aging-Cell-Morphology/blob/master/Prediction.ipynb>

7 FUTURE WORK

We presented geometric and texture features derived from mice liver histology images using FCM-based image clustering and Scale-Space Filtering-based blob detection that were reasonably strong predictors of biological age, but there is much work left to complete. We suggest that future investigators, inspired by our work, look for biological pathways associated with the features that identified. Knowledge of the molecular actors involved in the aging process could reveal candidate targets for drug therapies.

Our method is certainly not without faults. As was discussed, more data would better define the integrity of our claims. It would have been useful to have images from many more mice from a broader age range. Furthermore, we have yet to discover a proper reason for the spurious behavior of images taken at age 1 month and recommend its investigation in a follow-up study. Finally, our two-pronged approach to segmentation indicates the complexity of the segmentation task. In this project, we were unable to segment the cell cytoplasm well enough for proper characterization. Future attempts to do so might want consider a convolutional neural network approach.

REFERENCES

- [1] [n.d.]. Nuclei Segmentation. <https://digitalslidearchive.github.io/HistomicsTK/examples/nuclei-segmentation.html>
- [2] Eileen M Crimmins. 2015. Lifespan and healthspan: past, present, and promise. *The Gerontologist* 55, 6 (2015), 901–911.
- [3] Guillaume Gay. [n.d.]. Segmenting nuclei with skimage. <http://morphogenie.fr/segmenting-nuclei.html>
- [4] Trevor Hastie and Robert Tibshirani. 1998. Classification by pairwise coupling. In *Advances in neural information processing systems*. 507–513.
- [5] Hee Kim, Tatiana Kisseleva, and David A Brenner. 2015. Aging and liver disease. *Current opinion in gastroenterology* 31, 3 (2015), 184.
- [6] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. 2017. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE transactions on medical imaging* 36, 7 (2017), 1550–1560.
- [7] Pamala D Larsen. 2019. Geroscience: The Intersection of Basic Aging Biology, Chronic Disease, and Health.
- [8] Teresa Niccoli and Linda Partridge. 2012. Ageing as a risk factor for disease. *Current Biology* 22, 17 (2012), R741–R752.
- [9] S Oja, P Komulainen, A Penttilä, J Nystedt, and M Korhonen. 2018. Automated image analysis detects aging in clinical-grade mesenchymal stromal cell cultures. *Stem cell research & therapy* 9, 1 (2018), 6.
- [10] Platt J Sequentia1 Minima Optimizat. 1999. ion: A Fast Algorithm for Training Support Vector Machines. *Advances in Kernel Methods SuppOrt Vector I earring*, MIT Press, Cambridge, MA (1999), 185–208.
- [11] J Ross Quinlan. 1996. Improved use of continuous attributes in C4. 5. *Journal of artificial intelligence research* 4 (1996), 77–90.
- [12] Arnout C Ruifrok, Dennis A Johnston, et al. 2001. Quantification of histochemical staining by color deconvolution. *Analytical and quantitative cytology and histology* 23, 4 (2001), 291–299.
- [13] Douglas L Schmucker. 2005. Age-related changes in liver structure and function: implications for disease? *Experimental gerontology* 40, 8-9 (2005), 650–659.
- [14] Douglas R Seals, Jamie N Justice, and Thomas J LaRocca. 2016. Physiological geroscience: targeting function to increase healthspan and achieve optimal longevity. *The Journal of physiology* 594, 8 (2016), 2001–2024.
- [15] Khin Win, Somsak Choomchay, Kazuhiko Hamamoto, and Manasanan Raveesunthornkiat. 2018. Detection and Classification of Overlapping Cell Nuclei in Cytology Effusion Images Using a Double-Strategy Random Forest. *Applied Sciences* 8, 9 (2018), 1608.
- [16] Ian H Witten, Eibe Frank, Mark A Hall, and Christopher J Pal. 2016. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- [17] Jacob M Zahn, Suresh Poosala, Art B Owen, Donald K Ingram, Ana Lustig, Arnell Carter, Ashani T Weeraratna, Dennis D Taub, Myriam Gorospe, Krystyna Mazan-Mamczarz, et al. 2007. AGEMAP: a gene expression database for aging in mice. *PLoS genetics* 3, 11 (2007), e201.