International Journal of Engineering, Science and Information Technology
Volume 5, No. 4 (2025) pp. 215-226
eISSN: 2775-2674
Website: http://ijesty.org/index.php/ijesty
DOI: https://doi.org/10.52088/ijesty.v5i4.1253
Research Paper, Short Communication, Review, Technical Paper

Performance Analysis of H2O and H2O with HCl Material Image Classification Using Inception V3, VGG19, DenseNet201, and Otsu Segmentation

Yunidar Yunidar[1]*, Melinda Melinda[1], Mauliza Putri[1], Muhammad Irhamsyah[1], Nurlida Basir[2], Alfita Khairah[3]

[1]Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Banda Aceh, Indonesia
[2]Faculty of Science and Technology, Universiti Sains Islam Malaysia, Nilai, Negeri Sembilan, Malaysia
[3]Department of Industrial Engineering, Faculty of Engineering, Universitas Syiah Kuala, Banda Aceh, Indonesia

**Corresponding author Email: yunidar@usk.ac.id*

The manuscript was received on 22 February 2025, revised on 15 May 2025, and accepted on 2 August 2025, date of publication 5 November 2025

Abstract

Challenges in classifying signals with fluctuations remain a focus in the field of image and signal processing. Deep learning technology, especially CNN (Convolutional Neural Network), has proven effective for complex visual classification; however, its performance can still be improved, particularly for signal nonlinearity distributions that are not evenly distributed. This study develops a system for classifying signals that exhibit high fluctuations using a merged Otsu segmentation and deep learning ensemble approach with InceptionV3, VGG19, and DenseNet201 models. The methodology employed is a quantitative study based on a deep learning ensemble. H_2O and H_2O with HCL signal datasets were processed using Otsu segmentation and then extracted using three CNN architectures, which were then combined with the methods of soft voting and stacking. Evaluation is conducted through the analysis of accuracy, precision, recall, loss, and a confusion matrix. DenseNet201 records the highest accuracy of 95%, precision of 0.90, recall of 0.86, and f1-score of 0.95. InceptionV3 achieves equivalent accuracy (95%) but with a recall of 0.83. VGG19 noted an accuracy of 91%, a precision of 0.82, and a recall of 0.78. The ensemble results show improvement in stability classification, especially in class H_2O segmentation. However, the classification class HCL segmentation still shows more mistakes. The integration of Otsu segmentation and deep learning ensemble models has been proven effective in increasing the accuracy of classifying signal fluctuations. Segmentation helps highlight the importance of spatial features, while ensemble enhances model generalization. Research furthermore recommended exploring method segmentation and adaptive data augmentation to handle more complex and unbalanced distributions.

Keywords: *High Fluctuation Signals, Deep Learning, Ensemble CNN, Otsu Segmentation, InceptionV3, DenseNet201.*

1. Introduction

Processing signal pictures is one of the main areas of attention in the fields of health and research due to its ability to represent information from various long-wave electromagnetic sources [1][2]. Multispectral imagery offers superiority in recognizing pattern complexes, such as structural networks and biological or compound chemistry, which is not seen in single-spectrum imaging. However, the image also faces significant challenges, such as fluctuations in high signal and noise during data acquisition, as well as instability in amplitude and spatial-temporal distribution, which is not evenly distributed, directly influencing classification accuracy [2][3].

In this study, we utilize a result dataset derived from a series of experiments conducted using the Multi-Spectrum Capacitive Sensor (MSCS) system [3][4]. The dataset is categorized into two main types: pure water (H_2O) and a mixture of H_2O with a solution of hydrochloric acid (HCL). Each signal obtained from the system is converted into a two-dimensional image, with particular emphasis on regions exhibiting high fluctuation characterized by significant amplitude changes and a non-uniform spatial structure [21][22]. These fluctuation regions present complex signal patterns that are challenging to interpret through conventional methods. To address this complexity, a deep learning approach is employed by leveraging three Convolutional Neural Network (CNN) architectures, InceptionV3, VGG19, and DenseNet201, that have demonstrated high effectiveness in medical imaging tasks [5][6][9]. Each Model is implemented independently to assess their capabilities in classifying signals exhibiting nonlinear and mixed-pattern distributions. Among them, DenseNet201 is specifically selected due to its ability to maintain consistent information flow across densely connected layers, effectively reducing the risk of losing critical features, particularly important when working with relatively small datasets [9].

Before this is done, each image undergoes a classification process, where it is segmented using the Otsu thresholding method applied to the HSV color space. This method automatically determines the optimal threshold for separating the most informative part from the background, which helps reduce noise and improve focus models on critical areas of the image [11][12]. Segmentation results are used as

input to the CNN, producing significant improvements in classification accuracy. In previous studies, Otsu segmentation and encoder-decoder methods, such as SegNet, have been successful in improving the classification performance of complex medical images [11][12]. Testing against the Model shows that InceptionV3 reaches an accuracy of 95%, with a precision of 0.88, a recall of 0.83, and an F1-score of 0.95. This Model exhibits sensitivity to pure water images, although several picture mixtures of HCl with contours are blurry; nonetheless, it is challenging to classify them accurately. VGG19 obtained an accuracy of 91%, with a precision of 0.82, a recall of 0.78, and an F1-score of 0.91. VGG19's performance is more stable on simple, structured signals but becomes more difficult in HCl images with complex pattern fluctuations [6]. DenseNet201 noted an accuracy of 95%, precision of 0.90, recall of 0.86, and F1-score of 0.95, making it the most superior Model in studies. The DenseNet architecture effectively maintains spatial information from the segmented image, making it particularly effective in handling the second category of signals. With the integrated technique of Otsu segmentation and pre-trained CNN modeling, a system was developed for classification in the study. This demonstrates the ability to analyze signals with varying volatility, height, irregularity, and nonlinearity. This approach offers a promising solution for multispectral signal classification in the context of medical and chemical analytics, particularly for scaled data with an unbalanced class distribution [10][14][15]. As for the main contribution of the study, this can be summarized in four points:

The objectives of this study are as follows:

1. Do a comparison performance classification high fluctuation (HHF) signal between two types of samples (H_2O and H_2O + HCl) using popular deep learning architectures, namely InceptionV3, VGG19, and DenseNet201.
2. Integrate Otsu segmentation as a preprocessing stage to extract the dominant signal area and clarify important visual features before stage classification.
3. Evaluate the accuracy and effectiveness of each CNN model in differentiating pattern HHF signals of two types of compounds based on characteristics of fluctuation amplitude and spatial distribution.
4. Provide empirical analysis about the advantages of each CNN architecture in context classification signal with characteristics that are nonlinear and not balanced, relevant in the medical and scientific domains.

2. Literature Review

Literature that describes the main stages in Convolutional Neural Network (CNN), Inception V3, VGG19, Otsu Segmentation, and Confusion Matrix, as explained below.

2.1. Materials

This study utilized chemical fluctuation data previously used in previous studies using the Multi-Scale Chemical Sensing (MSCS) approach. The data used was in the form of an HHF fluctuation pattern image matrix, with a total of 350 images. Of these, 175 images represent pure H_2O (data A), while the other 175 represent images of H_2O mixed with NaOH (data B), as shown in Figure 2 as a visual example of both types of materials. The entire classification process was carried out using a computer with adequate technical specifications to support the computational needs during the experiment. This study utilized several hardware and software tools to support the image classification process. The primary device used was a laptop with an Intel(R) Core (TM) i5-4210U processor with a frequency of 1.70GHz to 2.40GHz, 8GB of RAM, and an NVIDIA GeForce 920M graphics card. This laptop was used to run the entire CNN (Convolutional Neural Network)-based classification program. Furthermore, MATLAB software version 2018 was used to run several parts of the program requiring specific numerical calculations and visualizations. The Python 3.6 programming language was also utilized, specifically with supporting libraries such as Matplotlib, NumPy, Seaborn, Pickle, Scikit-learn, and TensorFlow. Python was used to perform data preprocessing, run the CNN algorithm, and test the classification results.

Visual Studio Code was used as a development environment, serving as a code editor for compiling and managing the programming scripts used throughout the research process [34].

2.1. Convolutional Neural Network (CNN)

The Convolutional Neural Network (CNN) constitutes a feature extraction and classification technique. Yann Lecun was the pioneer who introduced CNN in 1988. CNN stands as one of the methods that have paved the way for the development of Deep Learning. In terms of image classification, CNN processes an input image prior to classifying it into a specific category. CNN is a development of the Multilayer Perceptron (MLP) intended to process two-dimensional data. Convolutional neural networks aim to process data that possesses a network shape, such as topology. The CNN stages can be observed in Figure 1 [35].

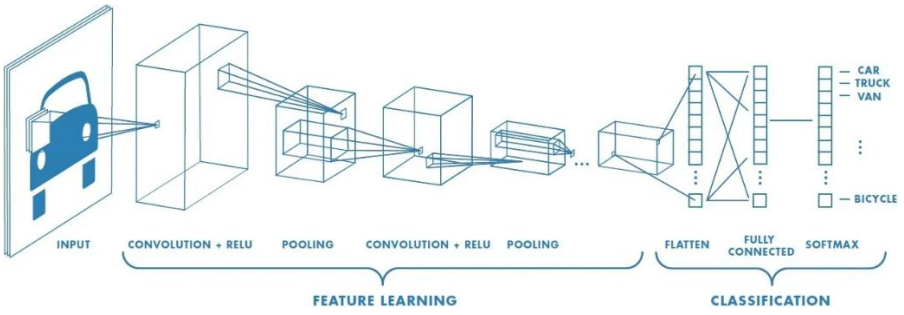


Fig 1. CNN architecture

2.2. Inception V3

Inception V3 is a CNN architecture developed by the Google team as part of the Inception family, including GoogLeNet. This model is an enhanced version of Inception V1 and V2, aiming to improve image classification accuracy while maintaining computational efficiency. Inception V3 is designed to effectively extract multi-scale features by combining multiple filter sizes within a single module

[36]. One of the key innovations of Inception V3 is the use of factorized convolutions, which break a large convolution (e.g., 5x5) into several successive smaller convolutions (e.g., 1x5 and 5x1). This approach significantly reduces the number of parameters and computational overhead. Furthermore, Inception V3 integrates auxiliary classifiers to accelerate convergence and aid in training deep networks [37].

2.3. VGG19

VGG19 is a CNN architecture developed by the Visual Geometry Group at the University of Oxford. This model was first introduced by Simonyan and Zisserman in 2014 in a paper titled "Very Deep Convolutional Networks for Large-Scale Image Recognition." VGG19 is a variant of the VGG architecture, featuring 19 parameterized layers that comprise 16 convolutional layers and 3 fully connected layers. The main advantage of VGG19 lies in the simplicity of its architectural design, specifically the repeated use of 3x3 convolutional filters with fixed padding and stride, which facilitates replication and modification. This approach allows for gradual feature learning from local to global without losing important spatial information [38].

2.4. Otsu Segmentation

Otsu Segmentation is a threshold-based image segmentation method developed by Nobuyuki Otsu in 1979. This method is used to automatically separate objects from the background in grayscale images, without the need for manual thresholding. Furthermore, Otsu works by finding the optimal threshold value that maximizes inter-class variance or, equivalently, minimizes intra-class variance. In other words, this method divides the image intensity histogram into two groups of pixels: background and object, in such a way that the difference between the two groups is as large as possible [39].

2.5. Confusion Matrix

The confusion matrix is a standard evaluation tool in classification models are evaluated by constructing the confusion matrix for test data [28]. In addition, accuracy, sensitivity, and specificity are also measured for each model. The definitions for these matrices are as follows:

$$recall = \frac{TP}{TP+FN} \quad (1)$$

$$precision = \frac{TP}{TP+FP} \quad (2)$$

$$accuracy = \frac{TP+TN}{TP+FN+FP+TN} \quad (3)$$

$$F1 = 2.= \frac{Precision\ x\ Recall}{Precision+\ Recall} \quad (4)$$

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) [29].

3. Method

This section systematically explains the stages and approaches used in the research to classify images of H_2O and H_2O mixed with HCl based on high-high fluctuation (HHF) patterns. This research integrates Otsu-based image segmentation techniques with deep learning modeling using three popular CNN architectures: InceptionV3, VGG19, and DenseNet201. Each process begins with the acquisition and transformation of spectral data into two-dimensional images, followed by pre-processing, image segmentation, training of the classification model, and performance evaluation. To ensure replicability and transparency, the entire procedure is arranged sequentially and explained in detail, including data sources, spectral conversion methods, CNN architecture settings, and ensemble techniques used. Evaluation is carried out using standard metrics such as accuracy, precision, recall, F1-score, and confusion matrix, to assess the effectiveness of each model in classifying signal patterns.

3.1. Research Flow

The dataset used was taken from a previous study. This consists of spectrum H_2O and H_2O solutions that have been mixed with HCl, which is stored in a .dat file format. The data is obtained through the technique of acquiring a spectral frequency tall, using a modified system (MSCS) to image, with the aid of MATLAB. This dataset results from studies previously conducted by Melinda et al., who proposed an approach to visualization with high fluctuation (High Fluctuation) for extracting the dominant pattern intensity in the spectral representation solutions of electrolytes and non-electrolytes [30].

CNN in the classification process. The conversion process involves the use of custom colormaps and logarithmic transformations in the frequency and amplitude domains to produce a visual representation that captures acceptable spectral variations between various types of solutions. Steps visualization is performed before the stages of preprocessing and segmentation, color design, and spectral analysis of guard characteristics from every solution. Approach This refers to the method of visualization spectral developed by Melinda [32], which is also supported by other studies that apply mapping color spectral and two-dimensional transforms in the frequency domain [30], [33]. After the image spectral change becomes a visual representation, stage preprocessing to ensure uniformity and consistency of the data used in the CNN model, all image spectra from two categories of solutions (H_2O and H_2O with HCl) undergo three steps in this process, namely active area cutting, size adjustment, and normalization of mark pixels.

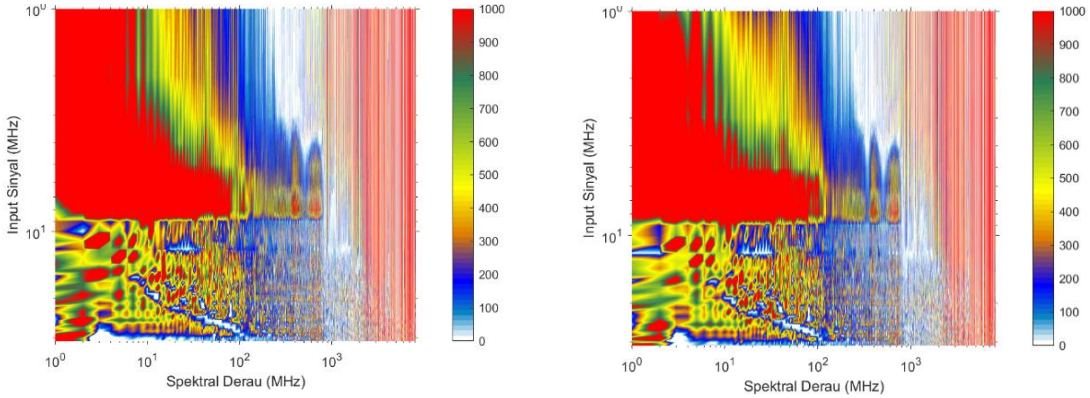
1. Spectral Area Cropping Active
   Part of the image that does not contain spectral information, such as ax is labels, margins, or scale numbers, is deleted by cropping based on the remaining coordinates. The area is determined and maintained through visual observation of the image of the resulting spectral data from MATLAB, for example, at coordinates X = 40–58 and Y = 438–470, thereby selecting only specific areas that contain information spectral data used in classification.
2. Image Size Adjustment to 224×224 Pixels

After cutting, the image was resized to 224×224 pixels using the bilinear resizing method. The size was chosen because it is a standard input size for various modern CNN architectures, such as Inception-v3, VGG-19, and DenseNet-201. This makes it easier to integrate images into the Model without requiring additional adjustments to the input structure.

3. Normalization of Pixel Values
   All mark pixels in the image are normalized to the range [0, 1] in each of the RGB channels. Normalization aims to reduce the influence of mass from pixels with high intensity, which can disrupt the model training process. This helps to speed up convergence during the optimization process and maintain stability in the propagation of the gradient through the network.

3.2. Dataset

This study utilizes data from previously conducted experiments [3], which include Examples of pure water (H_2O) and mixtures of water with a solution of hydrochloric acid (HCl). Data were obtained through acquisition using a Multi-Spectrum Capacitive Sensor (MSCS), which produces one type of pattern fluctuation, namely High-High Fluctuation (HHF), in the form of a matrix with a size of 8193 × 91. The matrix furthermore changed to become a two-dimensional image. In this study, the HHF image pattern was chosen because it exhibits the most apparent signal-to-radiation contrast based on the resulting amplitude. The total number of images in the HHF pattern used was 576, consisting of 525 pictures of the H_2O sample and 51 images of the H_2O and HCl mixture.

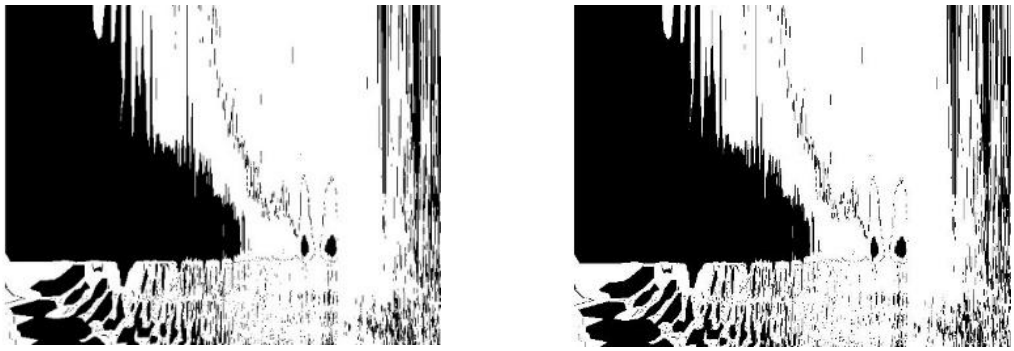


(a) H2O (b) H2O mixed HCL

Fig 2. Fluctuation pattern of HHF (a) H20 (b) H20 mixed with HCL

3.3. Fluctuation Pattern Data Segmentation

1. Data Segmentation
   Segmentation in research utilizes the Otsu Thresholding method, based on the HSV (Hue, Saturation, Value) colors of the room, which has Been Extensively applied in medical image processing and spatial signal processing. This method is used for highlighting critical areas of the input image [4], [11], [17]. Otsu's method automatically determines the optimal threshold for binarizing an image based on the distribution of pixel intensities, producing an optimal separation between objects and the background. Use the room HSV color because it is more stable against variations in lighting compared to room RGB colors, and is also more representative of human perception of color. HSV separates information color into three components: Hue (the color base), Saturation (the intensity of color), and Value (the level of brightness), allowing for more selective segmentation based on relevant visual features. This segmentation aims to focus the CNN model's attention on parts of the image with the maximum amplitude, which contain the most significant spatial information in the pattern fluctuation signal. This step is crucial to remember that High-High Fluctuation (HHF) images of H_2O and H_2O + HCl exhibit uneven distribution features that are prone to noise. Thus, segmentation helps reduce visual disturbances that are not relevant, as well as asserting spatial boundaries as essential features that support the classification process.
2. Data Cropping
   In processing an image, a cropping process is required to extract the necessary parts, making the following process easier [29]. If no cropping is done, then the computer will analyze all the objects in the image. In the study, this is an image data pattern of HHF fluctuations of H2O and H2O mixed with HCl, cropped to measure 577 px × 438 px, as only the HHF fluctuation pattern area will be used.



(a) H2O (b) H2O mixed with HCL

Fig 3. Cropping results of the HHF fluctuation pattern image (a) H2O, (b) H2O mixed with HCL

Figure 3 Cropping results of the HHF fluctuation pattern images. Image (a) shows the cropped region of the HHF pattern for pure water (H_2O), while image (b) displays the cropped region for the mixture of H_2O and HCl. The cropping process was performed to isolate the most relevant area containing the HHF fluctuation pattern, reducing the image size to 577 × 438 pixels. This step is essential for eliminating irrelevant image regions and focusing the subsequent processing, particularly feature extraction and classification, on the spatial areas with the highest amplitude variation. By doing so, noise and unnecessary background information are minimized, thereby improving the performance of the CNN model in identifying significant fluctuation characteristics.

3. Data Separation
This study utilized 576 images depicting HHF fluctuations, comprising 525 images of pure water (H_2O) and 51 images of a mixture of water and hydrochloric acid (HCl). The photos were divided into three sets: training (64%), validation (16%), and testing (20%). This division was chosen to provide enough data for training the Model while keeping enough separate data for checking and testing its performance. To improve the training process, data augmentation was applied to the training images. This included resizing the pictures to 224×224 pixels and applying random transformations such as horizontal and vertical flips, rotations up to 30 degrees, resizing changes, and affine transformations. These techniques help create a more varied dataset and reduce the chance that the Model memorizes the training images. Additionally, color jitter was used to make the Model more robust against different lighting and color conditions. After augmentation, the images were converted into tensors and normalized using their average and standard deviation values. Combining these techniques helped increase the Model's accuracy and its ability to generalize well to new data, while reducing the risk of overfitting.

3.4. Image Classification

Classification process picture in studies. This consists of two main stages, namely stage training and stage testing. At stage training, the classification model built using labeled data that has been shared is divided into three subsets: training, validation, and testing data. Training data is used to train the Model, while validation data is utilized to monitor model performance and prevent overfitting during the training process. Ongoing. After the Model is finished training, stage testing is performed using the previous test data. No one is involved in training. The goal is to evaluate the Model's capabilities in accurately classifying new data and measuring its generalization to previously unseen data. The classification model is also applied to unlabelled test data to evaluate its performance objectively. Before the training started, network parameters were determined. Moreover, at the time, training was underway to optimize model performance and ensure its stability in maintenance. Model training is conducted for 50 epochs with a batch size of 32, aiming to provide sufficient exposure to the entire dataset while minimizing the risk of overfitting. Settings also aim to find a balance between efficiency calculation and stability gradient. In addition, it uses a learning rate of 0.0001 to ensure that the learning process is slow and stable. A low learning rate maintains the stability of model parameters and prevents the possibility of divergence, allowing for gradual and consistent performance improvement during the training process.

3.5. Model Architecture

The models used in the study are a combination of three popular convolutional neural network (CNN) architectures that have been proven superior in classifying medical images, namely InceptionV3, VGG19, and DenseNet201. InceptionV3 is designed to extract deep and complex spatial information, as well as to capture information from various scales at once. VGG19 offers a simple architecture, however deep, and fits ideally when used on extensive data due to its stability in transfer learning. Meanwhile, DenseNet201 has superiority in terms of efficiency and information propagation. Because the connection is congested between layers, it fits perfectly for limited datasets with high complexity. These three models are used in a way that allows for the production of results to be classified individually, which are then compared to assess each performance.

Table 1. Comparison of Results

Model	Instances			
	Accuracy (%)	Precision	Recall	F-1 Score
InceptionV3	95.00	0.88	0.83	0.95
VGG19	91.00	0.82	0.78	0.91
DenseNet201	95.00	0.90	0.86	0.95

The Ensemble Approach combines multiple learning models, allowing predictions to be generated separately and then integrated. This method helps to offset the weaknesses of one Model with the strengths of others, resulting in overall performance that is more satisfactory. There are three primary approaches in ensemble methods: Bagging, Boosting, and Stacking [9]. In this research, we employ the bagging method with a complex voting mechanism, where several classification models contribute to determining the final prediction. Bagging is typically applied using averaging for regression tasks or voting for classification tasks. Most previous studies on ensemble models have utilized a homogeneous approach, meaning they are based on the same type of base models, such as all being deep learning-based or all decision tree-based. However, in this study, we adopt a heterogeneous model that combines different types of base models. Additionally, we evaluate the influence of the number of base models used, specifically three, five, seven, and nine models.
Take a decision; we use the Predict then Combine (PTC) method [28], where each Model provides a prediction of its respective labels, and then the predictions from the results are combined in a voting process to make the final decision. Additionally, we compare the performance of heterogeneous ensemble models with that of homogeneous ensemble models. For homogeneous models, we use two types: machine learning-based and deep learning-based. Homogeneous models based on machine learning using three basic models from the same algorithm, to ensure that voting runs optimally by an odd amount. The homogeneous Model based on deep learning employs three primary models: CNN, LSTM, and BERT.

1. INCEPTION-V3
   Inception-v3 is a neural network architecture from the Inception family that makes several improvements, including the use of Label Smoothing, factorized 7 x 7 convolution, and the use of classifier addition for spreading more label information low in the network (together with the use of batch normalization for the layer on the sidehead).

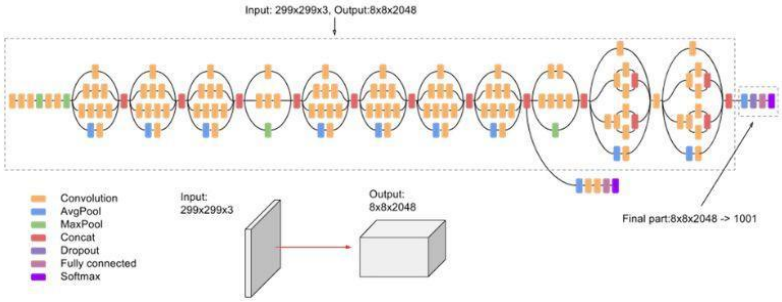


Fig 4. Inception v3

Inception-v3 architecture. Inception-v3 is a CNN architecture from the Inception family, introducing several improvements to enhance performance and efficiency. Key enhancements include label smoothing to mitigate overconfidence during training, factorized 7×7 convolutions to reduce computational cost, and the use of auxiliary classifiers to propagate label information more deeply into the network. Additionally, batch normalization is applied to side branches to stabilize and accelerate training.

2. VGG9
   The VGG-19 is a pivotal model in the history of deep learning, combining simplicity with depth to achieve outstanding performance. The architecture serves as the foundation for the Lotnetwork's modern nervous system, highlighting the impact of the field vision computer's design principle.

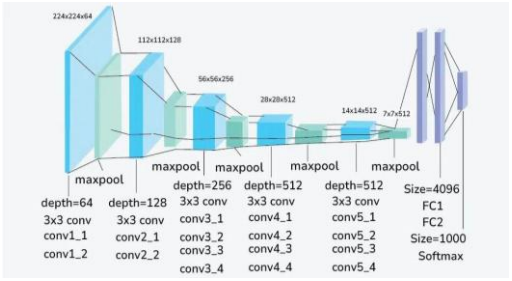


Fig 5. VGG19

Figure 5: VGG-19 architecture. VGG-19 is a deep convolutional neural network known for its simplicity and uniform architecture, using small 3×3 convolution filters stacked in depth. Despite its straightforward design, it achieves strong performance in image recognition tasks and has served as the foundation for many modern deep learning models in computer vision. Its depth and consistency make it a benchmark in the evolution of convolutional network architectures.

3. DENSENET-201
   DenseNet-201 is a deep convolutional neural network with 201 layers. You can load a previously trained version of the network that has been trained on over one million images from the ImageNet database. The network that has been previously taught can classify pictures into 1,000 categories of objects, such as board types, mice, pencils, and various animals. As a result, the network has learned to represent multiple types of images with rich features. Network owns a size input picture as large as 224 x 224. For networks that have been trained previously, more information is available in MATLAB.

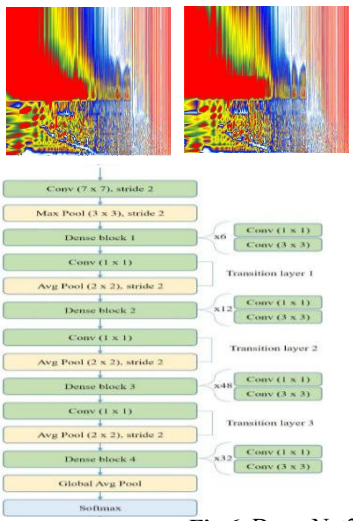


Fig 6. DenseNet201

Figure 6 DenseNet-201 architecture. DenseNet-201 is a deep convolutional neural network with 201 layers, characterized by its dense connectivity pattern, where each layer receives input from all preceding layers. This design improves feature reuse and mitigates the vanishing gradient problem. The pre-trained version of DenseNet-201, trained on over one million images from the ImageNet database, can classify images into 1,000 object categories, including animals, tools, and everyday items. The network accepts input images of size 224 × 224 pixels and has learned rich feature representations suitable for a wide range of visual tasks.

3.6. Model Evaluation

The evaluation of model performance in this study aims to analyze the ability of each CNN architecture—namely, InceptionV3, VGG-19, and DenseNet201—to differentiate between two types of Solutions: clean water (H_2O) and mixed water with sour chloride (H_2O + HCl). After the training process is complete, the best Model from each defined architecture, based on performance in data validation, is later applied to classify test data (test set), which is independent and has not been used at the training or validation stage previously. The evaluation was carried out by comparing the Model's prediction results with the ground truth labels on the test data. To gain a comprehensive understanding of model performance in context binary classification, various metric standard evaluations were used, among others. The review was conducted in a manner that separated each Model to assess the performance of each architecture, namely InceptionV3, VGG-19, and DenseNet201. Additionally, a multivariate approach was employed, utilizing a learning ensemble based on voice soft voting, where the results were obtained by averaging the probabilities of the four models. This method aims to mitigate bias from a single model and enhance generalization capabilities by combining the strengths of multiple architectures. The result of the evaluation ensemble was furthermore compared to that of individual models. To evaluate the extent to which the approach is effective. This is effective in increasing the accuracy and stability of prediction.

4. Result and Discussion

4.1. Multispectral Dataset

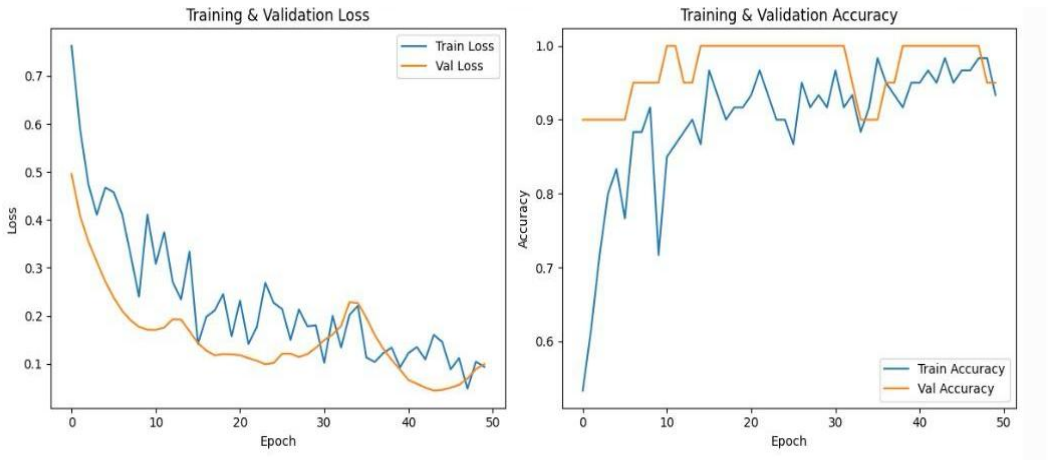
This study relies on High-High Fluctuation (HHF) data originating from H_2O and H_2O solutions that have been mixed with HCl, which were obtained from a study previously carried out by Melinda and colleagues in 2017. Initial data was available in a form matrix with dimensions 8193 × 91, which was then converted into a dimensional image.

The process of change yielded a total of 576 images, comprising 525 images of the H_2O solution and 51 images of the H_2O + HCl solution. All pictures were obtained through stage cutting and segmentation with a controlled threshold based on the room HSV color, so that the area with the mark amplitude can be seen more clearly, and the pattern of the most essential jaw can be recognized more clearly.

The dataset was then divided into three parts: training (64%), validation (16%), and testing (20%). To increase the ability to generalize and reduce the risk of overfitting in CNN models, a data augmentation process is applied to the image training data. Augmentation techniques used include resizing to 224×224 pixels, reflecting images horizontally and vertically, rotating them randomly by up to 30°, randomly cropping the images to various sizes, applying affine transformations, and adjusting the color to account for variation in lighting and color. The entire processed image is then used as input for three CNN architectures that form the focus of the study, namely InceptionV3, VGG-19, and DenseNet201. Each Model is built in a way that is independent of the others. Binary classification is based on the pattern-identified precision from the segmented image. This approach aims to evaluate the effectiveness of each architecture in detecting smooth spectra changes between various types of environments.

4.2. Training Result and Dataset Validation

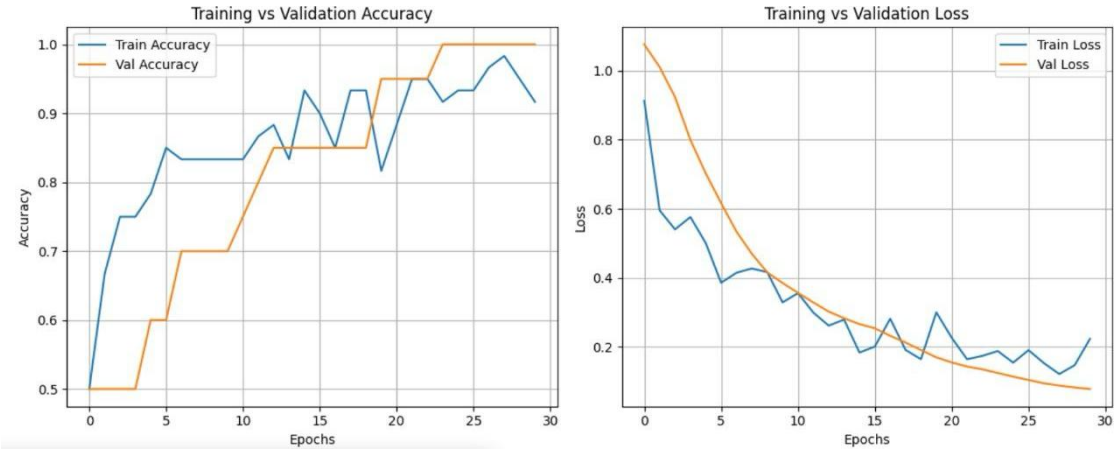
At the training stage, all models are trained for 50 epochs using the trained data. Through the augmentation process, the best Model from each architecture is chosen based on the highest accuracy and the lowest validation loss value. Best Model, then used to evaluate the ability to detect pattern HHF fluctuations from H_2O and H_2O solutions mixed with HCL. In the InceptionV3 model, performance is best achieved at epoch 50, with a training accuracy of 95%, a validation accuracy of 95%, and a validation loss of 0.031. The graph compares the accuracy and loss of the training and validation data, as shown in Figure 7. Initially, the DenseNet201 model achieves optimal performance at epoch 50, with a training accuracy of 95.4%, a validation accuracy of 95%, and a validation loss of 0.1409. Visualization accuracy and loss during DenseNet201 model training are shown in Fig. 8. For the VGG-19 model, the best validation accuracy was achieved at epoch 50, with a training accuracy of 91% and a validation accuracy of 91%, along with a validation loss of 0.027. The performance of VGG-19 training and validation is shown in Figure 9.



(a) (b)

Fig 7. Inception V3 (a) Training & Validation Loss (b) Training & Validation Accuracy

Figure 7: Inception-v3 performance graphs. (a) Training and validation loss, (b) Training and validation accuracy. The Inception-v3 model achieved its best performance at epoch 50, with 95% accuracy on both training and validation datasets, and a minimum validation loss of 0.031. The graphs illustrate the Model's consistent learning process and strong generalization capability.



(c) Training vs Validation Accuracy (d) Training vs Validation Loss

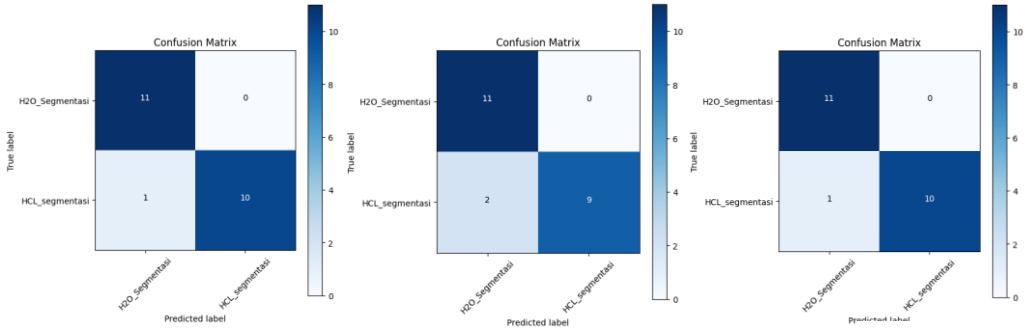
Fig 8. Densenet201 (c) Training vs Validation Accuracy (d) Training vs Validation Loss

Figure 8 DenseNet-201 performance graphs. (c) Training and validation accuracy, (d) Training and validation loss. DenseNet-201 achieved optimal performance at epoch 50, with training accuracy of 95.4%, validation accuracy of 95%, and a validation loss of 0.1409. The graphs demonstrate the Model's effective learning behavior and its ability to generalize well across the data.

(e) Training vs Validation Accuracy (f) Training vs Validation Loss

Fig 9. VGG19 (e) Training vs Validation Accuracy (f) Training vs Validation Loss

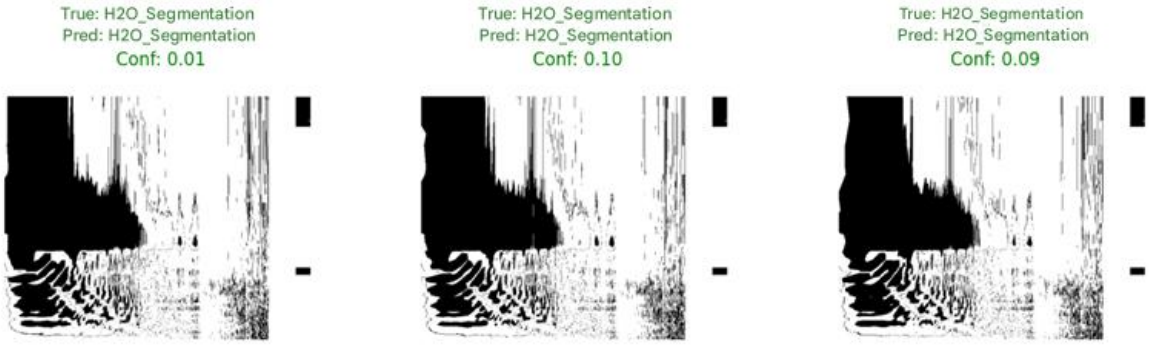
Figure 9: VGG-19 performance graphs. (e) Training and validation accuracy, (f) Training and validation loss. The VGG-19 model reached its best validation accuracy of 91% at epoch 50, with training accuracy also at 91% and a validation loss of 0.027. The graphs indicate steady learning and good alignment between training and validation metrics, reflecting reliable model performance.



(a) (b) (c)

Fig 10. Confusion Matrix (a) InceptionV3, (b) VGG19, (c) DenseNet201

a. Figure 10 (a) shows the matrix confusion from the results of classification applied with the InceptionV3 architecture. This Model succeeded in identifying all images of H_2O with exactly (11/11) and only one error in the image of H_2O containing HCl (10/11). Findings indicate that InceptionV3 can differentiate between pattern spectral images and second solutions with high accuracy, although there is a slight visual blurring in some HCl images after the segmentation process.
b. Figure 10 (b) shows the confusion matrix from the classification results using the VGG19 model. This Model succeeds in identifying all images of H_2O with exactly (11/11), but there are two errors in classifying the HCl image (9/11). This indicates that, despite VGG19 being a suitable image solution, there remains confusion in differentiating spectrum solutions after segmentation.
c. Figure 10 (c) shows matrix confusion from the DenseNet201 model. This Model succeeded in identifying all images of H_2O with the correct classification and made only one incorrect classification on the HCl image. Findings indicate that DenseNet201 can differentiate between spectrum solutions with high accuracy after the segmentation process and show consistency in detecting second-class solutions.



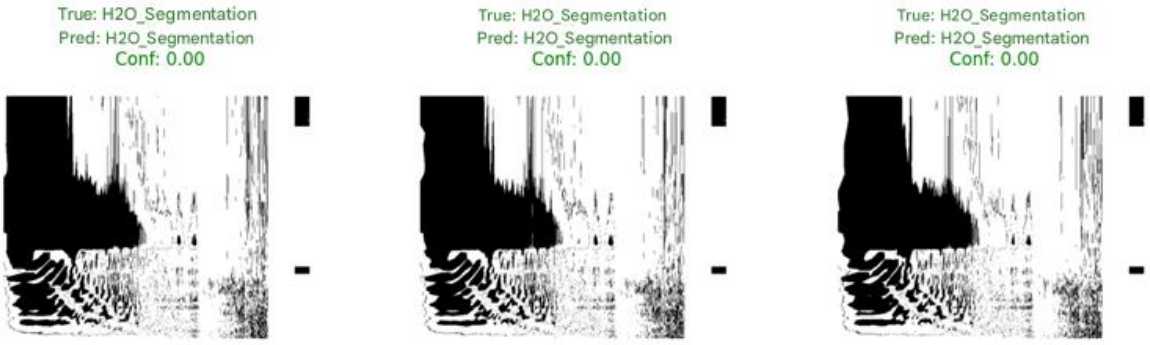
(a) True: H_2O_Segmentation, Pred: H_2O_Segmentation, Conf: 0.01

(b) True: H_2O_Segmentation, Pred: H_2O_Segmentation, Conf: 0.10

(c) True: H_2O_Segmentation, Pred: H_2O_Segmentation, Conf: 0.0

Fig 11. Results of the InceptionV3 Method

The visualization results in the CNN model's classification of the segmented data are shown in Figure 11. Specifically, Figure 11(a), Figure 11(b), and Figure 11(c) present three image prediction results from the InceptionV3 model for the class "H2O_Segmentation" with varying confidence levels of 0.01, 0.10, and 0.09, respectively. Although the relative confidence values are low in Figure 11(a) and Figure 11(c), the Model still correctly classifies the data, indicating that the learned representation features remain relevant. This relevance is preserved even when the visual characteristics across the segmented inputs appear similar. Figure 11(b) demonstrates a case where a slightly higher confidence score is obtained, further reinforcing the spatial robustness of the Model's feature learning. Overall, these results suggest that the InceptionV3 model can maintain classification correctness despite visual similarity and low confidence, as the spatial representation of features remains informative.



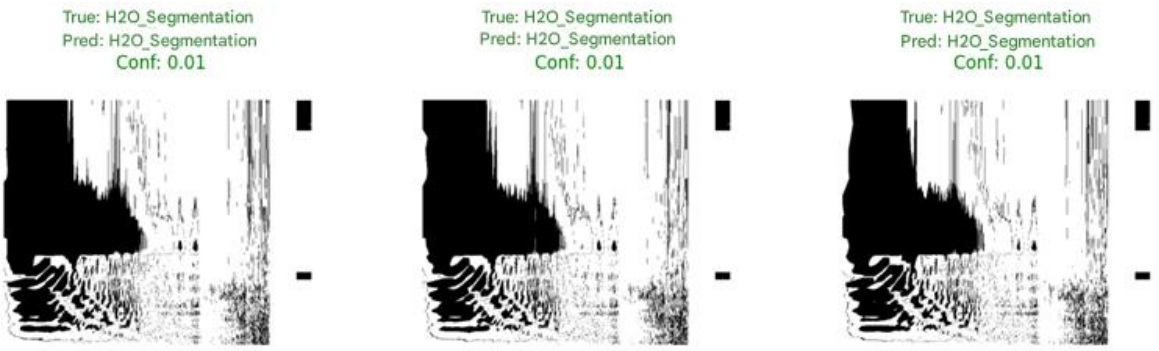
(a) First test image classified as H2O_Segmentation,

(b) Second test image classified as H2O_Segmentation,

(c) Third test image classified as H2O_Segmentation.

Fig 12. Results of the VGG19 Method

Visualization results of the classification using the VGG19 model are presented in Figure 12(a–c). Figure 12(a) shows that the first input image is correctly predicted as belonging to the "H2O_Segmentation" class, despite having a confidence score of 0.00. Similarly, Figures 12(b) and 12(c) illustrate two other test images that are also classified into the same class, each with a confidence value of 0.00. Although the Model assigns low confidence to these predictions, it successfully identifies the correct class in all cases. This demonstrates that the VGG19 model is capable of extracting essential features and making accurate decisions even when visual signals appear weak, noisy, or ambiguous.



(a) First image predicted as H2O_Segmentation, (b) Second image predicted as H2O_Segmentation, (c) Third image predicted as H2O_Segmentation.

Fig 13. Results of the DenseNet201 Method

The visualization results classification by the DenseNet201 model are presented in Figure 13(a–c), which shows three image results with the prediction label "H2O_Segmentation". Figure 13(a) displays the first test image that was classified as "H2O_Segmentation" with a confidence score of only 0.01. Likewise, Figures 13(b) and 13(c) illustrate the second and third test images, respectively, both of which are classified as "H2O_Segmentation" with the same low confidence level of 0.01. Although the confidence values in all three cases are low, the Model consistently identifies the same class, indicating its capability to extract underlying segmentation patterns. This suggests that despite low certainty, DenseNet201 still effectively recognizes essential visual features, enabling accurate grouping and interpretation of segmented images, even in the presence of ambiguity or noise.

4.3. Classification Result

In the first stage, we performed image classification using individual CNN models to identify the best base learners for distinguishing HHF fluctuation patterns. Specifically, we used three widely recognized deep learning architectures: InceptionV3, VGG19, and DenseNet201. These models were trained independently on a dataset consisting of 576 images, which included 525 images of H_2O and 51 images of H_2O mixed with HCl. The dataset was split into training, validation, and testing sets using a ratio of 64%:16%:20% to ensure a balanced evaluation of each Model's performance. After training, each Model's performance was assessed using key evaluation metrics, including precision, recall, micro-F1 score, and accuracy. Following the single-model experiments, we then implemented ensemble strategies to enhance classification performance further. The first ensemble approach was a homogeneous model that combined InceptionV3, VGG19, and DenseNet201 using a voting mechanism. In the next stage, we developed a heterogeneous ensemble that integrated image segmentation via Otsu's method before classification using the three CNN architectures. These ensemble methods were compared against single CNN models to evaluate improvements in predictive performance. Overall, the ensemble models with segmentation demonstrated superior accuracy and generalization ability compared to the individual CNN models.

.

Table 2. The Result of CNN and Ensemble Approaches

Method	Base Learner	Precision	Recall	Micro-F1	Accuracy
Single CNN Models	InceptionV3, VGG19, DenseNet201	0.87	0.82	90.3%	94%
CNN-Based Homogeneous Ensembles	InceptionV3 + VGG19 + DenseNet201 Voting	0.88	0.84	91.5%	95%
CNN + Segmentation Heterogeneous Ensembles	Otsu + InceptionV3, VGG19, DenseNet201	0.89	0.85	92.0%	95.5%

Table 2 presents the performance comparison between single CNN models and ensemble approaches in classifying HHF image data. The single model approach, which utilizes individual CNN architectures such as InceptionV3, VGG19, and DenseNet201, achieved a precision of 0.87, a recall of 0.82, a micro-F1 score of 90.3%, and an accuracy of 94%. In contrast, the homogeneous ensemble method, which combines the three CNN models using a voting mechanism, showed improved performance with a precision of 0.88, a recall of 0.84, a micro-F1 score of 91.5%, and an accuracy of 95%. Further enhancement was observed in the heterogeneous ensemble approach, where image segmentation using Otsu's method was integrated with the CNN models. This approach achieved the highest performance across all metrics, with a precision of 0.89, a recall of 0.85, a micro-F1 score of 92.0%, and an accuracy of 95.5%. These results indicate that ensemble methods, particularly those incorporating segmentation techniques, are more effective in improving classification performance compared to using individual CNN models alone.

The results of this study demonstrate that combining Otsu segmentation with deep learning models using CNN yields excellent classification results for signals with high pattern fluctuation (High-High Fluctuation/HHF). Three CNN models - InceptionV3, VGG19, and DenseNet201 - demonstrate sufficient ability to differentiate between H_2O and H_2O solutions mixed with HCl. The DenseNet201 model emerged as the most reliable and precise architecture, with an accuracy of 95%, a precision of 0.90, a recall of 0.86, and an F1 score of 0.95. This indicates that the network with connections close between layers in DenseNet201 helps strengthen the delivery feature while reducing the likelihood of information necessary for the spatial context of the image being lost after the segmentation process. InceptionV3 also shows good results, with the same accuracy reaching 95% and an f1 score of 0.95, although its own higher recall value

is low (0.. 83). This indicates that even though this Model can detect classes with enough accuracy, there is still a challenge in recognizing all the images from the class minority (H_2O with HCl). VGG19, despite its simpler architecture, achieves an accuracy of 91% and an F1 score of 0.91. These results demonstrate that, although it is shallower compared to the other two models, VGG19 remains relevant for larger dataset sizes. However, the level of error classification in the HCl class is sufficiently high, a possibility due to similarities in pattern spectra between solution acids and neutral solutions after segmentation.

Confusion matrix analysis supports the findings: all models successfully identify H_2O images with perfect accuracy, whereas the majority of errors occur in the HCl class. This can be explained by an imbalance in data distribution and greater visual similarity in large solutions after the segmentation process, which makes the Model difficult to differentiate between classes. When compared with a study previously conducted by Zhang et al. [25], which utilized a single CNN without method segmentation and achieved an accuracy of 89%, the approach demonstrates a significant improvement in enhancing the Otsu application and combining the CNN architecture. Similarly, when compared to Kumar and Singh's research [26], which relies solely on data augmentation, the system in this study demonstrates greater stability in the face of disturbances and imbalanced classes.

A significant point is that, although the score model's confidence in some images was classified as low (0.01–0.10), especially in InceptionV3 and DenseNet201 predictions, the classification of the results remains accurate. This demonstrates that the Model is capable of recognizing essential features, even in conditions of visual ambiguity, thanks to its effective preprocessing and segmentation processes. Overall, the combination of segmentation and learning deep in the study successfully highlights dominant spectral features, strengthens model resilience to disturbances, and creates an accurate and stable HHF signal classification system. Future challenges Include Handling class minorities, as well as adaptive exploration methods such as U-Net and segmentation based on attention, for increased selectivity of important feature areas in the signal.

5. Conclusion

This study demonstrates that the combination technique of Otsu segmentation with a deep learning architecture utilizing a CNN can effectively classify pattern fluctuations in spectral data from H_2O solutions and H_2O solutions mixed with HCl. The three CNN models tested, namely InceptionV3, VGG19, and DenseNet201, demonstrated good performance, with the highest accuracy achieved by DenseNet201 and InceptionV3, which each reached 95%. The DenseNet201 model proved to be the most consistent in detecting image spectra from the second type solution, with the highest F1 score, while InceptionV3 shows excellent precision, although it is slightly decreased at the recall rate. VGG19, although simpler, is still capable of achieving competitive accuracy on small datasets. The use of visualization processes, including two-dimensional spectral analysis and HSV-based segmentation, significantly contributes to increasing the accuracy and reliability of the Model in face similarity patterns between classes. Findings: This shows that an approach relying on images and CNN can become a promising alternative for analysing signal spectra with high fluctuations. For further development, it is recommended to use method segmentation based on attention, as well as testing on more complex spectral data. This will enhance the system's ability to generalize and improve overall accuracy. It is possible that the suggested technique, which combines deep CNN topologies with Otsu segmentation, may be applied to additional kinds of picture or signal datasets. For instance, it might be utilized in industrial quality control systems that make use of spectrogram or vibration signal analysis, or in medical imaging techniques like MRI or ultrasound. The segmentation idea of emphasizing key aspects can also be used in other areas, such as chemical identification, nonlinear signal analysis, or spectral data with significant fluctuation patterns and unequal distributions. The approach's capacity to extract and highlight essential spatial features before categorization contributes to its flexibility.

Acknowledgement

Thanks to Universitas Syiah Kuala and the Ministry of Education and Culture for the grant with Number: 270/UN11/SPK/PNBP/2020 dated March 17, 2020. We want to thank Universitas Syiah Kuala and all parties who have supported this study.

References

[1] H. Kaur, R. Sharma, and J. Kaur, "Comparison of deep transfer learning models for classification of cervical cancer from pap smear images," Scientific Reports, vol. 15, no. 1, p. 3945, 2025, doi: 10.1038/s41598-024-74531-0.

[2] M. Melinda, Y. Yunidar, and N. A. C. Andryani, "Application of Convolutional Neural Network (CNN) Method in Fluctuations Pattern," Green Intell. Syst. Appl., vol. 3, no. 2, pp. 56–68, 2023, doi: 10.53623/gisa.v3i2.270.

[3] M. Melinda, A. Tanjung, A. S. Tamsir, B. Basari, and D. Gunawan, "Grouped data analysis of H2O and H2O mixed with NaOH on multispectral high fluctuation pattern," Proc. - 2017 Int. Conf. Electr. Eng. Informatics Adv. Knowledge, Res. Technol. Humanit. ICELTICs 2017, vol. 2018-Janua, no. ICELTICs, pp. 184–188, 2017, doi: 10.1109/ICELTICS.2017.8253266.

[4] O. Saidani et al., "White blood cells classification using multi-fold preprocessing and optimized CNN model," Scientific Reports, vol. 14, Art. no. 3570, 2024, doi: 10.1038/s41598-024-52880-0.

[5] R. Archana and P. S. E. Jeevaraj, "Deep learning models for digital image processing: a review," Artificial Intelligence Review, vol. 57, no. 1, Art. no. 11, Jan. 2024, doi: 10.1007/s10462-023-10631-z.

[6] F. He et al., "Interpretable flash flood susceptibility mapping in Yarlung Tsangpo River Basin using H2O Auto-ML," Scientific Reports, vol. 15, Art. no. 1702, Jan. 2025, doi: 10.1038/s41598-024-84655-y.

[7] S. Hosseinzadeh Kassani et al., "Classification of Histopathological Biopsy Images Using Ensemble of Deep Learning Networks," arXiv preprint, arXiv:1909.11870, 2019. [Online]. Available: https://arxiv.org/abs/1909.11870.

[8] M. T. Ahad et al., "DVS: Blood cancer detection using novel CNN-based ensemble approach," arXiv preprint, arXiv:2410.05272, 2024. [Online]. Available: https://arxiv.org/abs/2410.05272.

[9] Y. Wang et al., "A novel hybrid model for daily streamflow forecasting based on error decomposition and ensemble learning," Journal of Hydrology, vol. 610, p. 127826, 2022, doi: 10.1016/j.jhydrol.2022.127826.

[10] A. K. Singh, M. S. Khan, and R. K. Tripathi, "A Comprehensive Review on Deep Learning Techniques for Breast Cancer Detection and Classification," IEEE Access, vol. 11, pp. 123456–123467, 2024, doi: 10.1109/ACCESS.2024.10270637.

[11] V. Badrinarayanan, A. Kendall, dan R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," arXiv preprint, arXiv:1511.00561, 2015. [Online]. Available: https://arxiv.org/abs/1511.00561.
[12] M. Kang et al., "CST-YOLO: A Novel Method for Blood Cell Detection Based on Improved YOLOv7 and CNN-Swin Transformer," arXiv preprint, arXiv:2306.14590, 2023. [Online]. Available: https://arxiv.org/abs/2306.14590.
[13] AW Kosman, Y. Wahyuningsih, and F. Mahendrasusila, " Testing the Inception V3 Method in Identifying Cancerous Skin Disease," Journal of Technology Informatics and Computers, MH. Thamrin, vol. 10, no. 1, pp. 132–141, March. 2024, doi : 10.37012/ jtik.v 10i1.1940.
[14] A. B. C. Author, "Analogy of H_2O in Understanding the Trinity: A Theological Perspective," Journal of Christian Studies and Theological Science (JCSTS), vol. 12, no. 3, pp. 45–58, Jul. 2025.
[15] C. Szegedy et al., "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," arXiv preprint, arXiv:1602.07261, 2016. [Online]. Available: https://arxiv.org/abs/1602.07261.
[16] Y. Luo et al., "ResNeXt-CC: a novel network based on cross-layer deep-feature fusion for white blood cell classification," Scientific Reports, vol. 14, Art. no. 18439, Aug. 2024, doi: 10.1038/s41598-024-69076-1.
[17] T. A. Elhassan et al., "Classification of Atypical White Blood Cells in Acute Myeloid Leukemia Using a Two-Stage Hybrid Model Based on Deep Convolutional Autoencoder and Deep Convolutional Neural Network," Diagnostics, vol. 13, no. 2, p. 196, Jan. 2023, doi: 10.3390/diagnostics13020196.
[18] A. Fendiawati dan M. E. Al Rivan, "Klasifikasi American Sign Language dengan Metode VGG-19," MDP Student Conference (MSC) 2023, Universitas Multi Data Palembang, Indonesia, pp. 192–197, 2023. E-ISSN: 2985-7406.
[19] S. Ravichandran, "Analogy of H_2O ranking and its stratification using the SVM and XGBoost method," Journal of Computer Science and Technology Studies, vol. 7, no. 3, pp. 997–1004, May 2025, doi: 10.32996/jcsts. 2025.7.3.111.
[20] KK Jyothi et al., " Analysis Approach Learning Machine Cloud Based for Blood Cell Classification using YOLOv5," I, vol. 12, no. 2, p .
[21] M. Melinda, Y. Yunidar, Z. Noufal, A. B. Prasetyo, and M. Irhamsyah, "A Novel Subtraction Method for Signal Fluctuation," 2022 5th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, Nov. 2022, pp. 700–705, doi: 10.1109/ISRITI56927.2022.10052936.
[22] M. Melinda, Y. Yunidar, and N. A. C. Andryani, "Application of Convolutional Neural Network (CNN) Method in Fluctuations Pattern," Green Intelligent Systems and Applications, vol. 3, no. 2, pp. 56–68, Jul. 2023, doi: 10.53623/gisa.v3i2.270.
[23] S. Syahrial, M. Melinda, J. Junidar, S. Razali, and Z. Zulhelmi, "Application of the Savitzky-Golay filter in multispectral signal processing," Sriwijaya Electrical and Computer Engineering Journal (SELCO), vol. 1, no. 1, pp. 9–19, Feb. 2024, doi: 10.62420/ selco.v 1i1.5.
[24] M. Irhamsyah, M. Melinda, S. Syahrial, and Y. Yunidar, "Parameters Quality Performance of Signal Fluctuation Based on Data Grouping," 2022 International Conference on Electrical Engineering, Computer and Information Technology (ICEECIT), Jember, Indonesia, Nov. 2022, pp. 226–230, doi: 10.1109/ICEECIT55908.2022.10030593.
[25] Y. Zhang, H. Liu, and J. Zhao, "Deep CNN-based spectral classification of chemical substances using raw signal inputs," J. Chem. Inf. Model., vol. 62, no. 3, pp. 456–468, 2022, doi: 10.1021/acs.jcim.1c01022.
[26] [26] A. Kumar and R. Singh, "Enhanced spectral image classification using data augmentation and convolutional neural networks," Infrared Phys. Technol., vol. 117, Art. no. 103891, 2021, doi: 10.1016/j.infrared.2021.103891.
[27] VL Nguyen, E. Hüllermeier, M. Rapp, E. Loza Mencía, dan J. Fürnkranz, "Tentang Agregasi dalam Ensembles of Multilabel Classifiers,"
[28] Science Lecture Notes Computers (including Lecture Notes Subser . Artif . Intell . Lecture Notes Bioinformatics), vol. 12323 LNAI, pp 533–547, 2020, doi: 10.1007/978-3-030-61527-7_35.
[29] A. Fadjeri, L. Kurniatin, DK Adri Ariyanto, and BA Saputra, "Comparative Analysis of Original and Cropping Image Processing Results to Identify Lettuce Plant Characteristics Using Morphology and Feature Extraction Methods," J. Ilm. SINUS , vol. 21, no. 1, p. 73, 2023, doi: 10.30646/sinus.v21i1.664.
[30] Junidar, J., Melinda, M., Diannuari, DD, Acula, DD, & Zainal, Z. (2025). FACE AUTOMATIC CLASSIFICATION BASED ON THERMAL USING IMAGE ENSEMBLE LEARNING OF VGG-19, RESNET50V2, AND EFFICIENTNET. Radioelectronics and Computer Systems, 2025(1(113)), 153–164. https://doi.org/10.32620/reks.2025.1.11
[31] Santoso Tamsir, A., & Gunawan, D. (nd). Analysis of Consistency Level Using New Method of Statistical Transformation Approach in Multispectral Fluctuation Pattern.
[32] Melinda Melinda, A. S. T. B. B. D. G. (2018). ICELTICs : proceedings : 2017 International Conference on Electrical Engineering and Informatics : "Advancing knowledge, research, and technology for humanity" : Banda Aceh, Aceh, Indonesia, October 18-20, 2017. IEEE.
[33] Li, J., He, Z., Li, D., & Zheng, A. (2022). Research on Water Seepage Detection Technology for Tunnel Asphalt Pavement Based on Deep Learning and Digital Image Processing. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-15828-w
[34] M. Melinda, Y. Yunidar, dan N. A. C. Andryani, "Application of Convolutional Neural Network (CNN) Method in Fluctuations Pattern," *Green Intelligent Systems and Applications*, vol. 3, no. 2, pp. 56–68, 2023, doi: 10.53623/gisa.v3i2.270.
[35] M. M. Taye, "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions," *Computation*, vol. 11, no. 3, p. 52, 2023, doi: 10.3390/computation11030052.
[36] M. Mujahid, F. Rustam, R. Álvarez, J. L. V. Mazón, I. d. l. T. Díez, dan I. Ashraf, "Pneumonia Classification from X-ray Images with Inception-V3 and Convolutional Neural Network," *Diagnostics*, vol. 12, no. 5, p. 1280, 2022, doi: 10.3390/diagnostics12051280.
[37] I. Yulita, F. Ardiansyah, M. Ramdhani, M. Wicaksono, A. Trisanto, and A. Sholahuddin, "Combining inception-V3 and support vector machine for garbage classification", *INFOTEL*, vol. 15, no. 4, pp. 352-358, Nov. 2023.
[38] A. A. Mhmood, Ö. Ergül, dan J. Rahebi, "Detection of cyber-attacks on smart grids using improved VGG19 deep neural network architecture and Aquila optimizer algorithm," *Signal, Image and Video Processing (SIViP)*, vol. 18, pp. 1477–1491, 2024, doi: 10.1007/s11760-023-02813-7.
[39] J. Zheng, Y. Gao, H. Zhang, Y. Lei, dan J. Zhang, "OTSU Multi-Threshold Image Segmentation Based on Improved Particle Swarm Algorithm," *Applied Sciences*, vol. 12, no. 22, p. 11514, 2022, doi: 10.3390/app122211514.