Automated Bubble Detection in Contact Lenses Using a Hybrid Deep Learning Framework

Chee Chin Lim[1], Yen Fook Chong[2], Vikneswaran Vijean[3], Gei Ki Tang[4]
Faculty of Electronic Engineering & Technology, Universiti Malaysia Perlis, Perlis, Malaysia[1, 3, 4]
Sports Engineering Research Centre (SERC), Universiti Malaysia Perlis, Perlis, Malaysia[1, 2, 3]

***Abstract*—This study presents a hybrid deep learning approach for automated detection of bubbles in contact lenses, aiming to enhance quality assurance in the manufacturing process. A hybrid AlexNet+SVM model was developed using transfer learning, where AlexNet's convolutional features were leveraged for binary classification (bubble vs. normal) via a Support Vector Machine (SVM) classifier. The dataset consisted of 320 images (160 bubbles, 160 normal) pre-processed using median filtering, local histogram equalization, and circular masking to improve image clarity and consistency. Through systematic hyperparameter tuning, the model achieved 100% testing accuracy and 97.92% validation accuracy, with perfect precision (100%) and high recall (96%). Comparative evaluation against ResNet and VGGNet demonstrated that the AlexNet+SVM model offered superior generalization and robustness, particularly for small-scale datasets. While VGGNet also achieved 100% testing accuracy with 95.83% validation accuracy, ResNet underperformed in recall (89%), likely due to its deeper architecture and data limitations. The findings underscore the suitability of hybrid models for binary classification tasks in limited-data scenarios. Identified challenges, including dataset size and risk of overfitting, point to future research directions involving expanded datasets and more advanced pre-processing techniques. This research contributes to the advancement of automated defect detection systems for contact lens manufacturing, offering a reliable and efficient quality control solution.**

Keywords*—Bubble detection; contact lens quality assurance; deep learning; transfer learning; Support Vector Machine (SVM); AlexNet; image pre-processing; binary classification; defect detection

I. INTRODUCTION

The eye is one of the most sensitive and vital organs in the human body, serving as the primary means of gathering information from the external environment. Humans rely heavily on vision, more than any other sensory modality, for perception and interaction with their surroundings [1]. The development of contact lenses has significantly advanced vision correction since their inception. In 1888, the first glass contact lenses were developed by Adolf E. Fick and Edouard Kalt, opticians based in Paris. However, glass contact lenses were associated with several limitations, including excessive weight, high cost, and safety concerns. Their fragility made them prone to breaking upon impact, necessitating careful handling. These drawbacks led to the development of alternative materials, culminating in the invention of the first hydrophilic hydrogel soft contact lenses by Czech chemists Drahoslav Lim and Otto Wichterle in 1959 [2].

Despite these advancements, the manufacturing process of soft contact lenses is not without challenges. One significant issue is the formation of bubbles during production. The manufacturing process involves heating the lens material until it reaches a molten state, followed by injection into computer-designed molds under high pressure. After cooling, the lenses are removed from the mold and undergo polishing to achieve smooth edges and surfaces. The lenses are then hydrated and subjected to quality assurance testing. However, improper handling or suboptimal manufacturing conditions can lead to the entrapment of bubbles during the heating and molding stages. These bubbles compromise the structural integrity and optical performance of the lenses.

Current inspection methods for detecting bubbles in contact lenses often depend on subjective human judgment, resulting in inconsistencies that can compromise quality control and allow defective lenses to reach end-users, potentially leading to discomfort or vision issues. While significant advancements have been made in contact lens classification and quality assessment using deep learning, the automated and highly accurate detection of specific manufacturing defects—particularly air bubbles—remains insufficiently addressed in both academic research and industrial practice. Existing systems frequently struggle with precision, especially in scenarios involving small datasets or variable imaging conditions and tend to overlook the importance of preprocessing techniques in improving model reliability.

To bridge this gap, this study introduces a hybrid deep learning model that combines AlexNet with a Support Vector Machine (SVM) classifier, specifically optimized for robust bubble detection in constrained data environments. Enhanced by image quality-focused preprocessing strategies, the proposed approach aims to improve inspection consistency, reduce reliance on manual evaluation, and support the development of reliable, fully automated quality assurance systems for contact lens manufacturing. The remainder of this paper is structured as follows: Section II presents a comprehensive review of related work. Section III describes dataset acquisition and preprocessing techniques. Section IV details the proposed methodology, including model architecture and tuning strategies. Section V discusses the results and comparative analysis. Section VI presents the conclusion, study limitations, and recommendations for future research.

II. Related Work

This section reviews key studies relevant to bubble detection in contact lenses. Gautam and Mukhopadhyay [3] employed AlexNet as a deep convolutional neural network (CNN) architecture, combined with an SVM-based Error-Correcting Output Code (ECOC) classifier. They evaluated Linear, Quadratic, Cubic, and Gaussian SVM models, focusing on validation accuracy to assess performance. Their work demonstrated the effectiveness of deep CNNs for contact lens detection. Raghavendra, Raja, and Busch [4] extended SVM applications by using it for a three-class classification problem: no lens, soft contact lens, or textured contact lens. This approach highlighted SVM's utility in multi-class lens detection tasks.

In image preprocessing, Arvind Kumar [5] compared Median and Gaussian filters for denoising, concluding that the Median filter achieved better results in less time, emphasizing its efficiency for image enhancement. Mandalapu, Ramachandra, and Busch [6] utilized the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) to extract texture features from textured contact lenses. This method provided insights into image quality and texture patterns across different manufacturers, showcasing the potential of quality metrics for lens analysis.

In deep learning, H. Hou et al. [20] using a new detection network of YOLOv8-BGA model as a baseline, which can achieve effective identification of leakage bubbles and bubble images collected under different lighting conditions in a practical industrial inspection environment. Convolutional neural networks have been used to detect and classify defects in colored contact lenses. Kim Ty et al. [21] employed CNN GoogLeNet V4 to detect defective contact lenses in the AOI system that yielded a good product recognition rate of 83.3% despite the limitations of computational resources and data.

While these studies reflect advancements in contact lens classification and quality assessment, most existing approaches rely heavily on large datasets and lack targeted solutions for specific defects such as bubbles. Furthermore, limited attention has been given to preprocessing strategies and model robustness under small-scale data and varied lighting conditions. To address these challenges, this paper proposes a lightweight, generalizable deep learning framework—integrating AlexNet with SVM—enhanced by quality-focused preprocessing techniques, specifically designed for accurate and automated bubble detection in contact lens manufacturing.

III. Methodology

The workflow of the bubble detection system is illustrated in Fig. 1. The process begins with data collection, where images of contact lens specimens are captured using the experimental setup described earlier. No Reference Image Quality Assessment (NR-IQA) is performed on the raw images to evaluate the quality of images captured with and without additional lighting. Following this, data augmentation and image cropping are applied to prepare the dataset for pre-processing.

The pre-processing stage involves several image enhancement techniques, including median filtering, Gaussian filtering, power law transformation, local histogram equalization, circle detection, and circle masking. To validate the effectiveness of these preprocessing methods, Image Quality Assessment (IQA) metrics such as Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), and Absolute Mean Brightness Error (AMBE) are employed. After preprocessing, the images are resized to a uniform dimension to ensure consistency before classification. The classification stage utilizes the AlexNet convolutional neural network (CNN) model with transfer learning, incorporating Support Vector Machines (SVM) for enhanced performance. Parameter tuning is conducted to optimize hyperparameters such as epoch count, batch size, learning rate, and the number of workers. The model's performance is evaluated using a confusion matrix to assess its accuracy in detecting normal lenses and lenses with bubbles.

Finally, the proposed AlexNet model with SVM transfer learning is compared against two other state-of-the-art models, AlexNet, VGGNet and ResNet, to benchmark its effectiveness in bubble detection. This comprehensive approach ensures a robust and reliable system for identifying defects in contact lenses.

A. Dataset and Data Collection

The dataset used in this study was curated from real-world contact lens manufacturing samples provided by an industrial supplier. It includes 20 images of lenses with bubbles and 16 images of normal lenses for analysis. These images represent actual production defects, making the dataset highly relevant for developing industrial-grade defect detection systems. Though small in size, the dataset's authenticity and diversity in lighting conditions support the development of robust models using transfer learning and data augmentation techniques.

These specimens were imaged using a Logitech C922 Pro Stream Webcam for detailed inspection. The experimental setup included hardware (laptop, Logitech C922 camera, tripod stand, USB Type-A 2.0 cable) and software (PyCharm IDE) [7]. The camera, set to 1080p resolution, was mounted on a tripod at a 90-degree angle, 6 cm above the specimens as in Fig. 2. A black cardboard background was used to enhance contrast, facilitating bubble visualization against the white lenses. Two imaging modes were utilised: one without additional lighting and one with phone camera light. A custom camera app developed using PyQt5, QtGui, and Pyshine libraries, featured 'Start' and 'Take picture' buttons for image capture. All images were saved in PNG format at 1440×1080 resolution.

B. No Reference Image Quality Assessment

No Reference Image Quality Assessment (NR-IQA) was conducted using the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) on the raw images to evaluate the quality of images captured with and without additional lighting [8]. BRISQUE was selected for its ability to quantify image quality without a reference image. The imaging mode yielding lower BRISQUE values, indicating higher quality, was chosen for subsequent data augmentation.

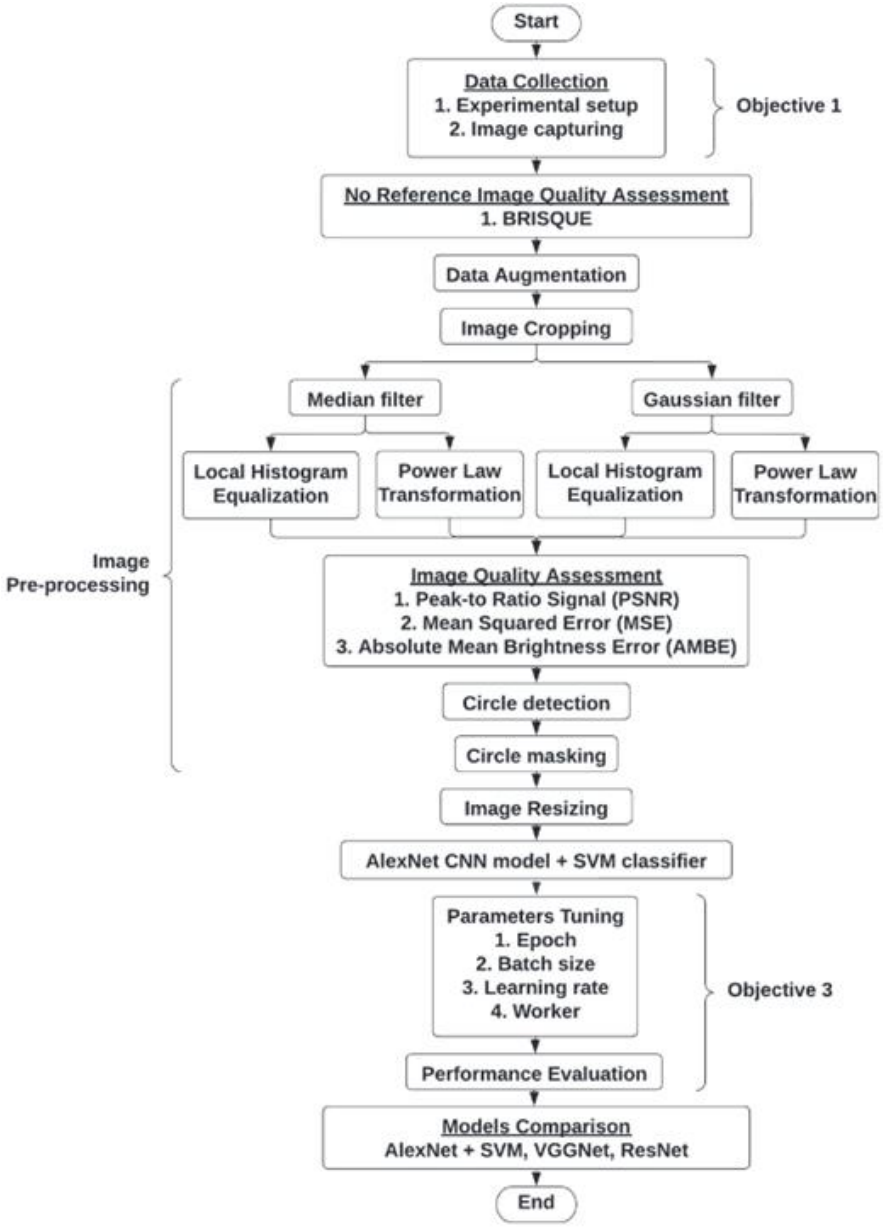

Fig. 1. Methodology workflow.

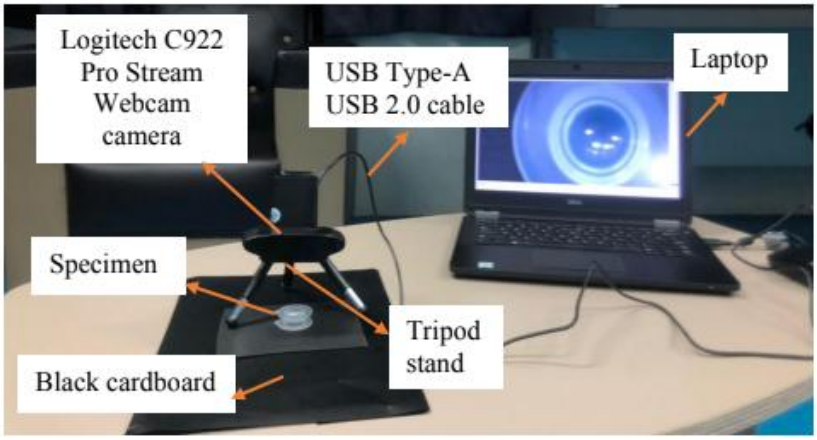

Fig. 2. Experimental setup.

C. Data Augmentation

Data augmentation was applied to enhance the dataset. For the 20 images of lenses with bubbles, each image was rotated at 0°, 90°, 180°, and 270°, resulting in 80 augmented images. For the 16 normal lens images, the first 10 images were rotated at 0°, 90°, 180°, and 270°, producing 40 images. The remaining 6 images were rotated at 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°, generating an additional 48 images. This resulted in a total of 88 normal lens images. Horizontal flipping was then applied to all augmented images, producing a final dataset of 320 images (160 with bubbles and 160 normal).

D. Image Cropping

The Region of Interest (ROI) for the dataset was the contact lens area, excluding the plastic mold and black cardboard background. Auto-cropping was performed on each image to remove unwanted regions, ensuring the background did not affect preprocessing. The cropped images varied in size between 240×240 and 350×350 pixels, as the ROI dimensions differed across specimens.

E. Image Pre-processing

Image pre-processing is essential for improving image quality, addressing factors like lighting intensity and camera quality. The process begins with segmenting the Region of Interest (ROI) within the boundary of interest, followed by focusing on the object of interest, categorized as either bubbles or normal.

1) Median filter: A non-linear filter used to reduce impulsive (salt-and-pepper) noise while preserving edges [9]. For a 5×5 filter, the centre pixel is replaced by the median value of the 25-pixel window, calculated by sorting pixel values and selecting the 13th value as in Eq. (1). This process is repeated across the image.

$$\hat{f}(x,y) = \underset{(s,t)\in S_{xv}}{median}\{g(s,t)\} \quad (1)$$

2) Gaussian filter: A linear filter applied to blur images or reduce noise. It performs a weighted average of surrounding pixels based on the filter size. A 5×5 kernel was used, with the sigma parameter controlling the blurring effect. The process iterates by shifting the window across the image by using Eq. (2).

$$g(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{(x^2+y^2)}{2\sigma^2}} \quad (2)$$

3) Local histogram equalization: Enhances contrast by adjusting image intensities using regional information. It spreads frequent intensity values, stretching the intensity range [10] to improve local contrast by using Eq. (3).

$$m_{s_{xy}} = \sum_{i=0}^{L-1} r_i P_{S_{xy}}(r_i) \quad (3)$$

4) Power law transformation: Maps narrow ranges of dark input values to wider output ranges or vice versa [11]. The transformation uses Eq. (4), where s and r are output and input pixel values, c is a constant, and γ controls the mapping. A γ value of 2 was chosen to darken the image, enhancing bubble edges.

$$s = c \cdot r^{\gamma} \quad (4)$$

F. Image Quality Assessment

Image Quality Assessment (IQA) evaluates distortions and degradations in images. Four pre-processing approaches were analyzed using Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), and Absolute Mean Brightness Error (AMBE). The approach yielding the best results across these metrics was selected for further processing.

1) PSNR: Measures the peak-to-noise ratio in decibels between the original and processed images. Higher PSNR values indicate better image quality [12]. It is calculated using Eq. (5), where R is the maximum fluctuation in the input image data type, and MSE is the mean squared error.

$$PSNR = log_{10}(\frac{R^2}{MSE}) \quad (5)$$

2) MSE: Represents the cumulative squared error between the original and processed images. Lower MSE values indicate smaller errors and better image quality [13]. Equation (6) calculates MSE, where n is the number of data points, Y_i is the observed value, and $\hat{Y}_i$ is the predicted value.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \quad (6)$$

3) AMBE: Assesses the preservation of original brightness. Lower AMBE values indicate better brightness preservation. Eq. (7) calculates AMBE, where E(X) and E(Y) represent the mean errors of the original and processed images, respectively.

$$AMBE = E(X) - E(Y) \quad (7)$$

G. Circle Detection and Masking

Circle detection was performed using OpenCV's Hough Gradient Method, which leverages gradient information from edges. The function cv.HoughCircles (image, method, dp, minDist, param1, param2, minRadius, maxRadius) was utilized with the following parameters:

- image: Pre-processed input images.
- method: cv2.HOUGH_GRADIENT for gradient-based edge detection.
- dp: Set to 1 to match the accumulator resolution with the input image.
- minDist: Set to 200 to avoid false detection of overlapping circles.
- param1: Set to 50 as the higher threshold for the Canny edge detector.
- param2: Set to 30 as the accumulator threshold for circle center detection.
- minRadius and maxRadius: Both set to 0 to detect circles of any size.

The circle detection process is defined by Eq. (8), where represents the circle's (x_{centre}, y_{centre}) and r is its radius. The output, HoughCircles, provides the detected circle parameters. After circle detection, masking was applied to isolate the Region of Interest (ROI). A mask was created by hiding the outer region of the detected circle. The original image was then combined with the mask, resulting in an output image where only the contact lens ROI is visible, and the background is black. This step ensures focus on the lens area while eliminating irrelevant background information.

$$(x - x_{centre})^2 + (x - y_{centre})^2 = r - HoughCircles() \quad (8)$$

H. Convolution Neural Network AlexNet + SVM Model

This study employs a transfer learning approach by integrating the AlexNet architecture with a linear Support Vector Machine (SVM) for bubble detection in contact lenses. The output images after the circle masking were resized to 227 by 227 as this is the optimum image size acceptable by the AlexNet CNN model. AlexNet, chosen for its robust performance in visual recognition tasks [3], consists of 8 layers: 5 convolutional and 3 fully connected. To adapt it to the binary classification task (bubbles vs. normal), the original 1000-class output layer [14] is replaced with a linear SVM classifier. Key modifications include:

- ReLU Activation: Replaces Tanh to accelerate training speed by sixfold [15].
- Dropout (Rate: 0.5): Mitigates overfitting, though doubling training time.
- Overlap Pooling: Reduces network size [16] and error rates (0.4% top-1, 0.3% top-5).

The model leverages transfer learning to address limitations of small datasets and binary classification. The dataset (320 images: 160 bubbles, 160 normal) is partitioned into 70% training, 15% validation, and 15% testing. The SVM classifier is applied to the final fully connected layer (fc8), enabling precise two-class categorization as shown in Fig. 3. This hybrid approach optimizes AlexNet's feature extraction capabilities with SVM's efficiency in linear classification, ensuring robust performance despite dataset constraints.

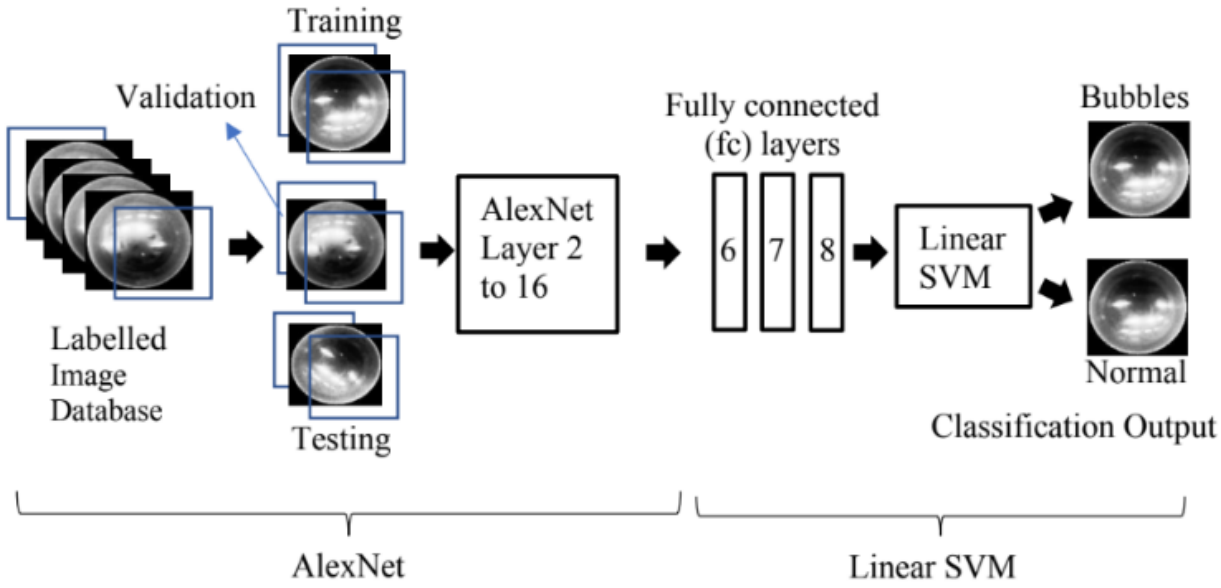

Fig. 3. Classification model of AlexNet + SVM on contact lens images.

IV. Result and Discussion

This section presents a comprehensive analysis of the experimental results, focusing on three critical phases of the study: (1) the evaluation of image pre-processing techniques to enhance dataset quality, (2) the optimization of hyperparameters for the AlexNet+SVM classification model, and (3) a comparative assessment of the performance of AlexNet, ResNet, and VGGNet architectures under identical conditions. First, the effectiveness of pre-processing methods is validated through No Reference Image Quality Assessment (NR-IQA) metrics, ensuring optimal input data for model training. Next, systematic parameter tuning—spanning epochs, batch size, learning rate, and workers—identifies configurations that maximize classification accuracy while mitigating overfitting. Finally, the optimized model is benchmarked against ResNet and VGGNet to determine the most robust architecture for bubble detection in contact lenses. The results highlight the interplay between data quality, model configuration, and architectural design, providing insights into achieving high precision and generalizability in defect detection tasks.

A. Result for Image Pre-processing

The BRISQUE scores for images captured without and with camera light were analyzed in Table I. For bubble specimens, the average BRISQUE values were 36.67 (without light) and 45.39 (with light). For normal specimens, the values were 43.89 (without light) and 48.98 (with light). Lower BRISQUE scores indicate better perceptual quality. Since images captured without camera light consistently showed lower BRISQUE values, they were selected for further pre-processing.

Four pre-processing approaches were evaluated using PSNR, MSE, and AMBE metrics as shown in Table II. The results showed that the median filter combined with local histogram equalization achieved the highest PSNR (27.870) and the lowest MSE (106.725) and AMBE (0.10518), indicating superior image quality. Consequently, this method was selected for pre-processing before model implementation.

TABLE I. BRISQUE VALUE FOR RAW DATA WITH AND WITHOUT CAMERA LIGHT

Data	BRISQUE (Average values)	
	Without camera light	With camera light
Bubble	36.67259	45.38722
Normal	43.89439	48.9829

TABLE II. AVERAGE IMAGE QUALITY ASSESSMENT OF THE ORIGINAL IMAGE AND IMAGE AFTER PRE-PROCESSING

Approach	Pre-processing Method	PSNR	MSE	AMBE
1	Median filter + Local histogram equalization	27.87000	106.72500	0.10518
2	Median filter + Power law transformation	26.75500	144.39200	0.18921
3	Gaussian filter + Local histogram equalization	27.86900	106.79500	0.10520
4	Gaussian filter + Power law transformation	26.75100	144.29800	0.18874

B. *Optimizing Turning Parameters for Deep Learning Classification Model*

This section analyzes the performance of the AlexNet+SVM model by evaluating the impact of varying parameters. The training and validation accuracy and loss were recorded and analyzed based on graph patterns. The results were compared between testing and validation accuracy for each tuning to identify optimal configurations.

1) Turning on Epoch: Epoch values of 40, 60, 80, 100, 120, and 140 were tested while keeping other parameters constant (learning rate: 0.0001, batch size: 6, workers: 8, optimizer: Adam, model: AlexNet, binary classification). The highest testing and validation accuracy (89.58%) were achieved at 100 epochs, indicating optimal performance as in Fig 4. Fig. 5 shows minimal gaps between training and validation accuracy and loss at this epoch, confirming model stability.

2) Turning on batch size: Batch sizes of 2, 4, 8, 16, and 32 were evaluated with other parameters fixed (epochs: 100, learning rate: 0.0001, workers: 8, optimizer: Adam, model: AlexNet, binary classification). A batch size of 32 yielded the highest testing accuracy (93.75%) and validation accuracy (83.33%). Larger batch sizes improved model performance, but further increases were impractical due to the validation size limit. Fig. 6 and Fig. 7 show the smallest gap between training and validation accuracy and loss at this batch size.

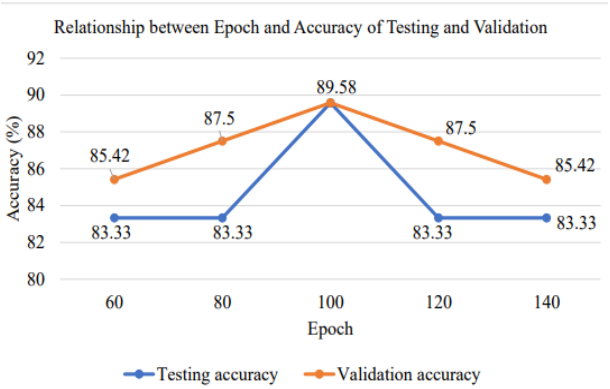

Fig. 4. Relationship between epoch and accuracy of testing and validation.

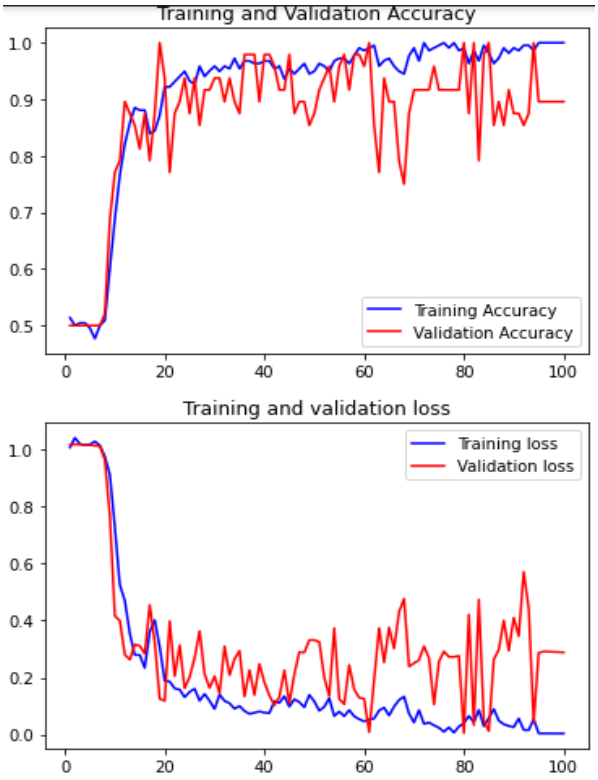

Fig. 5. Training and validation for accuracy and loss for tuning on epoch of 100.

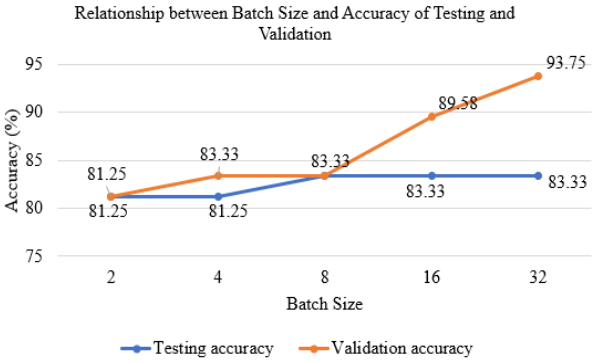

Fig. 6. Relationship between batch size and accuracy of testing and validation.

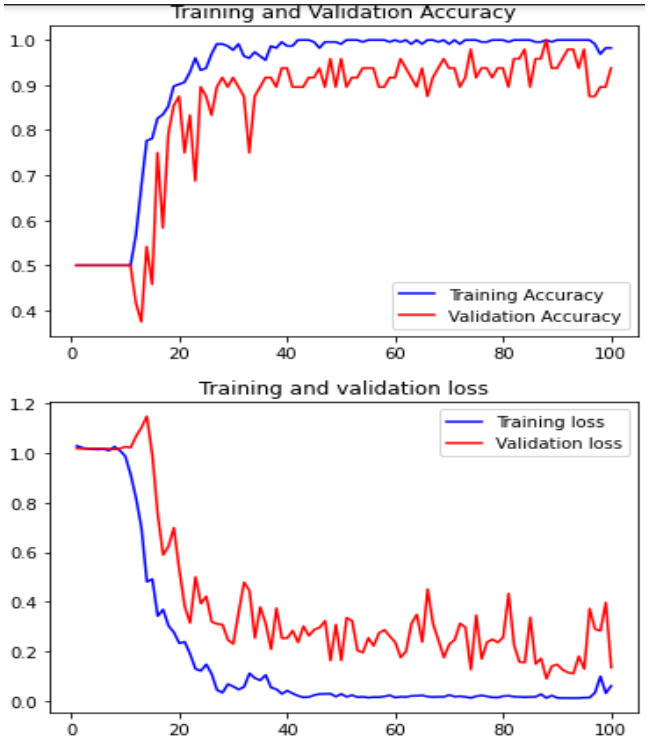

Fig. 7. Training and validation for accuracy and loss for tuning on batch size of 32.

3) Turning on learning rate: Learning rates of 0.1, 0.01, 0.001, 0.0001, 0.00001, and 0.000001 were tested while keeping other parameters constant (epoch: 100, batch size: 32, workers: 8, optimizer: Adam, model: AlexNet, binary classification). A learning rate of 0.0001 achieved the highest testing accuracy (87.5%) and validation accuracy (91.67%). Extremely low or high learning rates degraded performance, with 0.0001 providing the best balance. Fig. 8 and Fig. 9 demonstrate minimal gaps between training and validation accuracy and loss at this learning rate.

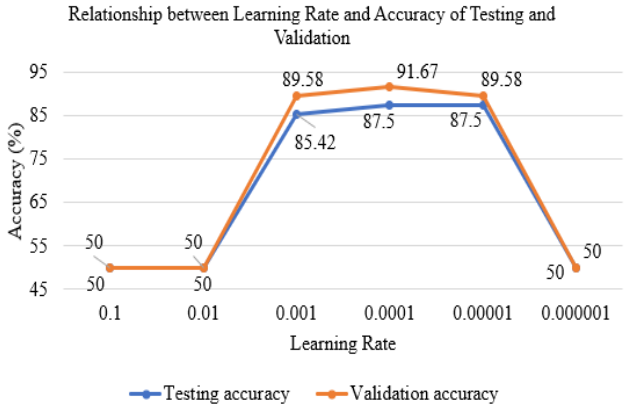

Fig. 8. Relationship between learning rate and accuracy of testing and validation.

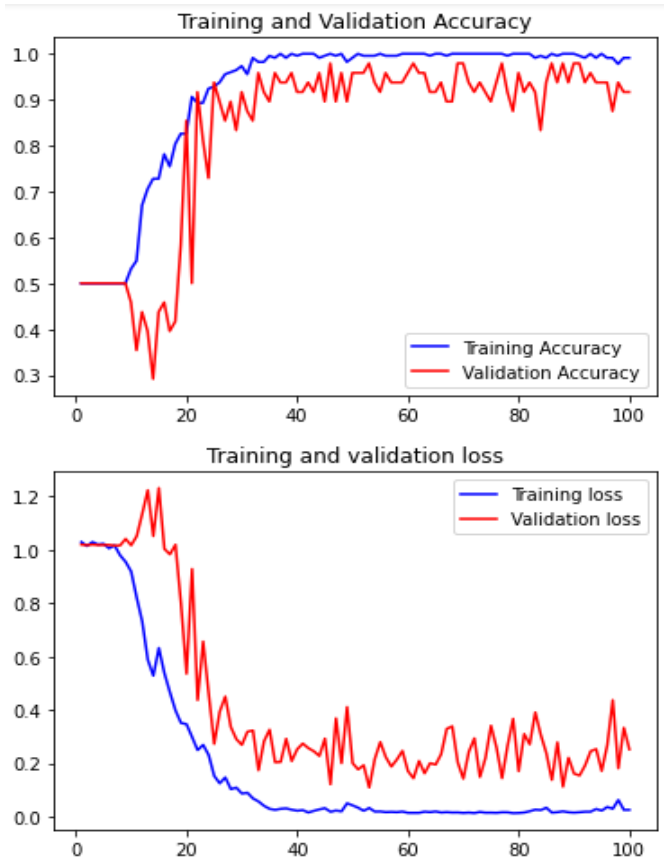

Fig. 9. Training and validation for accuracy and loss for tuning on learning rate of 0.0001.

4) Turning on worker: Worker values of 1, 2, 4, and 8 were tested with other parameters fixed (epochs: 100, learning rate: 0.0001, batch size: 32, optimiser: Adam, model: AlexNet, binary classification). A single worker achieved the highest testing and validation accuracy (93.75%). Additional workers did not improve performance, as the first worker efficiently managed batch loading and processing. Fig. 10 and Fig. 11 show the smallest gap between training and validation accuracy and loss with one worker.

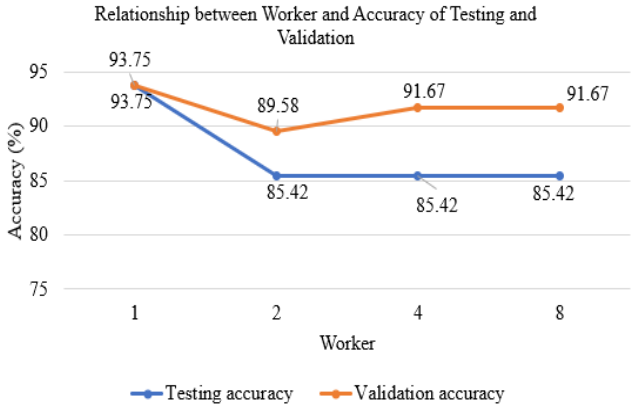

Fig. 10. Relationship between workers and accuracy of testing and validation.

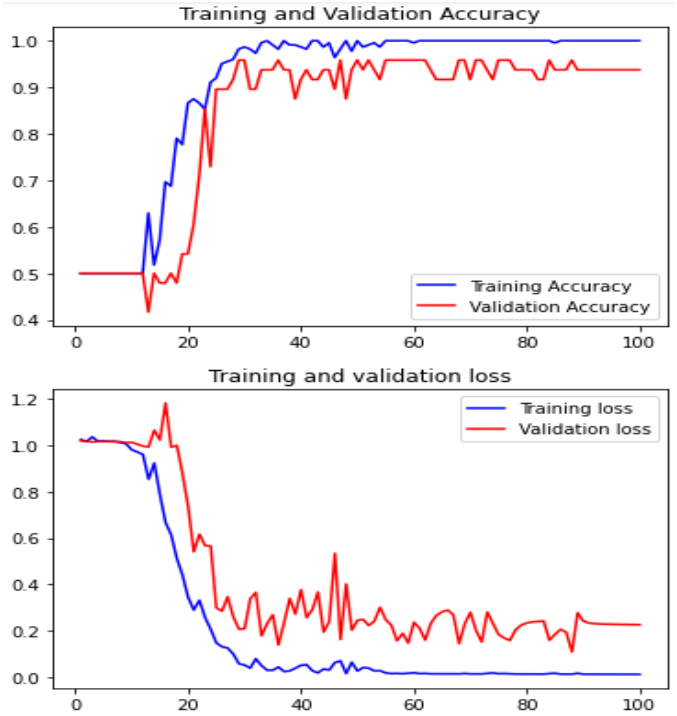

Fig. 11. Training and validation for accuracy and loss for tuning on workers of 1.

C. Comparison of Evaluation performance for AlexNet, ResNet and VGGNet and AlexNet+SVM at Optimum Tuning Parameters

After obtaining the optimal configuration, it was applied to AlexNet, ResNet, VGGNet and AlexNet+SVM to identify the best performing deep learning model for detecting bubble in contact lenses. This comparative analysis aimed to evaluate the effectiveness of each architecture under the same conditions, ensuring a robust and reliable solution for bubble detection. Table III presents the confusion matrix and its metrics of testing and validation accuracy results for AlexNet, ResNet, VGGNet and AlexNet+SVM models using identical parameters.

The AlexNet+SVM model achieved the highest validation accuracy (97.92%), outperforming VGGNet (95.83%) and ResNet (100%), while matching VGGNet's perfect testing accuracy (100%). All models demonstrated flawless precision (100%) and specificity (100%), confirming robust discrimination between bubbles and normal lenses. However, recall varied significantly: AlexNet and ResNet exhibited lower sensitivity (testing: 89%; validation: 89% and 100%, respectively), whereas VGGNet and AlexNet+SVM achieved perfect recall (100%) in testing, with AlexNet+SVM maintaining superior validation sensitivity (96%).

The AlexNet+SVM hybrid model's enhanced performance stems from integrating SVM's linear classification efficiency with AlexNet's feature extraction capabilities, mitigating overfitting risks inherent in deeper architectures like ResNet (138M parameters) [17]. VGGNet's intermediate parameter count (60.3M) and hierarchical feature learning likely contributed to its strong testing accuracy, though validation performance lagged slightly due to dataset variability [18]. AlexNet's moderate parameter size (62M) and SVM-driven classification optimized generalization, as evidenced by its balanced F1 scores (testing: 94; validation: 97.96). These results align with theoretical expectations: larger models (ResNet) risk overfitting on small datasets, while hybrid approaches (AlexNet+SVM) enhance stability by combining convolutional feature learning with linear discriminative power [19].

TABLE III. COMPARISON OF MODEL PERFORMANCE

	Model	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	F1 score
Testing	AlexNet	24	0	21	3	93.75	100	89	100	94
	ResNet	24	0	21	3	93.75	100	89	100	94
	VGGNet	24	0	24	0	100	100	100	100	100
	AlexNet+SVM	24	0	24	0	100	100	100	100	100
Validation	AlexNet	24	0	21	3	93.75	100	89	100	94
	ResNet	24	0	24	0	100	100	100	100	100
	VGGNet	24	0	22	2	95.83	100	92.31	100	96
	AlexNet+SVM	24	0	23	1	97.92	100	96	100	97.96

V. CONCLUSION

In conclusion, this study successfully achieved its objectives of developing a hybrid deep learning framework that integrates AlexNet with an SVM classifier to detect bubbles in contact lenses. The AlexNet model, optimized with specific parameters, achieved testing and validation accuracies of 93.75%. Performance evaluation using confusion matrix metrics demonstrated high precision, recall, specificity, and F1 scores. Comparative analysis revealed that the hybrid AlexNet+SVM model outperformed conventional AlexNet, ResNet, and VGGNet models achieving 100% testing accuracy and 97.92% validation accuracy, along with superior recall (96%) and F1 scores (97.96%).

Image preprocessing using median filtering and local histogram equalization significantly enhanced the quality of input data, contributing to superior classification performance. The optimized hyperparameters further ensured stable training and validation behavior. Although the model achieved high accuracy and demonstrated strong robustness, its reliance on a

small dataset may limit generalizability to broader production environments. This constraint, along with the observed risk of overfitting in deeper architectures like ResNet, highlights the need for larger, more diverse datasets to further validate the model's suitability for real-time industrial deployment. Future recommendations include enhancing pre-processing techniques, such as sharpening, and expanding the dataset to improve model accuracy and robustness. Additionally, exploring other hybrid architectures or advanced data augmentation techniques could further enhance model generalization and performance; and integrating the system into a live quality control pipeline.

ACKNOWLEDGMENT

The authors would like to thank Universiti Malaysia Perlis for providing the opportunity to collaborate with the contact lens manufacturing industry for practical research.

REFERENCES

[1] K. Irsch, J. H. Medicine, D. L. Guyton, and J. H. Medicine, "Encyclopedia of Biometrics," Encycl. Biometrics, no. January, pp. 10–16, 2009, doi: 10.1007/978-0-387- 73003-5.

[2] Mana, N. A. M. A., Chin, L. C., Fook, C. Y., Yazid, H, Ali, Y. M.: A review on contact lens inspection, Indonesian Journal of Electrical Engineering and Computer Science, 31(2), pp. 700-712, (2023).

[3] Gautam, G. and Mukhopadhyay, S.: Contact Lens Detection using Transfer Learning with Deep Representations. In: Proc. Int. Jt. Conf. Neural Networks, vol. 2018-July, (2018), doi: 10.1109/IJCNN.2018.8489590.

[4] Raghavendra, R., Raja, K.B., and Busch, C.: Ensemble of statistically independent fil-ters for robust contact lens detection in iris images. In: ACM Indian Conference on Computer Vision Graphics and Image Processing (ICVGIP), pp. 24:1–24:7 (2014).

[5] Arvind, K: Comparative analysis of gaussian filter, median filter and denoise autoeno-coder. IEEE Access (6), 45–51 (2020).

[6] Mandalapu, H., Ramachandra, R., and Busch, C: Image quality and texture-based fea-tures for reliable textured contact lens detection. In: Proc. - 14th Int. Conf. Signal Image Technol. Internet Based Syst. SITIS 2018, pp. 587–594 (2018), doi: 10.1109/SITIS.2018.00095.

[7] Mana, N. A. M A, Chin, L. C., Yazid, H., and Fook, C. Y, Evaluation of contact lens data acquisition approaches using enhancement techniques. In: 2022 4th International Conference on Artificial Intelligence and Speech Technology (AIST), Delhi, India, pp. 1-6, (2022), doi: 10.1109/AIST55798.2022.10065211.

[8] Shabashevich, A., Pierce, J., Dugan, W., Penoyer, G., and Kukla, A.: Project number P06219: Bubble elimination on the surface of a contact lens submerged in de-ionized water. In: Proceedings of the KGCOE Multi-Disciplinary Engineering Design Confer-ence, no. May, pp. 06219, Kate Gleason College of Engineering, Rochester Institute of Technology (2006).

[9] Nosrati, M. and Karimi, R.: Detection of circular shapes from impulse noisy images us-ing median and laplacian filter and circular hough transform. In: CCE 2011 - 2011 8th Int. Conf. Electr. Eng. Comput. Sci. Autom. Control. Progr. Abstr. B., no. October (2011), doi: 10.1109/ICEEE.2011.6106710.

[10] Sudhakar, S: Histogram Equalization. Towards Data Science, (2017). https://towardsdatascience.com/histogram-equalization-5d1013626e64, last accessed 2025/02/25.

[11] Kang and Atul: Power Law (Gamma) Transformations. The AI Learner, (2019) https://theailearner.com/2019/01/26/power-law-gamma-transformations/, last accessed 2025/02/25.

[12] MathWorks: Compute peak signal-to-noise ratio (PSNR) between images - Simulink. https://www.mathworks.com/help/vision/ref/psnr.html, last accessed 2025/02/25.

[13] Sara, U., Akter, M., and Uddin, M. S.: Image quality assessment through FSIM, SSIM, MSE and PSNR — A comparative study. J. Comput. Commun., 7(03) pp. 8–18, (2019) doi: 10.4236/jcc.2019.73002.

[14] Noor, S. S., Michael, K., Marshall, S., and Ren, J.: Hyperspectral image enhancement and mixture deep-learning classification of corneal epithelium injuries. Sensors, 17(11), 2644, (2017). https://doi.org/10.3390/s17112644

[15] Saxena, S.: Alexnet architecture | introduction to architecture of Alexnet. Analytics Vidhya, (2021). https://www.analyticsvidhya.com/blog/2021/03/introduction-to-the-architecture-of-alexnet/, last accessed 2025/02/25.

[16] Hassan, M. U.: AlexNet - ImageNet Classification with Convolutional Neural Networks. Neurohive, (2018) https://neurohive.io/en/popular-networks/alexnet-imagenet-classification-with-deep-convolutional-neural-networks/, last accessed 2025/02/25.

[17] He, K., Zhang, X., Ren, S., and Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 770–778, (2016).

[18] Simonyan, K. and Zisserman, A., Very Deep Convolutional Networks for Large-Scale Image Recognition. In: arXiv preprint arXiv:1409.1556, (2014).

[19] Krizhevsky, A., Sutskever, I., and Hinton, G. E., ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems (NIPS), pp. 1097–1105 (2012)

[20] Kim, T.-y., Park, D., Moon, H., & Hwang, S.-s.: A Deep Learning Technique for Optical Inspection of Color Contact Lenses. Applied Sciences, 13(10), 5966, (2023). https://doi.org/10.3390/app13105966

[21] Hou, H., Chou Z., Jian, C., Target Detection of Leakage Bubbles in Stainless Steel Pipe Gas Airtighness experiments Based on YOLOv8-BGA, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 16, No. 2, pp. 627 – 640, (2025). https://doi.org/10.14569/IJACSA.2025.0160264