Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 22 No. 9 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v22i09.61857

PAPER

Morphological Phenotyping of Corneal Endothelial Cells Using Elliptic Fourier Descriptors

Ashwini Ramchandra Doke, Suja Palaniswamy (✉)

Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, Karnataka, India

p_suja@blr.amrita.edu

ABSTRACT

Corneal endothelial cell morphology is clinically important to understand corneal health, and is routinely examined using specular microscopy. Traditional metrics such as endothelial cell density (ECD), coefficient of variation (CV), and hexagonality percentage provide limited morphological information and are unreliable when used on large diverse datasets. This paper presents an automated phenotyping framework using Elliptic Fourier Descriptors (EFDs), Zernike moments, and geometric features combined with unsupervised K-Means clustering for endothelial cell classification. The framework was validated on two independent datasets. The primary dataset comprised 179 cells (1,253 post augmentation) from four specular microscopy images. The external validation set (SREP-18-33533B) comprised 35,890 cells across 385 images. EFD based features achieved a mean silhouette score of 0.6213 (SD = 0.056) on the primary dataset, representing a large effect size over traditional measures (Cohen's d = 5.290). Across datasets, traditional measures declined by nearly 40% (0.5986 to 0.3628), whereas EFD based scores remained stable with under 1% change (0.6213 to 0.6119), as confirmed by the Wilcoxon signed-rank test ($p < 0.001$). Three reproducible phenotypes were identified. Cluster 0, contained regular compact cells (64.4%). Cluster 1, contained moderately regular cells (29.2%) and Cluster 2, contained irregular or elongated cells (6.4%). These results indicate that EFD based shape analysis remains discriminative on a 35,890 cells where traditional clinical metrics fail. The framework can support real world analysis of corneal endothelial cell phenotypes when combined with clinical expertise of researchers and ophthalmologists.

KEYWORDS

corneal endothelial cells, Elliptic Fourier Descriptors (EFD), specular microscopy, unsupervised clustering, morphological phenotyping, K-Means, Zernike moments, statistical validation

1 INTRODUCTION

The corneal endothelium is a single layer of cells lining the inner surface of the cornea. The endothelial cell layer performs a critical pumping function that removes excess fluid from corneal tissue and keeps the cornea transparent [1]. A key

Doke, A. R., Palaniswamy, S. (2026). Morphological Phenotyping of Corneal Endothelial Cells Using Elliptic Fourier Descriptors. *International Journal of Online and Biomedical Engineering (iJOE)*, 22(9), pp. 186–204. https://doi.org/10.3991/ijoe.v22i09.61857

Article submitted 2026-04-07. Revision uploaded 2026-05-25. Final acceptance 2026-05-26.

characteristic that distinguishes this cell layer from other tissues in the body is its inability to self-repair, when individual cells die they are not replaced by new ones. Instead, surrounding cells stretch and migrate to fill the gap, which over time leads to changes in cell size and shape [2]. Two recognised changes are polymegethism, where the variation in cell size increases beyond normal levels, and pleomorphism, where cells deviate from the typical six sided hexagonal arrangement [3]. Specular microscopy is the standard non-invasive clinical technique used to photograph this cell layer and monitor these shape changes [4]. Assessment of the images obtained is important for evaluating patients before and after corneal transplantation, screening donor corneas for viability, and tracking the long term effects of procedures such as LASIK surgery, cataract removal, and extended contact lens wear, including the use of rigid PMMA lenses [5].

Software that uses specular microscopes typically reports three summary measurements for a given image viz: the cell density per square millimetre, the extent of variation in individual cell sizes expressed as a coefficient of variation (CV), and the proportion of cells that are six sided [6, 7]. Although these three measures have a long history of clinical use, they compress the rich information of individual cells into a small set of numbers, discarding details about cell boundary shape, local texture and structural regularity [8]. These traditional measures also perform poorly when used on large and diverse datasets, which makes them less suitable for large scale automated analysis [9].

Mathematical shape descriptors provide more detailed and scalable way to describe cell morphology [10]. Elliptic Fourier Descriptors (EFDs) describe the cell boundary using a series of elliptic curves, providing a compact representation that stays consistent under rotation and scaling [11]. Zernike moments capture overall shape within a circular region using a set of orthogonal functions. When these EFDs and Zernike moments are combined with standard geometric descriptors, they produce a feature vector that captures both fine boundary detail and broader region level properties. Such feature vectors can then be used with unsupervised grouping methods to identify cell phenotypes without requiring any manually labelled training data [12].

Building on this background, the present study addresses three explicit research questions (RQ). First, can a multi-feature representation based on EFDs, Zernike moments, and geometric descriptors produce reproducible morphological phenotypes of corneal endothelial cells through unsupervised clustering? Second, how do EFD based features compare with traditional clinical metrics and with other shape descriptors (Zernike moments, geometric descriptors) in terms of clustering quality and stability across datasets of differing size and acquisition conditions? Third, do the identified phenotypes align with known biological categories of endothelial cell morphology, particularly in relation to PMMA-induced changes? The corresponding objectives are: (i) To design and implement a scalable EFD based feature extraction and clustering pipeline, (ii) To compare five feature configurations using silhouette analysis, effect size measures, and non-parametric statistical tests on both primary and a large external dataset and (iii) To interpret the resulting clusters in light of established corneal endothelial biology.

This framework falls within the scope of online and biomedical engineering. The proposed pipeline is automated and suitable for integration into web or cloud based ophthalmic image analysis workflow. The principal contributions of the study are therefore as follows: First, a multi feature extraction pipeline combining EFDs with standard and custom normalisation, Zernike moments and geometric descriptors is proposed and evaluated across five feature configurations. Second, the framework

is evaluated on a primary dataset of four images with effect sizes reported and statistically validated on a 385 image external dataset using Wilcoxon signed rank tests and the Friedman test (chi-square = 1298.16, $p < 0.001$). Third, quantitative evidence is provided that traditional clinical metrics fail to generalise at scale (silhouette drop of 0.236, Wilcoxon $p < 0.001$), supporting the adoption of shape based descriptors in automated corneal cell analysis pipelines [13, 14].

2 RELATED WORK

Researchers have been working on automated tools for corneal endothelial cell analysis. Among the earlier contributions, Ruggeri et al. [15] developed a system that combined edge-based image processing with a region growing algorithm to count and segment cells in specular microscopy images, achieving accuracy levels comparable to manual counting on good quality images. Its performance heavily depends on manual parameter tuning, which reduces its reliability across images captured under different conditions.

Among image segmentation methods, watershed-based techniques have been widely adopted for dividing touching cells in specular microscopy images. Fabijańska [16] showed that using gradient images at multiple spatial scales before applying the watershed transform improved the accuracy of cell boundary detection compared with simpler threshold-based approaches. More recently, deep learning has produced strong results in this area: Vigueras-Guillén et al. [17] reported cell counts errors below 5% using a convolutional neural network. The similar deep learning and computer vision techniques have shown strong performance in broader biomedical and visual recognition applications [18]. These approaches typically require substantial volumes of labelled training images, and their internal representations are difficult to interpret clinically. The present study takes a different path, using mathematically defined shape features and unsupervised clustering, which requires no labelled data and produces results that can be directly related to established biological concepts.

Elliptic Fourier Descriptors were originally introduced by Kuhl and Giardina [19] and have since been used to characterise cell shapes in areas such as haematology and histological tissue analysis. Zernike moments, applied to cell classification tasks by Lowe et al. [20, 21], offer a different type of shape encoding based on global region properties. Both descriptors have been used independently in biomedical image analysis. Their joint use with standard geometric features for unsupervised corneal endothelial cell phenotyping, with cross dataset validation by Wilcoxon signed rank and Friedman tests on 385 images, has not previously been reported.

3 MATERIALS AND METHODS

3.1 Overall framework

Figure 1 presents the overall processing pipeline. Raw specular microscopy images are pre-processed and segmented to extract individual cell contours. Cell boundaries are reconstructed using EFD and validated using Hausdorff and RMS Chamfer distances. A multi modal feature vector combining EFD coefficients, Zernike moments, and geometric descriptors is constructed for each cell. K-Means clustering is applied and results are visualised using PCA and t-SNE for phenotype identification.

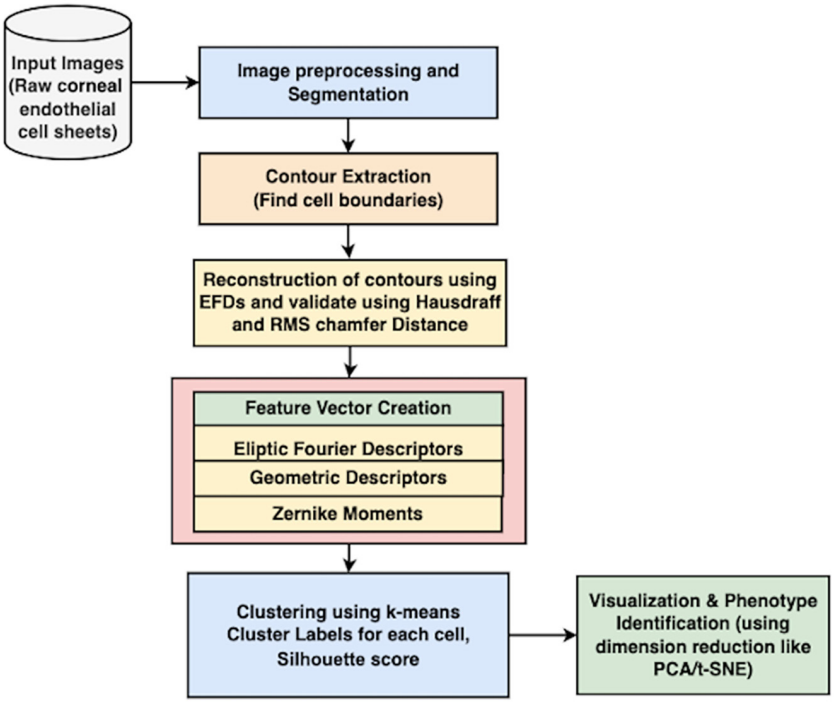

Fig. 1. Overall processing pipeline for corneal endothelial cell morphological phenotyping

Raw images are pre-processed and segmented, cell contours extracted and reconstructed using EFD, multi-modal features extracted, and K-Means clustering (k = 3) identifies distinct cell phenotypes visualised via PCA or t-SNE.

3.2 Datasets

Two datasets were used in this work. The primary dataset consists of four specular microscopy images: Control 1C, Control 4C, PMMA 1X, and PMMA 4X, acquired at 477 × 260 pixels (8-bit grayscale), resized to 256 × 256 pixels. A total of 179 cells were extracted after segmentation and quality filtering (42, 58, 45, and 34 cells for Control 1C, PMMA 1X, Control 4C, and PMMA 4X respectively) [22]. Following data augmentation [23], the total number of cell instances used for feature extraction and clustering was 1,253. Figure 2 shows all four images.

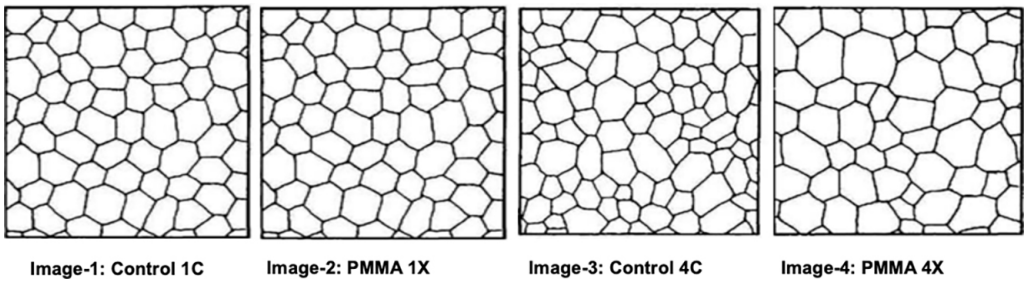

Fig. 2. Primary dataset: four specular microscopy images of corneal endothelium

Notes: Left to right: Control 1C, PMMA 1X, Control 4C, PMMA 4X. PMMA images show visibly larger and more variable cell morphology compared to Control images. Used with permission of Association for Research in Vision and Ophthalmology (ARVO), from 'Corneal Endothelial Polymegathism Induced by PMMA Contact Lens Wear,' Stocker EG and Schoessler JP, Investigative Ophthalmology & Visual Science, Vol. 26, No. 6, 1985; [22]. Permission conveyed through Copyright Clearance Center, Inc.

The external validation dataset consisted of 385 specular microscopy images from Daniel et al. [24]. Ground truth cell centre coordinates were provided for 324 images (42,147 annotated cell centres); the remaining 61 images, with empty annotation files, were excluded from count accuracy evaluation. The images in this dataset differ widely in cell density, contrast, and acquisition conditions, making it a challenging benchmark for generalisation for the proposed approach.

3.3 Pre-processing and segmentation

The images were first resized to 256 × 256 pixels. To improve local contrast, contrast limited adaptive histogram equalization (CLAHE) was applied with a clip limit of 3.5 and a tile size of 6 × 6. The bright and dark regions near the image borders were excluded, as they did not provide useful cell information [25, 26]. Cell contours were detected using a watershed algorithm based on a multi scale morphological gradient. After segmentation, regions that were too close to the image border or that had irregular shape were removed from the analysis. The output of segmentation on the primary dataset is shown in Figure 3.

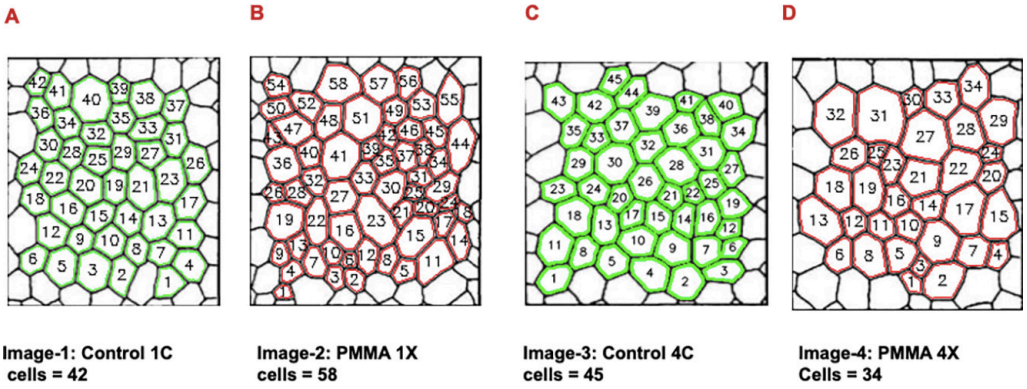

Fig. 3. Segmentation results on the primary dataset with numbered cell contours

Notes: Green = correctly segmented cells; red = border-excluded cells. Cell counts: Control 1C = 42, PMMA 1X = 58, Control 4C = 45, PMMA 4X = 34.

3.4 Elliptic Fourier Descriptors reconstruction validation

Figure 4 shows the reconstruction of cell shape from harmonics 1 to 20, illustrating how the contour progressively recovers its original form as the number of harmonics increases [27, 28]. The Hausdorff [29] and RMS-Chamfer [30] reconstruction errors decrease as the number of harmonics increases (see Figure 5), with diminishing improvement beyond the 15th harmonic. The errors at the 1st harmonic (Hausdorff distance = 1.46 pixels, RMS Chamfer distance = 0.79 pixels) were small enough to justify the use of the 1st harmonic for downstream clustering.

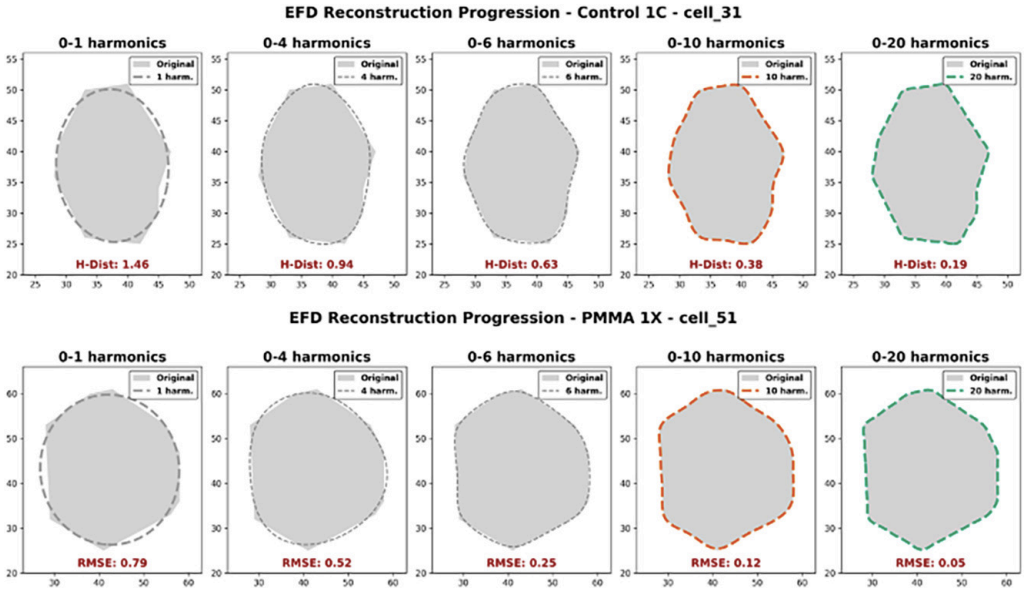

Fig. 4. Elliptic Fourier Descriptors reconstruction

Notes: Elliptic Fourier Descriptors reconstruction progression for a Control 1C cell (top) and PMMA 1X cell (bottom) at 1, 4, 6, 10, and 20 harmonics. H-Dist and RMSE confirm progressive shape recovery with increasing harmonic order.

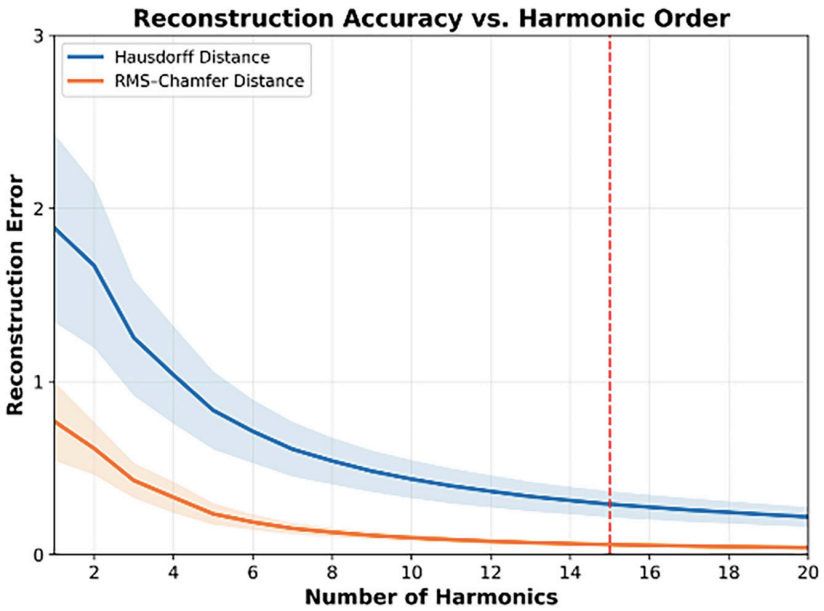

Fig. 5. Mean reconstruction error (Hausdorff and RMS Chamfer distance)

3.5 Feature extraction

Five feature configurations were evaluated in this work, summarised as shown in Table 1. The geometric only set comprised seven region descriptors (area, perimeter, eccentricity, solidity, circularity, compactness, and hexagonality). The traditional Clinical Metrics set comprised six traditional per cell metrics derived from clinical measures (area, area deviation, hexagonality score, circularity, solidity, eccentricity) [31].

EFD Standard (EFD std): denotes the b1 and d1 coefficients derived from the standard normalised first harmonic of the EFDs. EFD Custom: denotes the b1 and d1 coefficients derived from the phase corrected normalised first harmonic of the EFDs. The Zernike only set comprised 25 moments computed at a radius of 30 pixels [32]. The Combined sets concatenated the EFD, Zernike, and geometric descriptors into 34-dimensional feature vectors. No additional feature scaling was applied prior to clustering, as the individual feature types (EFD coefficients, Zernike moments, geometric descriptors) are internally normalised through their respective computation procedures. It is acknowledged that K-Means clustering is sensitive to feature scale, the decision not to apply further scaling was made to retain the natural units of each descriptor and to avoid introducing an additional arbitrary pre-processing step.

Data augmentation was applied to the primary dataset only, increasing the number of cell instances from 179 to 1,253 (a 7-fold increase). Each cell was transformed in six different ways. These comprised horizontal flip, vertical flip, rotations of 90º, 180º, 270º, and translation, producing six additional copies of each cell. These copies are instances of the same biological cells after augmentation and are not independent biological samples, but rather geometric transformations of the same 179 source cells from four images. Augmentation was used solely to provide a larger and more balanced set of contours for clustering.

Table 1. Feature sets evaluated in this work

Feature Set	Dim.	Description
Traditional Clinical Metrics	6	Area, area deviation, hexagonality score, circularity, solidity, eccentricity
Geometric Only	7	Area, perimeter, eccentricity, solidity, circularity, compactness, hexagonality
Zernike Only	25	Zernike moments at radius 30 px (25 moments)
EFD Std + Zernike + Geometric	34	Concatenation of EFD Std (b1, d1), Zernike, and geometric features
EFD Custom + Zernike + Geometric	34	Concatenation of EFD Custom (b1, d1), Zernike, and geometric features

3.6 Clustering and statistical evaluation

K-Means (k = 3, K-Means initialisation, 10 runs) was applied to each feature set. The value k = 3 was selected based on elbow plot analysis (see Figure 6). Clustering quality was evaluated using silhouette score [33, 34]. For the primary dataset (n = 4 images), statistical significance testing was not applied due to the small sample size (n < 10 required for reliable non-parametric testing); instead, effect sizes (Cohen's d) and descriptive statistics are reported. For the external validation dataset (n = 385 images), statistical significance was assessed using the Wilcoxon signed-rank test (non-parametric paired test) and the Friedman test (non-parametric repeated measures test across all feature sets). All tests used two sided alternatives with significance threshold alpha = 0.05. Cross feature set agreement was assessed using ARI [35].

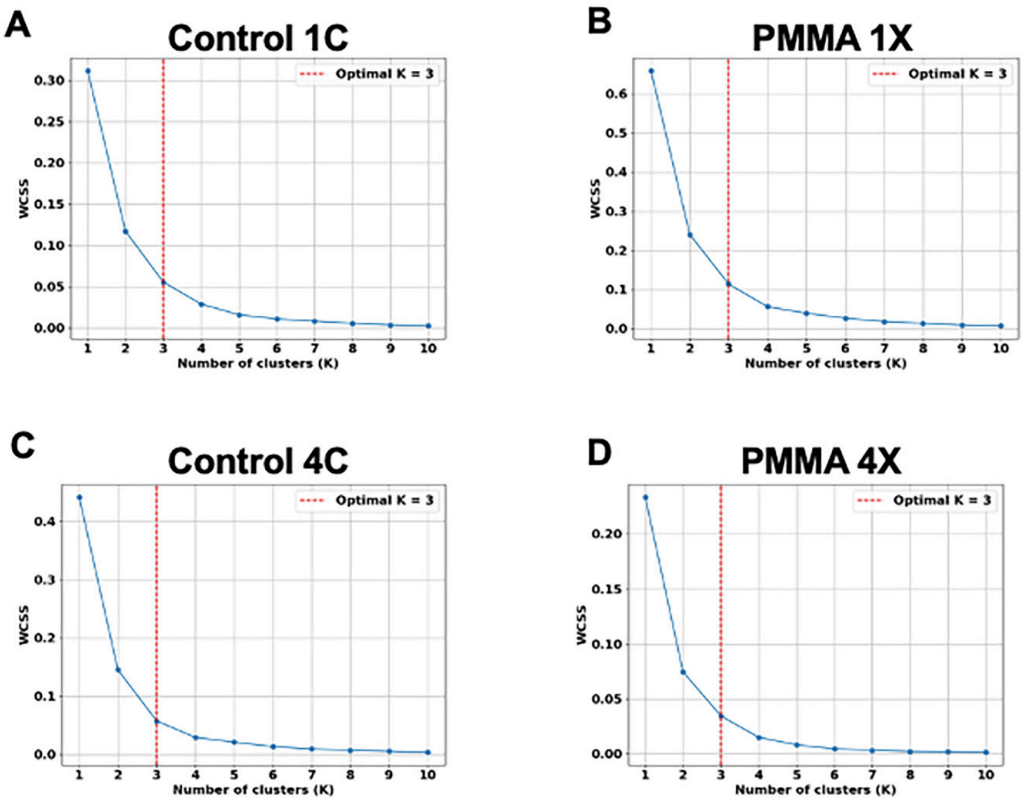

Fig. 6. Elbow plots (WCSS vs. k) for all four primary dataset images

Notes: (A) Control 1C, (B) PMMA 1X, (C) Control 4C, (D) PMMA 4X. All the images show an optimal k = 3 (red dashed line) confirms choice of three clusters.

3.7 Implementation details

The pipeline was implemented in Python 3.12.13 using OpenCV and scikit image for pre-processing and segmentation, pyefd for EFDs, mahotas for Zernike moments, scikit-learn for K-Means, PCA, t-SNE, silhouette, and ARI, SciPy for the Wilcoxon and Friedman tests, and pandas, NumPy, matplotlib, and seaborn for data handling and plotting. The primary dataset (4 images, 1,253 augmented cells) was processed on a laptop with an Intel Core i7 CPU and 16 GB of RAM in approximately 15 minutes. The external dataset (385 images, 35,890 cells) was processed on Google Colab with an NVIDIA L4 GPU in approximately 30 minutes, with segmentation accounting for the largest share of the runtime.

4 RESULTS

4.1 Primary dataset results

Table 2 reports the clustering performance for all feature sets across the primary dataset images based on silhouette scores. The EFD Standard + Zernike + Geometric set and the Geometric set alone achieved virtually identical mean silhouette scores of 0.6213 and 0.6215 respectively, a difference of only 0.0002. Zernike scored lowest at 0.4889 (SD = 0.049). EFD Standard and Geometric features consistently performed

best across all four images, followed by EFD Custom normalisation. The traditional features always performed below EFD and Geometric features, and Zernike only features ranked last in every single image. This ordering being preserved across four different images captured under different conditions. It confirms that the proposed method captures consistent and reproducible shape information rather than producing random results.

The Cohen's d effect sizes were computed to measure how meaningful the differences between feature sets are. The results are shown in Table 3. The EFD set scored substantially better than the traditional metrics (d = 5.290) and the Zernike set (d = 6.915). The results show EFD and geometric features showed virtually identical performance, with a difference of only 0.0002. Given the limited dataset size of the primary dataset, effect sizes were reported rather than statistical tests such as ANOVA, t-tests, or Wilcoxon.

Table 2. Silhouette scores for all feature sets on the primary dataset (4 images)

Feature Set	Img 1	Img 2	Img 3	Img 4	Mean	SD
Traditional Clinical Metrics	0.5795	0.5640	0.5668	0.6839	0.5986	0.057
Geometric Only	0.6067	0.5827	0.5929	0.7038	0.6215	0.056
Zernike Only	0.4864	0.4703	0.4421	0.5570	0.4889	0.049
EFD Std + Zernike + Geometric	0.6064	0.5824	0.5927	0.7036	0.6213	0.056
EFD Custom + Zernike + Geometric	0.5944	0.5730	0.5828	0.6982	0.6121	0.058

The results therefore describe about primary dataset only and should not be read as conclusions about the general population of corneal endothelial cells. Figure 7 shows silhouette scores for each feature set separately for the primary dataset images. Where, Image 1 refers to Control 1C, Image 2 to PMMA 1X, Image 3 to Control 4C, and Image 4 to PMMA 4X. EFD Standard and Geometric features consistently scored highest while Zernike alone scored lowest, indicating that the method behaves reliably across images.

Table 3. Effect size comparison of selected feature sets on the primary dataset

Comparison	Mean A	Mean B	Cohen's d	Effect Size
EFD (Std + Zernike + Geometric) vs Traditional	0.6213	0.5986	5.290	Large
EFD (Std + Zernike + Geometric) vs Zernike Only	0.6213	0.4889	6.915	Large
EFD (Std + Zernike + Geometric) vs Geometric Only	0.6213	0.6215	4.330	Equivalent

4.2 External validation dataset results (385 Images)

The statistical Friedman test was applied across the 385 images of the external dataset. It confirmed that clustering quality significantly differs between feature sets (chi-square = 1298.16, $p < 0.001$). This confirms that the choice of feature representation has a statistically significant impact on clustering quality across independent datasets. Table 4 shows the scores and variations between the two datasets. The EFD Std + Zernike + Geometric dropped only 0.0094 from the primary to the external dataset, and EFD Custom + Zernike + Geometric dropped by 0.006. By contrast,

the traditional metrics dropped sharply from 0.5986 to 0.3628, demonstrating their limitations on large, diverse datasets.

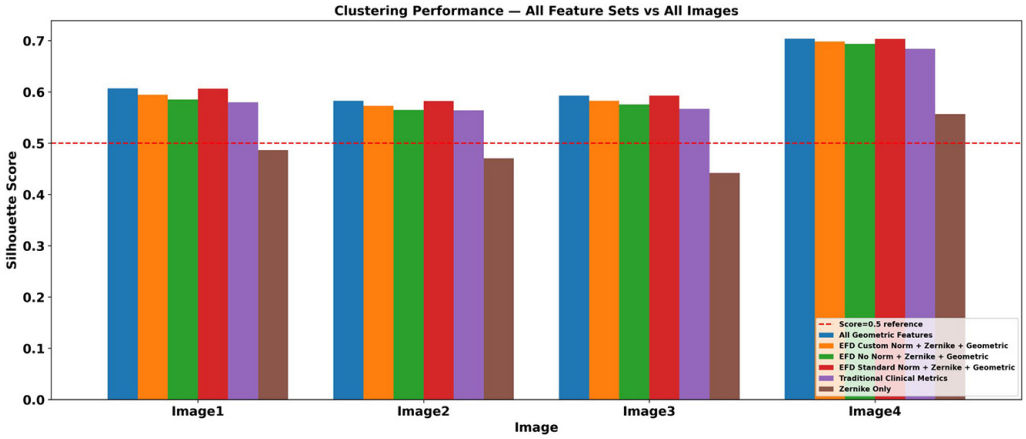

Fig. 7. Silhouette scores for primary dataset images for all six feature sets

This drop of nearly 40% for traditional metrics contrasts sharply with the less than 1% drop for EFD based features, as illustrated in Figure 8. Figure 8a compares the mean silhouette score for each feature set on both datasets, making it easy to see which features hold up and which fall off. Figure 8b shows the distribution of the 35,890 cells from the external dataset across the three identified cell shape groups. Together, Figure 8 confirm that EFD based and geometric features generalise well to unseen data, whereas traditional clinical metrics do not.

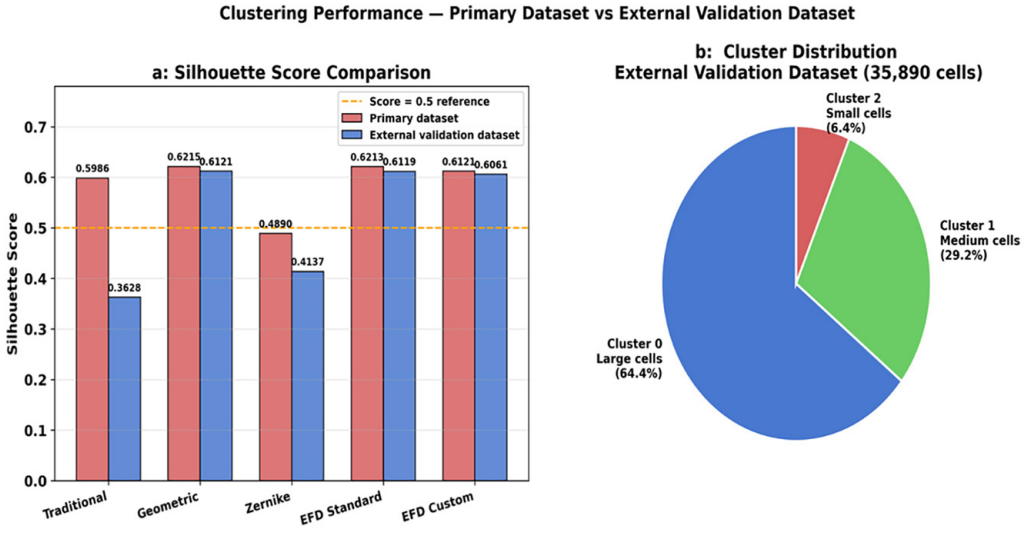

Fig. 8. Clustering performance

Notes: (a) Comparison of mean silhouette scores for all feature sets across the primary dataset (red bars) and the external validation dataset (385 images, blue bars). (b) Cluster size distribution across 35,890 cells shown as a pie chart. EFD based and geometric features show minimal performance drop across datasets (difference 0.006–0.009), while traditional clinical metrics show a large drop (difference 0.236), confirming that EFD features generalise well to unseen data.

Table 4. Clustering performance on primary (4 images) and external validation (385 images) datasets with cross-dataset generalisation assessment

Feature Set	Silhouette Score (4 Images)	Silhouette Score (385 Images)	Diff	Generalisation
Traditional Clinical Metrics	0.5986	0.3628	–0.236	Poor ($p < 0.001$)
Geometric Only	0.6215	0.6121	–0.009	Excellent
Zernike Only	0.4889	0.4137	–0.075	Moderate
EFD Std + Zernike + Geometric	0.6213	0.6119	–0.009	Excellent
EFD Custom + Zernike + Geometric	0.6121	0.6061	–0.006	Excellent

Table 5 presents the results of pairwise Wilcoxon signed-rank tests carried out on the external validation dataset. These tests compare the per image silhouette scores of each feature set, treating the 385 images as paired observations. All comparisons involving EFD based features against Zernike only features were highly significant with large effect sizes. The Friedman test was used to determine whether all feature sets produced similar clustering quality. A significant result indicates that the choice of feature set has a measurable effect, supporting the relevance of the feature representations evaluated here.

Table 5. Results of pairwise statistical tests between feature sets on the 385-image external dataset

Comparison	Mean A	Mean B	p-Value	Cohen's d	Effect
EFD Std vs Zernike Only	0.6238	0.4260	<0.001***	4.889	Large
EFD Std vs EFD Custom	0.6238	0.6174	<0.001***	1.300	Large
EFD Custom vs Traditional	0.6174	0.6238	<0.001***	1.305	Large
EFD Std Combined vs Std Only	0.6238	0.6182	0.013*	0.121	Negligible
Friedman (all feature sets)	–	–	<0.001***	chi-sq = 1298.16	–

Notes: Asterisks denote statistical significance of the two sided Wilcoxon signed-rank test, where * denotes $p < 0.05$ and *** denotes $p < 0.001$. The final row reports the Friedman test across all five feature sets, in which the value in the Cohen's d column is the Friedman chi-square statistic.

4.3 Cluster maps and phenotype identification

Figure 9 shows cluster coloured maps of cells from the primary dataset, where each colour represents a group of similar cell shapes. Three clusters were identified viz, blue cells are regular, compact and mostly hexagonal. Orange cells are moderately regular and of medium size. Green cells are elongated and irregular. The PMMA images show consistent morphological changes associated due to PMMA lens wear. On the Daniel external dataset, the cluster distribution was: Cluster 0 = 23,096 cells (64.4%), Cluster 1 = 10,480 cells (29.2%), Cluster 2 = 2,314 cells (6.4%).

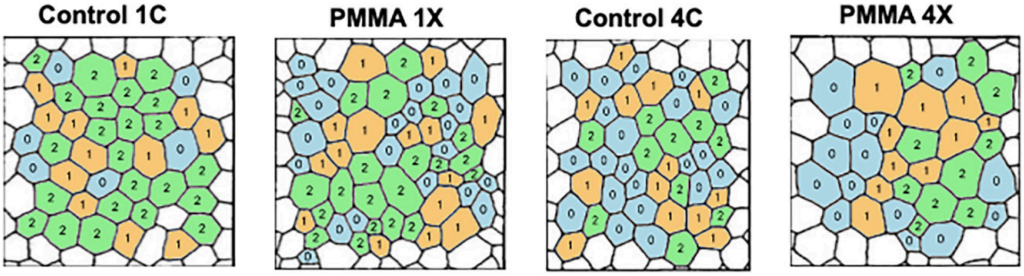

Fig. 9. Clustered cell maps for the primary dataset images

4.4 PCA and t-SNE visualisation

PCA and t-SNE projections show clear separation of the three clusters across all images as shown in Figure 10. It confirms that the identified phenotypes are distinct.

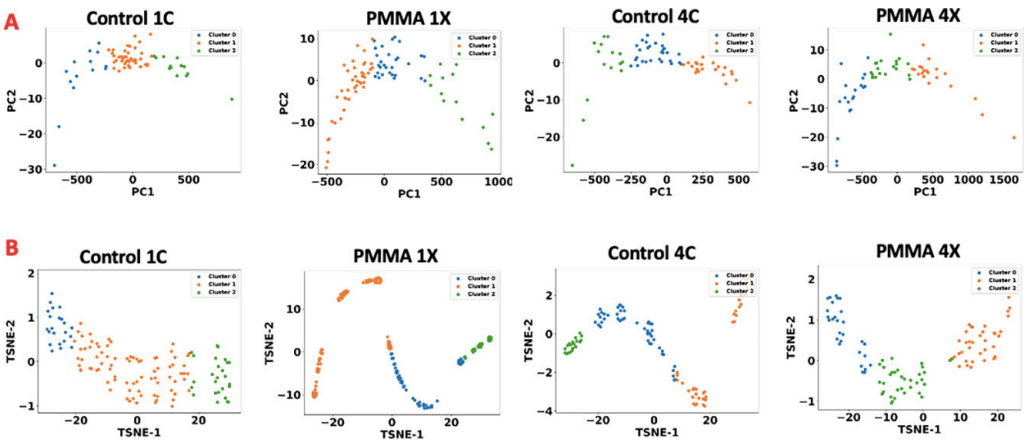

Fig. 10. (A) PCA and (B) t-SNE visualisations on primary dataset images

4.5 Geometric feature distributions

Figure 11 shows the distribution of geometric features across Control and PMMA groups using violin plots. PMMA cells shows greater variation in area, compactness, and perimeter, which indicates more pronounced morphological change. Control cells are more circular and closer to the hexagonal shape characteristic of healthy endothelium. These differences show that the proposed method is able to captures biologically meaningful morphological variation at the single cell level.

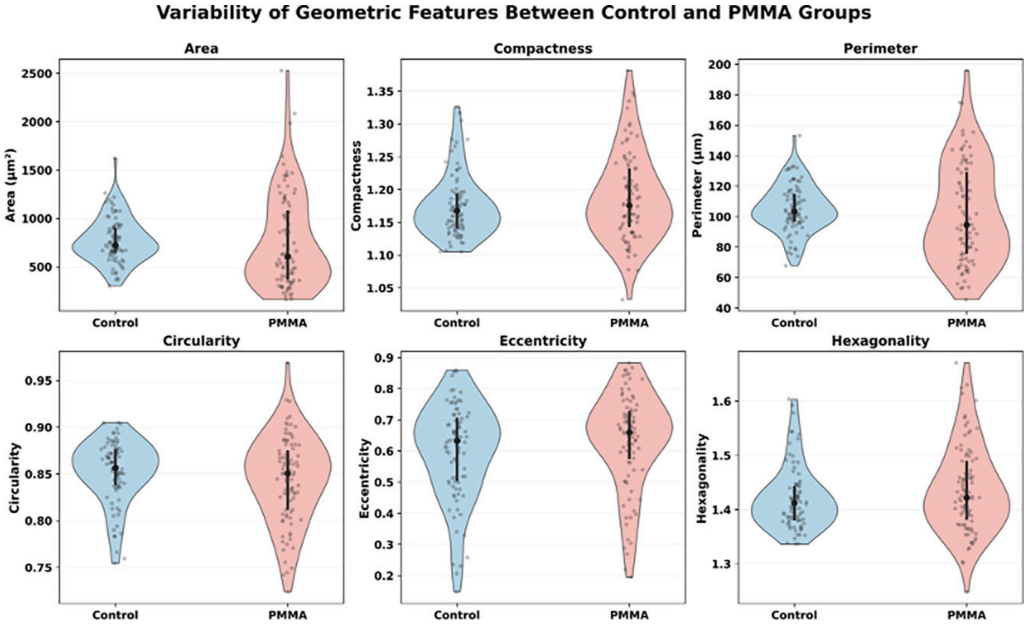

Fig. 11. Comparison of geometric feature distributions among Control and PMMA groups

4.6 Adjusted rand index (ARI) analysis

Table 6 compares ARI between different feature set clustering on the external dataset. The EFD Std + Zernike + Geometric and Geometric only feature sets produced almost identical groupings (ARI = 0.9961). Zernike features grouped the cells very differently from geometric features (ARI = 0.3754), indicating that they capture different aspects of cell shape. EFD only features produced a markedly different grouping from all other sets, reflecting their focus on boundary detail that is not captured by region based descriptors.

Table 6. ARI between feature set clustering assignments on the external dataset (35,890 cells)

Feature Set A	Feature Set B	ARI	Interpretation
Geometric Only	EFD Std + Zernike + Geometric	0.9961	Same grouping
EFD Std + Zernike + Geometric	EFD Custom + Zernike + Geometric	1.0000	Same grouping
Zernike Only	Geometric Only	0.3754	Different complementary info
EFD Std Only	Geometric Only	–0.003	Different boundary level info
EFD Std Only	Zernike Only	0.010	Different boundary level info

4.7 Segmentation count accuracy

Table 7 reports cell count accuracy on the 324 Daniel images with ground truth. The pipeline detected 30,272 cells against 42,147 ground truth (under count 28.2%). Median per-image count error was 50.76%, with 62 images (19.1%) achieving errors below 20%. This variability reflects the challenge of fixed parameter watershed on variable density images and represents a known limitation discussed in Section 5.

Table 7. Segmentation count accuracy on external validation dataset

Metric	Value
Images evaluated	324/385
Total ground truth cells	42,147
Total predicted cells	30,272 (under count 28.2%)
Median count error per image	50.76%
Images with count error < 20%	62/324 (19.1%)
Images with count error < 10%	32/324 (9.9%)

5 DISCUSSION

The most important finding of this work is that EFD based shape features produced stable and repeatable clustering results across two datasets that differed greatly in size and acquisition conditions. On the four image primary dataset, every feature set ranked exactly the same order across all four images, with EFD Standard and Geometric features achieving the highest silhouette scores (0.6213 and 0.6215 respectively) and Zernike only features scoring lowest (0.4889). The practical

significance of these differences was large in all cases: the effect size between EFD Standard and Traditional metrics shows Cohen's d = 5.290, and between EFD Standard and Zernike it was d = 6.915. Effect sizes of this magnitude confirm that the performance gaps are genuine and not the result of random variation in a small dataset.

When the same feature sets were applied to the 385 images external dataset, a Friedman test confirmed that clustering quality varied significantly depending on which features were used (chi-square = 1298.16, $p < 0.001$). EFD Standard and geometric features each showed only a small drop in silhouette score compared to the primary dataset results, with differences of 0.006 to 0.009 respectively. This small gap demonstrates that these feature sets describe aspects of cell shape that are consistent across different imaging conditions and cell density levels. Traditional clinical features, by contrast, fell sharply from a silhouette score of 0.5986 on the primary dataset to 0.3628 on the external dataset, a statistically significant drop (Wilcoxon, $p < 0.001$) that highlights a fundamental limitation of population level metrics when used outside the narrow conditions for which they were designed.

The large effect size (Cohen's d = 4.889) indicates a strong difference between EFD and Zernike features on the external dataset. Several alternative explanations for the weaker Zernike performance are possible and cannot be fully disentangled from the present data. Firstly, Zernike moments encode global region level intensity patterns over a circular support. They are less sensitive to fine boundary undulations than EFDs, which directly parametrise the contour. For endothelial cells, the diagnostically important variations (polymegethism and pleomorphism) manifest primarily as changes in boundary shape and side count rather than in interior intensity, and the encoding mismatch is therefore plausible. Secondly, Zernike moments are sensitive to small segmentation errors and to the choice of radius parameter. With variable cell sizes and the fixed 30-pixel radius used in this study, some of the underperformance may reflect a parameter selection effect rather than an intrinsic limitation of the descriptor. Thirdly, the orthogonality of Zernike moments holds only over a continuous unit disk, and discretisation on small cell images introduces approximation error that may further reduce their discriminative power on this task. Future work should explore adaptive Zernike radius selection and a denoised contour pipeline before drawing stronger conclusions about the descriptor itself.

It should be noted that on the primary dataset, geometric features alone achieved a silhouette score of 0.6215, virtually identical to the EFD combined feature set at 0.6213, a difference of only 0.0002. This means that on the small primary dataset, the EFD combined set does not visibly outperform the geometric descriptors alone. However, the advantage of the combined EFD representation becomes more apparent at scale on the large external dataset, where it performed consistently with performance difference 0.009. It suggests that the EFDs feature set captures stable morphological information.

ARI analysis indicates that EFD features capture different morphological information than the other feature sets. Future work is required to investigate whether EFD provides additional biological information beyond these metrics.

Regarding comparison with existing work, the Daniel et al. [24] dataset forms the external validation corpus of this study. The work focused exclusively on automatic cell segmentation and density estimation using U-Net, and the authors explicitly note that the cell centroids generated by their pipeline could be used in a downstream classification step to differentiate normal and pathological endothelial cells but did not themselves implement such a step. The present study addresses this open gap

by applying EFD based unsupervised clustering to the same dataset. Three morphologically distinct phenotypes are identified with a mean silhouette score of 0.6119. To the best of the author's knowledge, this is the pioneer study which perform unsupervised morphological clustering with silhouette score based internal evaluation on a corneal endothelial dataset of this scale (35,890 cells across 385 images). The EFD based silhouette scores changed only marginally between the primary (0.6213) and external (0.6119) datasets. It proves that the proposed method generalises well to new data.

The three clusters identified in this study shows direct biological interpretation based on the literature of corneal endothelial morphology. Cluster 0 comprising 64.4% cells indicates the majority of cells are regular, compact and nearly hexagonal same as the healthy cornea. In healthy cornea, approximately 70–80% of endothelial cells exhibit a hexagonal shape [36]. Cluster 1 comprising 29.2% of cells with moderate morphological irregularity. It indicates intermediate polymegethism arising from mild corneal stress or age. Cluster 2 comprising 6.4% of cells, captures elongated and irregular cells. It indicates severe polymegethism and pleomorphism. It is observed that PMMA images contain a higher proportion of Cluster 1 and Cluster 2 cells compared to Control images, consistent with prior findings [37]. This observation should nevertheless be treated with caution, since only two PMMA and two Control images were available in the primary dataset. The percentages reported are specific to the primary dataset and segmentation settings, and confidence intervals would require a larger dataset to compute reliably.

This observation is therefore presented as preliminary evidence requiring validation on a larger and independently acquired dataset.

The Carlson et.al [38] noted that prolonged PMMA contact lens wear is well documented to induce corneal hypoxia, leading to polymegethism and pleomorphism, without necessarily reducing overall cell density. MacRae et al. [39] illustrate that long term PMMA wearers show significantly increased polymegethism (CV 0.32 vs. 0.26 in age matched controls, $p < 0.01$) and significantly shows severe pleomorphism (frequency of hexagons < 40%), which corresponds precisely to the morphological profile of Cluster 2 identified in the PMMA images. The EFD based pipeline detected these three biologically meaningful phenotypes automatically, without per cell manual labelling [40, 41]. The pipeline processed 35,890 cells in approximately 30 minutes on the external dataset, a throughput that supports its use on datasets of a size that would be impractical to label by hand.

The segmentation pipeline showed a median per image counting error of about 50.76% on the external dataset. It is because of a single fixed watershed parameter configured across images with widely varying cell densities. The clustering results remained statistically reliable in spite of this error, indicating that the EFD based feature extraction and grouping is not heavily affected by the segmentation noise present in this dataset, which is important for practical use in clinical analysis. The Ruggeri et al. [15] and Fabijańska [16] focused on cell counting and segmentation. Vigueras-Guillén et al. [17] used a supervised CNN to achieve count errors below 5% but did not perform unsupervised morphological phenotyping or report cross dataset validation. Daniel et al. [24] released the U-Net-segmented external corpus used here but did not undertake downstream classification. The present work extends these studies in two respects. First, EFD based shape descriptors are combined with unsupervised clustering on a 35,890 cells dataset. Second, cross dataset silhouette stability is reported with Friedman and Wilcoxon statistical evidence.

6 CONCLUSION

This work developed and statistically validated a morphological framework for clustering corneal endothelial cells into biologically meaningful phenotypes. It tested across a total of 36,069 cells drawn from two independent datasets. On the primary dataset, EFD Standard features produced a mean silhouette score of 0.6213 and ranked consistently highest across all images, with a large practical advantage over traditional clinical metrics (Cohen's d = 5.290). Applying the same approach to the large external dataset confirmed these findings are statistically significant. The Friedman test showed that the feature set choice significantly affected clustering quality as chi-square = 1298.16, $p < 0.001$. EFD Standard features outperformed Zernike with Wilcoxon $p < 0.001$ and Cohen's d = 4.889. The traditional clinical metrics showed a large and statistically significant performance drop at scale with Wilcoxon $p < 0.001$. The three identified cell clusters align with published descriptions of normal, early stressed, and severely stressed endothelial cell morphology. PMMA images containing a higher proportion of the latter two groups (moderately regular cells and irregular or elongated cells) as expected from the known biological effects of that lens type. Future work would include developing adaptive segmentation that adjusts to varying cell densities, exploring deep learning based cell detection at a pre-processing step, and carrying out formal clinical validation of the identified phenotypes through collaboration with clinical experts to confirm its usefulness for practical real world applications.

7 ACKNOWLEDGMENTS

The authors gratefully acknowledge Sri Mata Amritanandamayi Devi (Amma), Chancellor, Amrita Vishwa Vidyapeetham, for her inspiration and for providing financial support for the article processing charges (APC) for this publication. The authors sincerely thank Dr. Sangly P. Srinivas, Indiana University, Bloomington, USA, and Dr. Surekha Paneerselvam, Department of Electrical and Electronics Engineering, Amrita School of Engineering, Bengaluru, Amrita Vishwa Vidyapeetham, India, for their valuable guidance, insightful discussions, and continued encouragement throughout this work.

8 REFERENCES

[1] L. J. Jeang, C. E. Margo, and E. M. Espana, "Diseases of the corneal endothelium," *Experimental Eye Research*, vol. 205, p. 108495, 2021. https://doi.org/10.1016/j.exer.2021.108495

[2] J. P. Vigueras-Guillén *et al.*, "DenseUNets with feedback non-local attention for the segmentation of specular microscopy images of the corneal endothelium with guttae," *Scientific Reports*, vol. 12, no. 1, p. 14035, 2022. https://doi.org/10.1038/s41598-022-18180-1

[3] N. Tananuvat and N. Khumchoo, "Corneal thickness and endothelial morphology in normal Thai eyes," *BMC Ophthalmology*, vol. 20, no. 1, p. 167, 2020. https://doi.org/10.1186/s12886-020-01385-1

[4] B. Mohd-Ali and L. Y. Chen, "The morphology of corneal endothelial cells in long term soft contact lens wearers in Kuala Lumpur," *Contact Lens and Anterior Eye*, vol. 44, no. 1, pp. 72–75, 2021. https://doi.org/10.1016/j.clae.2020.06.007

[5] A. Consejo, I. Trillo-Moreno, and L. Remon, "Corneal tissue changes following short-term soft contact lens wear of different materials," *Ophthalmic and Physiological Optics*, vol. 43, no. 1, pp. 35–45, 2023. https://doi.org/10.1111/opo.13062

[6] I. Gustafsson, "The assessment of disease progression in keratoconus and corneal cross-linking in thin corneae," *Ph.D. Dissertation, Lund Univ., Lund*, Sweden, 2021.

[7] A. R. Doke, S. Palaniswamy, and S. P. Srinivas, "Automated detection of tight junction damage in corneal endothelium using machine learning," in *Proc. 4th Int. Conf. Ubiquitous Computing and Intelligent Information Systems (ICUIS)*, 2024, pp. 675–679. https://doi.org/10.1109/ICUIS64676.2024.10867247

[8] P. Gain *et al.*, "Global survey of corneal transplantation and eye banking," *JAMA Ophthalmology*, vol. 134, no. 2, pp. 167–173, 2016. https://doi.org/10.1001/jamaophthalmol.2015.4776

[9] T. Ono, Y. Mori, R. Nejima, T. Iwasaki, T. Miyai, and K. Miyata, "Corneal endothelial cell density and morphology in ophthalmologically healthy young individuals in Japan: An observational study of 16842 eyes," *Scientific Reports*, vol. 11, no. 1, p. 18224, 2021. https://doi.org/10.1038/s41598-021-97776-5

[10] B. Selig, K. A. Vermeer, B. Rieger, T. Hillenaar, and C. L. Luengo Hendriks, "Fully automatic evaluation of the corneal endothelium from in vivo confocal microscopy," *BMC Medical Imaging*, vol. 15, no. 1, p. 13, 2015. https://doi.org/10.1186/s12880-015-0054-3

[11] B. A. Benetz *et al.*, "Automatic determination of endothelial cell density from donor cornea endothelial cell images," *Translational Vision Science & Technology*, vol. 13, no. 8, p. 40, 2024. https://doi.org/10.1167/tvst.13.8.40

[12] J. P. Vigueras-Guillén, J. van Rooij, A. Engel, H. G. Lemij, L. J. van Vliet, and K. A. Vermeer, "Deep learning for assessing the corneal endothelium from specular microscopy images up to 1 year after ultrathin-DSAEK surgery," *Translational Vision Science & Technology*, vol. 9, no. 2, p. 49, 2020. https://doi.org/10.1167/tvst.9.2.49

[13] Y. Wang and W. Jiang, "An automatic classification and clustering algorithm for online learning goals based on cognitive thinking," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 13, no. 11, pp. 54–66, 2018. https://doi.org/10.3991/ijet.v13i11.9587

[14] M. Amane, K. Aissaoui, and M. Berrada, "Multi-label classification of learning objects using clustering algorithms based on feature selection," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 17, no. 20, pp. 248–260, 2022. https://doi.org/10.3991/ijet.v17i20.32323

[15] A. Ruggeri, E. Grisan, and J. Jaroszewski, "A new system for the automatic estimation of endothelial cell density in donor corneas," *British Journal of Ophthalmology*, vol. 89, no. 3, pp. 306–311, 2005. https://doi.org/10.1136/bjo.2004.051722

[16] A. Fabijańska, "Segmentation of corneal endothelium images using a U-Net-based convolutional neural network," *Artificial Intelligence in Medicine*, vol. 88, pp. 1–13, 2018. https://doi.org/10.1016/j.artmed.2018.04.004

[17] J. P. Vigueras-Guillén *et al.*, "Fully convolutional architecture vs sliding-window CNN for corneal endothelium cell segmentation," *BMC Biomedical Engineering*, vol. 1, no. 1, p. 4, 2019. https://doi.org/10.1186/s42490-019-0003-2

[18] B. J. Sowmya *et al.*, "A visual computing unified application using deep learning and computer vision techniques," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 18, no. 1, pp. 59–74, 2024. https://doi.org/10.3991/ijim.v18i01.42673

[19] F. P. Kuhl and C. R. Giardina, "Elliptic Fourier features of a closed contour," *Computer Graphics and Image Processing*, vol. 18, no. 3, pp. 236–258, 1982. https://doi.org/10.1016/0146-664X(82)90034-X

[20] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004. https://doi.org/10.1023/B:VISI.0000029664.99615.94

[21] M. J. Doughty, S. Jonuscheit, and N. F. Button, "Assessment of the reliability of endothelial cell-density estimates in the presence of pseudoguttata," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 250, no. 1, pp. 111–121, 2012. https://doi.org/10.1007/s00417-011-1812-8

[22] E. G. Stocker and J. P. Schoessler, "Corneal endothelial polymegathism induced by PMMA contact lens wear," *Investigative Ophthalmology & Visual Science*, vol. 26, no. 6, pp. 857–863, 1985. https://pubmed.ncbi.nlm.nih.gov/3891666/

[23] S. Sanchez, N. Vallez, G. Bueno, and A. G. Marrugo, "Data augmentation via warping transforms for modeling natural variability in the corneal endothelium enhances semi-supervised segmentation," *PLoS ONE*, vol. 19, no. 11, p. e0311849, 2024. https://doi.org/10.1371/journal.pone.0311849

[24] M. C. Daniel *et al.*, "Automated segmentation of the corneal endothelium in a large set of 'real-world' specular microscopy images using the U-Net architecture," *Scientific Reports*, vol. 9, no. 1, p. 4752, 2019. https://doi.org/10.1038/s41598-019-41034-2

[25] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*, P. S. Heckbert, Ed. San Diego, CA, USA: Academic Press, 1994, pp. 474–485. https://doi.org/10.1016/B978-0-12-336156-1.50061-6

[26] A. Kucharski and A. Fabijańska, "CNN-watershed: A watershed transform with predicted markers for corneal endothelium image segmentation," *Biomedical Signal Processing and Control*, vol. 68, p. 102805, 2021. https://doi.org/10.1016/j.bspc.2021.102805

[27] F. P. Kuhl and C. R. Giardina, "Elliptic Fourier features of a closed contour," *Computer Graphics and Image Processing*, vol. 18, no. 3, pp. 236–258, 1982. https://doi.org/10.1016/0146-664X(82)90034-X

[28] H. Iwata and Y. Ukai, "SHAPE: A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors," *Journal of Heredity*, vol. 93, no. 5, pp. 384–385, 2002. https://doi.org/10.1093/jhered/93.5.384

[29] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 850–863, 1993. https://doi.org/10.1109/34.232073

[30] G. Borgefors, "Hierarchical chamfer matching: A parametric edge matching algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 6, pp. 849–865, 1988. https://doi.org/10.1109/34.9107

[31] W. M. Bourne, "Biology of the corneal endothelium in health and disease," *Eye*, vol. 17, no. 8, pp. 912–918, 2003. https://doi.org/10.1038/sj.eye.6700559

[32] E. Alizadeh, S. M. Lyons, J. M. Castle, and A. Prasad, "Measuring systematic changes in invasive cancer cell shape using Zernike moments," *Integrative Biology*, vol. 8, no. 11, pp. 1183–1193, 2016. https://doi.org/10.1039/C6IB00100A

[33] S. Tadepalli *et al.*, "Blink-induced artifacts in aqueous flare measurements by EOG-based spot fluorometer and their reduction using unsupervised clustering," *Biomedical Signal Processing and Control*, vol. 96, p. 106486, 2024. https://doi.org/10.1016/j.bspc.2024.106486

[34] D. Chicco, A. Campagner, A. Spagnolo, D. Ciucci, and G. Jurman, "The silhouette coefficient and the Davies-Bouldin index are more informative than Dunn index, Calinski-Harabasz index, Shannon entropy, and gap statistic for unsupervised clustering internal evaluation of two convex clusters," *PeerJ Computer Science*, vol. 11, p. e3309, 2025. https://doi.org/10.7717/peerj-cs.3309

[35] M. J. Warrens and H. van der Hoef, "Understanding the adjusted Rand index and other partition comparison indices based on counting object pairs," *Journal of Classification*, vol. 39, no. 3, pp. 487–509, 2022. https://doi.org/10.1007/s00357-022-09413-z

[36] R. F. Mutwaly, "Corneal endothelial polymegathism and pleomorphism induced by daily-wear soft contact lenses," *Cureus*, vol. 16, no. 11, 2024. https://doi.org/10.7759/cureus.74187

[37] E. Karmiris, A. Tsiogka, K. Tsiripidis, E. Papakonstantinou, I. Georgalas, and E. Chalkiadaki, "Correlations of corneal endothelial morphology and corneal thickness with anterior segment parameters in healthy individuals," *Cornea*, vol. 43, no. 6, pp. 764–770, 2024. https://doi.org/10.1097/ICO.0000000000003515

[38] K. H. Carlson, W. M. Bourne, and R. F. Brubaker, "Effect of long-term contact lens wear on corneal endothelial cell morphology and function," *Investigative Ophthalmology & Visual Science*, vol. 29, no. 2, pp. 185–193, 1988. https://iovs.arvojournals.org/article.aspx?articleid=2177982

[39] S. M. Mac Rae, M. Matsuda, S. Shellans, and L. F. Rich, "The effects of hard and soft contact lenses on the corneal endothelium," *American Journal of Ophthalmology*, vol. 102, no. 1, pp. 50–57, 1986. https://doi.org/10.1016/0002-9394(86)90209-6

[40] S. Akshay, S. Dhanush, G. N. Aswin, and J. Amudha, "Eyehelp: Predicting an AOI based on eye gaze for patient assistance," *Procedia Computer Science*, vol. 258, pp. 1486–1495, 2025. https://doi.org/10.1016/j.procs.2025.04.381

[41] S. Tadepalli *et al.*, "Reliability of aqueous flare measurements during uveitis by a spot fluorometer," *Journal of Ocular Pharmacology and Therapeutics*, vol. 38, no. 1, pp. 66–73, 2022. https://doi.org/10.1089/jop.2021.0053

9 AUTHORS

Ashwini Ramchandra Doke is currently a research scholar with the Department of Computer Science and Engineering, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Bengaluru Campus. Her research interests include Deep Learning, Meta learning, AutoML (Automated Machine learning), Machine learning, medical image analysis and related topics (E-mail: BL.EN.R4CSE20004@bl.students.amrita.edu).

Dr. Suja Palaniswamy is a faculty member with the Department of Computer Science and Engineering, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Bengaluru, Karnataka, India. Her research interests include computer vision, medical informatics, and deep learning for diagnostic imaging. She continues to work on interdisciplinary research integrating artificial intelligence with medical imaging and computer aided diagnostic systems (E-mail: p_suja@blr.amrita.edu).