Project Report

Fuzzy C-means Clustering and CNN-based Multiclass Skin Lesion Classification Technique Using Dermoscopic Images

ECE 624 A1: Fuzzy Sets in Human-Centric Systems

Term: Fall 2024

Submitted by **Md Touhidul Haque**

Student ID: 1862437

Contents

1 Introduction **2**

2 Dataset **3**

3 Methodology **3**

3.1 Image Preprocessing 3

3.2 FCM-based Image Segmentation 4

3.3 CNN-based Image Classification 5

3.4 Evaluation Metrics 6

4 Experimental Result and Analysis **7**

4.1 Segmentation Result 7

4.2 Classification Result 9

4.2.1 Proposed FCM-CNN Model 9

4.2.2 CNN Model 11

5 Discussion **11**

6 Conclusion **13**

7 Dataset and Source Code Repository **13**

1 Introduction

Skin diseases, including various types of skin cancer, pose a significant public health challenge, affecting millions globally and impacting both physical health and quality of life. Among these, melanoma is particularly fatal if not detected and treated early. According to the World Health Organization, skin cancer accounts for one-third of all diagnosed cancers worldwide [1], [2]. In the United States alone, approximately 5.4 million new skin cancer cases are reported annually, with melanomas contributing to over 10,000 deaths each year [3].

Traditional diagnostic methods, such as physical examination and biopsy, are time-consuming, invasive, and often subject to variability in interpretation [4]. While dermoscopy offers a non-invasive method to enhance diagnostic accuracy through high-resolution visualization of deeper skin structures [5], its effectiveness depends heavily on the expertise of the dermatologist. Studies show that diagnostic accuracy ranges from 62% to 80%, with lower accuracy observed among less experienced dermatologists [6], [7]. Moreover, the requirement for extensive training and limited exposure to the full diversity of skin cancer cases further complicates accurate diagnosis.

Recent advancements in artificial intelligence (AI) and machine learning (ML) have shown significant promise in addressing these challenges by providing automated, scalable, and reliable diagnostic tools. Deep convolutional neural networks (CNNs), in particular, have demonstrated exceptional performance in image classification tasks, including skin cancer classification, by detecting complex patterns and learning domain-specific features from dermoscopic images [8]–[11]. A key advantage of CNNs is their ability to automatically learn "deep features" from the data, eliminating the need for manual "feature engineering" by machine learning experts [12]. However, while CNNs excel at complex pattern recognition, their performance often depends on the quality of the input images. Raw dermoscopic images can include irrelevant background information, hair artifacts, and lighting inconsistencies, all of which can hinder the extraction of meaningful features for lesion classification. To address these challenges, image segmentation techniques can be integrated into the classification framework. By isolating the lesion region and removing extraneous background information, segmentation enables the CNN to focus on diagnostically relevant features. In particular, Fuzzy C-means (FCM) clustering has emerged as a promising segmentation method [13]. Unlike hard clustering approaches, FCM acknowledges gradual transitions at lesion boundaries by assigning membership values rather than definitive cluster labels. In [14], authors demonstrate that FCM-based segmentation can improve melanoma region identification, enhancing both the clarity of lesion boundaries and the accuracy of subsequent classification.

This project aims to develop and evaluate a multi-class skin lesion classification framework that integrates FCM clustering-based image segmentation with CNN-based classification, utilizing the HAM10000 dataset [15]. In this approach, FCM is employed to isolate lesion regions prior to classification, thereby potentially enhancing the CNN's ability to distinguish between different lesion types. The primary objectives of this project are as follows:

- Design, train, and evaluate an integrated FCM-CNN technique for multi-class skin lesion classification.
- Compare and contrast the performance of the proposed FCM-CNN approach against a baseline technique relying solely on CNN-based classification without segmentation.
- Analyze the strengths and weaknesses of using FCM-CNN technique, providing insights to guide future improvements in automated skin lesion classification and analysis.

The remainder of this report is organized as follows: Section 2 provides a detailed description of the dataset. Section 3 elaborates on the methods employed in this project, including preprocessing, image segmentation, classification models, fine-tuning, and performance metrics. Sections 4 and 5 present the results and discussions, respectively. Finally, Section 6 concludes the report, and Section 7 includes a link to the GitHub repository hosting the project.

2 Dataset

The HAM10000 (Human Against Machine with 10000 training images) dataset [15] is a comprehensive collection of 10015 dermoscopic images of pigmented skin lesions and is publicly available. It has been curated from multiple sources, representing diverse populations and acquisition modalities, to serve as a benchmark training set for academic machine learning research.
Each image is stored in JPEG format with a resolution of 600×450 pixels. The images are manually cropped and centered around the lesion, with adjustments for contrast and color reproduction. Metadata accompanying the images includes patient age, sex, lesion ID (unique identifier for each lesion), image ID, diagnostic type (dx type), anatomical location of the lesion, and diagnostic category.
The dataset is classified into seven diagnostic categories:

i. *Actinic Keratoses and Intraepithelial Carcinoma (akiec)*: Noninvasive squamous cell carcinoma (327 images).

ii. *Basal Cell Carcinoma (bcc)*: A type of epithelial skin cancer with low metastatic potential (514 images).

iii. *Benign Keratosis-like Lesions (bkl)*: Includes seborrheic keratoses, lichen-planus-like keratoses, and solar lentigo (1099 images).

iv. *Dermatofibroma (df)*: Benign growths or inflammatory responses to minor trauma (115 images).

v. *Melanoma (mel)*: Malignant melanocytic tumors treatable by early surgical intervention (1113 images).

vi. *Melanocytic Nevi (nv)*: Benign melanocytic neoplasms appearing in various morphologies (6705 images).

vii. *Vascular Lesions (vasc)*: Includes cherry angiomas, angiokeratomas, and pyogenic granulomas (142 images).

The HAM10000 dataset is highly imbalanced, with the *df* class containing only 115 images, compared to 6705 images in the *nv* class. To address this imbalance, data augmentation techniques are employed to create a balanced training dataset for the FCM-CNN model. Specifically, a balanced dataset of 21000 images is created by applying horizontal flips, random rotations, shearing transformations, adjustments to contrast, saturation, and hue, as well as the addition of Gaussian noise. This process results in approximately 3000 images per class (except for *nv* as it already has sufficient images). The HAM10000 dataset also includes an independent testing set (ISIC2018_Task3_Test_Images) of 1512 lesion images, which is used to evaluate the performance of the proposed FCM-CNN approach.

3 Methodology

The preprocessing as well as the CNN model architecture of this proposed work is primarily based on the methodology presented in [16], where a CNN-based multi-class skin lesion classification technique has been developed and a performance comparison is made with several well-known ML models. The methodological steps of the proposed project are shown in Fig. 1. All models are implemented in Python using *Keras*, *Tensorflow*, *OpenCV*, *Imutils*, *skfuzzy*, and *cv2Numpy* libraries. The key steps of the proposed methodology are outlined below:

3.1 Image Preprocessing

The image preprocessing module takes the HAM1000 dataset as input and conducts the following tasks-

- *Dataset Organization:* Lesion images are systematically categorized based on unique identifiers (lesion ID and diagnostic labels), ensuring clear organization for analysis.

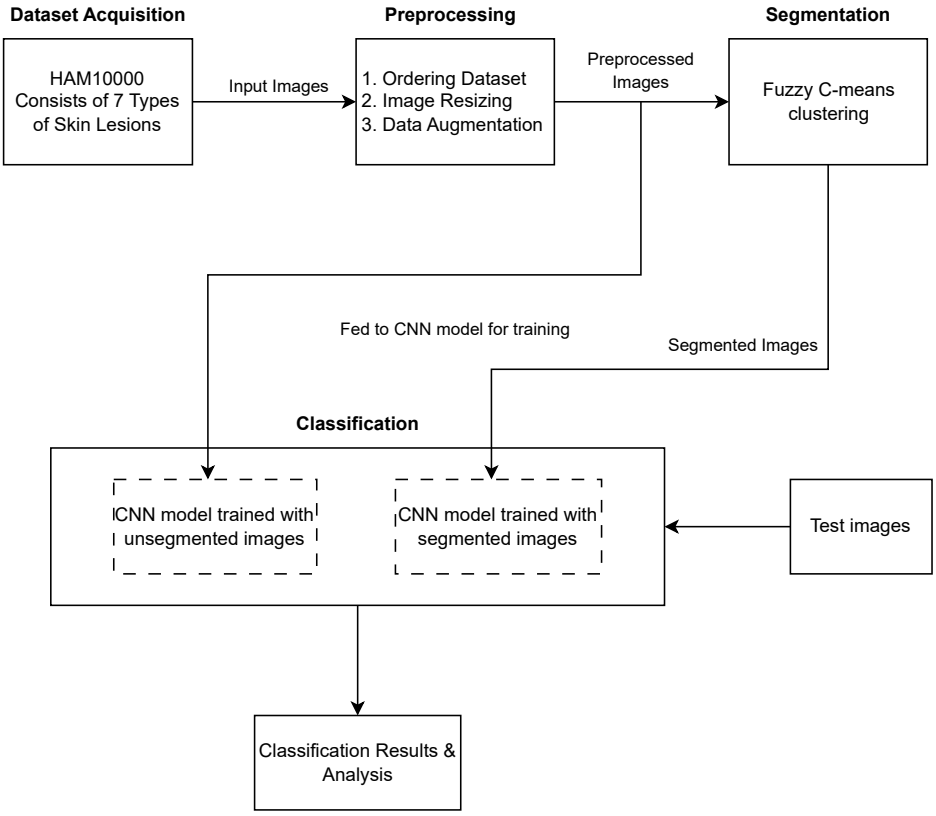


Figure 1: Overall methodology of the proposed FCM-CNN based multi-class skin lesion classification technique

- *Image Resizing:* Lesion images are resized to 96×96 pixels for CNN model, reducing memory requirements and computational overhead.
- *Data Augmentation:* The HAM10000 dataset is highly imbalanced. To address this challenge, data augmentation techniques are applied, including horizontal flips, random rotations, shearing transformations, contrast, saturation, and hue adjustments, and addition of Gaussian noise.

The preprocessing module outputs are passed to the segmentation module to segment the Region of Interest (ROI) of the skin lesion images. Outputs of the preprocessing module are also directly fed to the baseline CNN model for training.

3.2 FCM-based Image Segmentation

FCM clustering is a soft clustering technique widely used in image segmentation tasks. Unlike hard clustering methods (e.g., k-means) that assign each data point to a single cluster, FCM employs a degree of membership for each data point, thereby allowing partial belonging to multiple clusters. This membership-based approach often leads to smoother and more flexible segmentation results, as individual pixels can be associated with both the ROI and the background clusters with varying degrees of membership, rather than being forced into a single category. In FCM, the objective is to minimize the following cost function:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^{m} \|x_i - c_j\|^2,$$

where

- N is the total number of data points (pixels in the image),

- C is the number of clusters,
- x_i is the feature vector (e.g., pixel intensity value for this project) of the i-th data point,
- c_j is the centroid of the j-th cluster,
- u_{ij} is the membership degree of x_i in cluster j, satisfying $\sum_{j=1}^{C} u_{ij} = 1$ for all i,
- m is the fuzzification coefficient that controls the degree of cluster fuzziness (with $m > 1$).

In this project, FCM clustering is applied directly to the RGB values of skin lesion images. The algorithm proceeds by iteratively updating the membership values and recalculating the cluster centroids until convergence. The fuzzification coefficient m controls how memberships are distributed among clusters. Smaller m values yield results more similar to hard clustering, while larger values produce softer boundaries. Unlike threshold-based methods, FCM does not rely on predefined intensity thresholds. Compared to region-growing or edge-based techniques, it does not require clear boundary definitions, making it more flexible and suitable for complex lesion shapes.
In this project, the `skfuzzy` library in Python is used to apply FCM to segment skin lesion images into two clusters: one corresponding to the lesion (ROI) and the other to the surrounding background. After running FCM, the cluster with fewer pixels is assumed to represent the lesion, as lesions generally occupy a smaller portion of the image than the background. However, this assumption may not always hold. For example, if the lesion is large or the image contrast is poor, the cluster sizes may not reflect the actual ROI. Table 1 summarizes the main hyperparameters of the FCM-based segmentation used in this project.

Hyperparameter	**Value**
Number of clusters (C)	Set to 2 (ROI and background)
fuzzification coefficient (m)	Experimented values: 1.1, 2, 3.5
Max. number of iterations	1000
Error tolerance (ϵ)	1e-5
Initial cluster centers	Initialized automatically by `skfuzzy`

Table 1: Hyperparameters used for the FCM clustering algorithm

By comparing different values of m, we can investigate how increasing or decreasing the fuzzification coefficient affects segmentation quality. After the FCM algorithm is executed, its primary outputs are the cluster centroids and the membership degree matrices. To produce the final segmented image (a mask), we identify which cluster corresponds to the lesion (based on the smaller pixel count) and then generate a binary mask by selecting pixels that have higher membership in the lesion cluster. This binary mask highlights the ROI, enabling us to isolate the lesion from the background. This resulting mask is then used as input to the subsequent CNN-based classification stage.

3.3 CNN-based Image Classification

The adopted CNN architecture is shown in the Fig. 2. It is customized for classification of the described types of skin lesion images [16] and consists of the following key components:

(i) *Input Layer:* Takes images of size $96 \times 96 \times 3$ (RGB format).

(ii) *Convolutional Layers:* Three sets of convolutional layers with increasing filter counts: 32, 64, and 128 filters, each using 3×3 kernels. Each convolutional layer is followed by Batch Normalization to stabilize and accelerate training.

(iii) *Pooling Layers:*

- Max Pooling layers reduce spatial dimensions.
- 3×3 pooling after the first convolutional block.

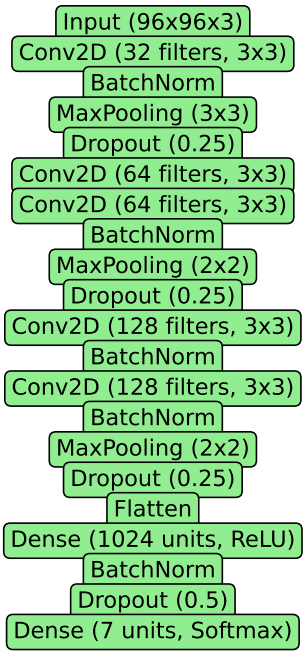


Figure 2: Architecture of the adopted CNN model

- 2×2 pooling after the second and third convolutional blocks.

(iv) *Dropout Layers:* Dropout regularization is applied with rates of 0.25 after the convolutional blocks and 0.5 before the final classification layer to prevent over fitting.

(v) *L2 Regularization:* L2 regularization is added to the convolutional layers and Dense layers to further mitigate over fitting by penalizing large weights.

(vi) *Flatten Layer:* Converts the feature map into a one-dimensional vector.

(vii) *Fully Connected Layers:*

- A Dense layer with 1024 units using ReLU activation and L2 regularization.
- A final Dense layer with 7 units and softmax activation for classification into 7 skin lesion classes.

This architecture combines feature extraction through convolutional and pooling layers with dense layers for classification, ensuring both spatial feature learning and robust generalization. We adopt $L2$ regularization in the model which would enhance the model's ability to generalize and prevents over fitting. The CNN model is optimized using the Adam optimizer and the categorical cross-entropy loss function. A hold-out validation approach is employed, where the dataset is split into separate training and validation sets, ensuring robust evaluation of the model's generalization. Training is conducted for a maximum of 140 epochs with a batch size of 32, and the learning rate is dynamically adjusted between 0.001 and 0.00001. Table 2 summarizes the hyperparameters used in the proposed CNN model.

3.4 Evaluation Metrics

To evaluate the performance of the FCM clustering-based segmentation across different values of the fuzzification coefficient (m), we employ four metrics: *Precision*, *Recall*, *F1-score*, and *Jaccard score*. The optimal value of m is determined based on these metrics. Furthermore, we assess the performance of the proposed FCM-CNN based multi-class skin lesion classification technique using four well-known metrics: *Accuracy*, *Precision*, *Recall*, and *F1-score*.

Hyperparameter	Value
Optimizer	Adam
Loss Function	Categorical Cross-Entropy
Epochs	140
Batch Size	32
Learning Rate	0.001–0.00001 (dynamically adjusted)
Holdout Validation	80-20 ratio
L2 Regularization	Weight Decay Factor (e.g., 0.001)

Table 2: Hyperparameters used for the CNN model

4 Experimental Result and Analysis

4.1 Segmentation Result

FCM clustering-based skin lesion segmentation for different values of the fuzzification coefficient m is evaluated with respect to the ground truth segmentation provided in the dataset. Table 3 summarizes the average segmentation performance across all seven lesion classes *(akiec, bcc, bkl, df, mel, nv, vasc)* for the three m values. Among the tested values, $m = 1.1$ yields the highest average precision, Jaccard, and F1-score. Although $m = 3.5$ provides slightly higher recall, its lower precision and overall balance make it less optimal. Consequently, $m = 1.1$ is chosen for further analysis.

m	**Precision**	**Recall**	**Jaccard**	**F1-score**
1.1	0.6378	0.5806	0.4393	0.6056
2.0	0.6178	0.5924	0.4345	0.6017
3.5	0.5863	0.6169	0.4288	0.5967

Table 3: Average segmentation performance across all classes for different values of the fuzzification coefficient m

One possible reason for the relatively low Jaccard values is that the Jaccard score, or Intersection-over-Union (IoU), imposes a more stringent criterion for segmentation quality. It heavily penalizes both false positives and false negatives, making even minor deviations from the ground truth significantly impact the score. Given that the ground truth masks are subjective and that our cluster selection assumes the smaller cluster corresponds to the lesion (which may not always hold), the final segmented region may include extraneous background pixels or miss parts of the lesion. As a result, the segmentation might overlap less precisely with the expert-annotated ground truth mask, leading to lower Jaccard scores compared to metrics like F1-score or precision.

To gain more insight, we examined the segmentation performance for each lesion class using $m = 1.1$. Table 4 presents the precision, recall, Jaccard, and F1-score for all seven classes. The results reveal notable variations among lesion classes. Lesion classes *mel* and *nv* achieve relatively higher segmentation performance, where as classes like *bcc* and *akiec* show lower performance. The overall mediocre segmentation performance and the variation of performances among classes can be attributed to the following major factors:

- *Subjective Ground Truth:* The reference or ground truth segmentations are provided by medical experts who manually delineate lesion boundaries. These annotations are inherently subjective, as different dermatologists may interpret subtle gradients and irregular shapes differently.
- *Heuristic Cluster Selection:* After applying FCM, we designate the cluster with fewer pixels as the lesion. While this heuristic often works, it is not infallible. In some images, the lesion may

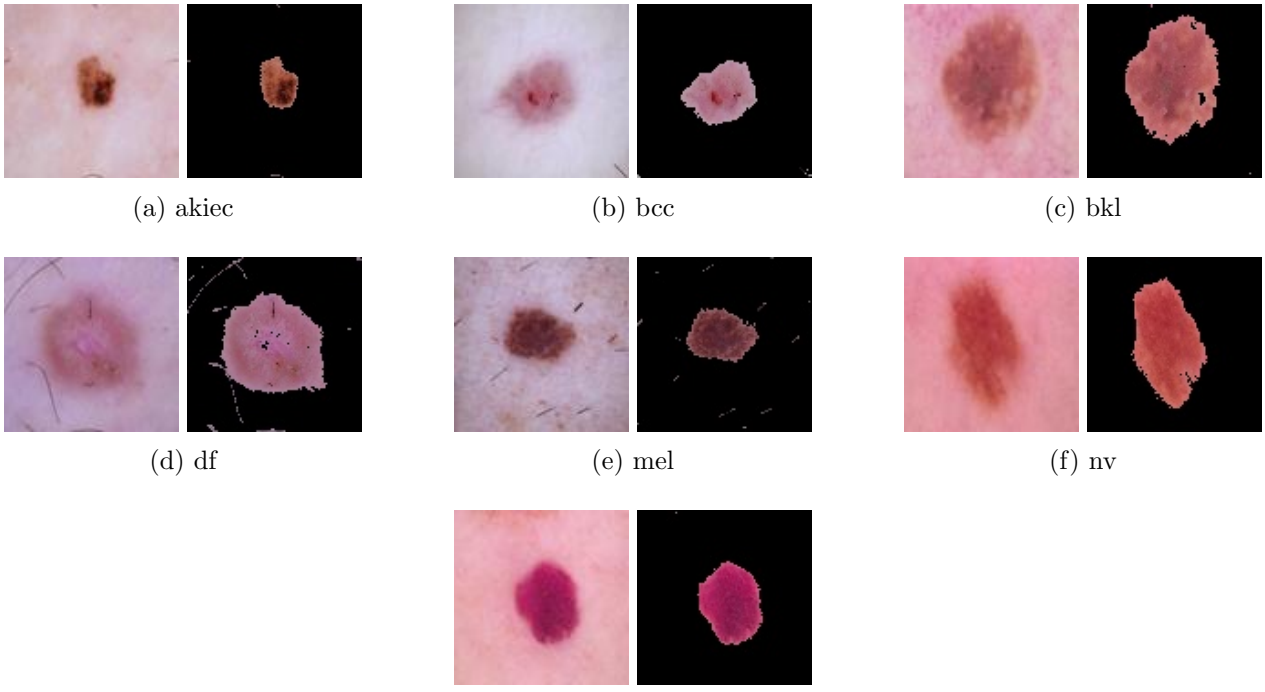
Class	Precision	Recall	Jaccard	F1-score
akiec	0.5813	0.4742	0.3534	0.5223
bcc	0.4610	0.4783	0.3067	0.4695
bkl	0.6156	0.5345	0.4007	0.5722
df	0.6553	0.6068	0.4600	0.6301
mel	0.7480	0.5964	0.4966	0.6637
nv	0.7865	0.6750	0.5705	0.7265
vasc	0.6166	0.6991	0.4873	0.6553

Table 4: Segmentation performance for each class using $m = 1.1$

occupy a relatively large area, or the background may contain small, dark regions that confound the assumption. As a result, the chosen cluster might not accurately represent the actual lesion, leading to performance drops.

- *Class-Specific Characteristics:* Certain lesion types are more challenging to segment due to their subtle contrast with the surrounding skin, irregularity in shape, or presence of additional artifacts. Such classes may require segmentation strategies that may incorporate color, texture, shape, or higher-level features in addition to pixel intensity values to distinguish the lesion from the background.

In summary, while FCM clustering offers a flexible and unsupervised approach to lesion segmentation, the observed performance limitations stem from subjective ground truths, a simplistic cluster selection heuristic, and class-dependent lesion characteristics.
A sample segmented lesion, alongside its corresponding unsegmented version, is shown for each class in Fig. 3.



(a) akiec (b) bcc (c) bkl

(d) df (e) mel (f) nv

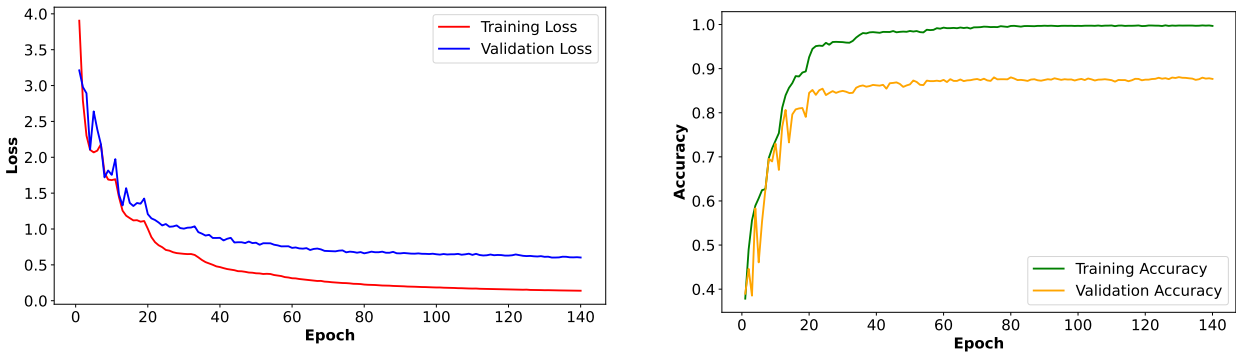
(g) vasc

Figure 3: Original and segmented skin lesion images for seven classes: akiec, bcc, bkl, df, mel, nv, and vasc. Each sub-figure contains the original image (left) and its segmented version (right).

4.2 Classification Result

4.2.1 Proposed FCM-CNN Model

The proposed FCM-CNN model is trained with 21000 segmented skin lesion images (3000 images per class) with holdout validation. Training and validation accuracy curves along with the corresponding loss curves of the FCM-CNN model are shown in Fig. 4. From the loss vs epoch curves, it is evident that both training and validation losses decrease rapidly during the initial epochs and gradually stabilize around epoch 100. Notably, the validation loss remains consistently higher than the training loss, indicating a potential gap between training and generalization performance. The accuracy vs epoch curves demonstrate a similar trend, with both training and validation accuracies increasing sharply during the early epochs and leveling off around epoch 100. The training accuracy approaches approximately 98%, while the validation accuracy stabilizes at around 86%. This suggests that the model achieves good generalization, with the testing accuracy expected to be close to the observed validation accuracy.



(a) Loss vs Epoch

(b) Accuracy vs Epoch

Figure 4: Training and validation performance of the FCM-CNN model: (a) Loss vs Epoch, (b) Accuracy vs Epoch.

Table 5 summarizes the overall performance of the FCM-CNN model on the testing dataset, providing the accuracy, as well as the weighted average of precision, recall, and F1-score. The model achieves a test accuracy of approximately 76.37%, which is notably lower than the validation accuracy of around 86% observed during training. This discrepancy is likely due to dataset augmentation, which may introduce some degree of similarities between the training and validation sets. Such similarities can inflate validation performance, making it less representative of the model's true generalization ability.

Metric	**Value**
Accuracy	0.7637
Precision	0.7527
Recall	0.7637
F1 Score	0.7550

Table 5: Performance metrics for the FCM-CNN model

The individual class performance metrics are shown in Table 6, while the confusion matrix is presented in Fig. 5. Among all classes, *nv* demonstrates the highest performance. Two key factors contribute to this outcome. First, the FCM clustering-based segmentation method performs particularly well for this class, resulting in high-quality segmented inputs that facilitate effective feature extraction by the CNN. Second, the *nv* class is well-represented in the dataset, eliminating the need for data augmentation for this class. The large and diverse sample set helps the model learn more generalized features for *nv*, thereby improving its performance on unseen test data.
Conversely, classes such as *akiec* and *bcc* exhibit the lowest performance. This shortfall can be

attributed to suboptimal segmentation results and high intra-class variability, both of which limit effective feature learning. Furthermore, the confusion matrix reveals that misclassifications for *akiec*, *bcc*, *bkl*, *df*, *mel*, and *vasc* often default to the *nv* class. This trend suggests that the augmentation performed on these classes—intended to counteract dataset imbalance—may inadvertently cause overfitting to certain intra-class variations. As a result, the model struggles to learn generalized features for these classes, leading to lower performance. These findings indicate that improved segmentation performance and the availability of sufficiently large, augmentation-free training sets could enhance the model's ability to generalize. By addressing these factors, it is possible that classification performance for those classes could approach that of the *nv* class.

Class	**Precision**	**Recall**	**F1-Score**	**Support**
akiec	0.5455	0.4186	0.4737	43
bcc	0.5616	0.4409	0.4940	93
bkl	0.6114	0.5438	0.5756	217
df	0.8056	0.6591	0.7250	44
mel	0.7014	0.5906	0.6413	171
nv	0.8217	0.9086	0.8630	908
vasc	0.7857	0.6286	0.6984	35
Weighted Average	0.7527	0.7637	0.7550	1511

Table 6: Performance Metrics for individual skin lesion classes for the FCM-CNN model

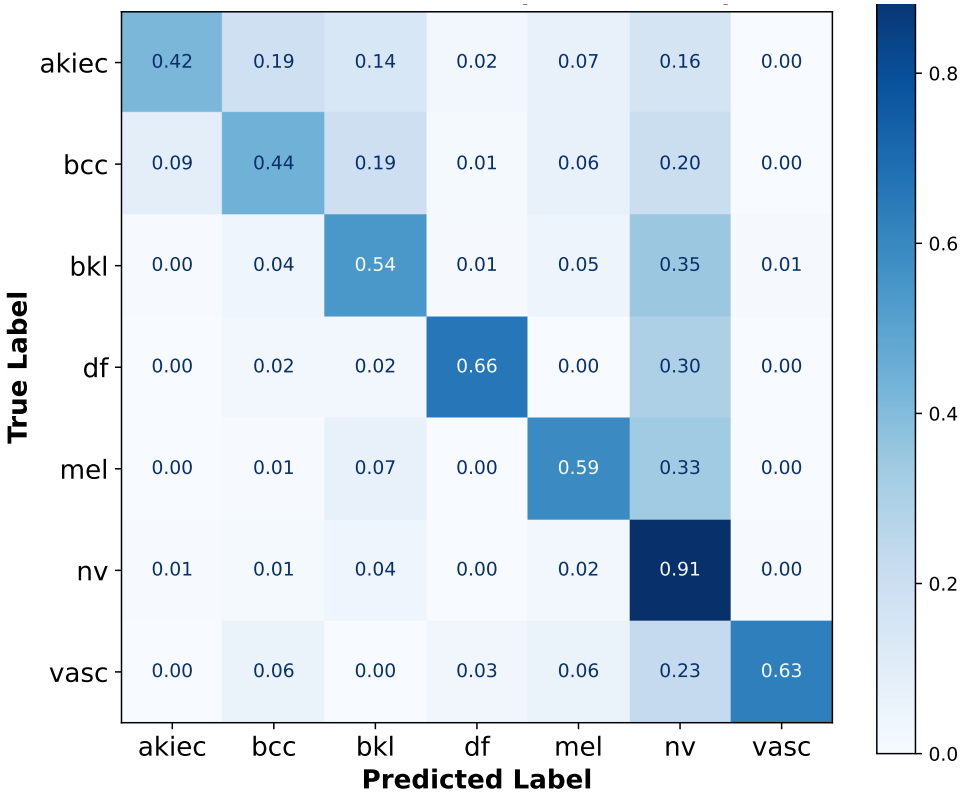


Figure 5: Confusion matrix (normalized) for the FCM-CNN model

4.2.2 CNN Model

We also evaluated a baseline CNN model, trained directly on the original skin lesion images without segmentation (as described in [16]), on the same test dataset for comparative analysis. The overall and per-class performance metrics for this baseline model are reported in Tables 7 and 8, respectively, while the corresponding confusion matrix is presented in Fig. 6. A visual performance comparison of the proposed FCM-CNN model against the baseline CNN model is provided in Fig. 7.
While the baseline CNN model achieves lower overall performance than the proposed FCM-CNN model, an interesting exception emerges when examining individual classes. Specifically, the baseline CNN demonstrates higher accuracy for the *akiec* and *bcc* classes compared to the FCM-CNN model. This discrepancy can be attributed to the segmentation step: among all the classes, *akiec* and *bcc* experience the weakest segmentation quality, potentially degrading the FCM-CNN's ability to extract meaningful features. In other words, the segmentation process, while generally beneficial, appears to have negatively affected performance for these two classes, allowing the simpler CNN approach (without segmentation) to outperform the proposed FCM-CNN model in those particular cases.

Metric	**Value**
Accuracy	0.6956
Precision	0.7171
Recall	0.6956
F1 Score	0.7028

Table 7: Performance metrics for the CNN model

Class	**Precision**	**Recall**	**F1-Score**	**Support**
akiec	0.3585	0.4419	0.3958	43
bcc	0.4747	0.5054	0.4896	93
bkl	0.4909	0.4977	0.4943	217
df	0.7826	0.4091	0.5373	44
mel	0.4148	0.5556	0.4750	171
nv	0.8690	0.8183	0.8429	908
vasc	0.6562	0.6000	0.6269	35
Weighted Average	0.7171	0.6956	0.7028	1511

Table 8: Performance Metrics for individual skin lesion classes for the CNN model

5 Discussion

In implementing the integrated FCM-CNN approach for multi-class skin lesion classification, several key insights are observed:

- *Limitations of Data Augmentation:* While data augmentation is essential for addressing class imbalance and improving model training, it can inadvertently lead to overfitting. By overly focusing on specific intra-class variations, augmented datasets may produce overly optimistic validation results that fail to generalize well to unseen test data. Therefore, it remains crucial to assess model performance using a truly independent test set.
- *FCM Clustering-Based Segmentation Technique:* Our chosen FCM clustering-based segmentation approach, which relies solely on pixel intensity and identifies the smallest cluster as the lesion, is a straightforward but not universally effective strategy. For certain classes, such as *akiec* and *bkl*, this simple heuristic underperforms. Achieving better segmentation may require more sophisticated cluster selection methods, incorporating additional image features, or integrating FCM with other segmentation techniques.

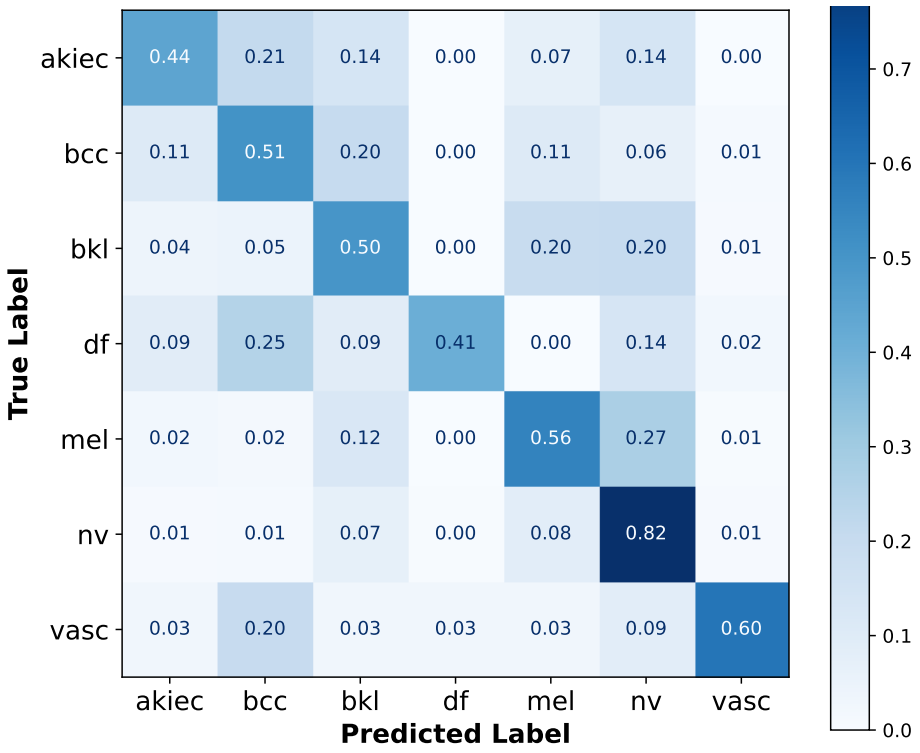


Figure 6: Confusion matrix (normalized) for the CNN model

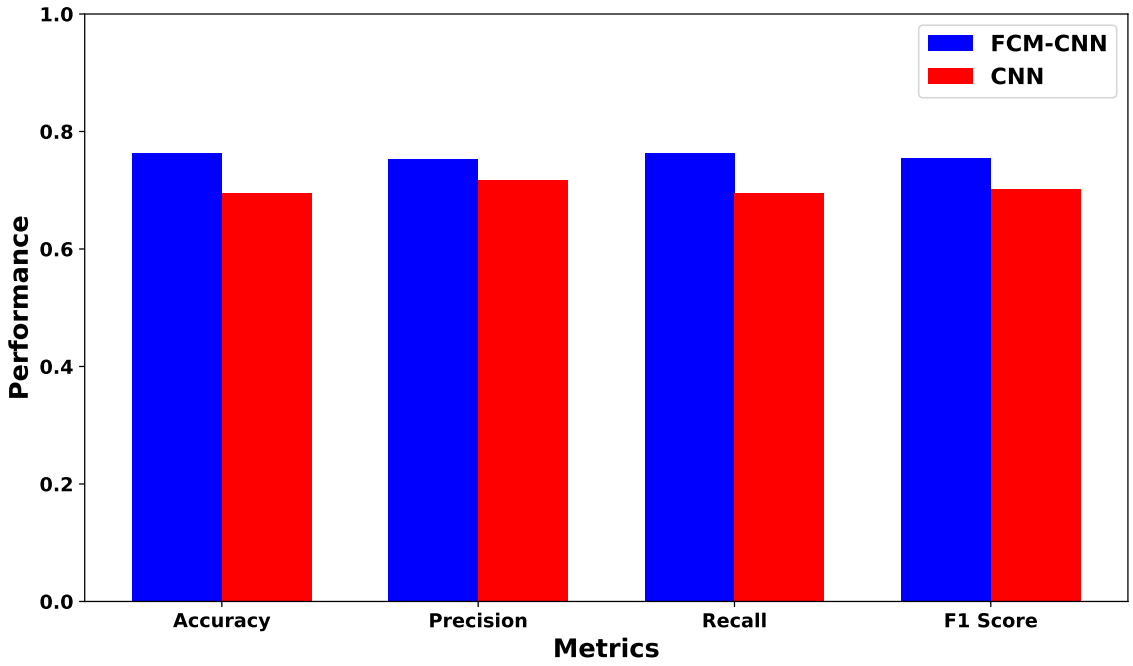


Figure 7: Performance comparison between the FCM-CNN model and only CNN-based model

- *Influence of Segmentation Quality on CNN Performance:* High-quality segmentation has the potential to significantly enhance the CNN's ability to learn discriminative features. Poor segmentation quality, as observed with the *akiec* and *bkl* classes, can hamper effective feature extraction and, consequently, can diminish overall model performance.

While the FCM-CNN approach outperforms a CNN-only model for multi-class skin lesion classification, this advantage hinges on the quality of the segmentation step. Future work may focus on refining segmentation strategies to fully realize the potential of integrated FCM-CNN techniques.

6 Conclusion

Skin cancer remains a pressing global health concern, with incidence rates rising significantly in recent decades. Deep CNNs have demonstrated remarkable potential for medical image classification, including skin cancer diagnosis, owing to their ability to automatically extract deep features from data. In this course project, we have implemented an integrated FCM-CNN approach for multi-class skin lesion classification. This method employs an FCM clustering-based technique for lesion segmentation, relying solely on pixel intensity values and selecting the smallest cluster as the lesion region, followed by CNN-based feature extraction and classification. Our findings indicate that the integrated FCM-CNN approach can outperform a CNN-only model. However, this performance gain is contingent on the quality of the segmentation step. Future work should emphasize refining the segmentation process, for instance, by employing more advanced cluster selection criteria, incorporating additional image features, or integrating FCM with other segmentation frameworks. Such advancements can further enhance the effectiveness of integrated FCM-CNN techniques, ultimately improving the accuracy and reliability of automated skin lesion classification.

7 Dataset and Source Code Repository

The dataset, source code, saved models, and results can be accessed from the following Github repository.
Link: https://github.com/MdTouhidulHaque/ECE624_Project

References

[1] B. C. Furriel, B. D. Oliveira, R. Prôa, *et al.*, "Artificial intelligence for skin cancer detection and classification for clinical environment: A systematic review," *Frontiers in Medicine*, vol. 10, p. 1 305 954, 2024.

[2] S. S. Chaturvedi, J. V. Tembhurne, and T. Diwan, "A multi-class skin cancer classification using deep convolutional neural networks," *Multimedia Tools and Applications*, vol. 79, no. 39, pp. 28 477–28 498, 2020.

[3] H. W. Rogers, M. A. Weinstock, S. R. Feldman, and B. M. Coldiron, "Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the us population, 2012," *JAMA dermatology*, vol. 151, no. 10, pp. 1081–1086, 2015.

[4] F. Grignaffini, F. Barbuto, L. Piazzo, *et al.*, "Machine learning approaches for skin cancer classification from dermoscopic images: A systematic review," *Algorithms*, vol. 15, no. 11, p. 438, 2022.

[5] M. Vestergaard, P. Macaskill, P. Holt, and S. Menzies, "Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: A meta-analysis of studies performed in a clinical setting," *British Journal of Dermatology*, vol. 159, no. 3, pp. 669–676, 2008.

[6] H. Kittler, H. Pehamberger, K. Wolff, and M. Binder, "Diagnostic accuracy of dermoscopy," *The lancet oncology*, vol. 3, no. 3, pp. 159–165, 2002.

[7] C. Morton and R. Mackie, "Clinical accuracy of the diagnosis of cutaneous malignant melanoma," *British Journal of Dermatology*, vol. 138, no. 2, pp. 283–287, 1998.

[8] T. G. Debelee, "Skin lesion classification and detection using machine learning techniques: A systematic review," *Diagnostics*, vol. 13, no. 19, p. 3147, 2023.

[9] A. R. Lopez, X. Giro-i-Nieto, J. Burdick, and O. Marques, "Skin lesion classification from dermoscopic images using deep learning techniques," in *2017 13th IASTED international conference on biomedical engineering (BioMed)*, IEEE, 2017, pp. 49–54.

[10] R. Baig, M. Bibi, A. Hamid, S. Kausar, and S. Khalid, "Deep learning approaches towards skin lesion segmentation and classification from dermoscopic images-a review," *Current medical imaging*, vol. 16, no. 5, pp. 513–533, 2020.

[11] A. K. Nugroho, R. Wardoyo, M. E. Wibowo, and H. Soebono, "Image dermoscopy skin lesion classification using deep learning method: Systematic literature review," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 2, pp. 1042–1049, 2024.

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

[13] T. Lei, P. Liu, X. Jia, X. Zhang, H. Meng, and A. K. Nandi, "Automatic fuzzy clustering framework for image segmentation," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 9, pp. 2078–2092, 2019.

[14] N. Nida, A. Irtaza, A. Javed, M. H. Yousaf, and M. T. Mahmood, "Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy c-means clustering," *International journal of medical informatics*, vol. 124, pp. 37–48, 2019.

[15] P. Tschandl, *The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions*, version V4, 2018. DOI: 10.7910/DVN/DBW86T. [Online]. Available: https://doi.org/10.7910/DVN/DBW86T.

[16] B. Shetty, R. Fernandes, A. P. Rodrigues, R. Chengoden, S. Bhattacharya, and K. Lakshmanna, "Skin lesion classification of dermoscopic images using machine learning and convolutional neural network," *Scientific Reports*, vol. 12, no. 1, p. 18 134, 2022.