Detecting age-related macular degeneration (AMD) biomarker images using MFCC and texture features

Yiyang Wang[a], Xufan Ma[a], Rob Weddell[a], Abum Okemgbo[b], David Rein[c], Amani A. Fawzi[d], Jacob Furst[a], and Daniela Raicu[a]

[a]DePaul University, Chicago, Illinois, United States
[b]University of Pennsylvania, Philadelphia, Pennsylvania, United States
[c]Duke University, Durham, North Carolina, United States
[d]Feinberg School of Medicine, Northwestern University, Chicago, Illinois, United States

ABSTRACT

Age-related macular degeneration (AMD) is the leading cause of irreversible vision loss in older individuals. Clinically, ophthalmologists visually inspect optical coherence tomography (OCT) volumes to diagnose the stage of AMD based on well-known biomarkers. An early characteristic of AMD is drusen, which appears as yellowish deposits under the retina. AMD is mainly categorized into two types: dry AMD (non-neovascular) and wet AMD (neovascular). Given the large number of OCT images in an individual volume, an efficacious computer-aided detection system can reduce the workload for ophthalmologists by automatically detecting biomarkers in the relevant images. Because the shape of the RPE is critical in defining the pathological changes caused by wet and dry AMD, we propose a novel approach to describe the RPE shape using Mel Frequency Cepstral Coefficients (MFCC). Our previous work indicates that Haralick texture features have the ability to distinguish drusen from healthy tissue on color photography, therefore, we also investigated Haralick texture features extracted from the region between Inner Limiting Membrane (ILM) and Bruch's Membrane (BM) layers in this study. We achieved a mean accuracy, sensitivity with respect to AMD image and specificity with respect to healthy image of 89.68%, 89.26% and 90.12% on testing sets and 69.22%, 67.40%, and 75.56% on new patient validation sets, respectively. Our binary classification results indicate that MFCC are uniquely suited for producing generalizable results to automatically detect AMD biomarker images.

Keywords: Age-related macular degeneration, Biomarker images, Computer-aided detection, Mel-frequency cepstral coefficients, Texture features

1. INTRODUCTION

Age-related macular degeneration (AMD) is the leading cause of irreversible vision loss in older individuals.[1] Clinically, ophthalmologists visually inspect optical coherence tomography (OCT) volumes to diagnose the severity and subtypes of AMD based on characteristic biomarker signatures in these images. An early characteristic of AMD is drusen, which appears as yellowish deposits under the retina. AMD is mainly categorized into two types: dry AMD (non-neovascular) and wet AMD (neovascular). Dry AMD spectrum is represented by drusen deposition, later evolving into confluent areas of regressed drusen and ultimately in the advanced dry stage presenting as loss of vision associated with retinal pigment epithelium (RPE) atrophy (clinically known as geographic atrophy, GA). Wet AMD is characterized by the leakage of fluid in the sub-RPE and subretinal spaces caused by neovascularization.[1] Fig. 1 shows three types of biomarker OCT images. OCT, a noninvasive imaging modality, is primarily used in AMD diagnosis and management.[2] Given the large number of OCT images in an individual OCT volume, an efficacious computer-aided detection system can reduce the workload for ophthalmologists by automatically detecting biomarkers in these images.

Further author information:
Y.W.: E-mail: ywang192@depaul.edu
D.R.: E-mail: draicu@cdm.depaul.edu

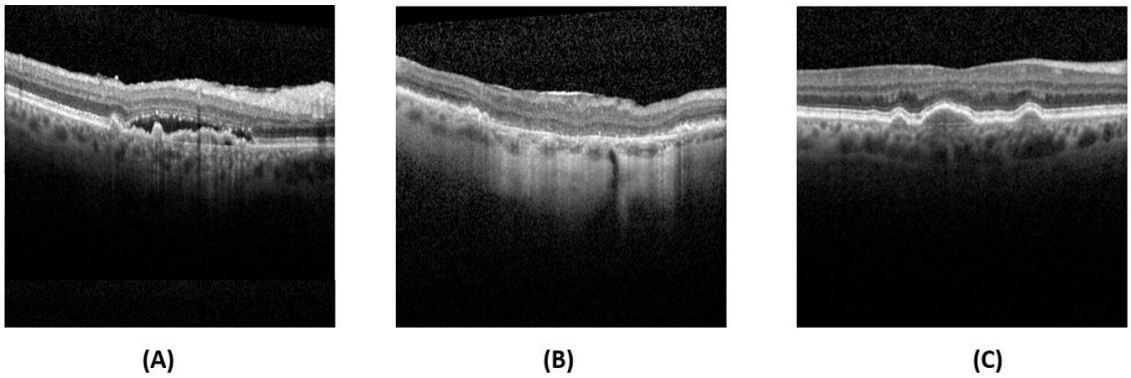

Figure 1: Three types of biomarker OCT images: (A) wet AMD; (B) dry AMD; (C) drusen

Although researchers have designed a few algorithms to automatically detect AMD biomarkers from OCT images, most previous works focus on classifying OCT images at the patient level or the OCT volumetric level. In other words, the classifier assigns one label (e.g. unhealthy) per patient or per entire volume of OCT images, although individual images within a volume can show a wide range of AMD biomarker signatures, including combinations of signatures within the same image. Farsiu *et al.*[3] investigated the thickness of RPE and other retinal layers to distinguish intermediate AMD patients from healthy subjects with 99% precision. Albarrak *et al.*[4] extracted oriented gradient local binary pattern histograms from OCT images and achieved a best area under the receiver operating curve (AUC) value of 94.4% at volumetric level. Instead of classifying whether a patient has AMD or whether an OCT volume belongs to an AMD patient, we aim to detect AMD biomarker in each image within an individual volume.

Furthermore, one of the limitations of some of the previous studies on OCT image classification is that they did not perform the validation on unseen new patient data, instead testing their models on data from patients seen in the training phase. Since same patient images are typically correlated, previous results may not be generalizable. For example, Wang *et al.*[5] classified AMD, diabetic macular edma (DME) and healthy macular OCT images and reported 99.3% overall accuracy by extracting linear configuration pattern (LCP) features on multiple scales. As for deep learning approaches, Treder *et al.*[6] utilized pretrained ImageNet to classify AMD and healthy images with 96% accuracy. Most recently, Kuwayama *et al.*[7] combined image augmentation techniques and convolutional neural networks (CNN) to detect AMD and other macular diseases. The precision and recall for wet AMD were 100% and 77%. To avoid the bias caused by correlated images in training and testing sets, we also report classification results on new patients data.

Our study innovatively combined Mel Frequency Cepstral Coefficient (MFCC) and Haralick texture features with a random forest classifier. While MFCC was originally designed for speech recognition[8,9] its application also includes hand gesture recognition,[10] phonocardiogram classification[11] and emotion detection.[12] Since the shape of RPE is critical in reflecting the pathological changes caused by both wet and dry AMD,[1] we propose a novel approach to describe the RPE shape using MFCC. Our previous work indicates that Haralick texture features have the ability to distinguish drusen from healthy tissue,[13] and we know drusen is the early sign of AMD,[1] therefore, we also investigated Haralick texture features extracted from the region between inner limiting membrane (ILM) and Bruch's Membrane (BM) layers in this study.

2. METHODS

Our methodology consists of human annotation, segmentation of the region of interest, feature extraction and classification as illustrated in Fig. 2. First, we asked experienced ophthalmologists to annotate images that contain either sub type of AMD (Wet and GA) as well drusen images; second, we segmented the RPE layer; third, we extracted MFCC using the segmented RPE layer and calculated Haralick texture features from the region between ILM and BM. Finally, we trained a random forest classifier on annotated images and tested the

classifier on new patient data. The classification labels in this study are 'AMD image' and 'healthy image'. We evaluated our method on testing sets and new patient validation sets.

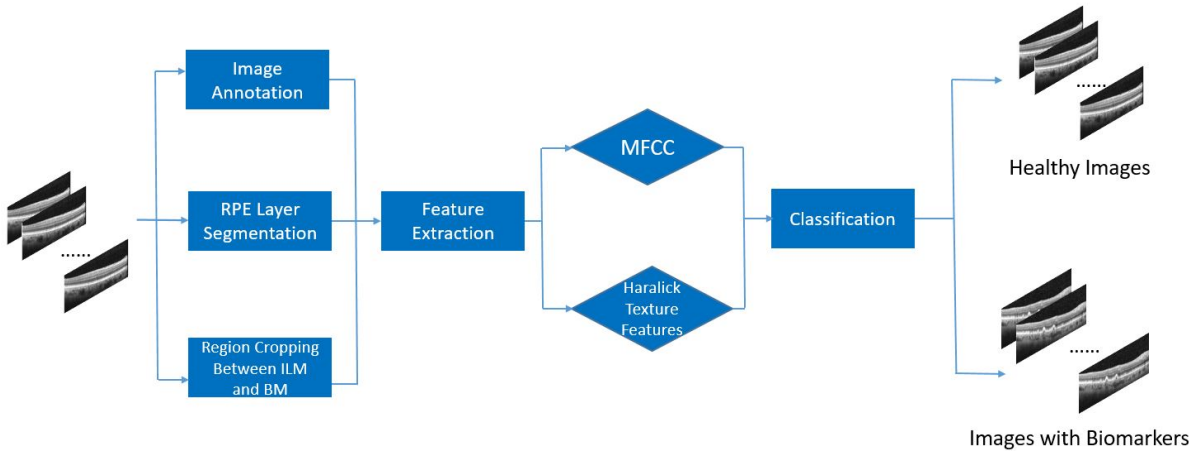

Figure 2: An overview of the methodology. The process consists of image annotation, ROI segmentation, feature extraction and classification.

2.1 Data

The data in this study was collected from Northwestern Memorial Hospital and all scans in the study were acquired with a spectral domain (SD)-OCT device (SPECTRALIS; Heidelberg Engineering, Inc., CA, USA). OCT images taken during each patient's follow-up visit were used to diagnose the spectrum of AMD progression. In this study, we choose to annotate and analyze OCT scans from the last visit that contains clear biomarker images and we also annotate 90 images without AMD biomarkers (Table 1).We implement a stratified method to split annotated images into training (70%) and testing datasets(30%). To evaluate our method on the validation sets, we remove images of two wet AMD patients and of two GA-AMD patients from the training set, treat the last visit OCT volumes of the removed four patients as validation sets. Our validation sets have 21 healthy images and 73 AMD images in total.

Table 1: A summary of the study data

Category	No. of Patients	No. of Annotated Images
Wet AMD	11	82
Advanced Dry AMD (GA)	16	121
Only Drusen	14	145
Healthy Images	17	90

2.2 RPE Layer Segmentation

Since the RPE layer is the most reflective layer in an OCT image,[14] we first iterate through all the columns in an image, and find the highest intensity pixels in each column as the candidate RPE pixels. Because the inner segment/outer segment (IS/OS) junction is highly susceptible to OCT signal strength[15] and the image acquisition process can bring noisy pixels, locations of highest intensity pixels of each column cannot guarantee the correct RPE segmentation. To solve this problem, we utilize a moving window to detect the potential wrongly selected pixels. Within a window, we calculate the first quartile ($Q1$), the third quartile ($Q3$) and the inter quartile range (IQR) of candidate pixels' y-coordinate values. If the y-coordinate value of a candidate pixel is smaller than

$Q1 - 1.5 * IQR$ or greater than $Q3 + 1.5 * IQR$, we removed this pixel from our candidate pixels. The height of the window is the interval between ILM and BM layers (the OCT device provided segmented coordinates of ILM and BM layers). We varied the window width and obtained the best RPE segmentation result with a window width of 200 pixels. Finally, we used linear interpolation to fill in the gaps caused by the outlier removal. Fig. 3 shows an example of the RPE layer segmentation. This image has clear bio-markers indicating GA. Blue dots are the candidate RPE pixels and the red line represents the segmented RPE layer after noise removal and smoothing.

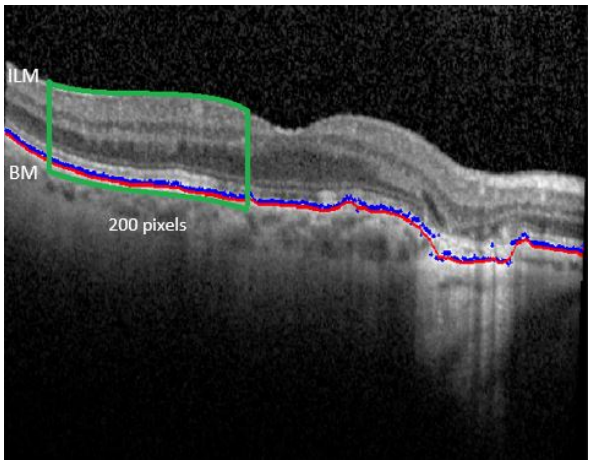

Figure 3: An example of RPE layer segmentation. Blue dots are the candidate RPE pixels and the red line represents the segmented RPE layer after noise removal and smoothing. The green box represents a window for outlier removal.

2.3 MFCC Features Extraction

From Sec. 2.2, we can obtain a vector of y-coordinates of RPE layer for each image and treat this vector as the input 'signal' of the MFCC extraction process (Fig. 4). First, we use Hamming window function (Equation (1))[16] in order to decrease the impact of spectral leakage and counteract the assumption of fast Fourier transform (FFT) that the data should be infinite.[17]

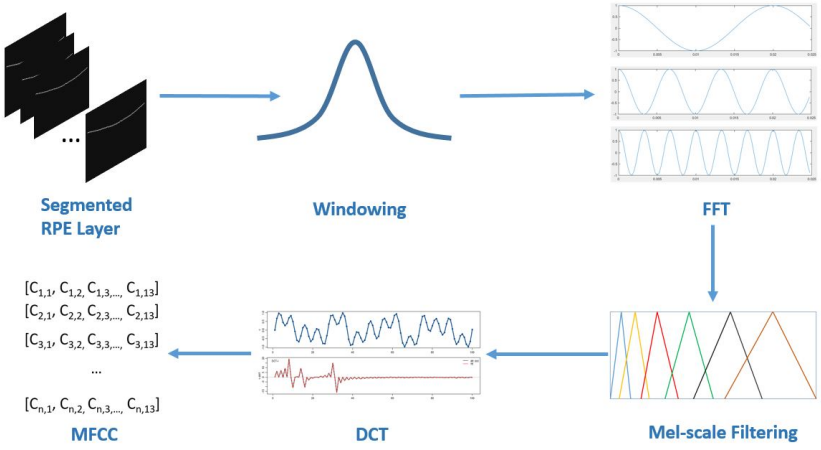

Figure 4: The process of MFCC feature extraction.

$$w[n] = 0.54 - 0.46cos(\frac{2\pi n}{N-1}), \tag{1}$$

where N is the total number of pixels on a segmented RPE layer and $0 \leq n \leq N-1$.

Second, we conduct N-point FFT (Equation (2)) and compute its power spectrum P (Equation (3)).[18]

$$FFT(x) = \sum_{n=1}^{N} xw[n]e^{-i2\pi n^2/N}, \tag{2}$$

Here, x represents a vector of segmented RPE y-coordinates and $e^{-i2\pi/N}$ is a primitive N^{th} root of 1.

$$P = \frac{|FFT(x)|^2}{N}, \tag{3}$$

Third, we multiply the Mel-Scale filter-banks H_m (Equation (4))[16] with the power spectrum computed in the previous step and added up the coefficients:

$$H_m(n) = \begin{cases} 0 & n < f(m-1) \\ \frac{n-f(m-1)}{f(m)-f(m-1)} & f(m-1) \leq n \leq f(m) \\ \frac{f(m+1)-n}{f(m+1)-f(m)} & f(m) \leq n \leq f(m+1) \\ 0 & n > f(m+1) \end{cases} \tag{4}$$

where m is the number of filters, $f()$ is $m+2$ mel-spaced frequencies. In our study, we chose $m = 26$, which is a standard number of filters applying to the power spectrum, and we obtained 26 filter-bank energies. Furthermore, we take the log of each of the energies and obtain the log of filter-bank energy $s(m)$:

$$s(m) = log_e(\sum_{n=1}^{N} |P|H_m(n)), \tag{5}$$

At last, in order to obtain the 26 cepstral coefficients, we utilize discrete cosine transform (DCT) (Equation (6))[19] to decorrelate the log of filter-bank energy. As the previous work suggested,[11] we discard coefficients that represent fast changes, therefore, only 13 of 26 coefficients are used.

$$c(i) = \sum_{i=1}^{m} s(m)cos(\frac{\pi i}{m}(i - \frac{1}{2})), i = 0, 1, 2, ...m \tag{6}$$

2.4 Haralick's Texture Feature Extraction

Besides considering the shape of RPE layer, we also examine the information between ILM and BM layers by extracting Haralick texture features[20] on ROIs. The 13 Haralick texture features are computed from a gray level co-occurrence matrix (GLCM) that encodes the spatial relationship of the gray-levels Lg across a certain direction and pixel displacement d.[13] In this study, we calculated 4 GLCM for $Lg = 8$, $d = 1$, and $\theta = 0°, \theta = 45°, \theta = 90°$, and $\theta = 180°$. We extract 13 Haralick texture features that were averaged with respect to the angle. The detail of all 13 texture features are listed in our previous study.[13]

2.5 Classification and Feature Importance Analysis

Random Forest is an ensemble of classifiers that creates a group of decision trees from the original data by bootstrapping and then randomly choosing features to build the tree.[21] The algorithm classifies an instance by a majority vote across all the classification outputs of the individual decision trees. We examine the "Out-of-Bag" (OOB) error[13] to determine the optimal number of features per split and the total number of trees. To determine the most relevant feature, we use Gini Index defined as:

$$Gini(node) = \sum_{i=1}^{m} p_i(1 - p_i), \tag{7}$$

where m is the number of distinct classes, $p_i = \frac{s_i}{s}$ and s_i is the number of samples in class C_i at that node with s samples in total.

To report more reliable results, we repeated the process 30 times, calculated the mean accuracy, sensitivity (AMD image is the positive case) and specificity (healthy image is the negative case), and report the classification performance under 95% confidence interval. To confirm that the differences between feature sets are statistically significant, we implement Welch's t test.[22] To avoid the curse of dimensionality, we only chose the top three important texture features in the 'MFCC and Texture Features' feature set.

3. RESULTS

Tab. 2 presents the classification results on both testing set and validation set. For these results, AMD image is the positive case and Healthy image is the negative case. The random forest classifier indicates that contrast, the first and the second measures of information correlation are the top three important texture features and these three features are included in 'MFCC and Texture Features' set. Our Welch's t test result shows that differences between 'Only Texture Features' and other two feature sets are significant ($p < 0.05$). However, differences between 'Only MFCC' and 'MFCC and Texture Features' are not significant for both testing and validation sets.

From Tab. 2, we can see that we achieved better classification performance using MFCC on testing set. For validation set, even though we achieved higher mean accuracy and sensitivity by using 13 texture features, the mean specificity is only 21.25%. This number shows that the classifier predicts almost all the cases as positive and since the majority of instances in our validation set are positive (73 out of 94), therefore, we can see a high mean accuracy but low mean specificity result by using just texture features. In contrast, the classification results using MFCC are uniquely suited for producing generalizable results on automatically detecting AMD bio-marker images.

Table 2: Random Forest Classification Result Under 95% Confidence Interval. 200 trees and 5 features per split in Random Forest algorithm.

Mean	Data Set	Only Texture Features	Only MFCC	MFCC and Texture Features
Accuracy	Testing	77.35% ± 2.10%	88.12% ± 1.59%	89.68% ± 1.75%
	Validation	71.49% ± 1.30%	69.22% ± 0.63%	70.11% ± 1.17%
Sensitivity	Testing	73.58% ± 3.40%	87.87% ± 2.75%	89.26% ± 2.88%
	Validation	85.80% ± 1.92%	67.40% ± 0.81%	68.58% ± 1.46%
Specificity	Testing	81.11% ± 2.98%	88.40% ± 2.21%	90.12% ± 2.03%
	Validation	21.25% ± 2.21%	75.56% ± 1.60%	75.40% ± 3.59%

4. CONCLUSION AND FUTURE WORK

This is the first study implementing MFCC in achieving AMD bio-marker images detection. We show that MCFF is suited for distinguishing the pathological shape of RPE layer in AMD images from healthy images. We achieve comparable classification result using our testing dataset to Albarrak *et al.*[4] (91.4% accuracy, 92.4% sensitivity and 90.5% specificity) and Kuwayama *et al.*[7] (77% sensitivity). We also attain generalizable results on four new patients OCT image dataset with 70.11% overall accuracy . As future work, we plan to improve our classification performance by including more patient data in the training set. We only considered a single visit image dataset in this study, therefore, to fully take advantage of the time information in our data sets, we will also use MFCC to build an AMD prediction model.

ACKNOWLEDGMENTS

This study is partly supported by NSF IIS-1659836 and DePaul CDM 2019 PhD summer research stipend. We would like to thank Hee Eun Lee and Gianna Marie Dingillo for helping collecting the data.

REFERENCES

[1] Ma, J., Desai, R., Nesper, P., Gill, M., Fawzi, A., and Skondra, D., "Optical coherence tomographic angiography imaging in age-related macular degeneration," *Ophthalmology and eye diseases* **9**, 1179172116686075 (2017).

[2] Gess, A. J., Fung, A. E., and Rodriguez, J. G., "Imaging in neovascular age-related macular degeneration," in [*Seminars in ophthalmology*], **26**(3), 225–233, Taylor & Francis (2011).

[3] Farsiu, S., Chiu, S. J., O'Connell, R. V., Folgar, F. A., Yuan, E., Izatt, J. A., Toth, C. A., Group, A.-R. E. D. S. . A. S. D. O. C. T. S., et al., "Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography," *Ophthalmology* **121**(1), 162–172 (2014).

[4] Albarrak, A., Coenen, F., Zheng, Y., et al., "Age-related macular degeneration identification in volumetric optical coherence tomography using decomposition and local feature extraction," in [*Proceedings of 2013 international conference on medical image, understanding and analysis*], 59–64 (2013).

[5] Wang, Y., Zhang, Y., Yao, Z., Zhao, R., and Zhou, F., "Machine learning based detection of age-related macular degeneration (amd) and diabetic macular edema (dme) from optical coherence tomography (oct) images," *Biomedical optics express* **7**(12), 4928–4940 (2016).

[6] Treder, M., Lauermann, J. L., and Eter, N., "Automated detection of exudative age-related macular degeneration in spectral domain optical coherence tomography using deep learning," *Graefe's Archive for Clinical and Experimental Ophthalmology* **256**(2), 259–265 (2018).

[7] Kuwayama, S., Ayatsuka, Y., Yanagisono, D., Uta, T., Usui, H., Kato, A., Takase, N., Ogura, Y., and Yasukawa, T., "Automated detection of macular diseases by optical coherence tomography and artificial intelligence machine learning of optical coherence tomography images," *Journal of ophthalmology* **2019** (2019).

[8] Ganchev, T., Fakotakis, N., and Kokkinakis, G., "Comparative evaluation of various mfcc implementations on the speaker verification task," in [*Proceedings of the SPECOM*], **1**(2005), 191–194 (2005).

[9] Kent, R. D. and Read, C., "Acoustic analysis of speech. 2nd," *Albany, NY: Singular/Thomson Learning* (2002).

[10] Baranwal, N. and Nandi, G. C., "An efficient gesture based humanoid learning using wavelet descriptor and mfcc techniques," *International Journal of Machine Learning and Cybernetics* **8**(4), 1369–1388 (2017).

[11] Nogueira, D. M., Ferreira, C. A., Gomes, E. F., and Jorge, A. M., "Classifying heart sounds using images of motifs, mfcc and temporal features," *Journal of medical systems* **43**(6), 168 (2019).

[12] Lalitha, S., Geyasruti, D., Narayanan, R., and Shravani, M., "Emotion detection using mfcc and cepstrum features," *Procedia Computer Science* **70**, 29–35 (2015).

[13] Wang, Y., Soetikno, B., Furst, J., Raicu, D., and Fawzi, A. A., "Drusen diagnosis comparison between hyper-spectral and color retinal images," *Biomedical optics express* **10**(2), 914–931 (2019).

[14] Fabritius, T., Makita, S., Miura, M., Myllylä, R., and Yasuno, Y., "Automated segmentation of the macula by optical coherence tomography," *Optics express* **17**(18), 15659–15669 (2009).

[15] Scarinci, F., Jampol, L. M., Linsenmeier, R. A., and Fawzi, A. A., "Association of diabetic macular nonperfusion with outer retinal disruption on optical coherence tomography," *JAMA ophthalmology* **133**(9), 1036–1044 (2015).

[16] Schafer, R. W. and Oppenheim, A. V., [*Discrete-time signal processing*], Prentice Hall Englewood Cliffs, NJ (1989).

[17] Lyon, D. A., "The discrete fourier transform, part 4: spectral leakage," *Journal of object technology* **8**(7) (2009).

[18] Welch, P., "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on audio and electroacoustics* **15**(2), 70–73 (1967).

[19] Khayam, S. A., "The discrete cosine transform (dct): theory and application," *Michigan State University* **114**, 1–31 (2003).

[20] Haralick, R. M., Shanmugam, K., et al., "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics* (6), 610–621 (1973).

[21] Breiman, L., "Random forests," *Machine learning* **45**(1), 5–32 (2001).

[22] Welch, B. L., "The generalization ofstudent's' problem when several different population variances are involved," *Biometrika* **34**(1/2), 28–35 (1947).