JURNAL INFOTEL
VOL. 18, NO. 2, MAY 2026, PP. 180–194. DOI:10.20895/INFOTEL.V18I2.1541

RESEARCH ARTICLE

Classification of H2O with HCl and H2O with NaOH Solution Images Using Otsu Segmentation and CNN

Mauliza Putri[1], Melinda Melinda[2,*], Siti Rusdiana[3], Aufa Rafiki[4], and Lailatul Qadri Zakaria[5]

[1,2,4]Faculty of Engineering, Universitas Syiah Kuala, Aceh, Indonesia
[3]Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Aceh, Indonesia
[5]Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia

*Corresponding email: melinda@usk.ac.id

Received: February 03, 2026; Revised: March 13, 2026; Accepted: April 06, 2026.

Abstract: The classification of the image of chemical solutions is crucial in laboratory automation and chemical industry applications; however, it remains challenging when solutions such as H2O with HCl and H2O with NaOH exhibit nearly identical visual characteristics under imaging conditions, particularly when their spectral fluctuation patterns are visually subtle. This study proposes an image classification framework that integrates Otsu-based segmentation in the HSV color space with convolutional neural network (CNN) models to classify High Height Fluctuation (HHF) images generated from a Multi-Scale chemical detection system (MSCS). The dataset consists of 102 HHF images, evenly distributed between the two solution classes. Transfer learning is applied using three CNN architectures, namely EfficientNetV2S, DenseNet201, and EfficientNetB0, and performance is evaluated using accuracy, precision, recall, F1-score, and confusion matrix analysis. The experimental results show that DenseNet201 achieves the best overall performance, while EfficientNetV2S provides competitive results with computational efficiency and EfficientNetB0 yields a lighter model with lower recall. These findings indicate that combining segmentation with modern CNN architectures can effectively improve classification robustness in chemically similar solutions. This study presents a practical framework that combines Otsu-based HSV segmentation with transfer-learning CNNs to classify chemically similar solutions, providing actionable insights for deep learning-based chemical sensing applications.

Keywords: otsu segmentation, convolutional neural network, high height fluctuation, chemical solution classification

 ISSN: 2085-3688; e-ISSN: 2460-0997

1 Introduction

The classification of chemical solution images is an important topic in the field of image processing and computer vision, particularly to support laboratory automation and chemical industry applications that require fast and accurate identification of solutions [1], [2]. However, chemical solution classification remains a challenging task due to the high similarity of visual characteristics between different solutions, especially mixtures with nearly identical optical properties, such as H2O with HCl and H2O with NaOH [3]. These very subtle visual differences make it difficult for conventional approaches based on manual inspection or handcrafted feature extraction to achieve reliable classification performance [4].

The ability to automatically distinguish between H2O with HCl (acidic solution) and H2O with NaOH (alkaline solution) is essential for several real-world applications. In pharmaceutical manufacturing, accidental misidentification of these two solutions can trigger hazardous chemical reactions with serious safety consequences. In industrial wastewater treatment, accurate solution identification is required before neutralization to prevent environmental damage. In laboratory automation, manual inspection of visually indistinguishable solutions is error-prone and time-consuming, limiting throughput and reliability.

Traditional approaches for chemical solution classification include manual visual inspection, pH indicator colorimetry, and handcrafted feature extraction methods such as color histograms and texture descriptors. Although these methods are simple and interpretable, they fail to reliably classify solutions with nearly identical optical properties, such as H2O with HCl and H2O with NaOH under MSCS imaging, where spectral fluctuations differ only slightly. Studies in related domains have reported that even conventional image-based methods can achieve less than 80% accuracy when input images lack discriminative visual features [5], [6]. Deep learning, particularly CNN-based transfer learning, overcomes these limitations by automatically learning discriminative hierarchical features directly from image data without manual feature engineering, demonstrating consistently superior classification performance in recent studies [7], [8].

Despite these advances, the performance of CNN models is highly dependent on the quality of the input features, highlighting the importance of image pre-processing and segmentation stages in the overall classification pipeline.

Among various segmentation techniques, Otsu segmentation is a widely used automatic thresholding method that determines an optimal threshold by maximizing the variance between classes in grayscale images [9], [10]. In the context of chemical solution spectral imaging, Otsu segmentation plays an important role in extracting regions of interest (ROIs) that contain dominant spectral information, allowing CNN models to focus on informative regions while reducing the influence of background noise or irrelevant visual patterns [11]. Previous studies have shown that integrating Otsu segmentation with CNN-based classifiers can improve the classification accuracy of solution images that show high fluctuation patterns [12].

Several related studies have explored the application of computer vision and deep learning techniques to chemical and laboratory image analysis. El-Khawaldeh et al. [1] demonstrated the use of computer vision for real-time monitoring and analysis of chemical experiments in laboratory environments, highlighting the potential of visual-based automation, while not focusing on classification tasks involving visually similar chemical solutions. Meshkov et al. [2] investigated the integration of computer vision and collaborative robotic systems for chemical laboratory automation, emphasizing system-level efficiency

rather than detailed evaluation of visually indistinguishable solution classes. In terms of segmentation, several studies have shown that automatic thresholding techniques such as Otsu segmentation can improve the robustness of image-based analysis by extracting informative regions of interest from spectral or chemical images [11]. Yunidar et al. [13] further combined Otsu segmentation with CNN models to classify fluctuation height patterns in chemical solutions; however, their experiments were limited to a single CNN architecture. Meanwhile, modern CNN architectures such as DenseNet [14] and EfficientNet [15] have demonstrated strong feature representation capability and high parameter efficiency in various image classification tasks. Nevertheless, their comparative performance when applied to the classification of chemically similar solutions has not been systematically investigated.

Based on the above discussion, a clear research gap can be identified in the lack of comprehensive evaluation of modern CNN architectures integrated with Otsu segmentation for classifying chemical solutions with highly similar visual characteristics. Therefore, this study proposes an image classification framework that integrates Otsu segmentation with CNN models to distinguish H2O with HCl and H2O with NaOH solutions. Specifically, DenseNet201, EfficientNetB0, and EfficientNetV2S are evaluated due to their feature reuse capability, computational efficiency, and proven performance in image classification tasks.

The main contributions of this study are summarized as follows:

1. The application of Otsu segmentation to extract informative spectral regions of interest from chemical solution images.
2. A comparative evaluation of three modern CNN architectures, namely DenseNet201, EfficientNetB0, and EfficientNetV2S, to classify solutions with highly similar visual characteristics.
3. A comprehensive performance analysis using standard evaluation metrics to assess the effectiveness of the proposed classification framework.

Practical insights and recommendations for implementing deep learning–based chemical solution classification systems in laboratory environments.

2 Research Method

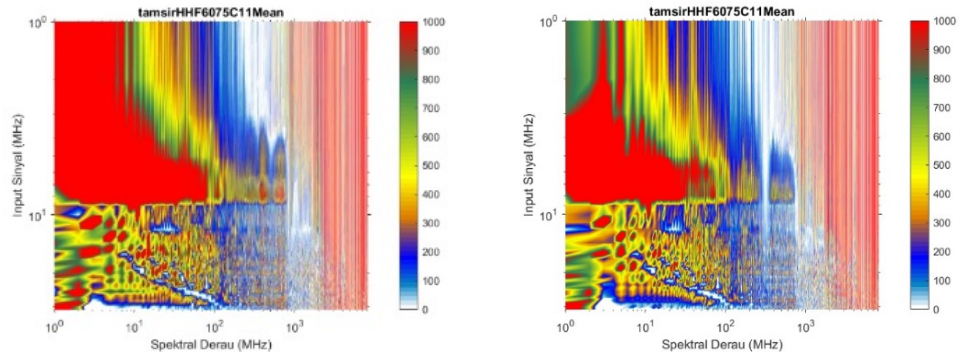
This section presents the theoretical foundations and research procedures used to develop an image classification system for H2O with HCl and H2O with NaOH solutions. The discussion covers the dataset characteristics, HHF image generation, pre-processing, Otsu-based segmentation, CNN-based classification using DenseNet201, EfficientNetB0, and EfficientNetV2S, as well as performance evaluation using standard metrics.

2.1 Dataset

The data used in this study are high-hitch fluctuation (HHF) pattern images generated from the Multi-Scale Chemical Sensing (MSCS) system. The MSCS system records the spectral responses of two chemical solutions, namely H2O with HCl and H2O with NaOH, which are measured separately. The acquired spectral data are stored in .dat file format and processed in MATLAB to generate two-dimensional visual representations.

The solution spectra are converted to images using the HHF method, which extracts and highlights the dominant spectral amplitude fluctuations. This approach has been introduced in previous studies to visualize intensity patterns from spectral signals of electrolyte

and non-electrolyte solutions [4], [16]. This dataset consists of 102 original HHF images, comprising 51 images of H2O with HCl solution and 51 images of H2O with NaOH solution. All images are original samples generated from the MSCS system. Data augmentation was only applied to the training set and was not included as additional dataset samples.



(a) (b)

Figure 1: Fluctuation pattern of HHF: (a) H_2O with HCl, (b) H_2O with NaOH.

2.2 Research Flow and Pre-processing

This study follows a systematic methodology consisting of dataset preparation, pre-processing, segmentation, CNN model training, and model evaluation. After HHF images are generated, the following pre-processing stages are applied prior to model training:

1. Active Spectral Area Cropping
   Image regions that do not contain spectral information, such as axis labels, margins, and scales, are removed by cropping based on fixed coordinates determined from HHF image observations. This step retains only the region that represents the spectral fluctuation patterns of the solution.
2. Image Size Adjustment
   Cropped images are resized to 224 $\times$ 224 pixels using bilinear interpolation. This resolution is selected because it is a standard input size for modern CNN architectures, allowing direct use of pre-trained networks without structural modification.
3. Normalization of pixel values
   Pixel values are normalized to the range [0,1] for each color channel. This normalization improves numerical stability during optimization, accelerates convergence, and helps prevent domination by specific intensity ranges.

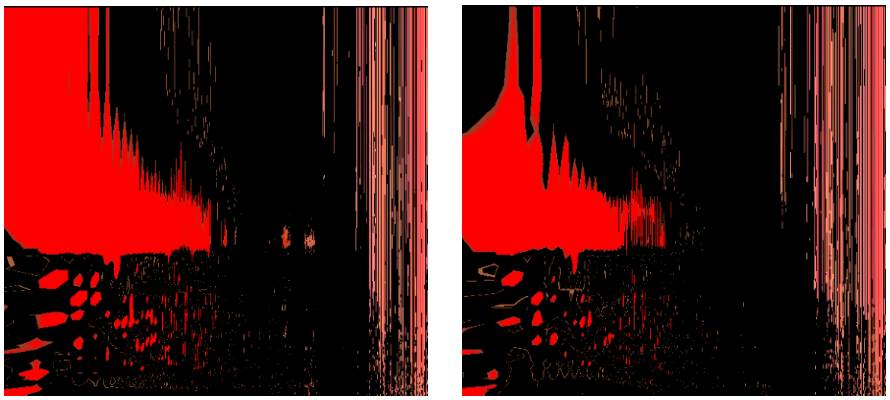
2.3 Otsu Segmentation

Otsu segmentation is an automatic threshold-based image segmentation method introduced by Nobuyuki Otsu [9]. The method determines an optimal threshold by maximizing inter-class variance in the image intensity histogram, or equivalently minimizing intra-class variance [10]. Its primary advantage is the ability to perform segmentation without manually defining threshold parameters.

In this study, Otsu thresholding is applied in the HSV color space to segment HHF images. HSV is used because it can better represent perceptual color information and is relatively robust to intensity variations compared to RGB. The Otsu method automatically separates the dominant spectral patterns from the background by maximizing the variance between classes in the pixel intensity distribution. The goal of this step is to extract the region of interest (ROI) that contains the most significant spectral fluctuations, thereby reducing background interference and improving feature learning during CNN training.

In the context of spectral images of chemical solutions, Otsu segmentation has been used to extract ROIs containing meaningful spectral information [11]. Research by Yunidar et al. [13] also indicates that the application of Otsu segmentation prior to CNN-based classification can improve accuracy by focusing the learning on informative regions while reducing the influence of noise and background.

After segmentation, a cropping stage is performed to further isolate the most relevant part of the image prior to classification [17]. Without cropping, the model processes the entire image area, including irrelevant regions, which can reduce efficiency and accuracy. In this study, the HHF pattern images are cropped to a size of 577 $\times$ 438 pixels prior to resizing, retaining only the HHF fluctuation region.



(a) (b)

Figure 2: Cropped ROI of HHF pattern: (a) H2O with HCl, (b) H2O with NaOH.

2.4 Convolutional Neural Network

Convolutional Neural Networks (CNNs) are widely used for image analysis tasks due to their ability to model spatial hierarchies in visual data through convolutional operations [18]. A typical CNN consists of convolutional layers, pooling layers, and fully connected layers that learn discriminative features from low-level patterns to high-level semantic representations. CNNs can automatically learn feature representations directly from raw image data, thereby reducing the reliance on handcrafted feature extraction.

CNN-based approaches have been applied to various image classification domains, including medical imaging and spectral image analysis, where they often achieve superior performance compared to conventional methods based on machine learning [7], [8]. Therefore, CNN is selected as the main classification method in this study to distinguish

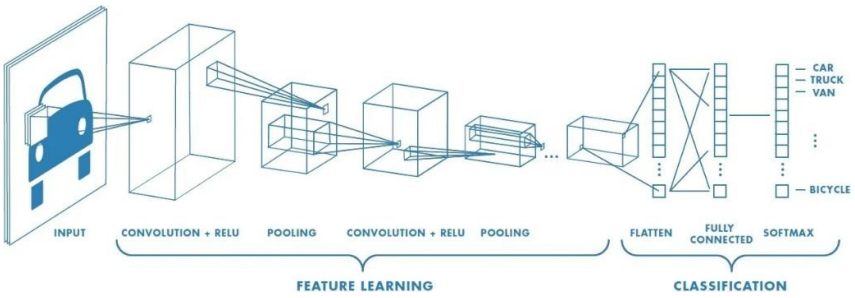


Figure 3: CNN architecture.

H2O with HCl and H2O with NaOH solution images, which exhibit highly similar visual characteristics. This study evaluates three CNN architectures, namely EfficientNetV2S, DenseNet201, and EfficientNetB0, implemented as separate models to compare their classification performance.

2.4.1 EfficientNetV2

EfficientNetV2 is an evolution of the EfficientNet family that emphasizes training efficiency and computational practicality through architectural refinements and training-aware design choices [15]. The variant EfficientNetV2S employs a combination of Fused-MBConv and MBConv blocks to improve computational throughput in the early stages of the network while maintaining strong representational capacity for image classification tasks [19]. In addition, EfficientNetV2 incorporates a progressive learning strategy that allows the network to adapt effectively during training, contributing to improved efficiency without a substantial loss of predictive performance [20].

In this study, the EfficientNetV2S model is utilized with ImageNet pre-trained weights, which were learned from a large-scale dataset containing over one million images across 1,000 object categories. This pre-training enables the model to extract rich and discriminative visual features and facilitates effective transfer learning for domain-specific image classification tasks. EfficientNetV2S adopts a standard input image size of 224×224 pixels, making it compatible with the commonly used CNN-based classification pipelines. Consequently, EfficientNetV2S is employed in this study to evaluate the performance of modern CNN architectures in classifying chemical solution images under conditions of high visual similarity.

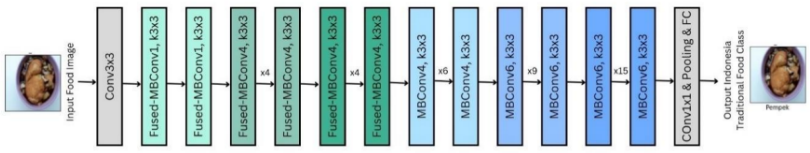


Figure 4: EfficientNetV2 architecture.

2.4.2 DenseNet201

DenseNet201 is a deep convolutional neural network architecture that employs a dense connectivity strategy, in which each layer receives feature maps from all preceding layers as its input. This dense connection pattern facilitates efficient feature propagation, improves feature reuse, and improves gradient flow throughout the network, thus stabilizing the training process in deep architectures [18], [21]. DenseNet201 consists of 201 layers organized into multiple dense blocks and transition layers, allowing the network to learn complex and hierarchical feature representations.

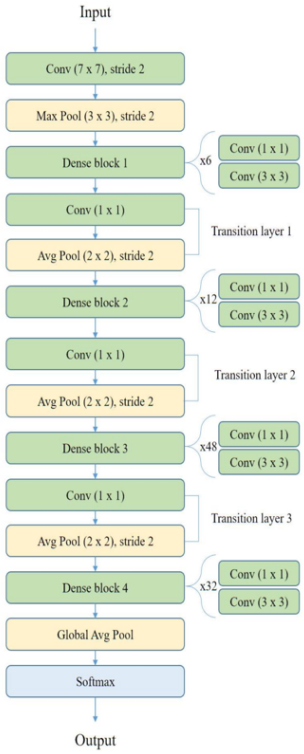


Figure 5: DenseNet201 architecture.

DenseNet201 is commonly used in transfer learning scenarios through a pre-trained model trained on the ImageNet dataset, which contains over one million images across 1,000 object categories. Pre-training on such a large-scale dataset enables the network to learn rich and generic visual features that can be effectively transferred to domain-specific classification tasks. In this study, the standard input image size of 224×224 pixels is used to ensure compatibility with the pre-trained DenseNet201 architecture and to allow direct utilization without modifying the network structure.

Recent studies have demonstrated that DenseNet-based models achieve strong performance and stable convergence in image classification tasks involving complex visual patterns, including scientific and medical imaging applications, due to their ability to preserve discriminative features across network layers [7], [8]. Furthermore, integration of DenseNet

architectures with appropriate image pre-processing or segmentation techniques has been shown to further improve classification performance by directing the network's attention to informative regions and reducing the influence of background noise. Based on these characteristics, DenseNet201 is evaluated in this study to assess its effectiveness in classifying HHF images of chemical solutions that exhibit high visual similarity.

2.4.3 EfficientNetB0

EfficientNetB0 is the base model of the EfficientNet family, developed using the compound scaling strategy to optimally balance the depth, width, and input resolution of the network [15]. This scaling approach enables EfficientNetB0 to achieve competitive classification performance while maintaining a relatively small number of parameters compared to conventional convolutional neural network architectures.

EfficientNetB0 employs Mobile Inverted Bottleneck Convolution (MBConv) blocks combined with Squeeze-and-Excitation mechanisms to enhance channel-wise feature representation and improve parameter efficiency. This architectural design allows the model to effectively extract representative visual features without incurring excessive computational costs [22].

In this study, EfficientNetB0 utilizes weights pre-trained on the ImageNet dataset, which contains over one million images across 1,000 object categories. The use of pre-trained weights facilitates transfer learning by enabling the model to leverage rich and generalizable visual features learned from large-scale data. EfficientNetB0 adopts a standard input image size of 224 × 224 pixels, making it well suited for image classification tasks with limited computational resources and relatively small datasets, such as the chemical solution image dataset used in this study.

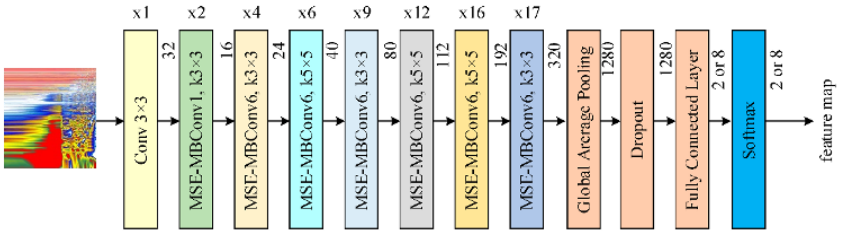


Figure 6: EfficientNetB0 architecture.

2.5 Data Splitting, Training Configuration, and Augmentation

The dataset was divided into training (75%), validation (15%), and testing (15%) sets using a stratification approach to ensure class balance between H2O with HCl and H2O with NaOH images. This splitting strategy was applied to each experiment to maintain a consistent evaluation protocol across all CNN models. It is important to note that the confusion matrix presented in the Results section is not derived from a single test split. Instead, it is generated by combining the prediction results from multiple evaluation experiments. As a result, the total number of samples represented in the confusion matrix exceeds 15%

of the total dataset. This pooling strategy was adopted to obtain more stable and reliable performance estimates, especially given the relatively small size of the dataset.

All images were resized to 224 × 224 pixels and normalized before being fed into CNN models. Transfer learning was applied using ImageNet pre-trained weights for EfficientNetV2S, DenseNet201, and EfficientNetB0. Training was performed for a maximum of 25 epochs with early stopping based on the validation loss. Adam Optimizer was used with a learning rate of 0.0001 and a batch size of 16. Real-time data augmentation, including rotation, flipping, and scaling, was applied only to the training set to improve model generalization and prevent overfitting.

2.6 Evaluation Metrics and Confusion Matrix

The performance of the model is evaluated using an independent test set. The confusion matrix is used to summarize the classification results based on true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) [16]. In addition, precision, recall, precision, and the F1 score are calculated as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

TP, TN, FP, and FN are True Positive (correctly predicted positive), True Negative (correctly predicted negative), False Positive (incorrectly predicted positive), and False Negative (incorrectly predicted negative) [17].

3 Results

This section presents the experimental results obtained from the evaluation of the three CNN architectures, namely EfficientNetV2S, DenseNet201, and EfficientNetB0, using the HHF image dataset of H2O with HCl and H2O with NaOH solutions.

3.1 Training and Validation Results

All CNN models were trained under the same configuration and evaluated using validation data. DenseNet201 achieved the highest validation accuracy of 97.56%, followed by EfficientNetV2S with 95.12% and EfficientNetB0 with 90.24%. The training and validation curves shown in Figure 7, Figure 8, and Figure 9 indicate stable convergence for DenseNet201 and EfficientNetV2S, while EfficientNetB0 exhibited a slightly higher validation loss.

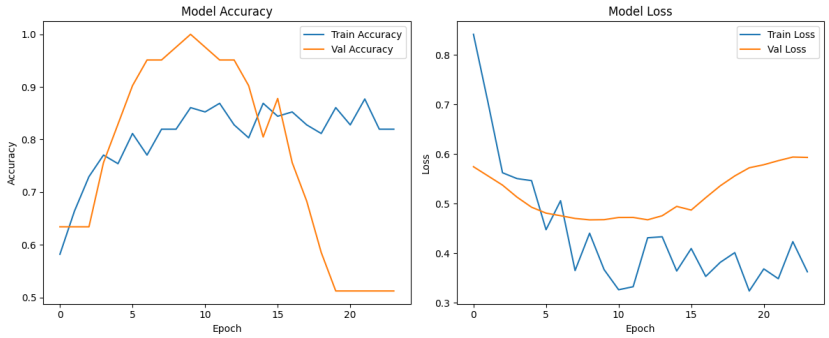


Figure 7: Training and validation learning curves of EfficientNetV2S.

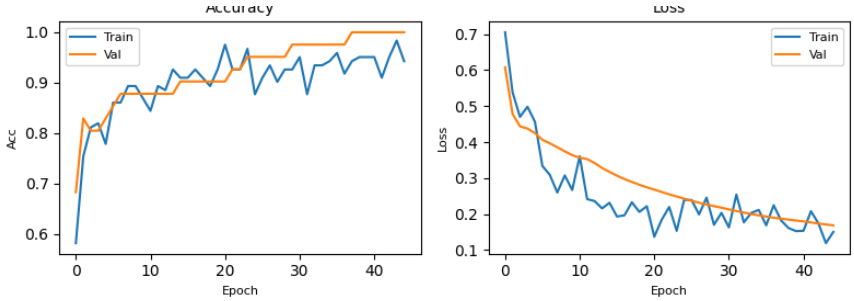


Figure 8: Training and validation learning curves of DenseNet201.

3.2 Confusion Matrix and Performance Analysis

Although the original dataset is perfectly balanced between the two classes, the number of samples for each class in the confusion matrix is not exactly equal. This is due to the data separation and aggregation process during the evaluation. This slight imbalance does not indicate dataset bias and does not affect the validity of the model's performance. Therefore, precision, recall, and F1-score have been reported to provide a comprehensive and fair evaluation of the performance of each CNN model based on precision, precision, re-

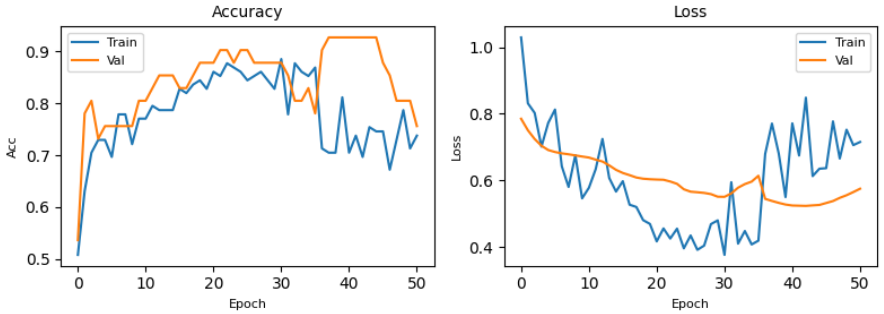
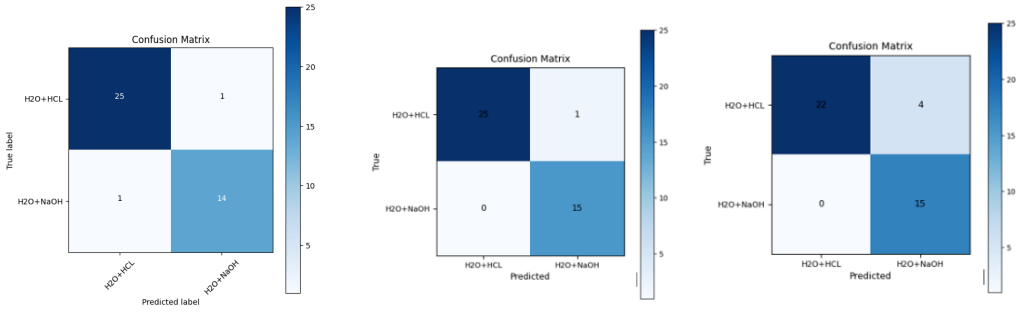


Figure 9: Training and validation learning curves of EfficientNetB0.

call, and F1-score, as summarized in Table 1. The confusion matrices for EfficientNetV2S, DenseNet201, and EfficientNetB0 are illustrated in Figure 10, which visualizes the distribution of correct and incorrect predictions for each solution class.



(a) (b) (c)

Figure 10: Confusion Matrix (a) EfficientNetV2, (b) DenseNet201 (c) EfficientNetB0.

For EfficientNetV2S, the model achieved an accuracy of 95.12%, with balanced precision and recall values of 0.96, resulting in an F1-score of 0.96, as reported in Table 1. As shown in Figure 10a, EfficientNetV2S produced a low number of misclassifications, indicating stable and reliable classification performance for both solution classes.

Table 1: Performance comparison based on confusion matrix

Model	Accuracy (%)	Precision	Recall	F1 score
EfficientNetV2S	95.12	0.96	0.96	0.96
DenseNet201	97.56	1.00	0.96	0.98
EfficientNetB0	90.24	1.00	0.85	0.91

DenseNet201 demonstrated the best overall performance among the evaluated models. As presented in Table 1, DenseNet201 achieved the highest accuracy of 97.56%, with a precision of 1.00, recall of 0.96, and an F1-score of 0.98. The corresponding confusion matrix in Figure 10b shows the dominance of true positive and true negative predictions with minimal false positives and false negatives, confirming the robustness of DenseNet201 in distinguishing HHF patterns.

EfficientNetB0 achieved an accuracy of 90.24%, with a precision of 1.00, recall of 0.85, and an F1-score of 0.91, as reported in Table 1. As illustrated in Figure 10c, the model exhibits a higher number of false negative predictions compared to the other architectures, which explains the lower recall value and reflects limitations in capturing subtle spectral variations due to its lightweight design. In general, the quantitative results in Table 1, together with the confusion matrix visualizations in Figure 10, confirm that DenseNet201 provides the most robust classification performance, followed by EfficientNetV2S, while EfficientNetB0 offers a trade-off between computational efficiency and classification accuracy.

4 Discussion

The experimental results demonstrate that the proposed framework, which combines Otsu-based segmentation and Convolutional Neural Network (CNN) architectures, is effective for classifying High Height Fluctuation (HHF) images of H2O with HCl and H2O with NaOH solutions that exhibit highly similar visual characteristics. The performance differences observed between EfficientNetV2S, DenseNet201, and EfficientNetB0 highlight the impact of architectural design choices on the capability of feature representation when dealing with subtle spectral variations.

Among the evaluated models, DenseNet201 consistently achieved the best performance in all evaluation metrics. This superior performance can be attributed to the dense connectivity mechanism employed by DenseNet, which enables the reuse of features across layers and improves the propagation of gradients during training. Such characteristics are particularly beneficial for relatively small datasets, such as the HHF dataset used in this study, where preserving fine-grained spectral features is critical for accurate classification. The minimal number of misclassifications observed in the confusion matrix further indicates that DenseNet201 is able to capture discriminative patterns associated with the spectral characteristics of the two solutions. EfficientNetV2S also demonstrated strong classification performance, indicating that training-aware architectural refinements can provide competitive accuracy while maintaining computational practicality. The use of Fused-MBConv and MBConv blocks supports efficient feature extraction without substantially reducing representational capacity, suggesting that EfficientNetV2S provides a favorable trade-off between accuracy and efficiency. This property is relevant for laboratory automation scenarios where computational constraints or deployment considerations may limit the use of heavier architectures.

In contrast, EfficientNetB0 exhibited comparatively lower performance, particularly in terms of recall. Although the model achieved high precision, the increase in the number of false negative predictions indicates limitations in detecting subtle spectral differences between the two solution classes. This behavior is consistent with the lightweight design of EfficientNetB0, which prioritizes parameter efficiency. Although this design can be advantageous for low-resource environments, it may reduce the ability to learn the highly discriminative features required to distinguish visually similar HHF patterns.

The confusion matrix analysis further reinforces the importance of integrating suitable pre-processing with deep learning architectures. Otsu segmentation contributes by isolating informative spectral regions and suppressing irrelevant background patterns, thereby allowing CNN models to focus on the most salient HHF features. This segmentation step is particularly important when the visual distinction between classes is subtle and the dataset size is limited.

Despite the encouraging results, several limitations should be acknowledged. First, the dataset size is relatively small and limited to two solution classes, which may restrict the generalizability of the findings to broader chemical solution categories. Second, the experiments rely on HHF image representations generated under a specific sensing setup and processing pipeline, and variations in acquisition conditions may influence the visual patterns learned by the models. Third, although transfer learning improves performance on limited data, the domain shift between ImageNet and HHF images may limit the optimality of feature representations. Addressing these limitations would support stronger evidence for deployment in practical laboratory environments.

5 Conclusion

This study developed an image classification framework to distinguish HHF images of H2O with HCl and H2O with NaOH solutions by integrating Otsu-based segmentation with CNN-based classifiers. The experimental evaluation indicates that the proposed pipeline can achieve reliable performance in chemically similar solutions, where conventional visual inspection and handcrafted feature extraction approaches are often insufficient. Among the evaluated architectures, DenseNet201 provided the most robust classification results, while EfficientNetV2S offered competitive performance with improved computational practicality, and EfficientNetB0 presented a lighter alternative with lower sensitivity to subtle spectral variations. Future work will focus on expanding the dataset to include additional solution types and broader acquisition conditions to improve generalization. Further investigation will also consider alternative segmentation strategies, domain-specific fine-tuning approaches, and additional validation protocols to strengthen robustness in real laboratory deployment scenarios.

Acknowledgments

The authors express their sincere gratitude to Universitas Syiah Kuala (USK) for the institutional support and facilities provided during the completion of this research.

References

[1] R. El-Khawaldeh, M. Guy, F. Bork, N. Taherimakhsousi, K. N. Jones, J. M. Hawkins, L. Han, R. P. Pritchard, B. A. Cole, S. Monfette, *et al.*, "Keeping an "eye" on the experiment: Computer vision for real-time monitoring and control," *Chemical Science*, vol. 15, no. 4, pp. 1271–1282, 2024.

[2] A. V. Meshkov, V. Y. Yurova, T. A. Aliev, V. V. Potapov, M. D. Rudakova, A. P. Ageev, and E. V. Skorb, "Collaborative robots using computer vision applications in a chemical laboratory," *Mendeleev Communications*, vol. 34, no. 6, pp. 769–773, 2024.

[3] X. Liu, H. An, W. Cai, and X. Shao, "Deep learning in spectral analysis: Modeling and imaging," *TrAC Trends in Analytical Chemistry*, vol. 172, p. 117612, 2024.

[4] M. Melinda, Y. Yunidar, Z. Noufal, A. B. Prasetyo, and M. Irhamsyah, "A novel subtraction method for signal fluctuation," in *2022 5th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, pp. 700–705, IEEE, 2022.

[5] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, p. 53, 2021.

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[7] S. K. Zhou, H. Greenspan, C. Davatzikos, J. S. Duncan, B. Van Ginneken, A. Madabhushi, J. L. Prince, D. Rueckert, and R. M. Summers, "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 820–838, 2021.

[8] C. Zhang, L. Zhou, F. Liu, J. Huang, and J. Peng, "Application of deep learning in laser-induced breakdown spectroscopy: a review," *Artificial Intelligence Review*, no. Suppl 2, pp. 2789–2823, 2023.

[9] N. Otsu *et al.*, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, 1979.

[10] M. Sezgin and B. l. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146–168, 2004.

[11] Z. Guo, J. Zhang, H. Dong, J. Sun, J. Huang, S. Li, C. Ma, Y. Guo, and X. Sun, "Spatio-temporal distribution patterns and quantitative detection of aflatoxin b1 and total aflatoxin in peanut kernels explored by short-wave infrared hyperspectral imaging," *Food Chemistry*, vol. 424, p. 136441, 2023.

[12] E. M. Bursalı, M. A. Özdemir, and M. Şen, "High-accuracy machine learning–based colorimetric ph quantification using a custom-built portable strip-imaging device and smartphone application," *Microchimica Acta*, vol. 192, no. 12, p. 850, 2025.

[13] Y. Yunidar, M. Melinda, M. Putri, M. Irhamsyah, N. Basir, and A. Khairah, "Performance analysis of h2o and h2o with hcl material image classification using inception v3, vgg19, densenet201, and otsu segmentation," *International Journal of Engineering, Science and Information Technology*, vol. 5, pp. 215–226, Sept. 2025.

[14] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, 2017.

[15] M. Tan and Q. Le, "Efficientnetv2: Smaller models and faster training," in *International Conference on Machine Learning*, pp. 10096–10106, PMLR, 2021.

[16] J. Junidar, M. Melinda, D. D. Diannuari, D. D. Acula, and Z. Zainal, "Face autistic classification based on thermal using image ensemble learning of vgg-19, resnet50v2, and efficientnet," *Radioelectronic and Computer Systems*, vol. 2025, no. 1, pp. 153–164, 2025.

[17] J. Li, Z. He, D. Li, and A. Zheng, "Research on water seepage detection technology of tunnel asphalt pavement based on deep learning and digital image processing," *Scientific Reports*, vol. 12, no. 1, p. 11519, 2022.

[18] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in vision: A survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 10s, pp. 1–41, 2022.

[19] B. Kim and S. Seo, "Efficientnetv2-based dynamic gesture recognition using transformed scalogram from triaxial acceleration signal," *Journal of Computational Design and Engineering*, vol. 10, no. 4, pp. 1694–1706, 2023.

[20] H. Zhao, X. Zhang, Y. Gao, L. Wang, L. Xiao, S. Liu, B. Huang, and Z. Li, "Diagnostic performance of efficientnetv2-s method for staging liver fibrosis based on multiparametric mri," *Heliyon*, vol. 10, no. 15, 2024.

[21] G. Ayana, K. Dese, A. M. Abagaro, K. C. Jeong, S.-D. Yoon, and S.-w. Choe, "Multistage transfer learning for medical images," *Artificial Intelligence Review*, vol. 57, no. 9, p. 232, 2024.

[22] M. Latha, P. S. Kumar, R. R. Chandrika, T. Mahesh, V. V. Kumar, and S. Guluwadi, "Revolutionizing breast ultrasound diagnostics with efficientnet-b7 and explainable ai," *BMC Medical Imaging*, vol. 24, no. 1, p. 230, 2024.