JOURNAL OF APPLIED SCIENCE AND SOCIAL SCIENCE
eISSN 2229-3113 pISSN 2229-3205
Volume 15 Issue 10, October 2025
Impact factor: 2019: 4.679 2020: 5.015 2021: 5.436, 2022: 5.242, 2023: 6.995, 2024 7.75

FORMATION AND NORMALIZATION METHODS OF MEDICAL IMAGE DATABASES RELATED TO EYE DISEASES

Musayeva Muxtasar Zayirjon kizi

Abstract: This article discusses the processes of creating and normalizing a medical image database to improve the effectiveness of artificial intelligence (AI) technologies in the early diagnosis of eye diseases. It analyzes errors caused by image inconsistency, the need for data standardization, and the practical application of various normalization methods. The results show that applying normalization techniques can improve the accuracy of deep learning models by 10–15%.
Keywords: eye diseases, medical imaging, retina, database, normalization, deep learning, artificial intelligence, data preparation.

Introduction. Digital image analysis technologies are becoming increasingly important in medical diagnostics. In particular, AI technologies play a vital role in the early detection of eye diseases. Conditions such as diabetic retinopathy, glaucoma, and macular degeneration are often diagnosed in later stages, leading to vision loss. Therefore, automated image analysis and evaluation using diagnostic models are among the most relevant tasks in modern medicine.
Retinal images obtained from different clinics, under various conditions, and using different devices often vary significantly in lighting, contrast, size, and color range. Such inconsistencies reduce the accuracy of deep learning models during training. Hence, normalization—bringing all images to a common standard—is necessary before training the model. This article examines, on a scientific basis, the processes of creating a medical image database, cleaning, labeling, and normalizing data.
Scientific Basis for Forming an Eye Image Database. A reliable image database is essential for automated analysis of eye diseases. Such a database is used in the stages of model training, testing, and validation. The following principles are observed when forming an image database:

- *Diversity of data sources.* Images are obtained from various clinics, medical centers, and educational projects, which increases the model's generalization ability.
- *Inclusion of different disease stages.* Each disease (e.g., diabetic retinopathy) should include images from stage 0 to stage 4.
- *Metadata attachment.* Each image should be accompanied by information such as:
  - patient's age and gender,
  - disease type and stage,
  - device model (camera type),
  - image quality (lighting, contrast),
  - diagnosis provided by a physician.
- *Data anonymization.* To protect patient privacy, personal identifiers are removed from the images.

International research often utilizes open datasets such as EyePACS, Messidor, DRIVE, STARE, ODIR-5K, RIM-ONE, and CHASE-DB1. However, in Uzbekistan, collaboration with local medical clinics to create a national eye image database can significantly advance AI research.

JOURNAL OF APPLIED SCIENCE AND SOCIAL SCIENCE
eISSN 2229-3113 pISSN 2229-3205
Volume 15 Issue 10, October 2025
Impact factor: 2019: 4.679 2020: 5.015 2021: 5.436, 2022: 5.242, 2023: 6.995, 2024 7.75

Image Processing and Preparation Procedures. Since medical images vary in quality, they must undergo several preprocessing steps before being fed into deep learning networks:

1. Resizing. Images captured by different devices have varying dimensions. They are standardized to sizes such as 224×224, 256×256, or 512×512 pixels:

$$I'(x, y) = I(\alpha x, \alpha y)$$

where α is the scaling coefficient.

2. Color balance correction. Retinal images often have high contrast in red-green channels. The green channel provides the most diagnostic information, so some studies base their analysis on it.
3. Noise reduction. Differences in optical device quality can introduce noise, which can be mitigated using median filtering or Gaussian smoothing.
4. Illumination correction. Some images are brighter in the center and darker at the edges. Techniques like shading correction or homomorphic filtering are used to fix this.
5. Region of Interest (ROI) detection. Key regions such as the optic disc, macula, and fovea are automatically extracted and centered for analysis.
6. Data augmentation. To prevent overfitting, images are artificially expanded through rotation, translation, flipping, and color intensity adjustments.

Scientific Analysis of Normalization Methods. Normalization reduces differences between medical images and stabilizes model training. The most common normalization methods are:

- *Z-score normalization:* Each pixel value is scaled based on its deviation from the image mean:

$$X' = \frac{X - \mu}{\sigma}$$

where μ is the mean brightness and σ is the standard deviation. This helps maintain activation function balance in deep networks.

- *Min-Max scaling:* Pixel values are rescaled to the [0, 1] range:

$$X' = \frac{X - Xmin}{Xmax - Xmin}$$

This method speeds up computation and stabilizes gradients.

- *Histogram Equalization and CLAHE:* These methods enhance image contrast. The *Contrast Limited Adaptive Histogram Equalization (CLAHE)* algorithm is particularly effective for retinal images, optimizing contrast in local tiles and eliminating uneven lighting.
- Color channel normalization: In RGB images, each channel is normalized separately, reducing color noise and improving visibility of optic disc and blood vessel boundaries.

Experimental Results and Analysis. For experimentation, 5000 retinal images were selected from the ODIR-5K and Messidor datasets. Different normalization methods were applied, and the ResNet-50 model was used to classify diabetic retinopathy stages. The following results were obtained:

JOURNAL OF APPLIED SCIENCE AND SOCIAL SCIENCE
eISSN 2229-3113 pISSN 2229-3205
Volume 15 Issue 10, October 2025
Impact factor: 2019: 4.679 2020: 5.015 2021: 5.436, 2022: 5.242, 2023: 6.995, 2024 7.75

Normalization Type	Accuracy	F1-Score	Loss
None	80.3%	0.79	0.42
Z-score	86.5%	0.85	0.31
Min-Max + CLAHE	91.9%	0.90	0.21
RGB normalization	89.7%	0.88	0.25

The results show that applying normalization algorithms improves model accuracy by 10–12%. The best performance was achieved with the combination of CLAHE and Min-Max methods.

Conclusion. The efficiency of AI technologies in diagnosing eye diseases largely depends on the quality and normalization level of medical image databases. The research concludes that:

- Standardizing images is crucial for effective model training.
- CLAHE and Z-score methods yield the most optimal results for retinal images.
- Neural network accuracy and stability improve significantly with normalized data.

Future work involves creating a national eye disease image database based on these methods, developing a local diagnostic model, and implementing automated diagnostic systems in medical institutions.

References

1. Abramoff M.D., Lavin P.T., et al. *"Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy."* NPJ Digital Medicine, 2018.
2. Pratt H., Coenen F., Broadbent D. *"Convolutional Neural Networks for Diabetic Retinopathy."* Procedia Computer Science, 2016.
3. ODIR-5K Dataset – https://odir2019.grand-challenge.org/
4. Tan N.M., Acharya U.R., et al. *"Automated diagnosis of diabetic retinopathy using digital fundus images."* Journal of Medical Systems, 2018.
5. Nguyen Q., et al. *"Deep learning in ophthalmology: A review."* Eye and Vision, 2020.
6. Gonzalez R., Woods R. *"Digital Image Processing."* Prentice Hall, 2018.