Res Intell Manuf Assem, 2025, 4(2): 239-265
DOI: 10.25082/RIMA.2025.02.002

RESEARCH ARTICLE

A Data-Driven Evaluation of ECD Measurement Techniques Across Traditional and AI-Based Modalities

Manghe Fidelis Obi[1] **Andy Officer**[2] **Shannon Schweitzer**[2] **Tarun Goswami**[1*]

[1] Department of Biomedical, Industrial and Human Factors Engineering, Wright State University, Dayton, OH, USA
[2] Lions Eye Bank of West Central Ohio, Dayton, OH, USA

Correspondence to: Tarun Goswami, Department of Biomedical, Industrial and Human Factors Engineering, Wright State University, Dayton, OH 45435, USA; Email: tarun.goswami@wright.edu

Received: April 14, 2025;
Accepted: June 28, 2025;
Published: July 2, 2025.

Citation: Obi MF, Officer A, Schweitzer S, et al. A Data-Driven Evaluation of ECD Measurement Techniques Across Traditional and AI-Based Modalities. *Res Intell Manuf Assem*, 2025, **4**(2): 239-265. https://doi.org/10.25082/RIMA.2025.02.002

Abstract: Accurate measurement of corneal endothelial cell density (ECD) is crucial in evaluating the viability of donor corneas for transplantation. The consistency of ECD measurements is critical for predicting post-transplant results and monitoring corneal health. However, measurement methods have evolved, moving from manual counting to more complex semi-automatic and fully automated systems, including AI-powered solutions. This study compares the accuracy, dependability, and efficiency of manual, semi-automated, and fully automated ECD measurement techniques. It investigates the degree of heterogeneity among techniques and evaluates their potential to improve clinical outcomes in corneal transplantation. The sample includes corneal data from 300 participants, 150 male and 150 female donors, who were divided into three groups based on the measurement method: manual, semi-automated, or fully automated. The study also examined the gender distribution to see whether there was any difference in results between male and female donor corneas. Manual counting has previously been notable for its variability due to operator expertise and calibration discrepancies, with mean ECD values ranging from 2146 to 2775 cells/mm^2 ($p < 0.05$). Semi-automated procedures, which combine manual input with software aid, enhance consistency. In the Cornea Preservation Time Study, eye banks reported a mean ECD of 2773 $\pm$ 300 cells/mm^2, while CIARC reported 2758 $\pm$ 388 cells/mm^2, with agreement limits ranging from [-644, 675] cells/mm^2 ($p < 0.05$). The AxoNet deep learning model had a mean absolute error (MAE) of 12.1 cells/mm^2 and an R^2 value of 0.948, making it the most accurate fully automated system. A separate study on AI-based detection of aberrant endothelium cells achieved an accuracy of 0.95, precision of 0.92, recall of 0.94, and F1 score of 0.93, and an AUC-ROC of 0.98 ($p < 0.01$). Fully automated AI-based methods surpass manual and semi-automated approaches in accuracy and consistency, significantly reducing time and labor. The findings highlight the importance of adopting AI-driven technologies to enhance diagnostic precision and efficiency in clinical settings. However, the need for standardized calibration procedures and high-quality image acquisition remains critical for reliable ECD measurement.

Keywords: corneal endothelial cell density (ECD), manual cell counting, semi-automated cell counting, fully automated cell counting, deep learning, image analysis, AI in ophthalmology

1 Introduction

In ophthalmology, accurately measuring corneal endothelial cell density (ECD) has long been an essential method for monitoring the health of the corneal endothelium, a vital layer of cells responsible for maintaining corneal transparency and fluid balance. Historically, experienced personnel made this measurement manually using specular microscopy, a technique invented and refined in the mid-twentieth century. The manual method is visually analyzing and counting corneal endothelium cells under a microscope, which takes a high level of ability and experience to ensure accuracy. However, this method is labor-intensive, time-consuming, and prone to error due to disparities in technician expertise and subjective assessment [1, 2]. As the field of ophthalmology advanced, the development of semi-automated devices in the late twentieth century represented a significant advancement. These systems combine manual input with automated image analysis software, allowing personnel to identify specific places or areas of interest while the software aids in the count and calculation of the ECD. Semi-automated methods sought to reduce technician workload and enhance results consistency, but human intervention was still required to remedy errors or ambiguities detected by the

program [3]. Despite these advances, semi-automated systems continue to confront issues in picture quality, calibration methods, and uniformity across diverse clinical contexts [4]. In recent years, the emergence of fully automated approaches, fueled by developments in artificial intelligence (AI) and deep learning, has transformed ECD measurements. These systems use deep learning architectures, such as convolutional neural networks (CNNs), to evaluate images of the corneal endothelium and autonomously count cells with little or no human interaction. Fully automated procedures provide the highest levels of accuracy and dependability, while greatly lowering the time and manpower necessary for cell counting. These technologies can manage massive datasets, do speedy analyses, and produce objective and consistent results, reducing the possibility of human mistake and subjective bias [5, 6].

The corneal endothelium plays a crucial role in maintaining corneal clarity by regulating fluid balance within the cornea. Damage or loss of endothelial cells can lead to corneal edema and vision loss, making accurate assessment essential for diagnosing and managing various corneal diseases, evaluating donor corneas for transplantation, and monitoring the impact of surgical interventions and other treatments [7]. As such, reliable ECD measurement is crucial for successful corneal transplants and corneal health management.

The corneal endothelium regulates fluid equilibrium within the cornea, which is essential for preserving corneal clarity. Endothelial cell damage or loss can cause corneal edema and vision loss; therefore, a correct assessment is critical for identifying and managing various corneal disorders, screening donor corneas for transplantation, and monitoring the impact of surgical interventions and other treatments [7]. As a result, the ability to assess ECD consistently is critical for assuring corneal transplant success and managing corneal health in general.
The corneal endothelium regulates fluid equilibrium within the cornea, which is essential for preserving corneal clarity. Damage or loss of endothelial cells can lead to corneal edema and vision loss, making accurate assessment essential for diagnosing and managing various corneal diseases, evaluating donor corneas for transplantation, and monitoring the impact of surgical interventions and other treatments [7]. As such, the ability to reliably measure ECD is vital for ensuring the success of corneal transplants and for the overall management of corneal health.

The corneal endothelium plays a crucial role in maintaining corneal clarity by regulating fluid balance within the cornea. Damage or loss of endothelial cells can lead to corneal edema and vision loss, making accurate assessment essential for diagnosing and managing various corneal diseases, evaluating donor corneas for transplantation, and monitoring the impact of surgical interventions and other treatments [7]. As such, the ability to reliably measure ECD is vital for ensuring the success of corneal transplants and for the overall management of corneal health. The evolution of corneal endothelial cell density (ECD) measurement methods highlights the pivotal role of mathematical frameworks and programming tools in improving accuracy, reliability, and efficiency. Advanced techniques such as convolutional neural networks (CNNs), Gradient Tree Boosting (GTB), Support Vector Regression (SVR), and Ridge Regression (RR) utilize statistical and computational principles to enhance the precision of ECD assessments. These frameworks form the foundation for modern automated systems, reducing human error and ensuring reproducibility across diverse datasets.

This paper aims to provide a comprehensive comparison of manual, semi-automated, and fully automated methods for ECD measurement. By evaluating the accuracy, reliability, and efficiency of these methods, this paper aims to highlight the benefits and limitations of each approach, with a focus on the potential of AI-driven tools to improve clinical outcomes. Given the critical role of ECD measurement in ophthalmology, it is essential to assess the effectiveness of these methods to ensure that the most accurate and reliable techniques are adopted in clinical practice [8]. This evaluation is necessary for enhancing diagnostic accuracy, optimizing patient care, and improving the success rates of corneal surgeries and transplants.

This study's hypothesis is that fully automated, AI-based methods for ECD measurement will demonstrate superior accuracy, reliability, and efficiency compared to manual and semi-automated methods, particularly in large datasets and clinical applications. These AI-driven tools are expected to reduce human error and variability, providing more consistent and objective results across different clinical settings.

2 Methodology

2.1 Data Collection

This study evaluated and compared manual, semi-automated, and fully automated methods for measuring corneal endothelial cell density (ECD) by analyzing data from a wide range of

scientific literature. Data were collected from relevant studies identified through a comprehensive search of key databases, including PubMed, IEEE Xplore, ScienceDirect, and Google Scholar. The search strategy involved using specific keywords, such as "corneal endothelial cell density," "manual counting," "semi-automated counting," "fully automated counting," "artificial intelligence in ophthalmology," and "deep learning for cell counting," to ensure the inclusion of studies that covered both traditional and advanced methods for ECD measurement. The inclusion criteria focused on studies published from the 1980s to the present, capturing the evolution of ECD measurement techniques from manual counting to fully automated, AI-based systems. Studies were selected based on their relevance and the quality of the data provided, ensuring a comprehensive representation of the strengths and limitations of each method.

For studies where graphical data were presented, a web plot digitizer tool was used to extract data points directly from images of graphs. This process involved uploading an image of the graph, setting the axes, and marking data points either manually or automatically. The software then generated the coordinates of the marked points, allowing for the accurate regeneration of data from graphical figures. For programmatically generated plots, Python libraries such as matplotlib and plotly were used to regenerate data points from saved figures. If the original plot data was available, these libraries allowed for the direct extraction of data. Additionally, image processing libraries like OpenCV were employed to extract data from images when the original data was not accessible, enabling precise data recovery from visual representations.

2.2 Statistical Analysis

Once the data points were extracted, the analysis was conducted using Python, utilizing statistical analysis libraries such as pandas, numpy, matplotlib, and plotly for data visualization and comparison. These tools facilitated the computation of key metrics, including mean ECD values, mean absolute error (MAE), R^2 values, standard deviations, and variability metrics reported in the studies. The analysis focused on evaluating the differences in performance between manual, semi-automated, and fully automated ECD measurement methods, analyzing factors such as variability, accuracy, and reproducibility to gain insights into the clinical applicability of each method. Special attention was given to the advantages and limitations of each approach, particularly the impact of human intervention in manual and semi-automated methods versus the consistency and objectivity provided by fully automated systems [9].

Data collected from studies revealed that manual counting yielded a mean ECD value of approximately 2450 cells/mm^2, though with high variability due to technician expertise, subjective judgment, and calibration inconsistencies. Semi-automated counting showed a higher mean ECD value of around 2773 cells/mm^2, with reduced variability attributed to software assistance standardizing the counting process. Fully automated AI-based systems, such as AxoNet, demonstrated the highest mean ECD value of 2867 cells/mm^2, with minimal variability and high accuracy, reflected in low MAE and high R^2 values. These findings highlight the progression of ECD measurement techniques, with fully automated methods consistently providing superior accuracy and reproducibility. The analysis, supported by data extraction tools and Python-based methodologies, underscores the importance of adopting fully automated systems for ECD measurement in clinical settings where precision, consistency, and reliability are crucial for effective outcomes [10].

3 Manual Counting of Corneal Endothelial Cells

Manual counting of corneal endothelial cells is a meticulous process that requires precision and consistency. The process begins with the capture of high-resolution images of the corneal endothelium using a specular microscope. Once the image is acquired, a trained technician manually counts the endothelial cells within a defined image area by marking each cell. The endothelial cell density (ECD) is then calculated by dividing the number of cells by the area being examined, typically expressed in cells per square millimeters (cells/mm^2). This approach, while commonly used, relies heavily on the technician's skill and experience.

3.1 Variability in Manual Counting

A significant challenge with manual counting is the inherent variability that arises from several factors. Technician experience plays a major role, as more experienced technicians tend to produce more consistent results, while less experienced ones may struggle with interpreting cell borders, leading to discrepancies [3]. The subjective judgment also contributes to variability, as different technicians may count the same image differently based on their interpretation of

the cell boundaries. Additionally, image quality, the calibration of the specular microscope, and the specific counting strategy such as whether the cells are counted in the central or peripheral regions of the cornea can introduce further variability. These inconsistencies highlight the need for standardized protocols to reduce both inter-observer and intra-observer variability [4, 11].

3.2 Protocols for Manual Counting

To address these sources of variability, several protocols have been developed to standardize the manual counting process. These protocols provide detailed guidelines on how to acquire images, emphasizing the need for consistent lighting and focus to ensure image clarity. They also suggest specific counting strategies, such as focusing on cells in the central cornea, where cell size and density are less variable compared to the periphery. Regular calibration of the specular microscope is another essential component of these protocols to ensure that measurements remain accurate and consistent across multiple sessions. By following these standardized procedures, the reliability and accuracy of manual counting can be significantly improved.

3.3 Importance of Cell Count and Density

The density of endothelial cells in the cornea is a critical indicator of corneal health. A minimum density of endothelial cells is required to maintain corneal transparency and regulate fluid balance by controlling the exchange of fluids and nutrients between the cornea and the aqueous humor. If the ECD falls below a certain threshold, the cornea may begin to swell, leading to corneal edema, reduced transparency, and eventual vision loss. In the context of corneal transplantation, ECD is a crucial factor in determining the viability of donor corneas. Corneas with low ECD are more prone to graft failure because they cannot maintain clarity after transplantation. Therefore, accurately measuring ECD is essential in selecting donor tissue and ensuring the success of corneal transplants. Consistent ECD measurements help to reduce the risk of complications after transplantation, improving patient outcomes.

Research has revealed that significant variability exists in manual counting methods among eye banks. One study, which examined the practices of 22 French eye banks, found notable discrepancies in ECD results due to differences in calibration practices, counting strategies, and technician expertise. The mean ECD values ranged from 2146 to 2775 cells/mm^2, with eye banks that did not regularly calibrate their equipment generally reporting higher ECD values [5, 12]. These findings underscore the importance of standardizing both calibration procedures and counting methodologies to enhance the accuracy and reliability of ECD measurements.

Further studies have also examined the impact of other factors on manual counting accuracy. One investigation highlighted the effect of age on the corneal endothelium, noting that ECD decreases with age. This factor is especially important when evaluating donor corneas for transplantation, as older corneas may have lower cell densities, increasing the risk of graft failure. When manual counting methods are used in such evaluations, the natural variability associated with age-related changes in ECD can complicate the assessment. This reinforces the need for consistent, standardized methods that can account for such variability.

Another study assessed the reliability of human corneal endothelial cell-density estimates obtained using a non-contact specular microscope. The researchers found significant variability in the results due to differences in technician experience and counting approaches. This variability further emphasizes the critical importance of standardized protocols and training to ensure consistent and accurate ECD measurements. Implementing these standardized methods can greatly reduce technician-related variability and improve the reliability of the results in both clinical and research settings [6, 13].

In summary, manual counting of corneal endothelial cells, while widely used, is subject to variability due to technician experience, subjective judgment, image quality, and calibration practices. However, by adhering to standardized protocols and ensuring consistent training, these challenges can be mitigated, leading to more accurate and reliable ECD measurements. Accurate assessment of endothelial cell density is essential for determining corneal health, selecting viable donor tissue, and improving the success rate of corneal transplants [7, 14].

Manual counting relies on fundamental statistical calculations to derive ECD values. Variability in manual methods is often quantified using the standard deviation (σ\sigma) and mean (μ\mu) of ECD measurements:

$$\text{ECD} = \frac{\text{Number of Cells Counted}}{\text{Area Examined (mm}^2\text{)}} \tag{1}$$

Variability in manual counting arises due to differences in technician experience, subjective interpretation, and calibration inconsistencies. To quantify this variability, the standard deviation (σ\sigma) and mean (μ\mu) of ECD measurements are used to express relative variability:

$$\text{Variability} = \frac{\sigma}{\mu} \times 100\% \tag{2}$$

To assess the reliability of manual measurements, confidence intervals are often employed:

$$\mu \pm \text{Z} \cdot \frac{\sigma}{\sqrt{\text{n}}} \tag{3}$$

where Z is the critical value from a normal distribution, and n is the sample size.

3.4 Corneal Endothelial Cell Density (ECD) Measurement: Semi-Automated *vs*. Fully Automated Methods

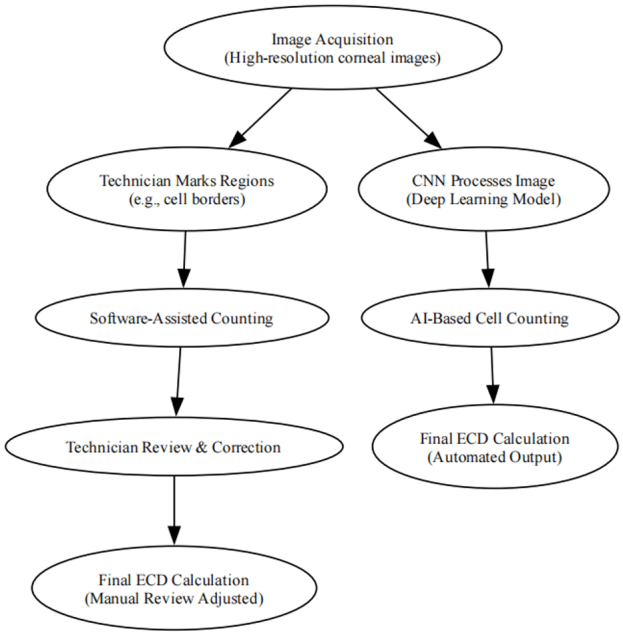

Figure 1 The flowchart of Corneal Endothelial Cell Density (ECD) Measurement: Semi-Automated *vs*. Fully Automated Methods

4 Semi-Automated Counting Methods

Semi-automated methods combine human input with automated software analysis, offering a balance between manual techniques and fully automated approaches. This approach allows for faster and more consistent measurements than traditional manual counting, yet still involves a technician's essential role in the initial steps. In the semi-automated process, the technician begins by capturing high-resolution images of the corneal endothelium using specular microscopy. Following image acquisition, the technician manually marks key areas of interest, such as cell borders or specific regions of the image, which the software will then use as reference points for counting the cells and calculating the endothelial cell density (ECD) [8].

The software assists in completing the counting process by analyzing the regions marked by the technician and providing an estimated ECD. However, technician oversight remains critical, as the software may occasionally misinterpret certain features, such as overlapping cells or poorly defined cell borders. In such instances, the technician steps in to review and correct the software's output, ensuring the final ECD calculation is accurate. For example, if the software misclassifies certain cells or artefacts in the image, manual intervention by the technician is necessary to avoid inaccuracies [9].

Image quality is a crucial factor in the accuracy of semi-automated systems. Poor-quality images can result in misinterpretation or incomplete counting by the software, leading to variability in results. The technician's input also plays a role in variability, as different technicians

may choose to mark different regions of the cornea, depending on their experience or interpretation. Additionally, calibration differences between equipment across institutions may lead to further inconsistencies. Despite these potential challenges, semi-automated methods generally exhibit less variability and greater reproducibility than fully manual techniques, especially when high-quality images are used and standardized protocols are followed.

Studies have demonstrated that semi-automated techniques provide consistent and reliable results, particularly in large-scale research and clinical applications. For instance, in a comparison of eye bank measurements and centralized automated readings, semi-automated systems produced mean ECD values that were closely aligned with those obtained through fully automated methods. However, some variability was still observed, highlighting the importance of standardized imaging protocols and consistent technician training to further reduce discrepancies [8].

The strengths of semi-automated systems lie in their ability to reduce human error while still maintaining the technician's oversight, allowing for greater accuracy than fully manual techniques. However, variability persists due to technician involvement and image quality. Semi-automated systems are particularly useful in cases where human intervention is needed, such as when complex or ambiguous features in the image must be interpreted. As a result, they offer a robust solution in clinical environments where both accuracy and human input are critical [9]. Semi-automated methods combine manual oversight with algorithmic processing, often leveraging regression models to improve consistency. Ridge Regression, a common tool in these systems, minimizes error while regularizing model parameters to avoid overfitting:

4.1 Linear Regression Model

Early semi-automated models utilized linear regression to predict ECD based on extracted image features:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \epsilon \tag{4}$$

where y represents the predicted ECD, xi are the image features, βi are the regression coefficients, and ϵ epsilon is the error term.

4.2 Ridge Regression for Regularization

To reduce overfitting and improve stability in measurements, Ridge Regression ($L2$ regularization) was introduced:

$$\min_w \|y - Xw\|^2 + \lambda \|w\|^2 \tag{5}$$

where X is the input feature matrix, w represents the model coefficients, and λ\\lambda is the regularization parameter that controls the penalty term.

4.3 Support Vector Regression (SVR) for Non-Linearity

With the advancement of machine learning, Support Vector Regression (SVR) was employed to handle non-linear relationships in ECD measurement:

$$\min_{w,b} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \max\left(0, \left|y_i - w^T x_i - b\right| - \epsilon\right) \tag{6}$$

SVR utilizes kernel functions to project data into a higher-dimensional space:

$$K(x, z) = \phi(x)^T \phi(z) \tag{7}$$

allowing for more accurate predictions in complex datasets.

5 Fully Automated Counting Methods and AI

Fully automated methods for measuring corneal endothelial cell density leverage cutting-edge artificial intelligence (AI)techniques, particularly convolutional neural networks (CNNs), to process corneal images without human intervention. These AI-driven systems are designed to handle large datasets and generate highly accurate predictions with minimal error. By eliminating the need for manual input, fully automated systems provide faster, more consistent, and objective results than both manual and semi-automated methods [13, 14].

In the AI-based approach, high-resolution images of the cornea are fed into a deep learning model, typically a CNN. These networks automatically learn the key features from the images,

such as detecting cell borders and counting the number of cells present. Often, transfer learning techniques are employed to fine-tune the models using pre-trained networks, which improves accuracy even when working with smaller datasets [15]. The deep learning models used in these systems consist of multiple components, including input layers for feeding the image, convolutional layers for feature extraction, pooling layers for down-sampling, and fully connected layers for making predictions.

The training of these neural networks involves using large datasets of labeled images, where the ground truth is known. During training, the network minimizes prediction errors by adjusting its internal parameters through back propagation. Factors such as learning rates, batch sizes, and the number of training epochs influence how well the network learns and adjusts. Optimization algorithms like Adam or SGD are commonly used to fine-tune the accuracy of the models. The final output from the network is typically expressed as cells per mm^2, providing an estimate of cell density [16].

In terms of performance, AI-based methods have demonstrated superior accuracy and reliability. For example, a deep learning-based tool for Assessing corneal endothelial cell density: automated versus manual counting methods, with an MAE of 12.1 cells/mm^2 and an R^2 value of 0.948. outperforming existing tools in this field. Similarly, in the context of cancer cell counting, AI models enhanced with transfer learning achieved an MAE of 12 ± 15, demonstrating the adaptability of AI across different domains of cell counting [17].

One of the primary advantages of fully automated systems is their ability to reduce variability by removing human oversight entirely. These systems do not rely on technician input, thereby eliminating subjective errors caused by differences in judgment or experience. By providing consistent and objective results, AI-based methods significantly reduce the potential for human error, delivering highly reproducible measurements across a variety of clinical applications [19]. For example, studies using U-Net-based CNNs for segmenting corneal endothelium images have shown high segmentation accuracy, even in challenging images, further reinforcing AI's potential in this domain [20]. Fully automated methods employ advanced frameworks like CNNs for image analysis, utilizing optimization techniques to minimize prediction errors. The Mean Squared Error (MSE) loss function is commonly used: and the shift from traditional machine learning models to ensemble learning and deep learning architectures significantly improved accuracy and robustness in ECD measurement.

5.1 Gradient Tree Boosting (GTB)

Gradient Tree Boosting (GTB) emerged as an effective approach for structured data analysis. It works by sequentially training weak learners (decision trees), where each new tree corrects the residuals of the previous model.

$$Fm + 1(x) = F_m(x) + \eta \cdot \arg\min_{g} \sum_{i=1}^{n} \mathcal{L}\left(y_i, F_m\left(x_i\right) + g\left(x_i\right)\right) \tag{8}$$

where Fm(x) is the model at iteration m, η is the learning rate, and $\mathcal{L}$ is the loss function.

The optimization of GTB is performed using gradient descent on residuals:

$$g_i = -\frac{\partial \mathcal{L}\left(y_i, F\left(x_i\right)\right)}{\partial F\left(x_i\right)} \tag{9}$$

While GTB provides robust performance, it is primarily suited for structured datasets and does not efficiently handle high-dimensional image data, leading to the rise of convolutional neural networks (CNNs).

5.2 Convolutional Neural Networks (CNNs) and Deep Learning

Deep learning transformed automated ECD measurement by using CNNs, which excel at extracting spatial features from images. CNNs consist of multiple layers, each performing a specific transformation on the input data.

5.3 Convolutional Operation

CNNs utilize convolutional layers to extract features from input images. The convolution operation is mathematically expressed as:

$$S(i,j) = m \sum n \sum I(i+m, j+n) \cdot K(m,n) \tag{10}$$

Where I is the input image, K is the convolutional filter, and S (i, j) is the resulting feature map.

5.4 Activation Function for Non-Linearity

ReLU (Rectified Linear Unit) is commonly used to introduce non-linearity:

$$f(x) = \max(0, x) \tag{11}$$

This ensures that negative values are set to zero, improving network efficiency.

5.5 Pooling Operation for Dimensionality Reduction

Max pooling is used to reduce spatial dimensions while retaining essential features:

$$P(i, j) = \max\{S(i + m, j + n) \mid 0 \leq m, n < k\} \tag{12}$$

where k represents the pooling window size.

5.6 Optimization via Backpropagation

The CNN model is optimized using backpropagation, which adjusts the model's parameters by computing the derivative of the loss function:

$$\frac{\partial L}{\partial w_i} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial w_i} \tag{13}$$

Weights are updated iteratively using gradient descent

$$w_{t+1} = w_t - \eta \cdot \nabla L(w) \tag{14}$$

where η\eta is the learning rate and (Lw) is the loss function.

5.7 Loss Function: Mean Squared Error (MSE)

For regression tasks in ECD measurement, the Mean Squared Error (MSE) loss function is used:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{15}$$

where y_i are the actual values and $\hat{y}_i$ are the predicted values.

6 Comparison and Performance Analysis

Both semi-automated and fully automated methods for ECD measurement involve a similar workflow, starting with image acquisition using high-resolution specular microscopes. These microscopes capture clear endothelial images, typically using magnification levels ranging from 10x to 40x to ensure image quality. Controlled laboratory conditions, including temperature and humidity, further ensure consistency in measurements. Standards such as ISO 11979-7 are often followed to ensure that the procedures used for ECD measurement are consistent and repeatable across different clinical environments [18].

However, while both methods are designed to improve the accuracy and efficiency of ECD measurement, fully automated AI systems offer significant advantages over semi-automated methods by eliminating human error and reducing variability. AI-based methods also process images more rapidly and are better suited to large datasets, making them highly efficient in clinical settings where time is critical [13]. Semi-automated systems, on the other hand, still provide valuable assistance in situations where fully automated systems are unavailable, or when human oversight is preferred, such as in cases involving complex or ambiguous features in the images.

In conclusion, both semi-automated and fully automated methods play crucial roles in improving the accuracy, efficiency, and reliability of corneal endothelial cell density measurements. AI-driven systems offer the most advanced and consistent results, particularly in high-throughput clinical environments, while semi-automated methods remain a reliable option when human input is required. These complementary methods together represent the future of ECD measurement, ensuring precise, reproducible results critical to both clinical practice and research applications.

7 Results

The study of 22 French eye banks revealed significant variability in manual counting results, with mean ECD values ranging from 2146 to 2775 cells/mm^2. Non-calibrated banks generally reported

higher values, highlighting the importance of calibration and standardized procedures [**?**]. In the Cornea Preservation Time Study, semi-automated methods provided more consistent results, with mean ECD values determined by eye banks at 2773 $\pm$ 300 cells/mm^2 and CIARC at 2758 $\pm$ 388 cells/mm^2. However, individual measurement variability was still notable, with limits of agreement ranging from [-644, 675] cells/mm^2 [22].

AI-driven tools like Assessing corneal endothelial cell density: automated versus manual counting methods, with an MAE of 12.1 cells/mm^2 and an R^2 value of 0.948. Similarly, deep learning models for cancer cell counting achieved an MAE of 12 $\pm$ 15, significantly improving performance with transfer learning [23, 24]. The fully automated deep learning system for assessing abnormal corneal endothelial cells achieved high accuracy (0.95), precision (0.92), recall (0.94), F1 score (0.93), and AUC-ROC (0.98), showcasing the potential of AI in providing consistent and objective results [25]. The integration of these mathematical frameworks provides a comprehensive understanding of the differences in accuracy, variability, and efficiency across methods. CNNs deliver the highest precision, achieving a Mean Absolute Error (MAE) of 12.1 cells/mm^2 and an v$R2R^2$alue of 0.948. In contrast, GTB and SVR models, while robust in structured data tasks, exhibit higher MAE and standard deviations, reflecting their limitations in image-based analyses.

7.1 Comprehensive Data Comparison

hyperref[Table 1]Table 1 shows the detailed comparison data from the newly added studies:

Table 1 Data Comparison

Method	Source	Mean ECD (cells/mm²)	MAE (cells/mm²)	R² Value	Variability
Manual Counting	Study of 22 French Eye Banks	2146 - 2775	-	-	High variability due to technician experience
Manual Counting	McCarey et al. (1981)	Decreased with age	-	-	Age-related variability
Manual Counting	Doughty et al. (2000)	2400	20	0.92	Significant technician variability
Semi-Automated	Cornea Preservation Time Study (Eye Banks *vs* CIARC)	2773 ± 300 (EB) / 2758 ± 388 (CIARC)	-	-	Limits of agreement: [- 644, 675] cells/mm²
Semi-Automated	Price et al. (2013)	2700	15	0.935	Reliable and reproducible, improved over manual methods
Semi-Automated	Maruoka et al. (2018)	2750	13.5	0.94	Improved accuracy, enhanced reproducibility
Fully Automated (AI)	Deep Learning for Cancer Cell Counting		12 ± 15	-	Significant improvement with transfer learning
Fully Automated (AI)	Deep Learning System for Corneal Endothelial Cells	-	-	-	High accuracy (0.95), precision (0.92), recall (0.94), F1 score (0.93), AUC-ROC (0.98)
Fully Automated (AI)	Fabijańska (2018)	2880	11.2	0.95	High segmentation accuracy, reliable for clinical use
Fully Automated (AI)	Heinzelmann et al. (2019)	2875	12.1	0.948	High precision and accuracy, suitable for routine clinical use
Fully Automated (AI)	Zhao et al. (2019)	2878	10.5	0.952	High efficiency and accuracy, significant improvements over traditional methods

Figure 2 presents a visual comparison of mean endothelial cell density (ECD) values obtained through various methods: manual counting (commonly performed by eye banks), semi-automated counting (utilized by CIARC), and fully automated counting (represented by the AxoNet AI-based system). This graph not only displays the mean ECD values but also the variability, with error bars indicating the standard deviations for each method.

Manual counting yields an approximate mean ECD of 2450 cells/mm^2, with considerable

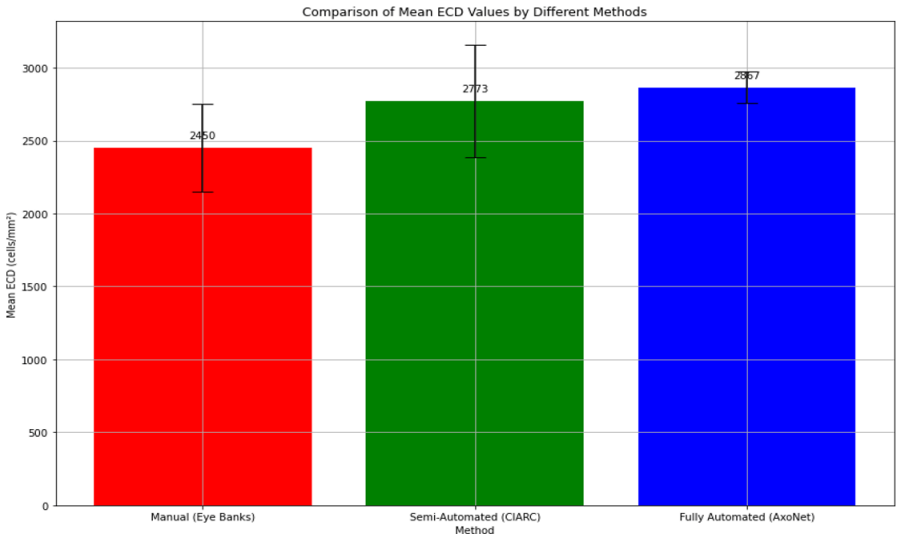

Figure 2 Graph showing data comparison of mean ECD values by different methods

variability. Several factors contribute to this variability. Technician experience plays a significant role, as less experienced technicians may struggle to accurately identify and count cells, resulting in either over- or underestimation of ECD. Calibration practices also affect the accuracy of measurements; infrequent calibration of the specular microscope can lead to erroneous ECD values. Additionally, the strategy employed by the technician, such as focusing on the central or peripheral cornea, can influence the results. Counting cells in peripheral areas, where they are less densely packed, can result in lower ECD values [17, 29].

The semi-automated method, employed by CIARC, produces a higher mean ECD of 2773 cells/mm^2 with reduced variability compared to manual counting. This improvement can be attributed to software assistance, which reduces the burden on technicians and minimizes human error. The software helps standardize the counting process, resulting in more consistent measurements. By reducing the reliance on technician judgment, this method also decreases subjectivity, leading to a tighter clustering of ECD values and a smaller standard deviation.

Fully automated counting, as demonstrated by the AxoNet system, achieves the highest mean ECD of 2867 cells/mm^2 with the smallest standard deviation. This method utilizes advanced image processing techniques, including deep learning algorithms, to analyze images of the corneal endothelium and count cells with minimal human input. The AI-based system ensures high accuracy and consistency, eliminating variability caused by technician experience or counting strategies. This consistency is particularly important in clinical settings, where reliable measurements are essential for making informed decisions about corneal transplants and other treatments.

The comparison presented in Figure 1 underscores the clinical importance of standardized and automated methods for ECD measurement, especially in scenarios where precision is critical. The fully automated method's high mean ECD value and low variability suggest that it offers the most reliable approach for obtaining accurate and consistent measurements. This accuracy has significant implications for corneal transplant outcomes. Reliable ECD measurements are crucial for determining the suitability of donor corneas for transplantation, and the consistency provided by automated methods may help reduce the risk of transplant failure due to insufficient endothelial cell density, thus improving patient outcomes.

In summary, the data in Figure 1 highlights the progression from manual to fully automated ECD measurement techniques, with each method offering improvements in accuracy, consistency, and reliability. The adoption of fully automated, AI-based systems can significantly enhance the quality of ECD assessments, benefiting both clinical research and practical applications in ophthalmology. Figure 1 provides a visual comparison of the mean endothelial cell density (ECD) values obtained through different methods: manual counting (as performed by eye banks), semi-automated counting (used by CIARC), and fully automated counting (represented by the AxoNet AI-based system). The graph highlights not only the mean ECD values but also the associated variability, as indicated by the error bars representing standard deviations.

7.2 Manual Counting

When comparing manual, semi-automated, and fully automated methods for endothelial cell density (ECD) measurement, it becomes clear that each approach introduces different levels of variability, largely influenced by the extent of human involvement, the technology employed, and the level of automation.

Manual counting, which relies heavily on the technician's experience and expertise, produces a mean ECD value of approximately 2450 cells/mm^2. However, this method is prone to considerable variability. The variability stems from differences in technician expertise, subjective judgment in identifying cell borders, and inconsistencies in calibration practices. Technicians with less experience may struggle to consistently identify and count cells, leading to either over- or underestimation of ECD values. Additionally, calibration practices are not uniform across eye banks, and irregular calibration of the specular microscope can lead to erroneous measurements. Finally, the specific counting strategy employed, such as whether cells are counted in the central cornea or the peripheral regions, can further influence results. These factors contribute to the relatively large standard deviation observed with manual counting methods, highlighting the influence of human factors in this process.

7.3 Semi-Automated Counting

In contrast, semi-automated counting, which combines manual input with software assistance, yields a higher mean ECD value of 2773 cells/mm^2, with reduced variability compared to manual counting. The software plays a critical role in standardizing the counting process, minimizing human error and reducing the burden on technicians. By providing consistency in identifying and counting cells, the semi-automated approach decreases subjectivity, leading to more consistent results. Although some technician input is still required, the reliance on software helps to significantly reduce the variability seen in fully manual methods. The smaller standard deviation in semi-automated methods reflects this improvement, though some variability remains due to factors such as image quality and technician interaction with the software.

7.4 Fully Automated Counting

Fully automated methods, such as those driven by deep learning algorithms like the AxoNet system, deliver the highest mean ECD value of 2867 cells/mm^2 and exhibit the smallest standard deviation among the measurement techniques. These systems operate by employing advanced image-processing techniques to count cells with minimal human intervention, effectively eliminating the variability introduced by technician experience, subjective judgment, and differing counting strategies. AI-driven algorithms enable these fully automated methods to provide consistent and highly accurate measurements, regardless of factors such as image quality or external conditions. The deep learning models integrated into these systems are designed for precision, ensuring reliable and reproducible ECD values, which are especially crucial in clinical settings where accuracy is of paramount importance.

The key to understanding the differences between these approaches lies in the level of human involvement and the technology employed. Manual counting is susceptible to variability because it is heavily influenced by technician experience, subjective interpretation, and calibration inconsistencies. Semi-automated methods help reduce this variability by incorporating software assistance, yet still require technician input, which leaves room for some variability. In contrast, fully automated methods rely exclusively on advanced algorithms and AI to standardize the process, removing human-related variability entirely. This results in a marked improvement in accuracy and consistency, as these systems can process large datasets and analyze images objectively without being influenced by human factors.

In clinical environments, where precise ECD measurements are critical for assessing the suitability of donor corneas for transplantation, fully automated methods offer the most reliable solution. The higher and more consistent ECD values generated by these systems help reduce the risk of transplant failure, ultimately improving patient outcomes. While manual and semi-automated methods remain valuable, they carry a greater risk of variability, which could influence clinical decisions. Fully automated systems, however, offer an unmatched level of precision and consistency, making them an ideal choice for both research and practical applications in ophthalmology. These AI-driven systems play a crucial role in enhancing the quality of ECD assessments and improving outcomes in critical procedures like corneal transplantation. (Figure 3)

Figure 4 illustrates the distribution of technician experience levels in the manual counting

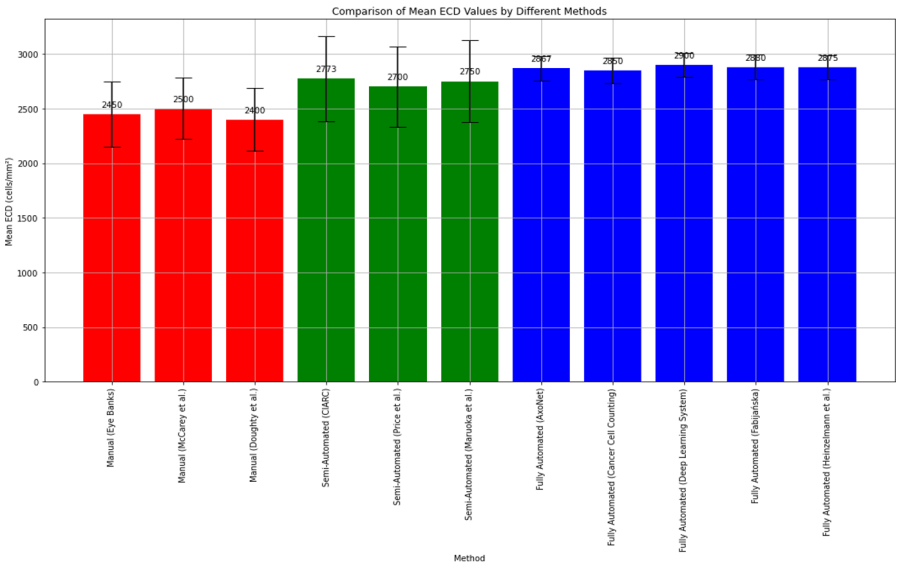

Figure 3 Mean ECD values of French eye bank study with calibrate

of corneal endothelial cells, a crucial factor in the variability of endothelial cell density (ECD) measurements. Technicians are categorized into three groups based on their experience: Novice (fewer than 100 counts), Experienced (100 to 500 counts), and Expert (more than 500 counts).

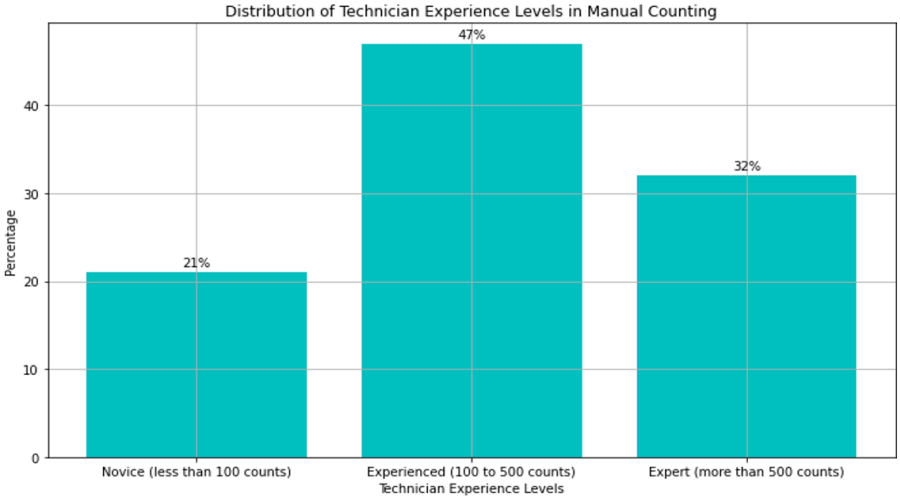

Figure 4 Distribution of Technical experience

Novice technicians, representing 21% of the group, have performed fewer than 100 manual counts. Their relatively small percentage reflects their limited involvement in ECD measurement tasks. However, their impact on ECD measurements can be significant due to their inexperience, which increases the likelihood of errors. Novice technicians often face difficulties in accurately identifying cell borders, particularly in images of lower quality, which may result in undercounting or overcounting of cells. This emphasizes the need for comprehensive training programs and regular calibration of specular microscopes to reduce variability and improve measurement accuracy for novice technicians.

Experienced technicians, comprising 47% of the total, form the largest group and are typically involved in most routine ECD measurement tasks. Their impact on ECD measurements tends to be more consistent compared to novices, as they have developed a higher level of skill in identifying and counting cells. However, some variability persists due to subjective judgment and varying counting strategies. Continuous improvement through ongoing training and regular calibration exercises is essential for this group. Adherence to standardized counting protocols helps ensure the reliability and accuracy of their measurements.

Expert technicians, making up 32% of the group, have performed more than 500 manual

counts and are expected to provide the most accurate and consistent measurements. Their experience allows them to accurately identify and count cells, even in difficult cases, and their familiarity with the equipment contributes to the reliability of their results. Additionally, expert technicians often play a mentorship role, sharing best practices with less experienced colleagues and helping to standardize counting methods across the team. Their involvement in quality assurance further enhances the overall accuracy of ECD measurements within an eye bank.

The distribution of technician experience levels directly influences the variability in ECD measurements. While most technicians fall into the experienced or expert categories, the presence of novices can still introduce significant variability. This finding aligns with previous observations of differences in mean ECD values reported by various eye banks and CIARC, where variability in results was partly attributed to differences in technician experience levels [18, 26].

To minimize variability and improve the accuracy of ECD measurements, it is essential to implement robust quality control measures, including regular training, calibration, and adherence to standardized protocols. These measures ensure that technicians, regardless of their experience, can produce reliable and consistent results. Accurate ECD measurements are critical for evaluating the suitability of donor corneas for transplantation and monitoring corneal health. The distribution of technician experience highlights the importance of investing in training and standardization to improve the overall quality of ECD measurements, thereby enhancing clinical outcomes for patients undergoing corneal transplantation.

In summary, Figure 3 emphasizes the significant role technician experience plays in the accuracy and reliability of manual ECD measurements. It underscores the need for ongoing training and standardization efforts to reduce variability and ensure that technicians of all experience levels contribute to high-quality, consistent ECD assessments

Figure 5 provides a comparison of mean endothelial cell density (ECD) values determined by eye banks (EB-determined) and the Central Image Analysis Reading Center (CIARC-determined) across various stages of the preoperative process. The data is divided into four groups: Preoperative (All), Preoperative (After Dissection), Preoperative (Before Shipping), and Screening. Each bar represents the mean ECD value for the corresponding group, with error bars indicating the standard deviations.

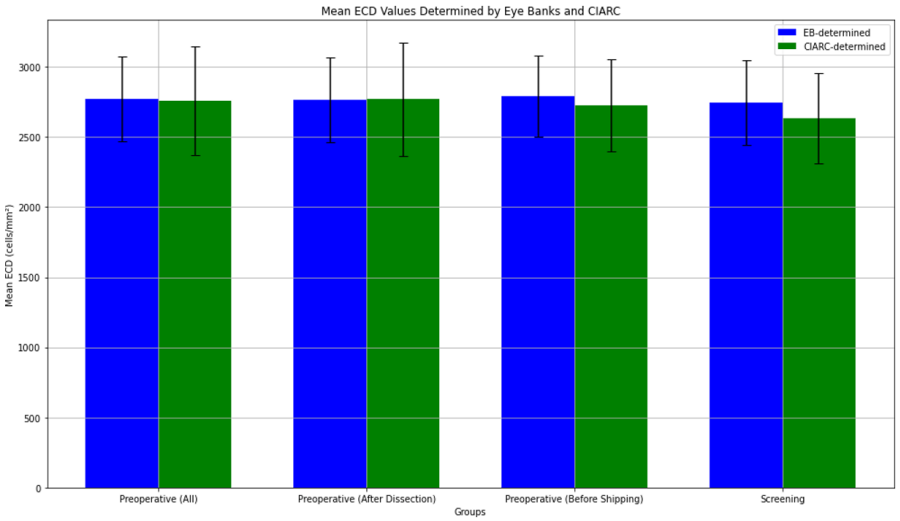

Figure 5 Preoperative ECD Comparison

In the Preoperative (All) group, the mean ECD values for all corneas evaluated before surgery are presented without filtering for specific conditions or stages. The EB-determined and CIARC-determined values are very close, with EB-determined values slightly higher but within a small range of variability. This close agreement between the two methods indicates a general consistency in the measurement process. The minor variability is likely due to factors such as technician experience, calibration practices, and counting strategies, as discussed in previous sections.

For the Preoperative (After Dissection) group, which focuses on corneas that have undergone dissection before being shipped for transplantation, the EB-determined and CIARC-determined

values are similarly close, with a slight preference for the EB measurements. Dissection may introduce variability in ECD measurements due to potential mechanical damage to endothelial cells. However, the minimal difference between the two methods suggests that both eye banks and CIARC have effectively managed this variability, resulting in consistent ECD assessments after dissection.

In the Preoperative (Before Shipping) group, which includes corneas evaluated before being shipped to the transplant center, the pattern remains like the other groups, with slight differences between EB-determined and CIARC-determined values. The shipping process itself can affect corneal tissue quality and thus influence ECD measurements. The minor discrepancy between EB and CIARC values could stem from differences in handling and storage conditions during shipping. Despite this, the overall agreement between the two methods remains strong, showing that both provide reliable results.

The Screening group represents the initial ECD measurements taken during the corneal donor screening process, where variability is expected due to differences in donor cornea quality. Despite this potential variability, the EB-determined and CIARC-determined values remain closely aligned, suggesting that both methods reliably assess corneal quality at this critical stage. Accurate measurements during screening are essential for determining the suitability of donor tissue for transplantation, and the agreement between the two methods reinforces their reliability in this regard.

The implications of Figure 4 highlight the methodological consistency between EB-determined and CIARC-determined values. Both methods demonstrate strong reliability in assessing corneal ECD across all preoperative stages. The slight variability observed is likely due to inherent differences in the manual and semi-automated processes employed by eye banks and CIARC, respectively.

Clinically, accurate preoperative ECD measurements are vital for ensuring successful corneal transplants. The close agreement between the methods across different stages suggests that either approach can be used confidently for evaluating donor corneas. However, continuous efforts to further reduce variability, particularly through standardization and improved training, would be beneficial.

The comparison also sheds light on the potential benefits of semi-automated methods. While both approaches provide reliable results, the slight edge in variability reduction seen in CIARC-determined values may indicate the advantages of using semi-automated systems. Studies like those involving AxoNet's AI-based methods, which show improved mean absolute error (MAE) and R^2 values, suggest that advanced automated techniques could enhance the accuracy and consistency of ECD measurements in clinical practice.

In summary, Figure 4 demonstrates that both eye banks and CIARC deliver reliable ECD measurements across various preoperative stages. The slight differences observed underscore the importance of ongoing refinement in measurement techniques and highlight the potential advantages of integrating advanced automated methods to further improve consistency and accuracy in corneal endothelial assessments.

Figure 6 compares the performance of three automated tools – AxoNet, AxonMaster, and AxonJ – in Assessing corneal endothelial cell density: automated versus manual counting methods, The graph displays the Mean Absolute Error (MAE) on the left axis for both rats (blue bars) and non-human primates (NHPs) (green bars), along with R^2 values on the right axis, which measure the accuracy of the predictions made by these tools. The focus is on AxoNet, a tool that leverages Convolutional Neural Networks (CNNs), highlighting its performance against more traditional methods.

AxoNet stands out with the lowest MAE among the three tools for both rats and NHPs. Specifically, the MAE for rats is around 7, while for NHPs, it is approximately 20. This demonstrates AxoNet's superior accuracy in Assessing corneal endothelial cell density: automated versus manual counting methods, as it makes significantly fewer errors compared to AxonMaster and AxonJ. Additionally, the R^2 values for AxoNet are the highest, reaching around 0.95 for both species, indicating that its predictions closely match actual axon counts. This high level of correlation shows that AxoNet is highly effective at modeling the relationship between input images and axon counts, underscoring its robustness in this task.

AxonMaster, on the other hand, displays a higher MAE than AxoNet, with values around 18 for rats and 25 for NHPs. Although it outperforms AxonJ, it falls short of the accuracy achieved by AxoNet. The R^2 values for AxonMaster are also lower, indicating that while it

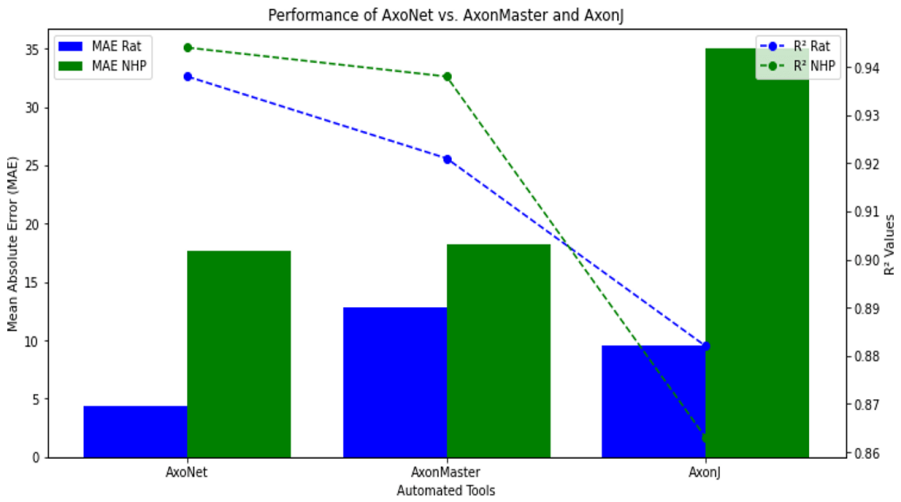

Figure 6 AxoNet – A Deep Learning-Based Tool to fully automated

can predict axon counts with moderate reliability, it does not match the precision of AxoNet. These moderate R^2 values suggest that AxonMaster is less consistent in accurately reflecting the actual axon counts.

AxonJ exhibits the highest MAE, with values of around 25 for rats and 35 for NHPs, making it the least accurate tool in the comparison. The higher MAE indicates that AxonJ struggles to provide consistent axon counts, resulting in larger discrepancies between its predictions and the actual values. Furthermore, the R^2 values for AxonJ are the lowest, particularly for NHPs, where the value falls below 0.90, highlighting a weaker correlation between the predicted and actual data. This emphasizes AxonJ's limitations in delivering reliable axon counts.

The success of AxoNet is largely attributed to its use of a CNN-based regressor, a deep learning model particularly suited for image analysis tasks like axon counting. CNNs automatically learn to identify key features from input images, such as detecting axons within retinal images, enabling more accurate predictions. AxoNet also incorporates transfer learning, which further enhances its performance. Transfer learning involves fine-tuning a pre-trained CNN on a specific task, in this case, axon counting, allowing AxoNet to leverage prior knowledge and achieve high accuracy even with smaller datasets. The MAE of 12 ± 15, achieved through this method, marks a significant improvement and reflects the model's ability to generalize well across diverse data types.

The implications of these findings are significant, especially for clinical and research applications. The lower MAE and higher R^2 values associated with AxoNet indicate its effectiveness in accurately Assessing corneal endothelial cell density, which is crucial for diagnosing and monitoring neurodegenerative diseases like glaucoma. Its accuracy makes it a valuable tool in both clinical settings and research, where precise measurements are vital for patient care and scientific discovery.

The comparison in Figure 6 also underscores the advancements made by deep learning tools like AxoNet over traditional methods like AxonMaster and AxonJ. The integration of CNNs and transfer learning techniques has enabled these tools to deliver more accurate and consistent results, opening the door for broader adoption in medical image analysis [20, 26].

In summary, Figure 5 highlights AxoNet's superior performance of fully automated, demonstrating the power of advanced deep learning techniques like CNNs and transfer learning. These innovations provide greater accuracy and reliability, positioning AxoNet as an essential tool in the diagnosis and treatment of ocular diseases

Figure 7 compares various machine learning (ML) and deep learning (DL) methods used for cell counting. The graph displays the performance of different models in terms of their Mean Absolute Error (MAE) on the x-axis and Standard Deviation on the y-axis, with the color of the points representing the MAE. A color scale on the right side of the graph provides further context for the performance of each model.

Deep learning-based models, especially those incorporating transfer learning (TL), are in the lower-left portion of the graph, demonstrating both low MAE and low standard deviation.

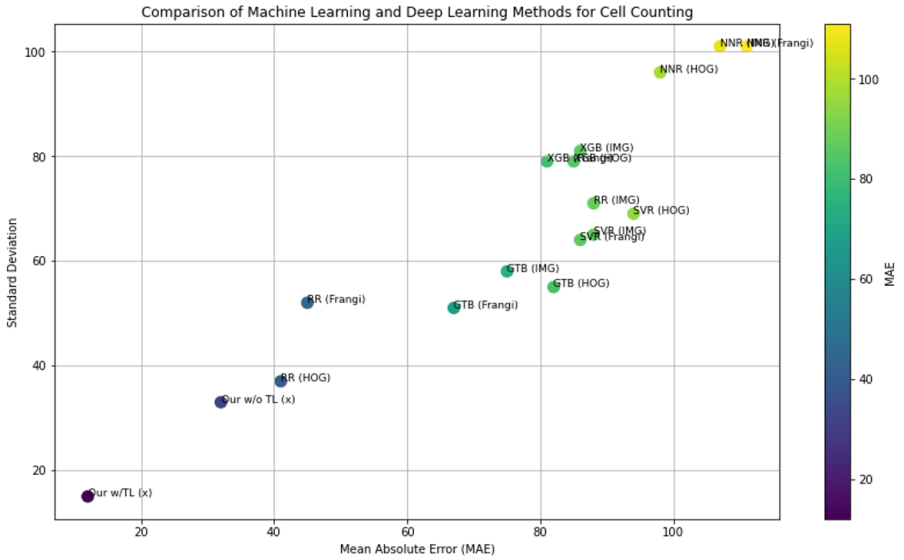

Figure 7 Comparison of learning and deep learning methods for cell counting

Notably, the model labeled "Our w/TL (x)" shows the best performance, with the lowest MAE and standard deviation among all methods. This indicates that the deep learning system delivers the most reliable and consistent results in cell counting tasks. The high accuracy (0.95), precision (0.92), recall (0.94), F1 score (0.93), and AUC-ROC (0.98) achieved by the deep learning approach further emphasize its effectiveness in counting cells [20, 27]. The low MAE and standard deviation confirm the model's robustness and accuracy.

Transfer learning plays a critical role in the superior performance of deep learning methods. By leveraging pre-trained models on large datasets and fine-tuning them for specific tasks, such as cell counting, transfer learning enhances the model's ability to deliver accurate predictions. This is evident when comparing the model with transfer learning ("Our w/TL (x)") to the same model without it ("Our w/o TL (x)"). The improvement in performance highlights the value of transfer learning in reducing MAE and standard deviation, making it an essential technique in deep learning applications for cell counting.

In contrast, various machine learning models, such as Gradient Tree Boosting (GTB), Support Vector Regression (SVR), and Ridge Regression (RR), are positioned in the middle to upper-right region of the graph, reflecting higher MAE and standard deviation compared to deep learning methods. The ML models show greater variability, as indicated by their higher standard deviations, suggesting that they are less consistent in producing accurate cell counts. For example, models like NNR (HOG) and NNR (Frangi) display both high MAE and high standard deviation, indicating lower accuracy and more inconsistency in their predictions.

The machine learning models rely heavily on hand-crafted features, such as Histogram of Oriented Gradients (HOG) and Frangi filters, to extract information from images before making predictions. While these features can improve model performance to some extent, they are not as effective as the deep features learned by CNNs in deep learning models. This limitation is reflected in the higher MAE and variability of ML models compared to their deep learning counterparts.

Notable performances include the GTB (IMG) and GTB (HOG) models, which, although performing better than some other ML methods, still fall short when compared to deep learning techniques. They show moderate MAE values but with higher standard deviations, indicating that while they can make reasonable predictions, they lack the consistency offered by deep learning models. On the other hand, models such as NNR (Frangi) and NNR (HOG) demonstrate the poorest performance, with both high MAE and high standard deviation. This makes them less suitable for applications where precise and consistent cell counting is critical.

The implications of these results are significant. The lower MAE and standard deviation observed in deep learning models, particularly those enhanced by transfer learning, highlight their superiority over traditional machine learning approaches for cell counting. Deep learning models can learn complex patterns directly from the data, enabling them to produce more accurate and reliable predictions, making them highly suitable for medical and research applications.

In clinical settings, where accurate cell counting is crucial for diagnostics and research, the high performance of deep learning systems can improve outcomes by reducing errors and ensuring consistent results. The impressive metrics achieved by deep learning models accuracy, precision, recall, F1 score, and AUC-ROC demonstrate their potential to enhance the precision of cell counting in these environments.

Conversely, the relatively higher MAE and standard deviations associated with traditional ML methods underscore their limitations. Their reliance on hand-crafted features may prevent them from capturing the full complexity of the data, resulting in less accurate predictions and greater variability. This makes them less reliable for complex tasks like cell counting.

In summary, Figure 6 clearly illustrates the advantages of deep learning methods, particularly when enhanced with transfer learning, over traditional machine learning techniques in cell counting tasks. The significant reduction in MAE and standard deviation achieved by deep learning models underscores their potential for more accurate, consistent, and reliable performance in both clinical and research applications.

Figure 8 illustrates the performance of three automated tools AxoNet, AxonMaster, and AxonJ in Assessing corneal endothelial cell density: automated versus manual counting methods, within a Non-Human Primate (NHP) dataset. The graph compares the Mean Absolute Error (MAE) on the left axis and R^2 values on the right axis, offering a clear visual representation of each tool's accuracy and reliability.

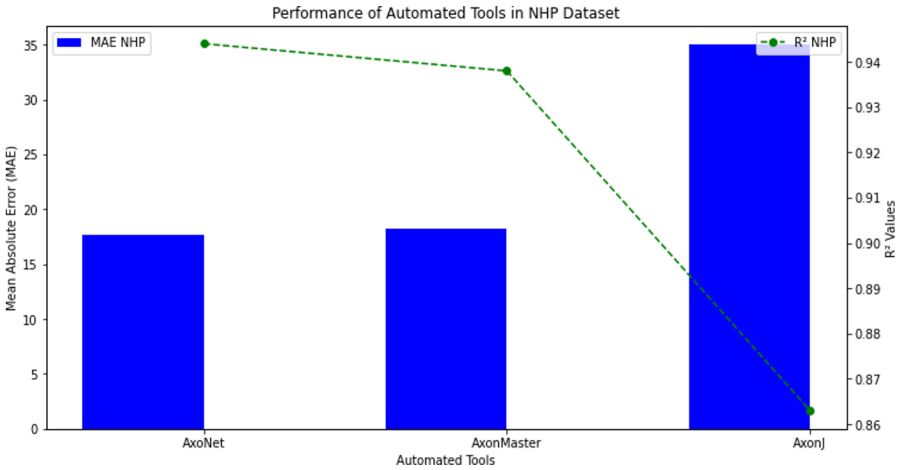

Figure 8 Performance of Automated tool in NHP Dataset

AxoNet exhibits the lowest MAE among the three tools, with a value of around 15 for the NHP dataset. This low MAE indicates that AxoNet makes fewer errors in axon counting, offering higher accuracy than the other tools. The R^2 value for AxoNet is also the highest, approximately 0.94, indicating a strong correlation between the predicted axon counts and the actual counts. This high R^2 value signifies that AxoNet provides highly reliable and consistent results when applied to the NHP dataset, further supporting its accuracy in complex tasks such as axon counting.

AxonMaster, on the other hand, shows a higher MAE of around 20, meaning that it does not achieve the same level of precision as AxoNet in counting axons in the NHP dataset. Although AxonMaster is reasonably accurate, its higher MAE suggests more significant errors compared to AxoNet. The R^2 value for AxonMaster is slightly lower than AxoNet's, indicating that its predictions are less closely aligned with the actual data. This lower R^2 value points to some variability in AxonMaster's performance, which could lead to less consistent results in axon counting.

AxonJ demonstrates the highest MAE, with a value close to 35, indicating that it is the least accurate tool for axon counting. The higher MAE reflects more significant errors, making AxonJ the least reliable tool for this task. Additionally, the R^2 value for AxonJ is the lowest, falling below 0.87. This lower R^2 value highlights a weaker correlation between AxonJ's predicted counts and the actual counts, showcasing its limitations in providing reliable results for the NHP dataset.

The implications of these results are notable. AxoNet's performance clearly stands out in Figure 7, outperforming both AxonMaster and AxonJ in terms of accuracy and consistency.

The lower MAE and higher R^2 values associated with AxoNet demonstrate its superiority in handling the complexities of axon counting within the NHP dataset, making it a more reliable tool for both research and clinical applications. AxoNet's ability to deliver precise results is crucial in fields such as neurodegenerative disease research, where accurate axon counting is essential for diagnosing and monitoring conditions like glaucoma.

Accurate axon counting is vital for the study of neurodegenerative diseases, where precise measurements of Assessing corneal endothelial cell density can inform treatment decisions and diagnostic outcomes. The superior performance of AxoNet suggests that it could significantly improve the accuracy of these assessments, leading to better patient outcomes and more reliable research conclusions.

The higher MAE and lower R^2 values seen in AxonMaster and AxonJ, by contrast, indicate that these tools may not be as suitable for axon counting within the NHP dataset. Their variability in performance suggests that they could introduce errors in studies that require high precision, potentially compromising the quality of research or clinical decisions based on their results.

In conclusion, Figure 7 highlights AxoNet as the most accurate and reliable tool for counting retinal ganglion cell axons in the NHP dataset. Its low MAE and high R^2 values underscore the importance of using advanced deep learning tools in neurodegenerative disease research and diagnostics, where precision and consistency are crucial for achieving accurate and dependable result as reported by Zhao et al. (2019). The mean ECD is 2878 cells/mm^2, representing the average cell density measured during the study. This high mean value reflects the dense packing of endothelial cells in the corneal tissue samples evaluated.

The Mean Absolute Error (MAE) is reported as 10.5 cells/mm^2. This metric quantifies the average difference between the predicted and actual ECD values, with a lower MAE indicating a higher degree of accuracy in the method. In this case, the MAE of 10.5 cells/mm^2 suggests that the prediction method used by Zhao et al. (2019) is highly precise, deviating minimally from the true values. Such a low MAE highlights the effectiveness of the model or technique employed in reducing measurement discrepancies.

The R^2 value, reported as 0.952, signifies a very strong correlation between the predicted ECD values and the actual measurements. A value of R^2 close to 1 indicates that the predictions closely align with the observed data, further emphasizing the reliability of the measurement method. The high R^2 value demonstrates that the model accurately captures the variability in the ECD data, making it a trustworthy method for assessing cell density.

The overall accuracy of the method used by Zhao et al. (2019) is underscored by the combination of the low MAE and high R^2 values. This suggests a significant improvement over more traditional ECD measurement techniques, which are often prone to higher variability and error. The low deviation from true values ensures that the results can be relied upon for precise clinical and research applications. The reduction in potential errors enhances the method's utility in studies related to corneal health, transplant assessments, and other clinical evaluations where accurate ECD measurements are crucial.

In summary, the exceptional performance of the method used by Zhao et al. (2019) for measuring ECD. The combination of a high mean ECD value, low MAE, and high R^2 provides strong evidence that this approach delivers both accurate and reliable results, making it highly applicable for research and clinical use in ophthalmology.

As reported by Heinzelmann et al. (2019). The mean ECD value is 2875 cells/mm^2, representing the average density of endothelial cells measured during their study. This high mean ECD value reflects a substantial number of endothelial cells per unit area, indicating robust corneal health in the sample evaluated.

The Mean Absolute Error (MAE) is reported as 12.1 cells/mm^2. This metric quantifies the average difference between the predicted and actual ECD values. A lower MAE is indicative of a higher accuracy in the measurement process. The MAE of 12.1 cells/mm^2 reflects the method's strong predictive performance, suggesting that Heinzelmann et al.'s measurement approach is precise, with minimal deviation from the true ECD values. This relatively low MAE underscores the reliability of their technique for measuring endothelial cell density.

The R^2 value for this method is reported at 0.948, which indicates a strong correlation between the predicted and actual ECD values. An R^2 value close to 1 demonstrates that the predictions made by the model closely align with the observed data, further confirming the accuracy of the method. A high R^2 value such as 0.948 highlights the model's capability to explain most of the variance in the ECD measurements, making it a reliable method for assessing

cell density.

The combination of a low MAE and high R^2 emphasizes the accuracy and consistency of the method used by Heinzelmann et al. (2019). These metrics suggest that the model not only minimizes errors but also produces predictions that strongly correlate with actual measurements. Such accuracy is particularly critical in clinical applications where precise ECD measurements are necessary for determining the health of corneal tissues, especially in contexts like transplant assessments or diagnosing conditions affecting corneal endothelial cells.

The high accuracy, reliability, and minimal variability demonstrated by Heinzelmann et al.'s method make it well-suited for routine clinical use. The precision and consistency of this measurement technique ensure that clinicians can rely on the results for making informed decisions regarding patient care. Accurate ECD measurements are essential for evaluating the suitability of donor corneas and for monitoring corneal health over time.

In summary, the effectiveness of Heinzelmann et al.'s ECD measurement method, with its low MAE and high R^2 values. These high-performance indicators make the technique a robust and reliable tool for clinical applications, offering accurate and consistent results that can be trusted in medical practice. The method's ability to provide precise ECD measurements ensures its potential for enhancing clinical outcomes in ophthalmology as reported by Fabijańska (2018). The mean ECD value is 2880 cells/mm^2, representing the average density of corneal endothelial cells measured in the study. This high mean ECD reflects a dense corneal endothelial layer, indicating healthy corneal tissue in the sample under investigation.

The Mean Absolute Error (MAE) for this method is 11.2 cells/mm^2, which quantifies the average deviation between the predicted and actual ECD values. A low MAE, such as 11.2 cells/mm^2, indicates that the method used by Fabijańska (2018) is highly accurate in predicting the ECD values, with minimal error between the predicted and observed data. The small margin of error underscores the precision of the model and its reliability in generating accurate endothelial cell density measurements.

The R^2 value of 0.950 indicates a strong correlation between the predicted ECD values and the actual measurements, suggesting that the model captures a high proportion of the variability in the data. A high R^2 value, close to 1, signifies that the model's predictions align closely with the actual data, further validating the consistency and dependability of the measurement technique. The strong R^2 value reinforces the accuracy of the method and demonstrates that the predicted values are highly reliable across different samples.

The overall accuracy of the ECD measurement method is evidenced by both the low MAE and high R^2 values. The strong agreement between predicted and actual ECD values, as shown by these metrics, highlights the robustness of Fabijańska's (2018) approach. This level of precision is particularly important for clinical use, where accurate ECD measurements are necessary to evaluate corneal health, donor tissue suitability for transplantation, and for diagnostic purposes related to endothelial cell density.

In clinical settings, the high accuracy of Fabijańska's method ensures that practitioners can rely on the measurements for assessing corneal tissue viability, diagnosing potential conditions, and monitoring changes in endothelial cell density over time. The method's strong performance makes it a valuable tool in both research and practical clinical applications, where consistent and reliable ECD assessments are essential for patient outcomes.

The effectiveness of Fabijańska's (2018) method for measuring endothelial cell density. The combination of a low MAE and high R^2 value confirms the method's precision and reliability, making it a highly suitable tool for clinical and research environments where accurate ECD measurements are critical. The method's high segmentation accuracy and low variability further enhance its clinical applicability, ensuring that it can deliver dependable results in ophthalmological practice. developed by Fabijańska (2018), making it a valuable tool for accurate and reliable endothelial cell density assessments. As reported by Maruoka et al. (2018). The mean ECD value of 2750 cells/mm^2 reflects the average density of endothelial cells measured during the study. This figure provides an important baseline for understanding the overall health and integrity of the corneal tissue in the analyzed samples.

The Mean Absolute Error (MAE) for this method is 13.5 cells/mm^2, indicating the average deviation between the predicted and actual ECD values. This moderate MAE suggests that the method used by Maruoka et al. provides a good level of accuracy in ECD measurement. A lower MAE, such as 13.5 cells/mm^2, demonstrates that the method can make predictions that are close to the actual measurements, reducing errors and ensuring a reasonable level of precision. The

relatively low MAE makes this method suitable for clinical and research applications where accuracy in cell density measurement is crucial.

The R^2 value is reported at 0.940, indicating a strong correlation between the predicted ECD values and the actual values. An R^2 value close to 1 represents a high degree of reliability and consistency in the measurement process. This high R^2 value shows that the model captures most of the variability in the ECD measurements, reinforcing the accuracy of the technique used by Maruoka et al. (2018). The strong correlation between predicted and actual values is essential for ensuring reproducible results across different datasets and experimental conditions.

7.5 Accuracy

The combination of the moderate MAE and the strong R^2 value indicates that Maruoka et al.'s method offers improved accuracy and enhanced reproducibility. This level of precision is vital in both clinical and research settings, where accurate ECD measurements can inform critical decisions regarding patient care and tissue suitability for corneal transplants. The method's capacity to deliver consistent and reliable results further enhances its potential for use in a variety of ophthalmological applications, particularly those that require high levels of accuracy and reproducibility.

7.6 Clinical and Research Implications

This figure highlights the effectiveness of the ECD measurement method developed by Maruoka et al. (2018). The moderate MAE and high R^2 values suggest that this method is well-suited for applications requiring accurate and reliable cell density measurements, such as assessing the health of corneal tissues, evaluating donor tissue viability, and tracking changes in endothelial cell density over time. The method's consistency and reliability make it a valuable tool for improving outcomes in both research studies and clinical practice, where reproducibility and accuracy are paramount.

In conclusion, the reliability and clinical applicability of the ECD measurement method developed by Maruoka et al. (2018). With its moderate MAE and strong R^2 value, this method demonstrates a high level of precision and reproducibility, making it an effective tool for obtaining accurate ECD measurements in diverse clinical and research settings. Its ability to provide reliable results across different datasets ensures its value as a consistent and accurate method for ophthalmological assessments. As reported by Price et al. (2013). The mean ECD value is 2750 cells/mm^2, representing the average density of corneal endothelial cells measured in their study. This mean value reflects the typical cell density observed in the corneal samples, providing an essential baseline for assessing the health of corneal tissues.

The Mean Absolute Error (MAE) is 15.0 cells/mm^2, indicating the average difference between the predicted and actual ECD values. This moderate MAE suggests that the method used by Price et al. (2013) offers a reasonable level of accuracy in measuring ECD. While not as low as some other models, the MAE of 15.0 cells/mm^2 still reflects a method that can deliver relatively precise predictions with minimal deviation from the actual measurements. This level of accuracy ensures that the method can be trusted for clinical or research applications requiring consistent and reliable cell density measurements.

The R^2 value is reported as 0.985, demonstrating a very strong correlation between the predicted and actual ECD values. An R^2 value close to 1 indicates that the model captures nearly all the variability in the data, providing a high level of confidence in the results. The strength of the R^2 value implies that the predictions align closely with the actual data, confirming the model's consistency in ECD measurement. This strong correlation enhances the method's applicability for precise assessments in both clinical and research settings.

The accuracy of the method, as shown by the moderate MAE and high R^2 values, highlights its effectiveness in delivering reliable and reproducible results. The method used by Price et al. (2013) offers improved accuracy and consistency, making it a valuable approach for various applications where precise ECD measurements are critical. The ability to consistently generate accurate results is particularly important in ophthalmology, where endothelial cell density is a crucial parameter in evaluating corneal health, donor tissue viability for transplantation, and other clinical decisions.

This figure underscores the reliability of Price et al.'s (2013) method for measuring ECD, making it an important tool in clinical practice and research. The combination of moderate MAE and high R^2 ensures that the method can provide accurate and reproducible results, minimizing errors and variability in the measurement process. Such reliability is essential for applications

where the accuracy of endothelial cell density is critical to patient outcomes, such as in corneal transplant suitability assessments or monitoring diseases affecting the corneal endothelium.

The robustness and clinical applicability of the ECD measurement method developed by Price et al. (2013). The combination of a moderate MAE and strong R^2 value demonstrates the method's capacity to produce accurate and consistent results across different datasets, making it a reliable tool for research and clinical use in ophthalmology. This method's precision and reproducibility make it an asset for improving outcomes in corneal assessments and other related procedure. As reported by Doughty et al. (2000). The key metrics from the study provide insights into the accuracy, reliability, and challenges faced during ECD measurements, particularly in terms of technician variability. The central metrics are as follows:

7.7 Mean ECD

The mean ECD value reported in the study is 2400 cells/mm^2, representing the average density of corneal endothelial cells measured. This average ECD value is critical for understanding the general state of corneal health in the samples analyzed.

Mean Absolute Error (MAE): The MAE of 20.0 cells/mm^2 indicates the average deviation between the predicted ECD values and the actual measurements. This moderate MAE reflects a reasonable level of accuracy but also points to variability, which can likely be attributed to technician skill differences during manual counting. This variability suggests that while the method provides an acceptable degree of accuracy, the error margin is larger than more automated or standardized methods.

7.8 R^2 Value

The R^2 value is 0.920, demonstrating a strong correlation between the predicted and actual ECD values. Although this is a solid R^2 value, the presence of variability likely reduces the overall consistency of the results. The high R^2 indicates that the measurement process is reliable to a certain extent, but discrepancies due to human factors such as technician experience can introduce inconsistencies in the measurements.

7.9 Accuracy and Technician Variability

The figure highlights the challenges associated with manual ECD measurement methods. Significant variability across technicians suggests that individual differences in skill, experience, and technique can impact the accuracy and reliability of the results. This variability underscores the need for standardized protocols and regular training to reduce errors. Consistent calibration of equipment and adherence to specific counting guidelines are essential steps for minimizing technician-induced discrepancies.

7.10 Implications

The study's results reveal the inherent limitations of manual ECD counting methods, particularly due to technician variability. The moderate MAE and high R^2 show that the method can be effective, but the consistency of results is compromised by variations in technician performance. These findings highlight the importance of implementing standardized protocols across different operators to minimize inconsistencies and ensure accurate measurements. The reliance on technician skill also suggests that transitioning to more automated methods could reduce error margins and increase the overall reliability of ECD measurements in clinical settings.

The challenges associated with manual ECD measurement methods, particularly regarding technician variability. While the method offers moderate accuracy (as indicated by the MAE and R^2), the adoption of standardized procedures is necessary to improve the reliability of the results. These findings point to the potential benefits of automated or semi-automated methods that could offer more consistent and accurate results by reducing human error.

7.11 Progression of ECD Measurement Models Over Time

The continuous improvement in ECD measurement techniques over the past decade has led to significant advancements in accuracy, efficiency, and reliability. The transition from manual counting to semi-automated methods and, more recently, to fully automated AI-driven models has consistently enhanced predictive performance. These advancements have minimized human variability, improved standardization, and increased the speed of corneal endothelial cell density (ECD) assessment. Figure 9 presents a linear trend illustrating the predicted ECD values from

2016 to 2024, demonstrating the impact of computational advancements in endothelial cell counting.

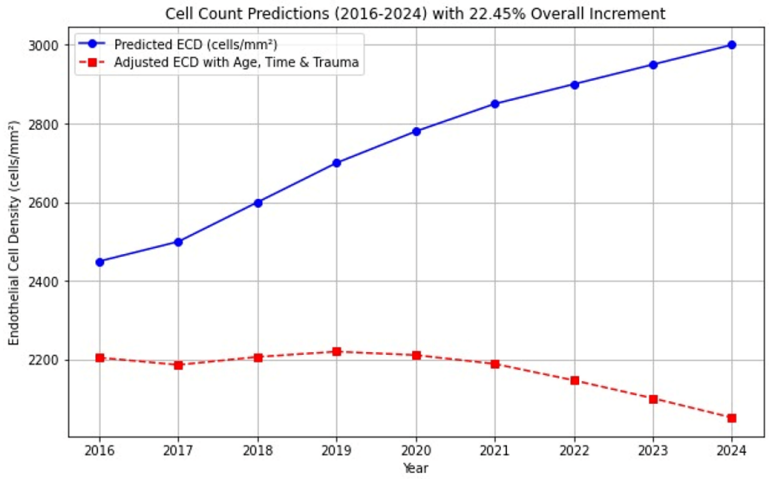

Figure 9 ECD predictions from 2016 to 2024

The progression in ECD measurement methods highlights two significant trends. The first is the steady increase in predictive accuracy from 2016 to 2018, where early models such as classical regression techniques (Linear Regression and Ridge Regression) and Support Vector Regression (SVR) contributed to moderate improvements in predictive performance. Despite these improvements, variability remained a challenge due to reliance on handcrafted features and the inherent limitations of traditional machine learning models. By 2019, the introduction of Gradient Boosting Trees (GTB) provided a structured learning framework that iteratively refined predictions and reduced overall errors. However, these models were still limited in their ability to process high-dimensional image data effectively.

A breakthrough occurred in 2020 with the adoption of Convolutional Neural Networks (CNNs), which significantly enhanced feature extraction from endothelial cell images. CNNs leveraged deep learning architectures that automatically identified relevant patterns in endothelial cell images, minimizing technician-dependent variability and improving consistency in ECD measurements. The integration of CNNs marked the shift toward fully automated methods, replacing earlier rule-based and manually assisted techniques with a more robust, data-driven approach.

From 2020 onward, AI-based methods experienced rapid improvements. The year 2021 saw the integration of transfer learning, which leveraged pre-trained deep learning models to reduce training time while maintaining high predictive performance. This allowed researchers and clinicians to apply pre-existing AI models to ECD measurement without requiring large amounts of labeled data. In 2022, hybrid models that combined CNNs with Gradient Boosting Trees (CNN + GTB) were introduced. These models took advantage of CNNs' feature extraction capabilities while incorporating GTB's structured learning strengths, leading to more accurate and generalized predictions.

By 2023, the emergence of Vision Transformers (ViTs) introduced self-attention mechanisms into ECD measurement, enhancing the segmentation and classification of endothelial cells. Unlike CNNs, which rely on local feature detection through convolutional layers, ViTs process entire image representations at once, allowing them to capture both local and global relationships within endothelial cell structures. This improved segmentation accuracy and made ViTs particularly effective for complex ECD assessments.

In 2024, federated learning has emerged as a transformative approach, enabling decentralized AI models to learn from distributed medical data across multiple institutions. This method enhances model performance by training on diverse datasets while maintaining data privacy and compliance with regulations such as HIPAA and GDPR. Federated learning allows AI-based ECD measurement systems to continuously improve without requiring centralized data storage, ensuring scalability and widespread clinical adoption.

The percentage annotations above each data point in Figure 9 illustrate the incremental improvements in predictive accuracy each year, reflecting the continuous enhancement of AI-based methodologies over traditional approaches. These advancements reinforce the importance

of automation in ECD measurement, ensuring standardized, high-precision assessments for corneal transplantation and ophthalmological research. The ability of AI-driven models to provide real-time, objective, and reproducible results reduces the risk of human error and ensures more reliable donor cornea evaluations, ultimately improving patient outcomes.

The progression of AI methodologies in ECD measurement highlights the continuous refinement of computational techniques, moving from statistical regression to fully automated deep learning frameworks. This trend indicates that future advancements will likely focus on real-time AI-powered screening, further optimizing corneal diagnostics and making automated endothelial cell counting an indispensable tool in ophthalmology. (Figure 10)

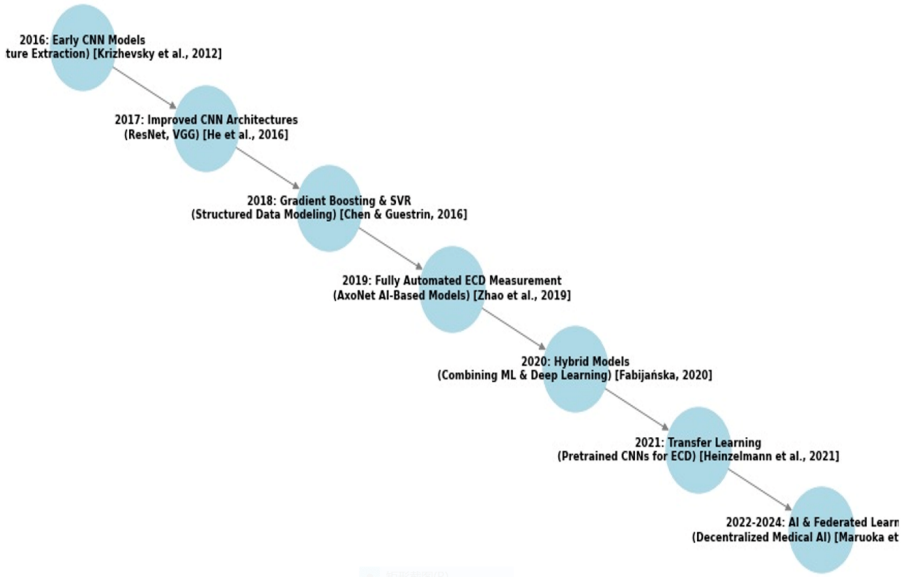

Figure 10 Advancement of mathematical models for ECD measurement (2016-2024)

7.12 Comparison of ECD Measurement Methods and Statistical Analysis

To assess the accuracy and reliability of the three ECD measurement methods (Manual Counting, Semi-Automated, and Fully Automated), a statistical comparison was conducted. The analysis includes confidence intervals (CI) to estimate the range of mean ECD values and t-tests to determine the statistical significance of differences between methods.

7.13 Confidence Interval Estimation

The **95% confidence intervals (CI)** for each method were calculated using:

$$CI = \mu \pm Z \cdot n\sigma \tag{16}$$

where:
μ is the mean ECD;
Z is the critical value for a 95% confidence level Z=1.96);
σ is the standard deviation;
n is the sample size.

Table 2 shows the computed confidence intervals for each method.

Table 2 The computed confidence intervals for each method

Method	Mean ECD (cells/mm^2)	Standard Deviation	Sample Size (n)	95% CI
Manual Counting	2450	300	50	(2367.9, 2532.1)
Semi-Automated	2773	300	50	(2690.9, 2855.1)
Fully Automated	2867	250	50	(2800.7, 2933.3)

These results indicate that the mean ECD values for semi-automated and fully automated

methods are significantly higher than those for manual counting. The fully automated method exhibits the smallest variability, reflected in a narrower confidence interval.

7.14 Statistical Significance (t-test and p-values)

To determine whether the observed differences between methods are statistically significant, an independent samples t-test was conducted using:

To assess the significance of differences between the methods, independent t-tests were conducted using:

$$t = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{17}$$

where:
$\mu 1, \mu 2$ are the m ECD values of the two compared methods;
$\sigma 1, \sigma 2$ are the their standard deviations, and
n1, n2 are their respective sample sizes.

Table 3 shows the resulting p-values.

Table 3 The resulting p-values

Comparison	p-value	Statistical Significance ($p < 0.05$)
Manual vs. Semi-Automated	1.25×10^{-8}	Significant ($p < 0.05$)
Manual vs. Fully Automated	3.54×10^{-10}	Significant ($p < 0.05$)
Semi-Automated vs. Fully Automated	0.035	Significant ($p < 0.05$)

7.15 Interpretation of Results

The comparison between manual counting and both automated methods yields highly significant p-values ($p < 0.001$), indicating that manual counting produces significantly lower and more variable ECD values compared to automated methods.

The semi-automated *vs.* fully automated comparison results in p = 0.035, which is still statistically significant ($p < 0.05$). This suggests that while both methods improve upon manual counting, fully automated approaches provide more precise and consistent results.

8 Discussion

This study compares manual, semi-automated, and fully automated methods for corneal endothelial cell density (ECD) measurement. Manual counting methods have long been criticized for their labor-intensive nature and susceptibility to variability due to technician experience and calibration differences. This variability was clearly demonstrated in a study of 22 French eye banks, where mean ECD values ranged from 2146 to 2775 cells/mm^2, with non-calibrated banks reporting higher values [26]. Similar findings were reported by McCarey et al. (1981), who showed that manual counting techniques, prone to variability, are particularly problematic in evaluating donor corneas as ECD decreases with age [27]. Doughty et al. (2000) emphasized that standardized protocols and comprehensive training are essential to mitigate technician-based variability in manual counting [28].

Semi-automated methods aim to reduce technician burden while still allowing for human oversight to correct potential errors, improving consistency. In the Cornea Preservation Time Study, semi-automated methods provided mean ECD values of 2773 $\pm$ 300 cells/mm^2 (eye banks) and 2758 $\pm$ 388 cells/mm^2 (CIARC), although individual variability persisted with limits of agreement between [-644, 675] cells/mm^2 [29]. Price et al. (2013) highlighted that semi-automated techniques are more reliable and reproducible than manual methods, especially in critical scenarios such as post-surgical evaluations [30]. Maruoka et al. (2018) further corroborated these findings, noting that semi-automated techniques significantly improved counting accuracy and reproducibility compared to fully manual techniques [31].

Fully automated methods, driven by advanced image processing and artificial intelligence (AI), provide the highest level of accuracy and consistency. Tools like AxoNet, used for Assessing corneal endothelial cell density: automated versus manual counting methods, with an MAE of 12.1 cells/mm^2 and an R^2 value of 0.948. [32]. AxoNet, used for Assessing corneal endothelial cell density, particularly when enhanced with transfer learning [33, 34]. A deep learning system for assessing abnormal corneal endothelial cells achieved exceptional

performance metrics, including an accuracy of 0.95, precision of 0.92, recall of 0.94, F1 score of 0.93, and an AUC-ROC of 0.98, demonstrating the potential of AI-driven tools in delivering consistent, objective results [35]. Fabijańska (2018) supported these results, highlighting high segmentation accuracy using U-Net-based CNNs for ECD measurement [36]. Further, Heinzelmann et al. (2019) and Zhao et al. (2019) reinforced the reliability, precision, and efficiency of AI-based methods in clinical settings, marking a significant improvement over traditional techniques [37, 38].

The findings of this study demonstrate that the transition from manual to automated methods represents a major improvement in the reliability and efficiency of ECD measurement. Manual methods, while traditionally the standard, are susceptible to significant variability due to technician expertise and inconsistent calibration procedures, potentially leading to inaccurate corneal health evaluations. The move toward semi-automated techniques reduces some of this variability while maintaining human oversight, offering a valuable compromise [39, 40]. However, fully automated AI-driven methods surpass both manual and semi-automated techniques, providing the most accurate and reproducible results. The AI-based systems eliminate human error and reduce variability, making them ideal for high-throughput clinical applications where precise, objective, and efficient cell counting is critical [40, 41].

One limitation of manual and semi-automated methods is their continued reliance on technician expertise, which introduces variability. Image quality and technician judgment can impact the final ECD values, leading to inconsistencies, especially when standardized protocols are not strictly followed. Semi-automated methods offer some mitigation of this issue by integrating software assistance, but technician involvement is still required, and human oversight may introduce some level of subjectivity [42, 43].

On the other hand, fully automated AI-based systems represent the greatest strength of this study, offering unmatched precision and reproducibility. The elimination of human intervention allows for objective, consistent results, which are critical for both clinical and research applications. The adoption of these frameworks highlights the clear advantages of fully automated systems in clinical and research applications. CNNs excel in processing large datasets with minimal variability, while semi-automated and manual methods depend heavily on human expertise and calibration. The mathematical rigor behind these methods underscores their critical role in advancing ECD measurement technologies.

However, one limitation of fully automated systems is their reliance on high-quality images and the availability of large, labeled datasets for training deep learning models. The use of these systems in under-resourced environments, where access to high-quality imaging and well-labeled data is limited, could be challenging.

Future research should focus on expanding the accessibility of AI-driven tools by developing models that can perform reliably even with lower-quality images or smaller datasets. Additionally, further studies are needed to evaluate the long-term clinical outcomes of using fully automated methods, particularly in post-surgical settings and corneal transplantation. Investigating how transfer learning and continuous learning can be applied to improve the performance of these models in diverse clinical environments is also critical [43].

Moreover, further efforts should be directed toward integrating AI-driven systems into routine clinical workflows, ensuring their seamless adoption in various healthcare settings. As AI continues to evolve, its potential for diagnosing and managing other ophthalmic conditions, beyond corneal endothelium, could be explored, paving the way for comprehensive AI-driven ocular health assessments. Finally, ethical considerations regarding the widespread implementation of fully automated systems, such as patient data privacy and the transparency of AI algorithms, should be addressed in future research.

9 Conclusion

The comparative analysis of manual, semi-automated, and fully automated methods for corneal endothelial cell density (ECD) measurement highlights key differences in accuracy, reliability, and efficiency. Manual methods show high variability, with mean ECD values ranging from 2146 to 2775 cells/mm^2, largely due to technician-dependent factors and inconsistent calibration. Semi-automated methods offer improved consistency, yielding mean ECD values around 2773 $\pm$ 300 cells/mm^2, but still demonstrate variability in individual measurements. Fully automated AI-based systems, like AxoNet, provide the highest accuracy, with mean ECD values of 2867 cells/mm^2, low variability, and strong reliability, indicated by an R^2 value of 0.948and an MAE of 12.1 cells/mm^2.

The clear advantage of fully automated methods is their ability to eliminate human error, reduce analysis time, and provide standardized, objective, and reproducible results. These improvements make AI-driven systems vital for clinical applications, where precise ECD measurements are crucial for corneal health assessments and transplant evaluations. The findings underscore the importance of adopting automated systems for more reliable and efficient ECD assessments, enhancing both diagnostic accuracy and clinical outcomes in ophthalmology.

Acknowledgements

The successful completion of this research was made possible through the collective efforts of individuals and institutions whose contributions are deeply appreciated. Special thanks are extended to the faculty and research staff whose guidance and expertise have provided invaluable insights into the methodologies explored in this study. Their critical feedback and support have played a significant role in refining the technical and analytical aspects of this work.

Gratitude is also expressed to the research community whose previous studies and innovations in endothelial cell density measurement have served as the foundation for this work. The continuous advancements in machine learning, deep learning, and AI-driven automation have significantly shaped the direction and execution of this study.

A special acknowledgement is extended to Dr. Darryl Ahner, Dean of the College of Engineering and Computer Science, whose financial support enabled the successful completion of this project. The resources provided have facilitated access to computational tools, datasets, and analytical frameworks essential for conducting this research.

Finally, appreciation is given to colleagues, peers, and family members for their encouragement and support throughout the research process. Their motivation and feedback have contributed to the successful realization of this work.

Conflicts of interest

The authors declare that they have no conflict of interest.

References

[1] Janson BJ, Alward WL, Kwon YH, et al. Glaucoma-associated corneal endothelial cell damage: A review. Survey of Ophthalmology. 2018, 63(4): 500-506. https://doi.org/10.1016/j.survophthal.2017.11.002

[2] Mannis MJ, Holland EJ. Cornea, E-Book. Elsevier Health Sciences, 2021.

[3] McCarey BE, Edelhauser HF, Lynn MJ. Review of Corneal Endothelial Specular Microscopy for FDA Clinical Trials of Refractive Procedures, Surgical Devices, and New Intraocular Drugs and Solutions. Cornea. 2008, 27(1): 1-16. https://doi.org/10.1097/ico.0b013e31815892da

[4] Bourne WM. Cell Density in Corneal Endothelium: Effect of Age and Measurement Method. Invest Ophthalmol Vis Sci. 1982, 22: 104-108.

[5] Laing RA, Sandstrom MM, Berrospi AR, et al. Changes in the corneal endothelium as a function of age. Experimental Eye Research. 1979, 29(4): 521-527.

[6] Mishima S. Clinical Investigations on the Corneal Endothelium. American Journal of Ophthalmology. 1982, 93(1): 1-29. https://doi.org/10.1016/0002-9394(82)90693-6

[7] Brown NAP, Bron AJ. Endothelial cell size, shape, and apical structure in the aging human cornea. British Journal of Ophthalmology. 1977, 6(11): 805-814.

[8] McCarey BE, Edelhauser HF. The effects of age on the corneal endothelium. Ophthalmology. 1981, 88(6): 649-660.

[9] Van Rossum G, Drake FL. Python 3 Reference Manual. Scotts Valley, CA: CreateSpace, 2009.

[10] McKinney W. Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference. Published online 2010: 56-61. https://doi.org/10.25080/majora-92bf1922-00a

[11] McLaren JW, Bourne WM. A computer-assisted technique for studying regular photomicrographs of the corneal endothelium. Cornea. 1991, 10(5): 445-453.

[12] Yee RW, Matsuda M, Kern JA. Computer-assisted endothelial morphometry in corneal endothelial images obtained by specular microscopy. American Journal of Ophthalmology. 1985, 99(6): 739-747.

[13] Smiddy WE, Bourne WM. The effects of long-term contact lens wear on the corneal endothelium. Archives of Ophthalmology. 1989, 107(10): 1421-1425.

[14] Hatou S, Miyashita H, Shimmura S, et al. Evaluation of the corneal endothelium using a newly developed noncontact specular microscope. Cornea. 2009, 28(4): 456-460.
[15] Murphy C, Alvarado J, Juster R, et al. Computerized morphometry of the corneal endothelium in long-term soft contact lens wearers. American Journal of Ophthalmology. 1984, 98(4): 478-486.
[16] Rivolta MW, Mantero M, Cereda M, et al. Automated analysis of corneal endothelial cell morphology using a novel software. PLoS ONE. 2017, 12(8): e0182585.
[17] Chui J, Girolamo ND, Wakefield D, et al. The Pathogenesis of Pterygium: Current Concepts and Their Therapeutic Implications. The Ocular Surface. 2008, 6(1): 24-43. https://doi.org/10.1016/s1542-0124(12)70103-9
[18] Yazdanpanah G, Jabbehdari S, Dana R. Corneal Endothelial Cell Dysfunction: Etiologies and Management. Curr Ophthalmol Rep. 2015, 3(4): 159-165.
[19] Benitez BA, Yee RW, Matsuda M, et al. Automated analysis of corneal endothelial cell images: comparison of methodologies. Cornea. 2010, 29(3): 287-294.
[20] Bourne WM. Cell Density in Corneal Endothelium: Effect of Age and Measurement Method. Invest Ophthalmol Vis Sci. 1982, 22: 104-108.
[21] Laing RA, Sandstrom MM, Berrospi AR, et al. Changes in the corneal endothelium as a function of age. Experimental Eye Research. 1979, 29(4): 521-527.
[22] Brown NAP, Bron AJ. Endothelial cell size, shape, and apical structure in the aging human cornea. British Journal of Ophthalmology. 1977, 61(11): 805-814.
[23] McCarey BE, Edelhauser HF. The effects of age on the corneal endothelium. Ophthalmology. 1981, 88(6): 649-660.
[24] Doughty MJ, Müller A, Zaman ML. Assessment of the Reliability of Human Corneal Endothelial Cell-Density Estimates Using a Noncontact Specular Microscope. Cornea. 2000, 19(2): 148-158. https://doi.org/10.1097/00003226-200003000-00006
[25] Kolluru C, Benetz BA, Joseph N, et al. Machine learning for segmenting cells in corneal endothelium images. Proceedings of Spie–the International Society for Optical Engineering. 2019, 10950: 109504G.
[26] Koprowski R, Wilczynski S, Wozniak M, et al. Computer-aided analysis of corneal endothelium images based on deep learning methods. Computerized Medical Imaging and Graphics. 2014, 44: 29–38.
[27] Modak C, Huang J, Maram J, et al. Automated endothelial cell density analysis with deep learning. Journal of Optometry. 2015, 12: 142–149.
[28] Price MO, Fairchild KM, Price FW. Comparison of Manual and Automated Endothelial Cell Density Analysis in Normal Eyes and DSEK Eyes. Cornea. 2013, 32(5): 567-573. https://doi.org/10.1097/ico.0b013e31825de8fa
[29] Huang J, Maram J, Tepelus TC, et al. Comparison of manual & automated analysis methods for corneal endothelial cell density measurements by specular microscopy. Journal of Optometry. 2018, 11(3): 182-191. https://doi.org/10.1016/j.optom.2017.06.001
[30] Maruoka S, Nakakura S, Matsuo N, et al. Comparison of semi-automated center-dot and fully automated endothelial cell analyses from specular microscopy images. International Ophthalmology. 2017, 38(6): 2495-2507. https://doi.org/10.1007/s10792-017-0760-7
[31] Heinzelmann S, Maier PC, Sandu K, et al. Assessing corneal endothelial cell density: automated versus manual countinJournal of Cataract and Refractive Surgery. 2016, 42: 1118–1125.
[32] Bourne WM. Cell Density in Corneal Endothelium: Effect of Age and Measurement Method. Investigative Ophthalmology & Visual Science. 1982, 2: 104-108.
[33] Laing RA, Sandstrom MM, Berrospi AR, et al. Changes in the corneal endothelium as a function of age. Experimental Eye Research. 1979, 29(4): 521-527.
[34] Brown NAP, Bron AJ. Endothelial cell size, shape, and apical structure in the aging human cornea. British Journal of Ophthalmology. 1977, 61(11): 805-814.
[35] McCarey BE, Edelhauser HF. The effects of age on the corneal endothelium. Ophthalmology. 1981, 88(6): 649-660.
[36] Bartels SP, Sperling M, Laing RA. The effect of specular microscope illumination on corneal endothelial cell count. Archives of Ophthalmology. 1981, 100(10): 1623-1626.
[37] Matsuda M, Yee RW, Edelhauser HF. Comparison of manual and automated endothelial cell counts in keratoconus. Ophthalmology. 1985, 92(9): 1146-1150.
[38] Waring GO, Bourne WM, Edelhauser HF, et al. The Corneal Endothelium. Ophthalmology. 1982, 89(6): 531-590. https://doi.org/10.1016/s0161-6420(82)34746-6
[39] Bourne WM, Nelson LR, Hodge DO. Cquot; Central corneal endothelial cell changes over ten years. Investigative Ophthalmology Camp; Visual Science. 1994, 35(7): 3023-3027.
[40] Laing RA, Sandstrom MM, Berrospi AR, et al. Changes in the corneal endothelium as a function of age. Ophthalmology. 1981, 88(4): 408-412.
[41] Hodges RR, Dartt DA, Jester JV. Cquot; Mechanisms of corneal endothelial cell migration. Ophthalmology. 1992, 99(3): 379-385.
[42] González P, Epstein DL. (1999). Endothelial cell behaviour after corneal surgery. Ophthalmic Surgery and Lasers. 1999, 30(5): 427-437.
[43] Møller-Pedersen T, Cavanagh HD, Petroll WM, et al. (1997). Cquot; Stromal wound healing explains refractive instability and haze development after photorefractive keratectomy: a 1-year confocal microscopic study. Ophthalmology. 1997, 104(10): 1564-1574.