AFRL-RH-FS-TR-2023-0007

Analysis of Synthetic Ronchi and ZPC Datasets in Support of Distortion Computations (DisCo)

Robin J. Allen
University of Georgia

Brenda J. Novar
711th Human Performance Wing
Airman Systems Directorate
Bioeffects Division
Optical Radiation Bioeffects Branch

20 March 2023

Interim - August 2022 - March 2023

Air Force Research Laboratory
711th Human Performance Wing
Airman Systems Directorate
Bioeffects Division
Optical Radiation Bioeffects Branch
JBSA Fort Sam Houston, Texas 78234

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporations; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the AFRL Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

"Analysis of Synthetic Ronchi and ZPC Datasets in Support of Distortion Computations (DisCo)"

(AFRL-RH-FS-TR- 2023-) has been reviewed and is approved for publication in accordance with assigned distribution statement.

FERRIS.LYNDSEY.MARIE.1381070391 Digitally signed by FERRIS.LYNDSEY.MARIE.1381070391 Date: 2023.03.28 11:33:30 -05'00'

LYNDSEY M. FERRIS, Lt Col, USAF, BSC
Chief, Optical Radiation Bioeffects Branch

MILLER.STEPHANIE.A.1230536283 Digitally signed by MILLER.STEPHANIE.A.1230536283 Date: 2023.04.18 12:53:22 -05'00'

STEPHANIE A. MILLER, DR-IV, DAF
Chief, Bioeffects Division
Airman Systems Directorate
711th Human Performance Wing
Air Force Research Laboratory

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute an official position of the U.S. Government.

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
		START DATE	END DATE
20-MAR-2023	Interim	AUG 2022	MAR 2023

4. TITLE AND SUBTITLE

Analysis of Synthetic Ronchi and ZPC Datasets in Support of Distortion Computations (DisCo)

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
FA8650-19-C-6098		
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**
		H14B

6. AUTHOR(S)

Robin J. Allen, Brenda J. Novar

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
711th Human Performance Wing
Bioeffects Division Optical Radiation Bioeffects Branch
4141 Petroleum Dr
JBSA Fort Sam Houston TX, 78234

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Air Force Research Laboratory 711th Human Performance Wing Airman Systems Directorate Bioeffects Division Optical Radiation Bioeffects Branch	711 HPW/RHDO	AFRL-RH-FS-TR-2023-0007

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited. Cleared: AFRL PA Case Number: AFRL-2023-0423. The views expressed are those of the author and do not necessarily reflect the official policy or position of the Department of the Air Force, the Department of Defense, or the United States Government.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report discusses analysis that was conducted during the summer of FY22 in support of the 711 HPW/RHDO modeling project called Distortion Computations of Optics (DisCo), a machine learning effort currently under collaborative development with Applied Research Associates (ARA). Two methods have been developed to evaluate optical distortion patterns observed through transparent eyewear. One such method is the use of wavefront analysis to determine wavefront errors; the other is the use of an optical system that employs a fixed-frequency Ronchi grating to provide a quick visual evaluation, as per the MIL-DTL-43511D.

15. SUBJECT TERMS

Machine learning, Zernike polynomials, Zernike coefficients, Convolutional Neural Networks,fringe tracing, aberrations, optical distortion

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	c. THIS PAGE		
U	U	U	U	34

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(include area code)*
Brenda Novar	(210) 539-7215

This Page Intentionally Left Blank

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the following individuals for providing valuable technical expertise during the development of this report: Ms. Novar (DisCo Primary Investigator), Dr. Allen Harvey (Machine Learning SME) and Mr. Gary Noojin (SAIC).

Funding for this effort was provided by the Oak Ridge Institute for Science and Education (ORISE).

1.0 SUMMARY

The optical quality of United States Air Force (USAF) developed eyewear is determined by a number of optical quality parameters. Prototype eyewear can have optical quality issues related to levels of haze, transmittance, color shifts, and localized defects such as scratches, surface blemishes or coating imperfections which can potentially impact visual quality for the wearer. One parameter that is hard to quantify succinctly, is optical distortion. This report discusses analysis that was conducted in support of the 711th HPW/RHD modeling project called Distortion Computations of Optics (DisCo), a machine learning effort currently under collaborative development with Applied Research Associates (ARA). Two methods have been developed to evaluate optical distortion patterns observed through transparent eyewear. One such method is the use of wavefront analysis to determine wavefront errors; the other is the use of an optical system that employs a fixed-frequency Ronchi grating to provide a quick visual evaluation, as per the MIL-DTL-43511D. The use of a wavefront sensor allows the evaluator to decompose the distorted wavefront into lower and higher order coefficients that represent the primary distortion components. The use of a Ronchi grating allows the evaluator to visually assess the shape and orientation of fringe aberrations. Machine learning algorithms and computer vision metrics can be used to quickly determine the acceptability of these patterns. Zernike polynomial fitting algorithms have been developed to quickly fit images of Ronchigrams and wavefronts in order to analyze the coefficients and map to aberrations in lenses.

2.0 INTRODUCTION

During flight and ground operations aircrews may employ various types of eyewear to protect and aid their vision. The United States Air Force (USAF) developed eyewear must be measured for specific optical quality tests, as outlined in the MIL-DTL-43511D. All of the tests are provided with measurable metrics, with exception of optical distortion, which is evaluated using a modified Ann Arbor tester (Model C Ronchi-type Optical Tester) - Ref 1. The results of this test provide the evaluator with a visual observation image referred to as a Ronchigram. The MIL-DTL-43511D provides two categories of images that were deemed acceptable or unacceptable. Each Ronchigram contains light and dark fringes, of which the amount, shape, and orientation contribute to the classification as acceptable or unacceptable. The standard relies on subjective analysis of those performing physical evaluations and can lead to discrepancies in classifications of the given eyewear depending on an individual‘s tolerance. Elimination of these discrepancies can be achieved through the use of computer vision and machine learning techniques.

The following contains analysis that builds upon research that has been completed by RHDO in support of the Distortion Computations project which established data set of synthetically generated Ronchigrams. Each of these synthetic Ronchigrams is associated with image derived metrics and Zernike polynomial coefficients.

2.1 Image Metrics

To quantitatively describe fringes contained within the synthetic Ronchigrams, several Python libraries have been developed to capture the shape and orientation of each fringe. Polygons were fit to each fringe in every Ronchigram and various descriptor values were extracted. These descriptor values included number of fringes, perimeter, area, shape, along with other values to be used in

defining image based metrics. The Ronchigram data set that the image metrics were calculated for included 699 acceptable instances and 2869 unacceptable instances. The difference in the number of acceptable versus unacceptable instances is due to the vast amount of characteristics that can render a Ronchigram unacceptable.

The descriptor values were stored within a large data set, each descriptor contained a value for each fringe. This presented an issue in which Ronchigrams had data of different dimensions dependent on the number of fringes. To resolve this, when image metrics were calculated, minimum or maximum values were assigned to a specific instance of the Ronchigram, thereby reducing the dimensionality of the data. These considerations will be addressed in the discussion of each metric. Development of these metrics are described in detail in the technical report, AFRL-RH-FS-TR-2022-0011.

2.2 Classification

The synthetic Ronchigram data consists of two classes, acceptable and unacceptable, which requires a binary classification algorithm. Convolutional Neural Networks (CNNs) are a machine learning tool that has the capability to handle such a task. The CNN developed to handle this task was created using the Tensorflow package in Python 3.8.

The CNN was developed in order to determine how well the ML algorithms could be trained to: use the synthetic Ronchigram data set, compare the characteristics of Ronchigrams that are misidentified by both the sorting method and ML approach, and determine characteristics of synthetic Ronchigrams that are difficult to classify.

The CNN discussed in this report is separate from the machine learning methods used in DisCo. DisCo is the machine learning effort developed to evaluate the acceptability of Ronchigrams and includes both metric based models and computer vision based models. The CNN discussed in this report was created to sort out the Ronchigrams that are considered borderline cases for further analysis.

2.3 Zernike Polynomial Coefficients

Another method for determining the acceptability of eyewear using Zernike polynomial coefficients (ZPCs) is currently being studied. The ZPCs can be used to quantify lens aberrations and further applied to optical distortion. In this work, similar methods from the analysis of the image based metrics were implemented on ZPC data obtained from the same set of Ronchigrams used to generate the image based metrics. These sets provide a means to validate the method and ensure no information is lost when analyzing a wavefront rather than an interferogram.

3.0 IMAGE METRICS

3.1 Perimeter

A distorted fringe will produce irregularly shaped polygon fits with irregular sides that cause a larger perimeter per polygon when compared to a fringe from an acceptable Ronchigram. The perimeter per fringe metric was calculated by dividing the total perimeter of the polygon fits in a given Ronchigram by the number of fringes in that Ronchigram. However, as shown in Table 1

below, the average values for the perimeter per fringe for both cases is nearly identical. Although the perimeter per fringe for the unacceptable cases appears to have a higher spread, there are few instances that deviate greatly from the average. A threshold was set at the maximum acceptable perimeter per fringe value. Unacceptable instances were then sorted according to this value, which eliminated 93 instances from the data. This yield was low, thus the perimeter per fringe is not the best metric to use in distinguishing acceptable from unacceptable.

Descriptor	Acceptable	Unacceptable
Max Perimeter/Fringe	1392.132	4440.308
Min Perimeter/Fringe	1052.4 90	546.334
Avg. Perimeter/Fringe	1171.986	1164.694
Max Perimeter St. Deviation	472.746	223489.086
Min Perimeter St. Deviation	196.446	135.718
Avg. Perimeter St. Deviation	322.245	437.271

Table 1. Descriptive values of the perimeter metric

Due to the low yield of the perimeter per fringe metric a new metric was defined using the perimeter values. Another indicator of irregular fitted polygons is a large difference in the perimeter of each consecutive fringe in the Ronchigram which occurs in non-symmetric cases. To study this the standard deviation of the perimeters of the fringes in each Ronchigram was calculated. Since the Ronchigram is a circular image some spread in the standard deviation value is expected for the acceptable cases. However, for images with oddly shaped fringes this spread should be larger than that of the images with acceptable fringes. Another case that this method can handle is images that contain fringes that are cut off at the edge of the photo. Taking the maximum value of the standard deviation of the acceptable instances as a threshold value and then sorting the unacceptable cases eliminates 703 instances of the 2869 instances from the unacceptable data without losing any of the acceptable instances. This corresponds to 24.5 % of the unacceptable instances. Included in these eliminated instances are 80 of the 93 that would be sorted out via the perimeter per fringe method. The first 100 of these instances that are sorted out are shown below in Figure 1.

3.2 Apparent Area Ratio

Another method for determining the shape of fringes within a Ronchigram is the apparent area ratio metric. The apparent area ratio considers the ratio of a rectangle formed by the filling the minimum and maximum spatial components of the fringe and the area of the polygon fit to the same fringe. Fringes that are not distorted have values close to one, while distorted fringes have increased values.

To reduce the amount of descriptor values of each Ronchigram to one, the maximum apparent area ratio was taken, as images containing at least one distorted fringe are rendered unacceptable.

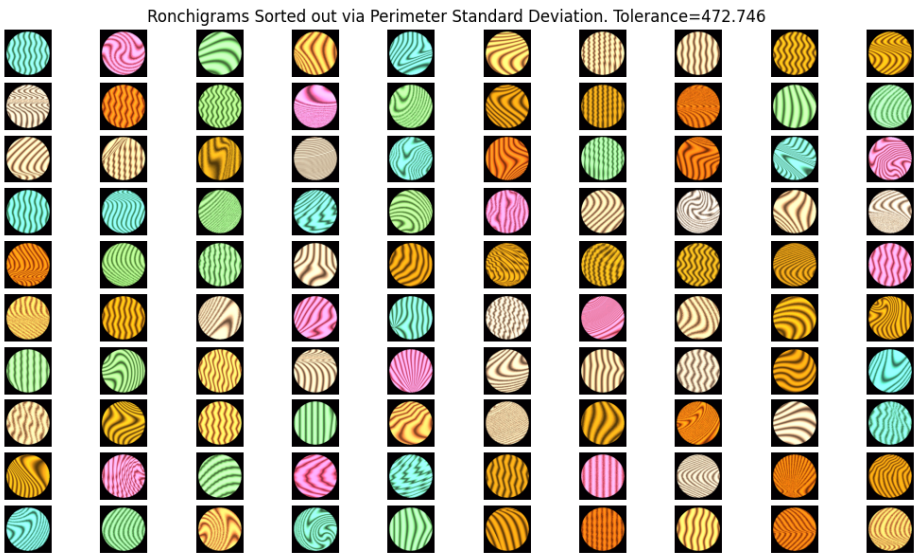

Figure 1. Sample Ronchigrams sorted out using the Perimeter Standard Deviation Method

Descriptor	Acceptable	Unacceptable
Minimum Apparent Area Ratio	1.100	1.197
Maximum Apparent Area Ratio	10.678	81.727
Avg. Apparent Area Ratio	1.501	6.026

Table 2. Descriptive values of the apparent area ratio metric

This led to a potential issue with images containing fringes cut off at the edge, however this only presented itself in a single instance. The minimum, maximum, and average value of the maximum apparent area ratio per Ronchigram is shown in Table 2. There is overlap between the two categories, but a significant amount of the unacceptable cases can be sorted out. If a threshold is set at the maximum value for the acceptable cases (10.678), only 451 unacceptable instances are sorted out (of which all were sorted via deflection method discussed later). This corresponds to 15.7 % of the unacceptable instances. Lowering the threshold to 4.0 led to sorting 1478 unacceptable instances out, while dropping only 4 acceptable instances. This eliminated 51.5 % of the unacceptable cases.

3.3 Convex Area Ratio

Tab Similar to the apparent area ratio, the convex area ratio considers the shape of the fringes within a Ronchigram. The difference between the two is the convex area ratio uses the area of a convex hull around the polygon fit to the fringe. As before, acceptable fringes have values near 1, while unacceptable values tend higher. Due to this, the maximum value of the convex area ratio

in each Ronchigram was again considered for this metric. Some descriptive values for the convex area ratio are given in Table 3. Utilizing a threshold of 1.790, the maximum of the acceptable cases led to sorting out 1168 unacceptable cases which is 40.7 % of the unacceptable cases.

Descriptor	Acceptable	Unacceptable
Minimum Convex Area Ratio	1.0	1.0
Maximum Convex Area Ratio	1.790	13.148
Avg. Convex Area Ratio	1.065	2.092

Table 3. Descriptive values of the convex area ratio metric

3.4 Angle From Nominal

Fringes that are oriented off the vertical axis also lead to unacceptable classification. A perfectly vertical oriented fringe will have a theta value of 90.0 (corresponding to a right angle), while tilted fringes can have a range of values above and below the aforementioned nominal value. This is caused by fringes being canted to both right and left sides, and in some cases appearing nearly inverted.

Since a tilt to either side of a nominally aligned fringe causes unacceptable classifications it is better to consider the angle from nominal (AFN). The AFN is simply the absolute value of difference between a given fringes theta value and the nominal value of 90.0. Since each Ronchigram contains multiple fringes the maximum and minimum value of theta were selected for each. The AFN of these values was calculated and the larger of the two was stored as the AFN for the selected Ronchigram. This consideration, like before, is to reduce the number of parameters per Ronchigram. Additionally it selects instances in which there are only a few tilted fringes, but not all. For acceptable cases AFN values are near 0, while for unacceptable cases they can span up to 90.0. Descriptive values for the AFN are listed in Table 4.

Descriptor	Acceptable	Unacceptable
Minimum AFN	0.119	0.109
Maximum AFN	16.546	89.965
Avg. AFN	2.321	26.337

Table 4. Descriptive values of the angle from nominal metric

When the threshold is set to the maximum AFN value of the acceptable instances (16.546), 1388 unacceptable cases are sorted out. With this threshold no acceptable cases are dropped. This eliminated 48.3 % of the unacceptable cases.

3.5 Deflection

Another method to measure the distortion within a Ronchigram is the deflection. The deflection metric compares the area of the polygon fit of a fringe to an ideal fringe. The ideal fringe is defined at the horizontal center of each distorted fringe and has the same area as the distorted fringe. Since this metric is defined as the ratio of the area of the actual fringe that is outside of the ideal fringe, nominal values are near 0.0 whereas highly distorted fringes have values near 1.0.

Similar to the previous metrics, the deflection metric is calculated for each fringe, giving each Ronchigram multiple values. As before, a single fringe being distorted is enough to render a classification of unacceptable, therefore the maximum value can be used to describe the Ronchigram as a whole.

Unlike before, using the maximum value of the acceptable cases does not provide a good threshold despite the clear separation of peaks in the histogram, Figure 2 (Top). As also shown in figure 2 (Bottom) the number of unacceptable instances that are sorted out drops drastically near the maximum deflection value.

When the threshold is chosen to be 0.4, 2226 of the unacceptable cases were sorted out which eliminates the most unacceptable cases at 77.5 %. This comes at the cost of dropping 61 acceptable cases, which is approximately 9 % of acceptable cases. A different threshold can be selected dependent on tolerance for dropping acceptable cases, however the steep drop off of unacceptable cases near a threshold of 0.8 should be noted.

3.6 Summary

When the synthetic Ronchigrams were sorted based upon the image based metrics overlap between the images sorted out occurred often. This behavior substantiates a given Ronchigrams classification of unacceptable as multiple indicators of distortion are present. These results are summarized in table 5.

	Perimeter Std.	AFN	Apparent	Convex	Deflection
Perimeter Std.	703	489	178	484	699
AFN	489	1388	432	908	1388
Apparent	178	432	451	416	451
Convex	484	908	416	1168	1166
Deflection	699	1388	451	1166	2226

Table 5. Overlap between each of the image based metrics

As shown in table 5 the deflection metric shows the greatest overlap with each of the other metrics. However, this comes at the cost of dropping approximately 9 % of the acceptable cases.

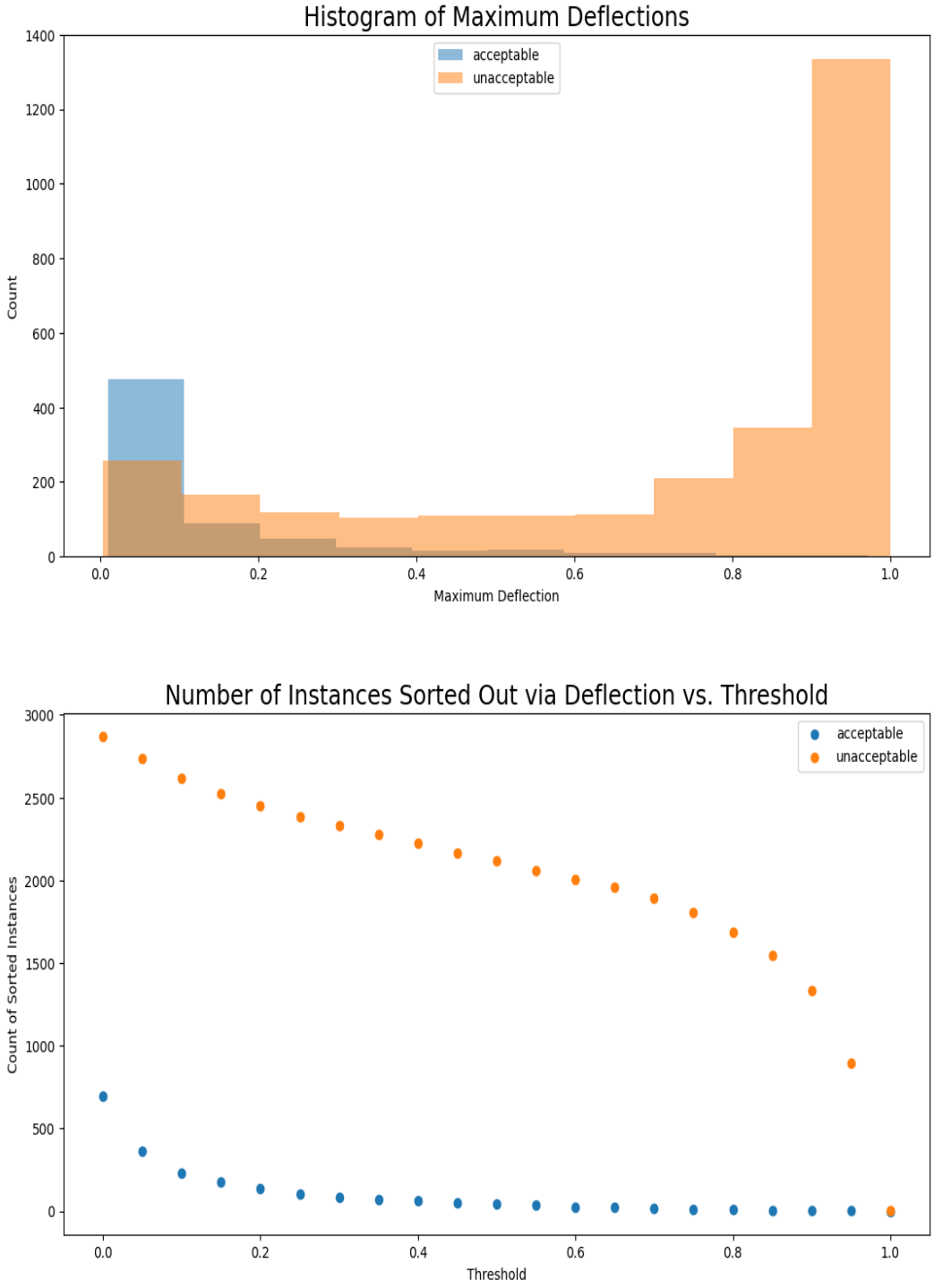

Figure 2. (Top) Histogram of Maximum Deflections (Bottom) and Number of Instances Sorted out vs. Maximum Deflection Threshold

4.0 CLASSIFICATION USING A CONVOLUTIONAL NEURAL NETWORK

4.1 Background

The synthetic Ronchigrams are images that can be read into Python as three dimensional arrays, making a convolutional neural network a clear candidate to perform the classification (LeCun, Bottou, Bengio, Haffner, 1998). To make classifications CNN's employ convolutional and pooling layers for feature extraction. These features are stored in new three dimensional arrays called feature maps until sufficient information has been extracted from the image and subsequent feature maps. The final feature maps are then flattened into one single fully connected layer to make the final classification. The basic structure of a CNN is shown in Figure 3.

4.2 Data Management

To train and test the CNN the data was split into three groups: Training, Validation, and Test. Training and Validation sets were used in the training epochs, that is a round of training. After the CNN attained reasonable accuracy training was stopped. To determine a reasonable point to stop training the categorical accuracy of the validation set was monitored. If this value did not improve within 5 training epochs then training would be stopped and the parameters from the best epoch were restored. The network was then applied to the test data.The training set was the largest containing 477 acceptable images and 2,192 unacceptable images. The validation data contained 65 acceptable and 320 unacceptable images. The test set contained 95 acceptable and 359 unacceptable images.

The unacceptable data had many more instances than the acceptable data. When machine learning algorithms are exposed to heavily one sided data sets there is the potential to over train to the dominant side, in this case the unacceptable classification. This manifested itself predominantly in the first few training epochs by the classification of all images as unacceptable, which would cause a categorical accuracy near 80 %.

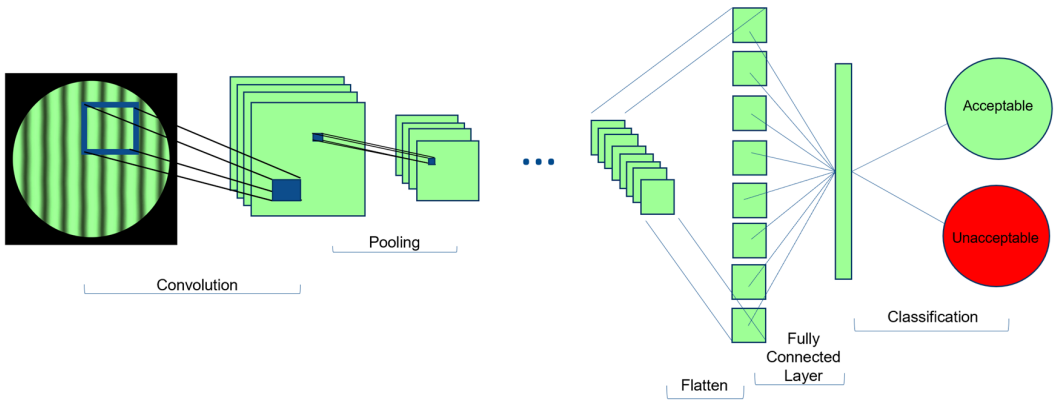

Figure 3. Schematic of a general CNN that classifies synthetic Ronchigrams

One way to combat the uneven amounts of data was to reduce the number of unacceptable cases to match that of the acceptable. This did prove feasible attaining categorical accuracies in the low 90‘s. However, this method had some drawbacks. The first was that the total data size decreased, meaning the CNN had less total instances to train itself. This made the CNN sensitive to which unacceptable instances that were selected. Another drawback was that the network on occasion would not converge in a reasonable amount of training epochs, and would trend towards the same accuracy as a random coin flip.

Another method for dealing with the mismatched data amounts was to input an early stopping step. This step served two purposes. First, to ensure that CNN continued to improve over the training epochs and second, to ensure the CNN did not over fit itself to this specific data set. This method was able to consistently attain categorical accuracies over 90 %, at times up to 96 % without ignoring any of the data.

Each synthetic Ronchigram was 756x756 pixels and there was a total of 3,508 images used in training and testing the CNN. The images are interpreted by the computer as a three dimensional array, storing the location of the pixel (horizontal and vertical locations) and the pixel value associated with the color of the pixel.This large amount of data required a significant amount of time to import and then eventually run on a standalone computer. Under the advice of the machine learning subject matter expert at ARA, the image size was reduced too 200x200 pixels. Additionally the pixel values were mapped from Red, Green, and Blue (RGB) to gray-scale values. Gray-scale images were employed for two main reasons. First, to reduce the amount of information to be stored about an image to make the CNN more efficient and second, to prevent the CNN from learning colors rather than fringe shapes and orientations. The latter consideration can be made as the color of a given Ronchigram does not impact the acceptability determination. This reduction in size did not alter the accuracy of the CNN and increased the speed the algorithm ran significantly. It is important to note that the images were not compressed until they were brought into the algorithm, meaning they were stored at full size.

4.3 Structure of CNN

Some classification tasks require vast networks with intricate structures which in turn need large amounts of computing power and time. The synthetic Ronchigram data set did not require a deep network which means sufficient information can be extracted from the images within a few layers without an intricate structure of the CNN. The structure of this network and the sizes of the intermediate feature maps is shown in Figure 4, and detailed below.

Four convolutional layers were used in the network to classify the synthetic Ronchigrams. Each Convolution layer used a filter, also known as a kernel that was moved across the image and compared each pixel value to that of the filter and then mapped to a feature map according to the activation function of that layer. These filters were also two dimensional arrays, however they were smaller than the image or feature map they were applied to, but every pixel in the filter maps to a new pixel in the following feature map. The convolution layers do not reduce the size of the image.

A pooling layer was added in between each convolutional layer in the network. The pooling layer served to reduce the size of the feature maps in between convolutional layers by moving

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 200, 200, 4)       40

 max_pooling2d (MaxPooling2D  (None, 40, 40, 4)        0
 )

 conv2d_1 (Conv2D)           (None, 40, 40, 5)         185

 max_pooling2d_1 (MaxPooling  (None, 8, 8, 5)          0
 2D)

 conv2d_2 (Conv2D)           (None, 8, 8, 6)           276

 max_pooling2d_2 (MaxPooling  (None, 1, 1, 6)          0
 2D)

 conv2d_3 (Conv2D)           (None, 1, 1, 7)           385

 flatten (Flatten)           (None, 7)                 0

 dense (Dense)               (None, 2)                 16

=================================================================
Total params: 902
Trainable params: 902
Non-trainable params: 0
_________________________________________________________________
```

Figure 4. Structure of the CNN used in this section

filters across the preceding feature map and generating a new feature map. The new feature map is smaller as the filter returns a single value for each location it analyzes in the preceding feature map rather than a value for every pixel within the filter. In this network only Max Pooling was employed, which means the filters returned only the maximum pixel value contained within the filter and mapped this to a new pixel in the new feature map.

The final component of the CNN flattened the feature maps and applied a fully connected layer. This layer is where the classification is made and the probabilities of being acceptable and unacceptable are calculated. This probability determines the final classification of an image by the network.

4.4 Training the CNN and Application to Validation Data

The training data was used to update the parameters within the layers of the CNN during the training process. After these parameters were updated, the network was then applied to the validation set and the categorical accuracy stored. The model would continue to train until the categorical accuracy of the validation set did not improve in the preceding five training epochs. When this condition was met, the training stopped and the best parameters within the layers were restored. The parameters which returned the best categorical accuracy on the validation set were

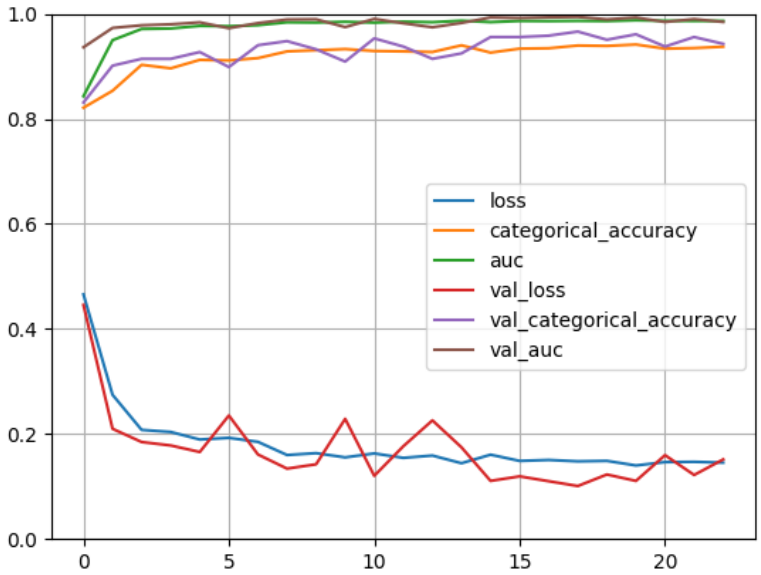

Figure 5. Sample learning curves from CNN

then reset as the default parameters for the network before continuing.

In addition to the early stopping method mentioned above, learning curves, Precision-Recall (PR) curves, and Receiver Operating Characteristic (ROC) curves were monitored. These plots were generated using the Sci-Kit Learn Python package (Pedregosa et al.).

The learning curves plotted the categorical accuracy, area under the curve (AUC) of the ROC plot, and the loss function value for both the training data and the validation data. This plot was monitored to ensure both loss functions reached a minimal plateau, and that the categorical accuracy and AUC values were maximized near 1.0. A sample learning curve is shown in Figure 5, which was generated during a trial to determine the characteristics of the synthetic Ronchigrams which are difficult to classify.

The PR curves were generated as another measure to monitor the learning of the CNN. This plot shows the trade-off between the precision and the recall, which compares how well the CNN performs in classification to how many of each classification it is able to find. A CNN that is well trained will have an area under the curve that is near one. This presents itself on the plot as a horizontal line near the top of the plot that turns sharply to a vertical line in the top right corner. An example PR curve from the CNN is shown in figure 6 (Top).

Similar to the PR curve the ROC curve was also analyzed to ensure the CNN performed well. In this plot the true positive rate was compared to the false positive rate. Well trained classifiers have ROC curves that are vertical along the left side of the plot and turn sharply to horizontal in the top left corner of the plot. As the turn of the curve gets closer to the top left corner the classifier gets

better. An example ROC from the CNN is shown in figure 6 (Bottom).

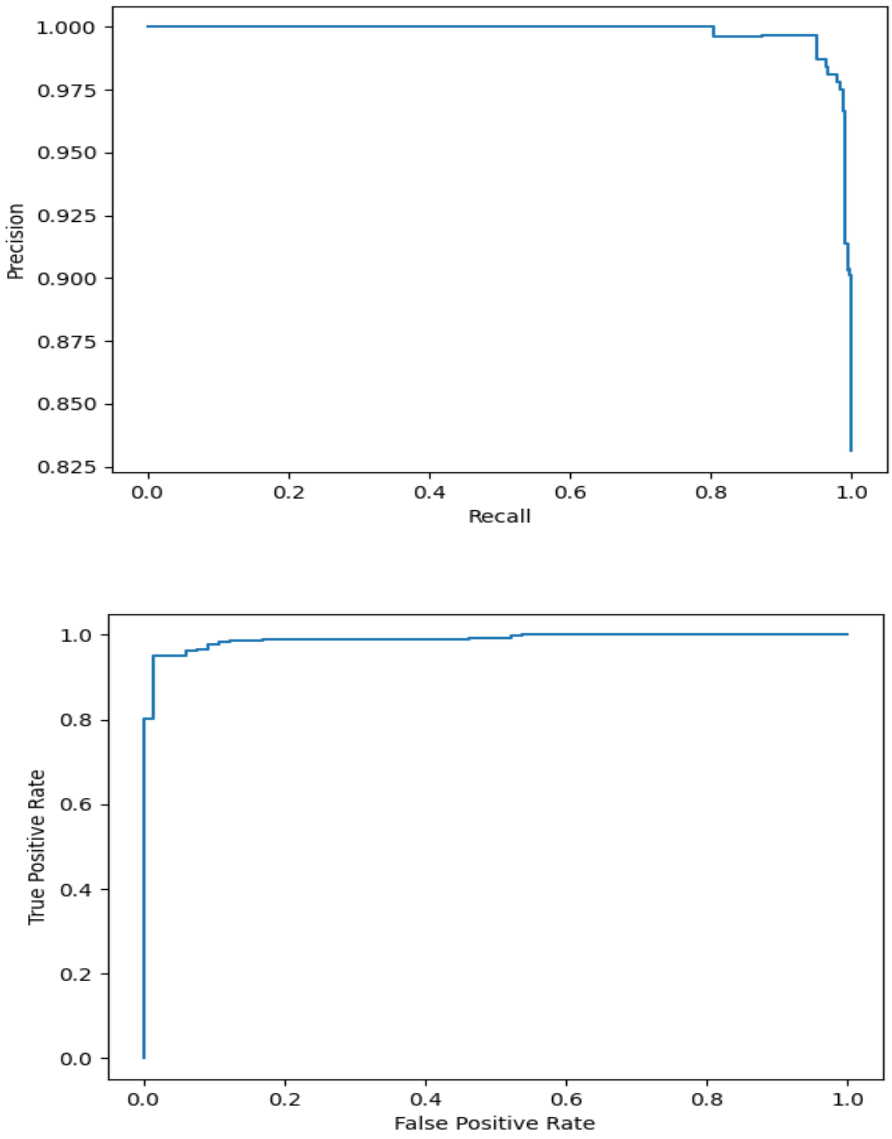

Figure 6: (Top) Precision Recall (PR) and (Bottom) Receiver Operating Characteristic (ROC) curves

4.5 Performance

Once the trained network was established it was applied to the validation data once again. This was done as an additional check to ensure that the high accuracy was not attained by random chance. Following this a confusion matrix and classification report would be generated (again using the scikit-learn package) to report the precision and recall for each class. The confusion matrix is shown in Figure 7. The classification report and confusion matrix were analyzed to determine if the network had any bias towards either of the classifications. In general the network would have slightly better accuracy when categorizing a truly unacceptable Ronchigram as compared to the

acceptable, though the difference was typically within a few percent. This difference in accuracy can be attributed to the class imbalance between the acceptable and unacceptable classes, and the presence of borderline instances in the training data.

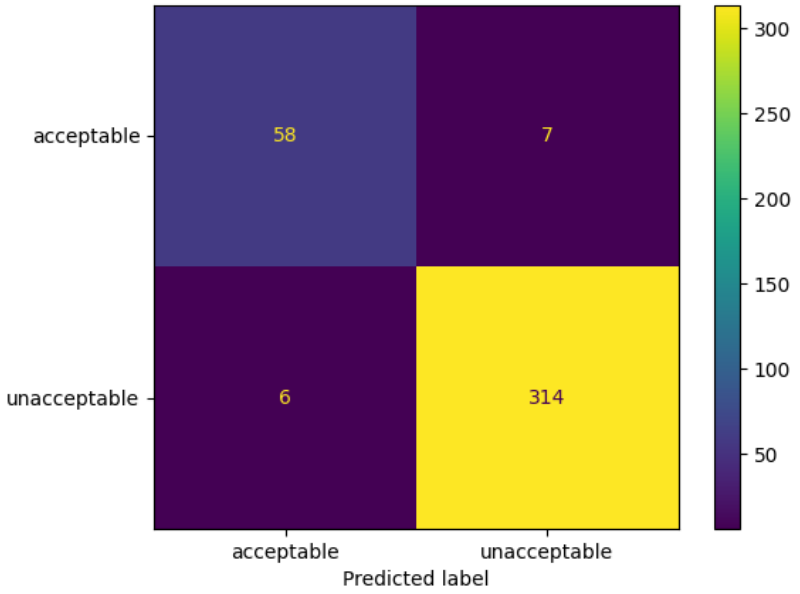

Figure 7. Sample Confusion Matrix generated from CNN results

To ensure the CNN was not over fitting to the unacceptable classification the misidentified instances in the validation set were plotted and are shown in Figure 8. The Ronchigrams that the CNN misidentified were typically comprised of the aforementioned borderline cases, meaning that the CNN had approached the maximum accuracy it could attain without over fitting to the specific data set. In short, the CNN misclassified the instances that a trained individual would struggle to classify. The benefit of the CNN over a trained individual is the borderline cases can be reproduced, allowing for a subset of the Ronchigram data to be created for analysis.

The CNN was also applied to the test data which had been withheld until after the analysis of the validation data was completed. Similar to the validation data a classification report was generated and the misidentified Ronchigrams were plotted shown in Figure 9. As with the validation data a slight bias towards the unacceptable classification appeared. However, once again upon further analysis of the misidentified Ronchigrams the bias was attributed to the presence of borderline cases in the data.

The overall categorical accuracy on the test data set was typically within 1 % of that of the validation data, meaning the CNN generalized to new data well. It should be noted that it was in general slightly lower than that of the validation data, but rarely more than 1 % lower.

4.6 CNN Sorted Data set

One important result of creating the CNN to classify the synthetic Ronchigrams was to find the borderline cases and separate them for examination. To do this, after the CNN was regularly achieving accuracies over 94 %, a subset of the synthetic Ronchigram data set was taken from the instances the CNN classified incorrectly, as these were the borderline cases. The CNN was trained five separate times and the file names of the Ronchigrams incorrectly classified by the CNN were stored and the duplicates removed. The training, validation, and test set remained the same for all five runs. The difference in each run was the initial conditions of the filters within the convolution layers. As a result of the different initial filters applied the number of training epochs changed from one network to the next. Despite this, the categorical accuracy did not vary by more than a few percent. The CNN sorted data set contained 98 instances of the total 3,508 Ronchigrams used to train and evaluate the CNN.

After the misidentified instances were stored similar analysis from section 3 was performed on the subset of the data. This was done to determine the overlap between the image metric analysis method and the computer vision method. The CNN out performed each of the image metric methods on their own. This behavior was expected since the image metrics were defined to capture a specific feature within the fringes and not the Ronchigram as a whole.

To further explore the effectiveness and overlap with the image metrics of the CNN, histograms of the image metrics were generated from the CNN sorted data set. Figure 10 shows the histogram of the convex area ratio and the apparent area ratio for the CNN sorted data, and highlights the overlap between the methods. The outlier was clearly unacceptable, having an excess of fringes that are compact and oriented improperly in the Ronchigram. This image is unrepresentative of real data and thus can be dropped from the CNN sorted data set. If this case were representative of a real Ronchigram then more images with similar spacings and orientation would need to be added to the training set in order to teach the CNN how to handle such cases.

Another advantage of the CNN over the image metric method was minimizing the number of acceptable instances that would be dropped from the acceptable classification. Metrics such as apparent area ratio and deflection require a threshold to be set lower than the maximum value of the acceptable cases, resulting in dropped acceptable instances that can be considered misidentified. The CNN takes the Ronchigram as a whole as the input and therefore can account for more relationships between the data features present. This led to the total number of acceptable cases that get misidentified in the CNN to be less than in the image metric method.

As shown in Figure 10, the overlap between acceptable and unacceptable cases in the CNN sorted data are indiscernible from one another. This overlap further illustrates the difficulty in classifying these cases as both the image metric method and CNN cannot adequately handle them. Moreover, any boundary or threshold set to determine acceptability will fall within this region, and will bias one classification. In order to set a quantitative metric to determine acceptability more information must be extracted from the Ronchigrams or a new method of testing eyewear must be developed.

Figure 8. Ronchigrams from the validation set that were miscategorized by the CNN

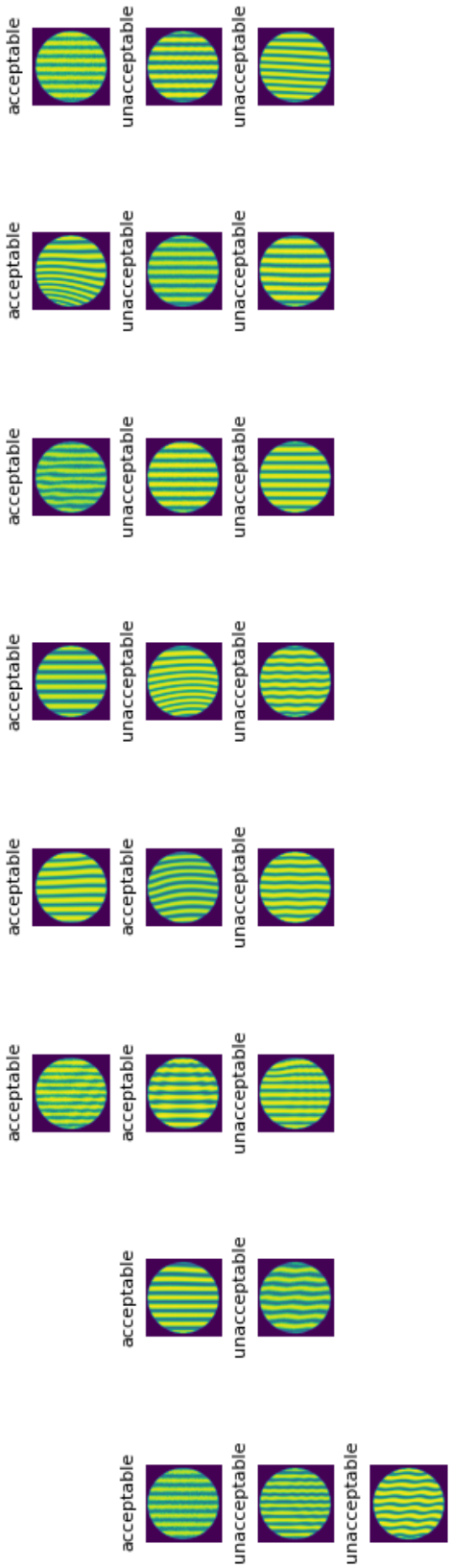

Figure 9. Ronchigrams from the test set that were miscategorized by the CNN

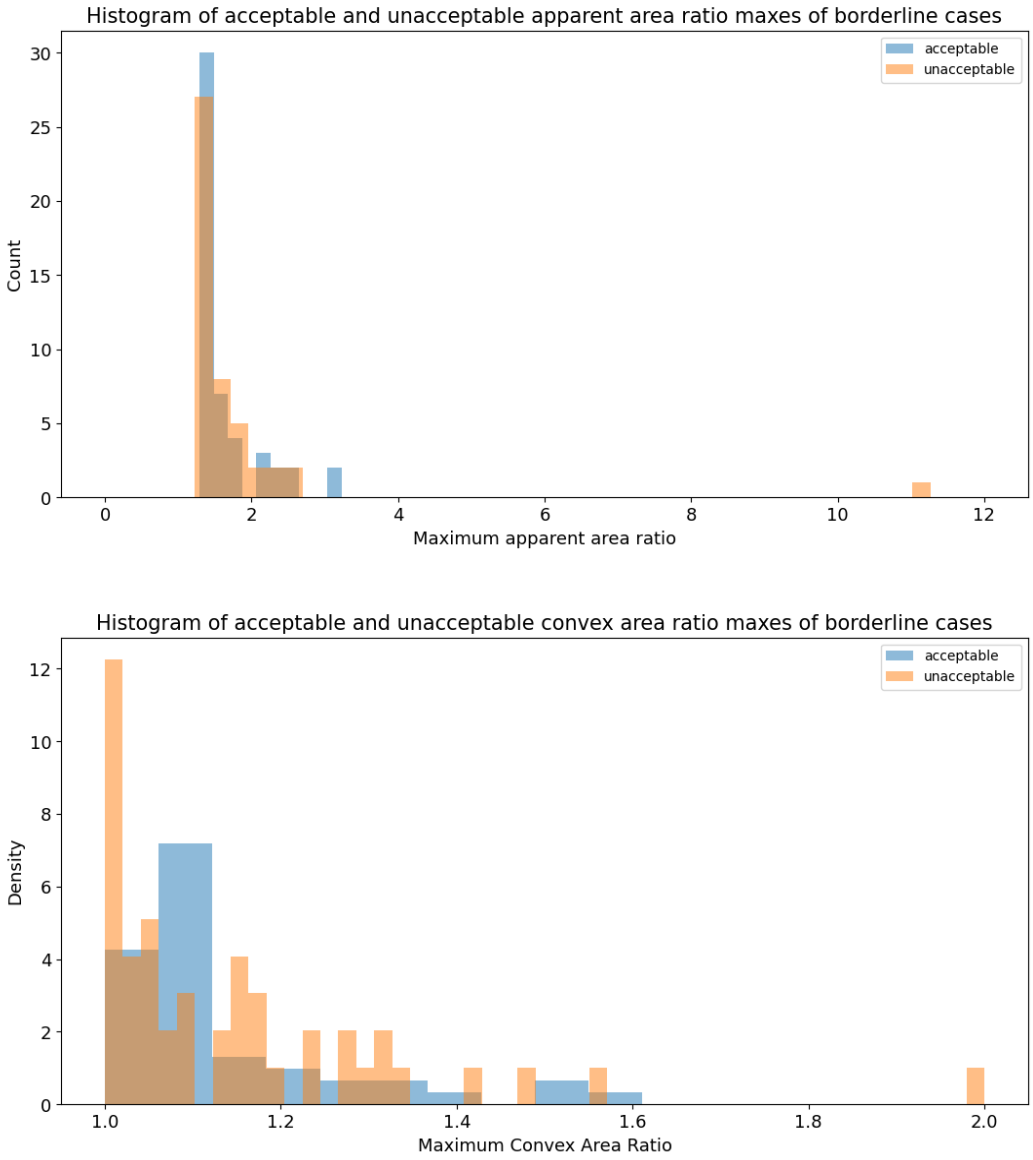

Figure 10. Histogram of convex area ratio of acceptable and unacceptable Ronchigrams from the borderline cases

5.0 ZERNIKE POLYNOMIALS

5.1 Background

Although the image metric analysis and CNN classification have proven effective in classifying Ronchigrams, they did not provided a clear metric to base acceptability on. Another way to quantify optical distortion is to examine the wavefront after it has passed through a sample. The wavefront can be mapped to Zernike polynomial coefficients (ZPCs) which has terms that match lens aberrations (Wyant Creath 1992). This furthers the effort towards the primary goal of DisCO, which is to provide a quantitative metric to determine the acceptability of AF developed eye wear. Accurately fit ZPCs can provide a metric for lens aberrations that cause optical distortion.

Two data sets containing ZPCs derived from the synthetic Ronchigrams by ARA were examined. These two data sets were generated using Fringe Tracing (FT) and Fast Fourier Transform (FFT) using the methods laid out in AFRL-RH-FS-TR-2022-0015 and AFRL-RH-FS-TR-2022-0017

5.2 Distribution of Zernike Polynomial Coefficients

The first of the two methods to fit ZPCs to the synthetic Ronchigram data employed a fringe tracing technique. This method was developed for use in Ronchigrams that contain higher amounts of noise. As shown in Figure 11 the distribution of the ZPCs is normal, with the exception of term 1 and term 2 in the acceptable instances which have the presence of right skew. Throughout all 49 terms the unacceptable ZPCs showed more spread, meaning thresholds could be set to sort the unacceptable instances out. The thresholds were set at the maximum and minimum values of each acceptable ZPC term. Figure 12 is the QQ-Plots for the first 2 terms. These plots emphasize the much larger range of values in the unacceptable data as well as point out the presence of outliers in the data.

The second method of ZPC fitting used Fast Fourier Transform (FFT). The histograms from the first four ZPCs are shown below in Figure 13. Term 2 has a bimodal distribution that is attributed to the FFT method checking the wavefront to determine if it is inverted. In such cases the algorithm inverts the wavefront to correct the values. Similar to the FT ZPCs, the unacceptable FFT ZPCs exhibited wider spread, allowing for thresholds to be set upon which a classification was made. These thresholds were taken to be the maximum and the minimum value of each acceptable ZPC term.

QQ-Plots for term 1 and term 4 FFT data are shown in Figure 14. Again the larger spread in the values of the unacceptable cases cause the widened shape in these plots. In the FFT data however, there are a lower number of outliers. With fewer outliers, the FFT ZPCs are believed to be closer to what actual wavefront measurements would yield.

5.3 Sorting Via ZPCs

As mentioned in section 5.2, the ZPCs from both fitting methods provided distinctions between the acceptable and unacceptable cases. To capture these differences and determine the the ability in which the ZPCs can be used to classify the synthetic Ronchigrams the maximum and minimum values for each term in the acceptable sets were calculated. Both the maximum and minimum values were used as thresholds as it allowed for better sorting of the unacceptable instances. It is

not sufficient to take the absolute value of the ZPCs and set a single threshold at the maximum value of the acceptable cases. This would result in some unacceptable cases not being sorted out as they would fall below the threshold. Additionally, taking the absolute value of a ZPC alters the actual wavefront that it is intended to describe.

Using the dual thresholds described above to sort out unacceptable cases provided excellent results. When sorting using the FT fit values for all 49 ZPC terms, 2547 of 2869 unacceptable cases were sorted out. Similarly, when using the FFT fit values for all 49 ZPC terms, 2524 of 2869 unacceptable cases were sorted out. This sorting method did not drop any of the acceptable cases like the deflection metric did when sorting out less unacceptable cases. These better results are likely a consequence of including 49 ZPC terms as compared to the 5 used in the image based metrics method.

Sorting via ZPCs also eliminated the majority of the same unacceptable cases that the image based metrics approach did. The large overlap between these two different methods shows that the ZPCs are able to capture the same optical distortion information from the lens that the image based metrics did. The benefit of the ZPC fits and wavefront data is the distortion is quantifiable and able to be sorted using a hard metric rather than a subjective test. The overlap with each metric is shown in Table 6.

	Overlap With ZPCs
Perimeter Std.	678
AFN	1383
Apparent	451
Convex	1168
Deflection	2156

Table 6. Overlap between the Ronchigrams sorted via the image based metric method and the ZPC method

5.4 ZPC Correlation Phenomena

As mentioned above the ZPCs were able to distinguish the acceptable cases from the unacceptable based upon a simple classification scheme (sorting method). To further investigate the effectiveness of the ZPC method in determining acceptability, the correlation values between every possible pair of terms were calculated. Since the ZPCs also describe lens aberrations, the correlations between the terms can aid in determining physical relationships between the lens aberrations and their contribution to a specific classification. It is important to note that the correlation values are calculated using a set of data that includes many instances, thus the value itself cannot be used as a measure to determine acceptability. Instead, the physical relationships discerned from the analysis can be used in a classification scheme.

The correlation values for all possible pairs of terms were calculated for four sets of data: the full set of Acceptable ZPCs fit using the FFT method, the full set of Unacceptable ZPCs fit using the FFT method, the CNN Sorted set of Acceptable ZPCs fit using the FFT method, and the CNN Sorted set of Unacceptable ZPCs fit using the FFT method.

These four sets were selected to find any relationships amongst the set as well as determine where the ZPC method performs better than the image based metric and CNN methods. The FFT fit data was selected to mitigate the impact of outlier values as these sets contained less outliers than the FT fit sets. Sets 3 and 4 contained only the ZPC terms for the borderline cases discussed in section 4.6.

Figure 15 shows the list of terms in each of the selected sets that have strong correlations. The starred terms appear in both the full and CNN sorted data sets meaning the ZPCs are extracting information from the wavefront that the image based metrics and CNN cannot extract from the Ronchigram. Terms 1 and 3 from the full unacceptable data are also the only two terms that have a strong (negative) correlation, which also appears in the CNN sorted data set. These terms correspond to piston and x-tilt respectively and are first order terms. The acceptable cases show more correlation with higher order terms, and all have positive correlations. These differences in the types of correlations and the terms that are correlated may lead to physical relationships in the wavefront that can be used to quantify the extent of optical distortion.

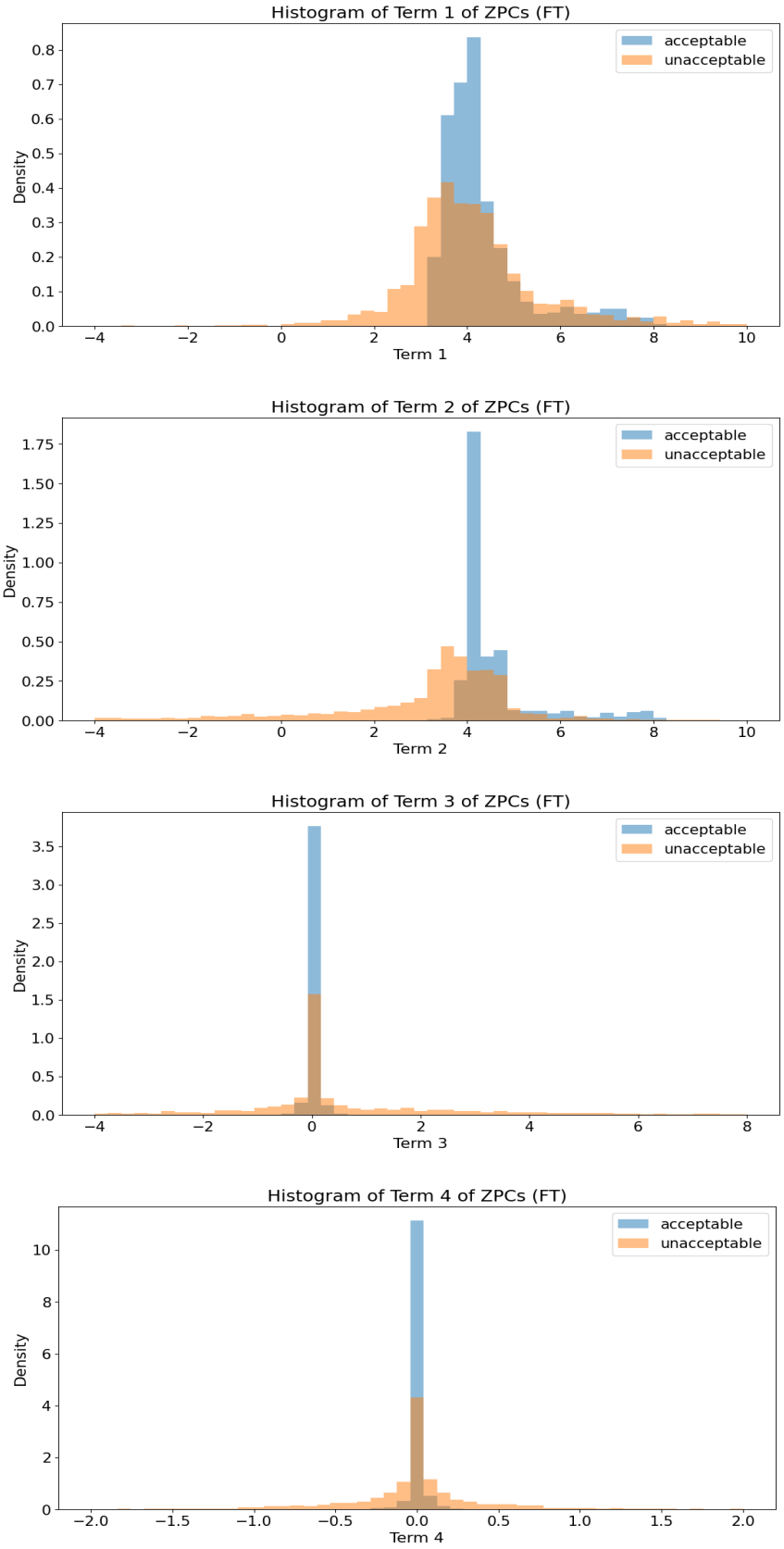

Figure 11. First four histograms of the Zernike polynomial coefficients fit using the fringe tracing method

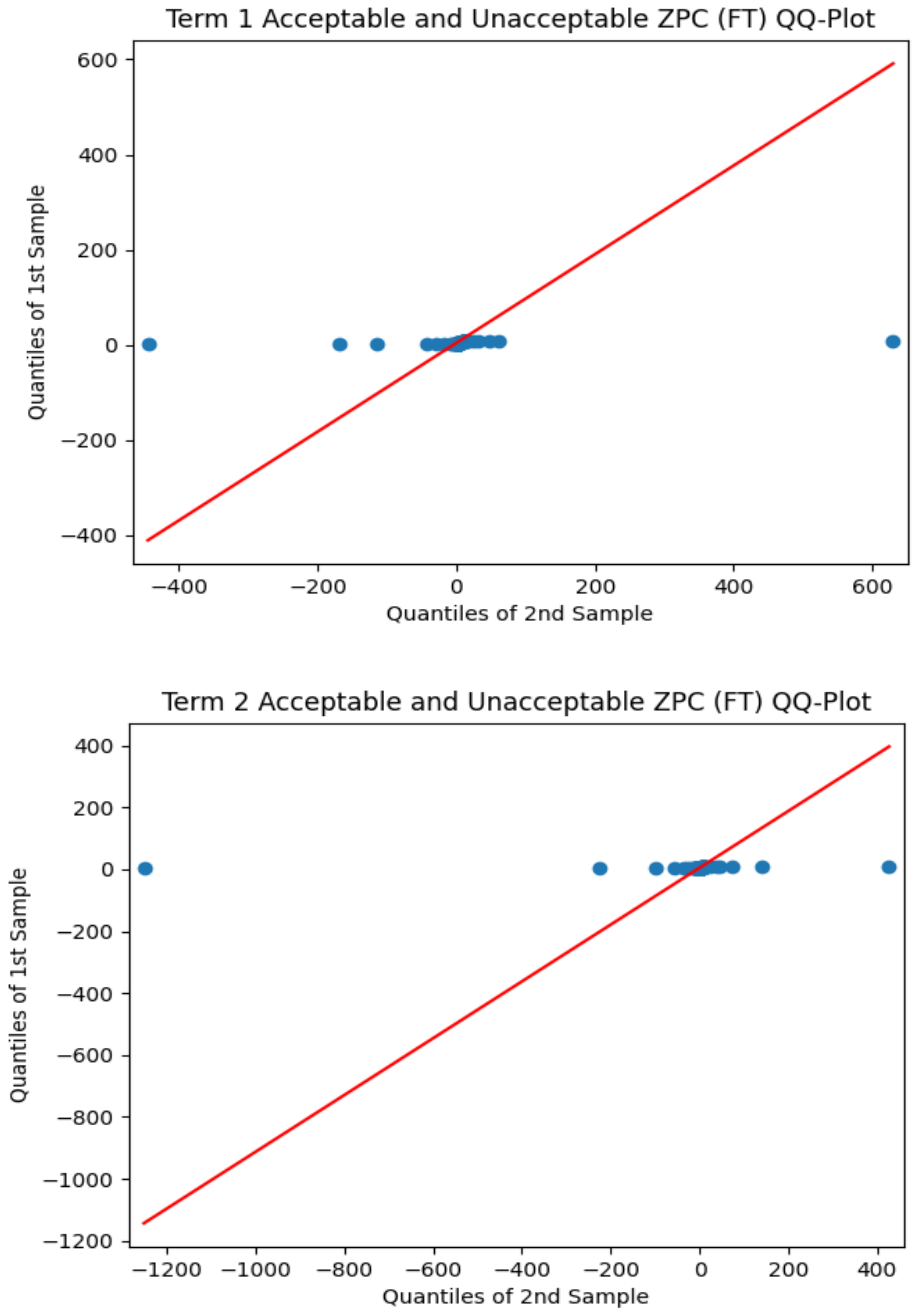

Figure 12. QQ-Plot comparing the acceptable ZPCs to the unacceptable ZPCs fit using the FT method

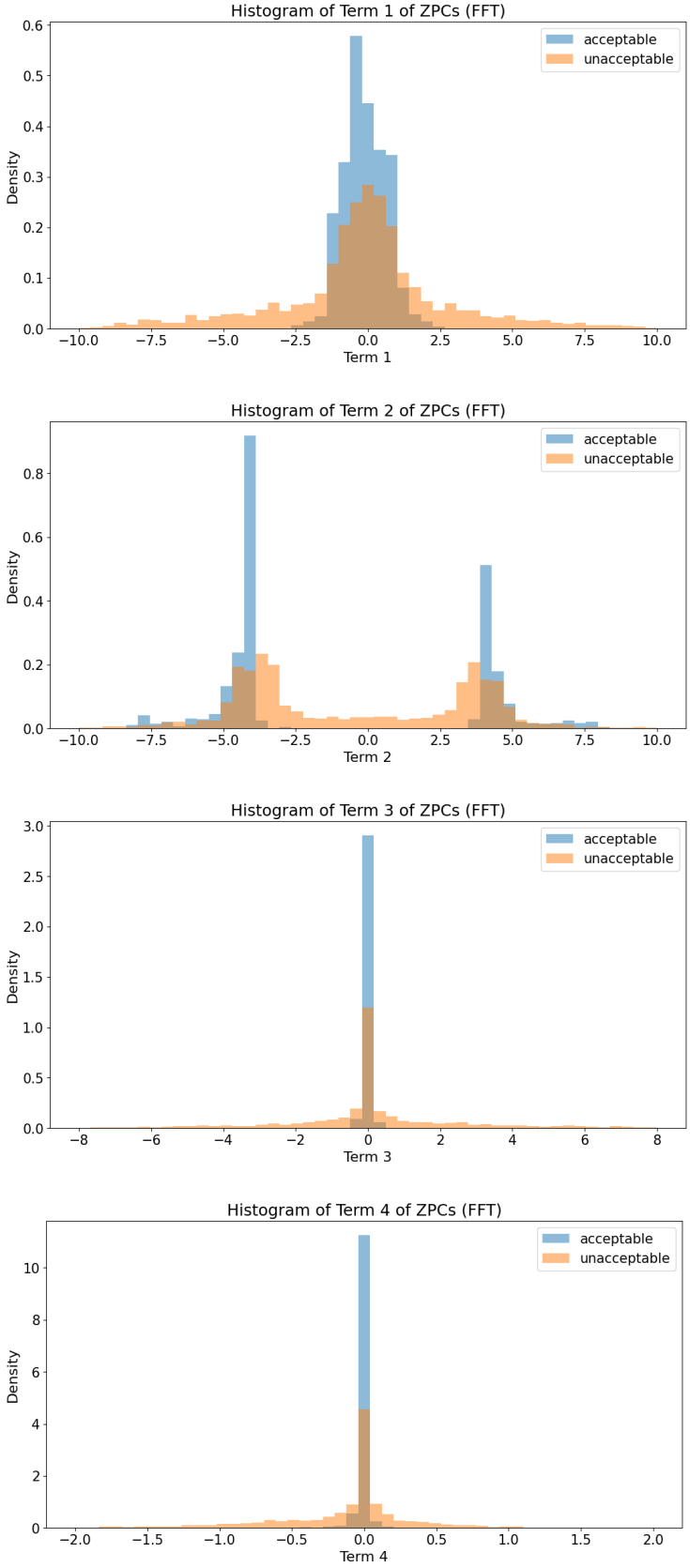

Figure 13. First four histograms of Zernike polynomial coefficient fits using the fast Fourier transform method

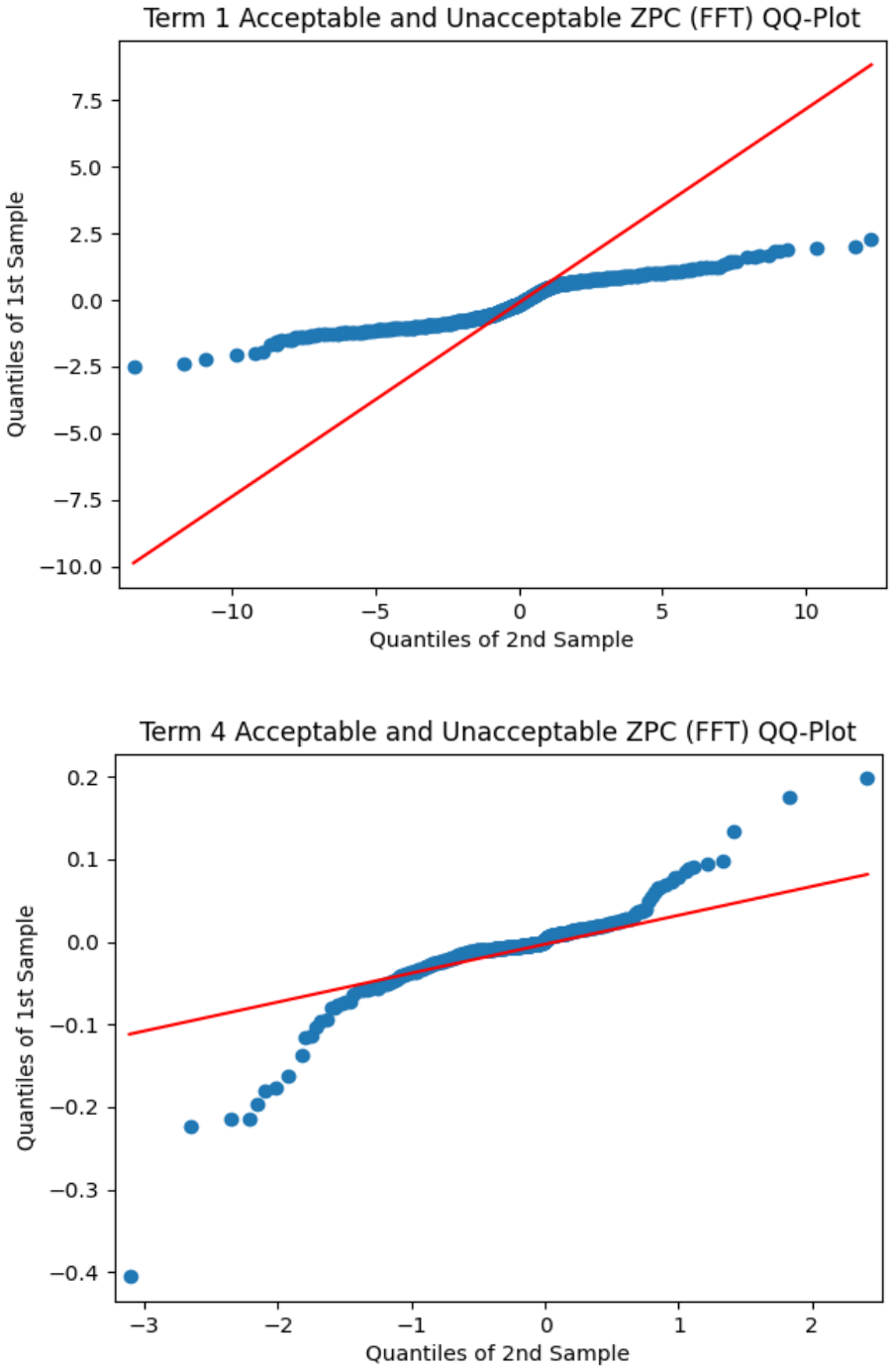

Figure 14. QQ-Plot comparing the acceptable ZPCs to the unacceptable ZPCs fit using the FFT method

```
Threshold Values of  Term 1 :  3.1442623389225264 , 8.113216223998124
Threshold Values of  Term 2 :  3.2774415189280792 , 8.119637089841792
Threshold Values of  Term 3 :  -0.5245463359223004 , 0.5074175778688661
Threshold Values of  Term 4 :  -0.2055732827794237 , 0.2070231690665373
Threshold Values of  Term 5 :  -0.5104040149434674 , 0.3542411305580708
Threshold Values of  Term 6 :  -0.3313106333283532 , 0.5359635188365977
Threshold Values of  Term 7 :  -0.2296497002382212 , 0.325690302058381
Threshold Values of  Term 8 :  -0.1228852164088193 , 0.124116468263548
Threshold Values of  Term 9 :  -0.0779321721130677 , 0.1379773256432575
Threshold Values of  Term 10 :  -0.1775855636661242 , 0.3892019392894618
Threshold Values of  Term 11 :  -0.1916992732732076 , 0.2501605322326347
Threshold Values of  Term 12 :  -0.1491626299318043 , 0.2193438234244697
Threshold Values of  Term 13 :  -0.0862328187037346 , 0.0879159011550149
Threshold Values of  Term 14 :  -0.1282875174511292 , 0.1040258741846899
Threshold Values of  Term 15 :  -0.0462000895451582 , 0.0587577717196559
Threshold Values of  Term 16 :  -0.161555512424019 , 0.0450507363255059
Threshold Values of  Term 17 :  -0.1587948044929648 , 0.2043823886244341
Threshold Values of  Term 18 :  -0.0739232754795843 , 0.0889778983376833
Threshold Values of  Term 19 :  -0.1148860106440788 , 0.113851742570482
Threshold Values of  Term 20 :  -0.0602421545417743 , 0.0635390852028049
Threshold Values of  Term 21 :  -0.3170324118370266 , 0.0963502043961902
Threshold Values of  Term 22 :  -0.0513491605892294 , 0.0421809876975748
Threshold Values of  Term 23 :  -0.2392487511072669 , 0.057315143044246
Threshold Values of  Term 24 :  -0.0772815130015581 , 0.0479473559902736
Threshold Values of  Term 25 :  -0.1374787533830305 , 0.1028115072781163
Threshold Values of  Term 26 :  -0.1089022387361272 , 0.1131777442114013
Threshold Values of  Term 27 :  -0.0742434343506449 , 0.0819942271300079
Threshold Values of  Term 28 :  -0.3176214466282145 , 0.0913324846411182
Threshold Values of  Term 29 :  -0.0813376319085696 , 0.041279945923762
Threshold Values of  Term 30 :  -0.2408558173694674 , 0.125958061295984
Threshold Values of  Term 31 :  -0.0704379026043929 , 0.0387375827428841
Threshold Values of  Term 32 :  -0.2615460398072968 , 0.2159048195306556
Threshold Values of  Term 33 :  -0.0460537424276927 , 0.0391986186310866
Threshold Values of  Term 34 :  -0.2107355768608317 , 0.0457024739211572
Threshold Values of  Term 35 :  -0.0478557512103203 , 0.0372323661667715
Threshold Values of  Term 36 :  -0.0689525452615583 , 0.0833876104375938
Threshold Values of  Term 37 :  -0.306636056313073 , 0.0864524994289803
Threshold Values of  Term 38 :  -0.0694548198925914 , 0.039110726878799
Threshold Values of  Term 39 :  -0.2421284563981202 , 0.063025490589321
Threshold Values of  Term 40 :  -0.0954308816048204 , 0.0964015327367424
Threshold Values of  Term 41 :  -0.2430599111814771 , 0.2185311874483056
Threshold Values of  Term 42 :  -0.0736666176619217 , 0.0393485368860402
Threshold Values of  Term 43 :  -0.2008976678766868 , 0.0654831300158565
Threshold Values of  Term 44 :  -0.0383563668684217 , 0.0274458468069188
Threshold Values of  Term 45 :  -0.1437027623284555 , 0.1753665494279907
Threshold Values of  Term 46 :  -0.0322317257151891 , 0.0349937040471978
Threshold Values of  Term 47 :  -0.1349460817628476 , 0.0222564889837239
Threshold Values of  Term 48 :  -0.0329145763318947 , 0.0332480392548543
Threshold Values of  Term 49 :  -0.0141041178091527 , 0.0427588443656216
```

Figure 15. Threshold values from sorting method using FT values.

```
Threshold Values of  Term 1 :  -2.514944591060584 , 2.263762153632176
Threshold Values of  Term 2 :  -8.11204298845179 , 7.972700492595617
Threshold Values of  Term 3 :  -0.5266094448126432 , 0.4978700913001648
Threshold Values of  Term 4 :  -0.4058072952112341 , 0.1980921182687077
Threshold Values of  Term 5 :  -0.7815248172492181 , 0.295527539128508
Threshold Values of  Term 6 :  -0.304407182155937 , 0.5378401283377301
Threshold Values of  Term 7 :  -0.1942631577952205 , 0.3258039505101874
Threshold Values of  Term 8 :  -0.1258347328738555 , 0.1122862123820887
Threshold Values of  Term 9 :  1.0205146501976 69e-06 , 0.06516165880896 56
Threshold Values of  Term 10 :  -0.1264933098742702 , 0.3212562769754309
Threshold Values of  Term 11 :  -0.2457567220781533 , 0.1505373477280419
Threshold Values of  Term 12 :  -0.1422952166355136 , 0.1185148058983944
Threshold Values of  Term 13 :  -0.0949518588907498 , 0.0847678049819498
Threshold Values of  Term 14 :  -0.2954133781629368 , 0.0930019357281157
Threshold Values of  Term 15 :  -0.0718691086331089 , 0.0526187067686559
Threshold Values of  Term 16 :  -0.0434569277401434 , 0.1063261890503527
Threshold Values of  Term 17 :  -0.132676978160064 , 0.132935559988978
Threshold Values of  Term 18 :  -0.0870806346818941 , 0.0733318456743155
Threshold Values of  Term 19 :  -0.1594324830589589 , 0.1098428792886231
Threshold Values of  Term 20 :  -0.072746537041778 , 0.0733842860112111
Threshold Values of  Term 21 :  -0.0844532306502844 , 0.1495626276206234
Threshold Values of  Term 22 :  -0.0521043066689343 , 0.0634613553334354
Threshold Values of  Term 23 :  -0.088217376618153 , 0.0815517024723028
Threshold Values of  Term 24 :  -0.057616661518138 , 0.090115751420774
Threshold Values of  Term 25 :  -0.0434761190826957 , 0.0338257802727012
Threshold Values of  Term 26 :  -0.1479128790229941 , 0.0859887894280482
Threshold Values of  Term 27 :  -0.090123204925021 , 0.0837911821293887
Threshold Values of  Term 28 :  -0.0839353635766803 , 0.0839440599774077
Threshold Values of  Term 29 :  -0.0820899898108362 , 0.093387488743341
Threshold Values of  Term 30 :  -0.0853055406149357 , 0.0821601044248837
Threshold Values of  Term 31 :  -0.0501944585617789 , 0.0379121300195472
Threshold Values of  Term 32 :  -0.1117056378541297 , 0.0479384042249614
Threshold Values of  Term 33 :  -0.0566080881770613 , 0.0522128515623181
Threshold Values of  Term 34 :  -0.0845928338899152 , 0.0837086881128067
Threshold Values of  Term 35 :  -0.06011725454572 , 0.0314175159134934
Threshold Values of  Term 36 :  -0.0282450614031713 , 0.0272112585820282
Threshold Values of  Term 37 :  -0.093811314986647 , 0.1060525829493851
Threshold Values of  Term 38 :  -0.10299330919665 , 0.0477993570206298
Threshold Values of  Term 39 :  -0.0524692307346593 , 0.0708112661916898
Threshold Values of  Term 40 :  -0.105446577078504 , 0.0520624629020824
Threshold Values of  Term 41 :  -0.0929732454311261 , 0.0455198560834367
Threshold Values of  Term 42 :  -0.0524835405975488 , 0.0478821636246423
Threshold Values of  Term 43 :  -0.0612060346254716 , 0.126531147498406
Threshold Values of  Term 44 :  -0.0262343834221791 , 0.0418553763075565
Threshold Values of  Term 45 :  -0.0361350329710746 , 0.0382632279747842
Threshold Values of  Term 46 :  -0.0493420381348428 , 0.0476625261906378
Threshold Values of  Term 47 :  -0.1452679839084033 , 0.0849682035631989
Threshold Values of  Term 48 :  -0.0312985136386922 , 0.0343698051381747
Threshold Values of  Term 49 :  -0.0221124743478768 , 0.0257811274542859
```

Figure 16. Threshold values from sorting method using FT values.

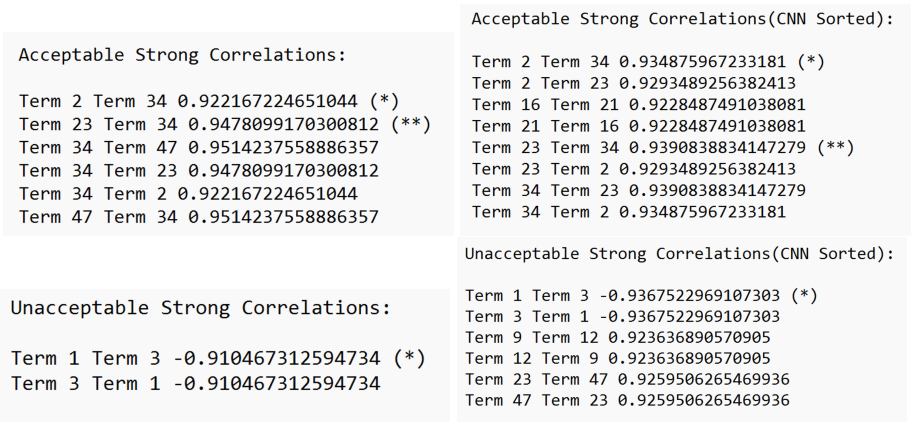

Figure 17. Correlation values between ZPC terms in acceptable and unacceptable data

6.0 CONCLUSION

Image based metrics give quantifiable metrics to Ronchigrams, CNN allows for classification and ZPCs quantify the distortion that can be mapped to physical characteristics in the wavefront and the lens system.

To address the subjectivity of the MIL-DTL-43511D, analysis was performed on a set of synthetic Ronchigrams. This analysis began with defining metrics to measure the extent of distortion present in the fringes in observed in synthetic Ronchigrams. These metrics gave measurable quantities with which thresholds could be set in order to classify the Ronchigrams. However, there is not a single image based metric that can be used to classify all Ronchigrams since each metric was defined to quantify a different distortion in the fringes. Many of these metrics contained significant overlap with the others, substantiating the overall classification of a given Ronchigram. Additionally, the image based metrics suffered on occasion as some thresholds were required to be set in a manner that eliminated a subset acceptable instances. Some of these eliminated acceptable instances contained some distortion and off parallelism, which indicates a subset of acceptable data that may require further consideration.

Another method developed to classify the synthetic Ronchigrams employed computer vision and a CNN. This method obtained sufficient accuracy that appeared to be limited by the quality of the training set provided to it. In short, to obtain higher accuracies with the CNN would lead to over fitting the model to the specific data set and would lead to a decrease in generalization ability. This is substantiated by the results from the test set which was withheld from the network during training and validation reaching within 1 % of the categorical accuracy as the training and validation set. The incorrectly classified instances from this CNN were compiled to create a new subset that contained borderline acceptable cases. These borderline cases did not posses any quantifiable differences when the image based metrics were applied to them.

Another approach to quantifying optical distortion involved fitting ZPCs to the synthetic Ronchigram data. The ZPCs can be used to describe lens aberrations and therefore are a promising candidate for use as a measure of optical distortion. Similar analysis to the image based metrics was performed on ZPC data sets that were fit with two different methods. This analysis produced promising results as carrying 49 ZPC terms allowed for the majority of the unacceptable instances to be sorted out without losing any acceptable instances. Inclusion of higher order terms can capture relationships not found in the lower order terms. Moreover, further analysis of the subset of incorrectly classified instances from the CNN which constitute the borderline cases revealed a relationship between ZPC terms that may be exploited as a method of classification.

Further work to validate these methods and determine the physical relationship between the ZPC terms is required. Validation can be achieved by creating a data set with real Ronchigrams and wavefront data from the same eyewear. The real Ronchigram data can then be used to generate image based metrics for comparison to the synthetic set as well as input into the CNN for classification. ZPCs will then need to be fit to the Ronchigrams using the two separate methods to determine if these methods are reliable. The wavefront data will be used to validate the fitting methods. Once the fitting methods are validated, analysis can be performed to determine if the

same relationships between the terms in the ZPCs appear in the real data and examined in greater detail.

Obtaining wavefront data also provides other quantities that can be measured and phase maps that can be used in image based classifiers. These classifiers can be added to ongoing projects that employ ML techniques to classify Ronchigrams or used to validate these initiatives.

7.0 SOURCES

Harvey, A., Eagan, C., Henry, A., Novar, B., Arakal, A., Cook, M., Ronchi Evaluator and Classifier of Imperfect Lenses, AFRL-RH-FS-TR-2022-0011, 8 Jan 2022.

Harvey, A., Eagan, C., Henry, A., Novar, B., Arakal, A., Cook, M., Development of Fringe Tracing and Zernike Polynomial Fitting Algorithms for Distortion Computations (Disco) Part 1, AFRL-RH-FS-TR-2022-0015, 8 Jan 2022.

Harvey, A., Eagan, C., Henry, A., Novar, B., Arakal, A., Cook, M., Development of Fringe Tracing and Zernike Polynomial Fitting Algorithms for Distortion Computations (Disco) Part 2, AFRL-RH-FS-TR-2022-0017, 8 Jan 2022.

Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.

Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

Wyant, James Creath, Katherine. (1992). Basic Wavefront Aberration Theory for Optical Metrology. Appl Optics Optical Eng. 11.