Journal of Theoretical and Applied Information Technology
15th March 2025. Vol.103. No.5

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

LEARNING-BASED APPROACH FOR AUTOMATIC SEGMENTATION OF IMAGES

[1]M SRIVIDYA [2]. DR. VENUBABU RACHAPUDI

[1]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India -522302

[2]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India-522302

Email-id[1]madugulasrividya@gmail.com [2]venubabu.r@gmail.com

ABSTRACT

One of the key fields of study for computer vision applications is image segmentation. Often, segmentation is used as a preprocessing step in various real-world applications. The existing literature has revealed that image segmentation can be done with heuristic and learning-based techniques. With the emergence of artificial intelligence, learning-based approaches became popular in computer vision applications. However, there is a need to explore ML and DL-based methods with optimizations geared towards image segmentation efficiency. In this document, we suggested an algorithm known as Spectral Clustering-Based Image Segmentation (SCB-IS), which exploits the spectral clustering process for efficient image segmentation. We also proposed an approach to segmentation based on deep learning that exploits SegNet, which is widely used for image segmentation. We proposed another algorithm known as Deep Learning-Based Image Segmentation (DLB-IS). Our empirical study with benchmark datasets revealed that the proposed deep learning-based algorithm outperforms many cutting-edge deep learning models, featuring the highest accuracy of 88.50%. The suggested learning-based approaches can be integrated into real-world computer vision applications for efficient segmentation of images.

Keywords - *Image Segmentation, Machine Learning, Deep Learning, Clustering, Artificial Intelligence, SegNet*

1. INTRODUCTION

As the adage " An image may convey a thousand words " says, examining an image might provide more information than reviewing a piece of writing. Computer vision research focuses on image segmentation, or dividing an image into its objects or regions of interest (ROI). Parts comparable to each other are formed by grouping the picture pixels. Preparing images for various image-based applications, including pattern recognition, medical imaging, object identification and classification, and biometric identity, is essential [1]. Here are a few of the well-known uses. Content-based image retrieval is analogous to searching big databases for digital photographs pertinent to a query. What is retrieved depends on the contents of the query picture. To retrieve an image's contents, segment the image. Robotic inspection and image analysis is known as machine vision and is mainly used in industry. This instance of segmentation involves extracting information from the gathered picture regarding a device or processed substance. Imaging in medicine: These days, picture division is helpful in many aspects of medical research, from medical operations to diagnostics. Examples include segmenting tumors to locate them, dividing a tissue to calculate the associated amounts, and segmenting units to carry out several digital pathology activities such as counting cells, classifying nuclei, etc.
Object recognition and detection: One significant use of computer vision is identifying and recognizing objects. Here, an item might be a person's face, a passing car, or any number of aerial items, such as highways, woods, farms, etc. Since the indented object must first be extracted from the picture, this program is necessary for segmenting images. Video surveillance: To accomplish an assigned goal, such as determining the activity being done in the recorded pictures or handling traffic, tallying the number of items, and numerous other tasks, the motions in the area of interest are captured by the video camera, which

then evaluates them. Segmenting the area of interest is the first step in doing the analysis.

From a computer vision standpoint, segmenting a picture into its constituent ROIs may seem simple to humans, but it is a pretty complicated process. A variety of issues might impact the effectiveness of an image segmentation method. Three main picture segmentation issues are identified. Illumination variation: It seriously affects pixels and is a fundamental issue with picture segmentation. This variance results from the various lighting conditions during which the picture was taken. Background complexity: A significant problem is an image with a complicated backdrop. When segmented, an image's region of interest may blend with its complicated surroundings and limitations. Such background complications, therefore, hamper the effectiveness of segmentation techniques. Furthermore, instead of using hundreds of pixels, displaying picture segmentation consists of producing a picture with a few essential segments. In addition, picture segmentation is considered a type of clustering: pixels meeting a set of criteria form a cluster, whereas pixels not meeting the requirements are divided into several groups. The standard method for addressing these issues and facilitating learning from unlabelled data is to categorize the data according to certain similarity or dissimilarity metrics and then label each group. The process of putting the data into groups is known as clustering.

The primary focus is to explore image segmentation using spectral clustering and deep learning-based techniques with a primary emphasis on SegNet. The work in the title primarily focused on benchmarking segmentation performance against benchmark datasets, applicable in medical imaging, object detection, and computer vision tasks. However, the study operates under the assumption that input images are relatively clean with few distortions, which may not always be accurate in the wild. Moreover, spectral clustering improves the segmentation performance but may have limitations in segmented high-complex and overlapped structures. While the automatic deep learning approach is highly effective, it is also resource-expensive and requires sufficient computational power, which is not always feasible in low-resource settings. These methods will be further improved to provide better adaptability in diverse real-world situations in future work.

Recent advances in deep learning and clustering-based techniques have significantly improved image segmentation, yet challenges remain in balancing efficiency, accuracy, and adaptability across diverse datasets. This research aims to bridge this gap by formulating **the following research objectives**: (i) to develop a spectral clustering-based segmentation algorithm (SCB-IS) for efficient and scalable image segmentation, (ii) to propose a deep learning-based segmentation approach (DLB-IS) leveraging SegNet to enhance segmentation accuracy, and (iii) to conduct a comparative evaluation with state-of-the-art methods to establish the efficacy of the proposed approaches.

The following are the things we contributed to this paper. We have developed two image segmentation algorithms: Spectral Clustering-Based Image Segmentation (SCB-IS) and Deep Learning-Based Image Segmentation (DLB-IS). SCB-IS uses spectral clustering for efficient image segmentation, while DLB-IS utilizes SegNet, a popular method for image segmentation. Our empirical study using benchmark datasets found that the DLB-IS algorithm achieved the highest accuracy of 88.50%, outperforming many advanced models for deep learning. These algorithms can be integrated into real-world computer vision applications for effective image segmentation. The following is how the remaining portions of the paper are arranged: The background knowledge needed to comprehend the suggested segmentation techniques is as stated in Section 3. Section 4 presents materials and techniques, including information on segmentation techniques. The experimental findings from our empirical investigation are shown in Section 5. Part 6 is where we conclude our investigation and offer recommendations for future research possibilities.

2. RELATED WORK

Clustering can be used to aid in the segmentation of image techniques and also deep learning methods. Das et al. [1] state that automatic detection is essential for an early diagnosis of leukemia. DL and ML techniques for acute lymphoblastic leukemia (ALL) are examined in this study. It talks about future avenues for research and classifies methodologies. Vatsavayi and Andavarapu [2] suggested classifying wild animals from movies with the DRCNN-TSO algorithm. Before features are extracted and classified, video frames are pre-processed and segmented. Metrics of performance evaluate the approach against current practices. Zolfaghari and Sajedi [3] identified and categorized WBCs and acute leukemia. The accuracy of conventional SVM and CNN classifiers has increased. The best performance metrics are obtained by combining CNN with SVM. CNN is suggested for use in future research for better categorization. Shandilya et al. [4] produced a collection of drone and bird photos that are used to train accurate UAV detection algorithms and correct misclassifications of birds. It supports the assessment of models using training and testing sets, which is essential for UAV safety. Abdigapporov et al. [5] used an encoder-decoder with feature fusion for enhanced performance, and a joint object identification and semantic segmentation approach is suggested. Superior to previous approaches when evaluated on the ICV22 dataset.

Naik et al. [6] examined for the identification of lentils. The Indian lentil identification system EfficientNetB0 was shown to be better. Possibility of both an intuitive user interface and local optimization. Zhang et al. [7] covered foundation models like SAM. Examination of SAM's medical application revealed issues and provided direction for further study. Jiang et al. [8] used AG-BCE loss, MicroSegNet, a transformer UNet model, to improve prostate segmentation on micro-ultrasound. They have achieved exceptional results, supporting the diagnosis of malignancy. Herdy et al. [9] suffered greatly from biocrusts. Austria and Utah saw consistent success using a semantic segmentation method that included domain adaptation. Chong et al. [10] intended to use ML/DL techniques to categorize microalgae. Outperforming textures and geometrical features produced accuracy. With a high accuracy, DL may be able to develop a universal model.

Lee et al. [11] required due to the growing number of vehicles and weather-related potholes. Machine learning techniques increase forecast accuracy by utilizing temperature and traffic volume. Luo et al. [12] revolutionized semantic segmentation, which helps with pest identification, crop analysis, and other tasks. Robustness and labeled samples continue to pose challenges. Methods can be improved by integrating multimodal data. Mobarak et al. [13] delved into using machine learning in materials research, transforming comprehension and creativity via image processing, synthesis, and forecasting. Ahammed et al. [14] accepted that skin disorders necessitate prompt diagnosis. The proposed technique achieves accuracy by combining digital hair removal, SVM classification, and lesion segmentation. Although Davydzenka et al. [15] take much time to segment manually, X-ray µCT pictures help with the material investigation. Machine learning greatly enhances segmentation accuracy, particularly when combined with enriched data.

Loverdos et al. [16] used DL to partition bricks and detect cracks, enhancing masonry analysis automatically. Future research will focus more on fault identification and drone surveillance. Singh et al. [17] enhanced by techniques inspired by nature, such as artificial neural networks and genetic algorithms. More studies are required to improve automatic segmentation methods and system performance measures. Sarraf et al. [18] examined four Deep Learning Convolutional Neural Networks for segmenting 3D nan tomography images. With its excellent performance and quick training times, U-Net++ provided effective data segmentation for X-ray microscopy. Accuracy might be increased with further improvements like pre-processing and data augmentation. Hou et al. [19] needed to guarantee operation and safety. Pavement analysis benefits from intrusive sensing, image processing, and machine learning. Sensor optimization, adaptive imaging algorithms, and growing datasets for machine learning are among the ongoing challenges. Walvekar et al. [20] state that early diagnosis is essential because COVID-19 threatens world health. The greater sensitivity of CT imaging compared to RT-PCR helps in diagnosis. Clinicians' judgments are aided by automated segmentation using U-Net architecture, which lessens their workload.

Minaee et al. [21], with algorithms ranging from CNN, RNN, attention models, and more, Deep Learning (DL) has revolutionized picture segmentation. A review of performance is done using datasets such as PASCAL VOC. Haque and Neubert [22] examined the effects of deep learning on the segmentation of biological images, emphasizing issues with generalization, dataset size, and data sharing. It highlights the necessity for more labeled pictures and investigates options for unsupervised learning. Hesamian et al. [23], with a summary of techniques and issues resolved, demonstrated that deep learning is an essential tool for medical picture segmentation. Garcia et al. [24] a lot of applications, such as virtual reality and autonomous driving, depend on semantic segmentation using deep learning. This work reviews data sets, strategies, difficulties, and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

potential paths forward. Guo et al. [25] enhanced the tumor segmentation using PET, CT, and MRI. Multi-modal fusion works well in networks, mainly when dealing with poor-quality pictures. More structures will be tested, and future studies will investigate unsupervised techniques. It is essential to register accurately across modalities.
Baumgartner et al. [26] state that automated cardiac segmentation is essential for evaluation. Using MR images as test data, 2D and 3D CNNs displayed Dice coefficients of 0.950 (LV), 0.893 (RV), and 0.899 (Myo). Liu et al. [27] examined modern and conventional DNN techniques, describing innovations and best practices. Segmenting pictures semantically is crucial for computer vision and image processing. Lateef and Ruichek [28] state that semantic segmentation is essential for computer vision applications such as medical research and autonomous cars. Different datasets are used to test and classify deep learning techniques, emphasizing unresolved issues. Wang et al. [29] suggested techniques to modify CNNs to increase precision and effectiveness. Interactive frameworks enhance medical picture segmentation using deep learning. Tajbakhsh et al. [30] recently helped to segment medical images more effectively, although this technique requires massive, well-annotated datasets. An emphasis on economically viable approaches is placed on solutions for weak and limited annotations.
Huang et al. [31] obtained this by merging FCM clustering with rough set theory in a novel picture segmentation technique that demonstrates anti-noise capacity and reduces mistakes in fuzzy boundary areas. Kim et al. [32] used feature extraction and differentiable clustering, and a novel CNN architecture was suggested for unsupervised picture segmentation. The introduction of extensions for user input and reference photos demonstrated improved accuracy and efficiency. Hoang and Kang [33] suggested using post-processing techniques, statistics computation, and feature embedding to create a pixel-level clustering architecture for segmenting pictures without supervision. This strategy works better than prior approaches, helping with different computer vision problems. Panic et al. [34] suggested a mixed model for the unsupervised mCT image segmentation of cellular metals, incorporating a novel merging method and spatial regularization. Promising findings are obtained after a thorough study of five materials. Ocana et al. [35] focused on heuristic electron tomography picture segmentation methods by solving the discrete ordered median issue. The goal is to minimize computation time without sacrificing image quality. Notable gains are shown when VNS is paired with the *P*-median heuristic.
Yadav et al. [36] evaluated the following methods for detecting outliers: DOFCM, CFCM, NC, and FCM. The best results are produced by DOFCM, which reliably detects outliers and correctly constructs clusters. Tseng and Tang [37] improved the XGBoost model's ability to detect brain tumors. PSO maximizes feature selection, K-Means aids in segmentation, and CLAHE enhances pictures. Oskouei and Hashemzadeh [38], with cluster weighting and group-local feature weighting, CGFFCM is a color picture segmentation technique that improves accuracy. Utilizing an imperialist competitive algorithm, it optimizes by combining eight characteristics. Mittal et al. [39] tagged pixels comprise a large portion of an image, and computer vision relies heavily on image segmentation to extract information. Performance measurements and datasets are covered, and partitional clustering techniques are reviewed. In the future, meta-heuristic techniques will be improved, and validity indices and algorithm stability will be compared. Qureshi and Ahamad [40] depend on images, and picture segmentation is essential for jobs like retrieval and pattern detection. Using neutrosophic logic to address uncertainties, this study presents a clustering algorithm that iteratively refines clusters for better segmentation. Deep learning also allows for the segmentation of images. Model innovations are also found for image processing innovations in [61] and[62]. More deep-learning optimizations are also found in [63] and [64]. Novel deep learning-based optimized ideas are also found in [65] and [66]. sLearning-based approaches with different optimizations are also found in [67], [68], and [69]. Many methods have been proposed for image segmentation based on clustering and deep learning, but they are limited. Conventional clustering methods like K-means and Fuzzy C-means are commonly sensitive to starting point selection and fail to accommodate complex and overlapping structures [39],[40]. Deep learning techniques, such as U-Net and Fully Convolutional Networks (FCN), have shown high Segmentation results but demand large amounts of labeled data and computational resources, which make them difficult to use in real-time [21], [29]. Although recent model upgrades like DeepLab and Transformer-based models have enhanced feature extraction aspects,

they are still subpar in fine-grain boundary separation and generalization on the same [7], [8]. The Spectral Clustering-Based Image Segmentation and Deep Learning-Based Image Segmentation algorithms proposed in this paper try to address these problems by combining spectral clustering and SegNet to deliver an efficient and accurate segmentation algorithm with low computational demands suitable for real-world computer vision problems.

3. PRELIMINARIES

Generally speaking, recommendation systems and market research are among the practical applications of clustering, which include photo segmentation, social network analysis, and online query search [42]. Optimizing similarity within clusters and reducing dissimilarity between them is the main objective of clustering methods, which group unlabeled pixels into homogenous groups with the highest similarity. The clustering process meets the following requirements on an image (X) of size (m × n), specified across d-dimensions, mathematically producing K clusters {C1, C2,…, Ck}:

$$C_i \neq \emptyset, for\ i = 1,2, \ldots, K$$
$$C_i \cap C_j = \emptyset, for\ i\ and\ j = 1,2, \ldots, K\ and\ i \neq j$$
$$U_{i=1}^{K} C_i = X$$

As guaranteed by the first requirement, every generated cluster must include at least one pixel. The following requirement states that no pixel will be allocated to multiple clusters; all generated clusters will be mutually exclusive. The last requirement is that the data values assigned to each cluster accurately depict the whole image. Several clustering techniques for picture segmentation are available in the literature. There isn't a set definition for a cluster, though, and several clustering approaches have their ways of organizing the data. Thus, the grouping approaches may be categorized into two primary categories, partitioned and hierarchical, based on the cluster formation.

3.1 Hard clustering methods

By applying a target function to the data, complex clustering algorithms repeatedly partition the data into clusters. The total squared Euclidean distance—which must be kept to a minimum—between the data and the associated centroid usually serves as the objective function. In these approaches, the centroid of the clusters is often defined as the center of the clustered data. Furthermore, hard clustering, as opposed to soft clustering, gives each data item a degree of belongingness of 0 or 1, assigning it to a single cluster alone. The challenging grouping method has high computational efficiency, is scalable, and is somewhat easy to use. Additionally, it can handle well-separated datasets with a spherical form. Nevertheless, it has some drawbacks, including the fact that the created cluster centroids are not very good cluster descriptors, that they are sensitive to the initial parameter choices, and that it is necessary to know in advance how many clusters will be created.

3.1.1 Kmeans-based methods

When employing Kmeans-based methods, the cluster centroid is updated using the average value following the assignment of each data item to the correct cluster. This process is carried out repeatedly until a specific convergence condition is satisfied. While techniques in this category provide advantages such as comparatively low time complexity, simplicity, and assured convergence, some drawbacks must be addressed. These restrictions include the need to know ahead of time how many clusters will be formed, the knowledge that the quality of the outcome is dependent upon how many clusters are formed initially, the fact that the strategy is inappropriate for data with non-convex distributions, the fact that it follows a hill-climbing path and thus frequently traps into local optima, and the reality that noise affects it somewhat, outliers, and the initialization phase. Generally speaking, every technique that draws inspiration from the most basic partitional clustering technique—K-means—falls under this category. It is the gold standard for literary clustering [42]. Below is a summary of the K-means algorithm.

K-means: There are several clusters (k) created from a collection of data points, X={x_1,...,x_n}. using K-means [43]. It divides data according to the similarity criterion, often the total of the squared errors, as stated in (4).

$$J = \sum_{i=1}^{k} sum_{x_j \in} X \left\| x_j - m_i \right\|^2$$
(4)

Since related clusters, $C = \{c_1, \cdots, c_k\}$, are represented collectively as $M = \{m_1, \cdots, m_k\}$, m_i is the designation for the centroid of cluster I. The criterion function J is continually reduced while using this method. Additionally, in compliance with (5) and (6), the created clusters C and related centroids M are updated.

$$x_i \in c_l, if\ l = argmin_{l=1}^{k} \left\| x_j - m_i \right\|^2 \quad (5)$$

$$m_i = \frac{\sum x_{i \in c_l} x_i}{|c_l|} \quad (6)$$

for $1 \leq i \leq N$ and $1 \leq l \leq k$.
The temporal complexity of the K-means method is O(nkt), where n, k, and t represent the number of data items, maximum iterations, and clusters to be created, respectively. Nevertheless, this approach typically traps into local minima and is biased towards the original cluster centroids. Moreover, the number of clusters affects the solutions.
Additional well-liked techniques in this area include partition around medoids (PAM) [44], K-means [45], sort-means [46], K-harmonic means [47], K-modes [48], and K-medoids [49] are divided into two halves, thus it is possible to cluster big applications using randomized search (CLARANS) [50]. Whichever data point is closest to the cluster center is the cluster centroid according to the K-medoids approach, which is a variation of K-means. For discrete data, this approach is appropriate. PAM bears similarities to the two-phase, medoid-based approach known as build and swap. Minimizing the difference between the objects with the nearest medoid is the goal of this greedy search strategy. A collection of data points are regarded as medoids in the first phase, and they determine the set of chosen objects. The leftover data points are stored in an additional collection known as unselected objects. This is in line with the construction phase's stages. The subsequent phase aims to enhance the cluster quality by replacing the data from the unselected objects set with the information from the selected objects group. A good clustering approach is CLARANS, especially for large datasets. Discovering the medoids for the supplied data takes advantage of the graph idea. It merely considers a portion of the potential swaps between the picked and unselected items, applying PAM to the entire data collection.

4. MATERIALS AND METHODS

This section presents the suggested procedure, which includes a clustering-founded approach and a deep-learning approach for efficient image segmentation.

4.1 Fast Spectral Clustering-based Segmentation

This section provides a thorough overview of FSC. It focuses on superpixel-level spectral grouping. Assume that there are m superpixels in picture I, or ., $I = \bigcup_{i=1}^{m} A_i$, where m is a considerably smaller number of pixels than n, and Ai is the i-th superpixel. The quad-tree's leaf node blocks are regarded as superpixels in FSC. Next, we use FSC to resolve the superpixel-based spectral image segmentation problem. The task involves dividing the superpixel set into k clusters, denoted as B = {B1,..., Bk}. The issue is comparable to the following minimization problem, which is Ncut's equivalent:

$$\min_{B_1,B_2,\dots,B_k} FSC(B_1, B_2, \dots, B_k),$$

where $FSC(B_1, B_2, \dots, B_k)$ Is defined as:

$$FSC(B_1, B_2, \dots, B_k) = \sum_{i=1}^{k} \sum_{j=1, j\neq 1}^{k} \frac{cut(B_i, B_j)}{vol(B_i)},$$

and

$$cut(B_i, B_j) = \sum_{A_i \in B_i} \sum_{A_j \in B_j} \frac{cut(A_i, A_j)}{\sqrt{|A_i||A_j|}}$$

$$vol(B_i) = \sum_{A_j \in B_i} \sum_{z=1}^{m} \frac{cut(A_j, A_z)}{\sqrt{|A_j||A_z|}},$$

where cut $(A_i, A_j) = \sum_{v_i \in A_i, v_j \in A_j} w_{ij}$ and $|A_i|$ is how many pixels there are in superpixel AI.
Superpixel-based unnormalized graph Laplacian and the cluster B_j indicator vector (j = 1,2,...,k) are needed to solve the issue given by Eq. (9). Define the superpixel indicator vector before building the unnormalized graph Laplacian using superpixels. A_j $(j = 1,2,\dots,m)$ as $h_j = (h_{1j}, h_{2j}, \dots, h_{nj})^T$ by

$$h_{ij} = \begin{cases} \frac{1}{\sqrt{|A_j|}} & v_i \in A_j \\ 0 & otherwise \end{cases} \quad (10)$$

$(i = 1,2,\dots,n)$,

straightforwardn the quantity of pixels in superpixel A_j is denoted by $|A_j|$. Next, we create a matrix $H \in \mathbb{R}^{n\times m}$ using the vectors that serve as indicate A_j (j = 1,...,m) as its columns. Matrix H is a mapping that takes place between the pixel space and the superpixel space. Observing that matrix H is an orthogonal matrix with columns is simple. Using matrix H and W, the superpixel-based similarity matrix $\tilde{W}$ is obtained as follows:

$$\tilde{W} = H^T W H$$

$$\begin{bmatrix} \frac{cut(A_1,A_1)}{|A_1|} & \cdots & \frac{cut(A_1,A_m)}{\sqrt{|A_1||A_m|}} \\ \vdots & \ddots & \vdots \\ \frac{cut(A_m,A_1)}{\sqrt{|A_m||A_1|}} & \cdots & \frac{cut(A_m,A_m)}{|A_m|} \end{bmatrix}$$

(11)

At this point, we may define the degree $\tilde{D}$ matrix using superpixels:

$$\tilde{D} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_m \end{bmatrix}, \quad (12)$$

$d_i = \sum_{j=1}^{m} \widetilde{W}_{ij}$.

The unnormalized graph Laplacian with matrices $\widetilde{W}$ and $\widetilde{D}$, defines, according to superpixels $\tilde{L}$, in this manner:

$\tilde{L} = \widetilde{D} - \widetilde{W}$. (13)

Afterwards, the indicator vector is defined $g_j = \left(g_{1j}, g_{2j}, \dots, g_{mj}\right)^T$ of B_j $(j = 1,2,\dots,k)$:

$$g_{ij} = \begin{cases} \frac{1}{\sqrt{vol(B_j)}} & A_i \in B_j, \\ 0 & otherwise \end{cases} \quad (14)$$

$$(i = 1,2,\dots,n),$$

Obtaining the following equations is simple:

$$g_j^T \widetilde{D}_{gj} = \frac{\sum_{A_i \in B_j} \sum_{z=1}^{m} \frac{cut(A_i, A_z)}{\sqrt{|A_i||A_z|}}}{vol(B_j)} = 1$$

$$g_j^T \tilde{L}_{gj} = \sum_{i=1, i \neq j}^{k} \sum_{A_a \in B_i} \sum_{A_b \in B_j} \frac{cut(A_a, A_b)}{vol(B_j)\sqrt{|A_a||A_b|}} = \sum_{i=1, i \neq j}^{k} \frac{cut(B_j, B_i)}{vol(B_j)}.$$

Then, construct a matrix $G \in R^{m \times k}$ in where the columns are the vectors that indicate B_j $(j = 1,2,\dots,k)$.

Consequently, The minimization issue in Eq. (9) may be reformulated as follows:

$$\min_{B} Tr\left(G^T \tilde{L} G\right) \quad (15)$$

$$\text{s.t. } G^T \widetilde{D} G = E$$

Tr r represents the matrix's trace. The aforementioned difficulty can be loosened by enabling any actual value to be entered into the matrix G elements. $P = \widetilde{D}^{\frac{1}{2}} G$ is substituted. We now have the following easy problem:

$$\min_{P \in R^{m \times k}} Tr\left(P^T \tilde{L}_N P\right) \quad (16)$$

$$s.t. \quad P^T P = E,$$

For a trace minimization problem, the typical version of Equation (16) is the minimization problem. The Laplacian $\tilde{L}_N$s initial k eigenvectors are the answer to the Rayleigh-Ritz theorem [51]. It is evident that $\tilde{L}_N$ is sparse and has a size of m × m. As a result, compared to calculating the Laplacian matrix's eigenvector, expressed in terms of n×n in Ncut, the computing cost of finding the eigenvectors of $\tilde{L}_N$is substantially smaller. After transforming G to G_p Depending on pixels from G based on superpixels, we can get the matrix that contains the pixel-based clustering information:

$G_p = HG$ (17)

where Eq. (10) defines H. Next, we handle every row in $G_p$as a point $\in R^k$ And employ the Fuzzy C-means method to create k clusters out of all the locations, deriving the clustering outcome based on pixels.

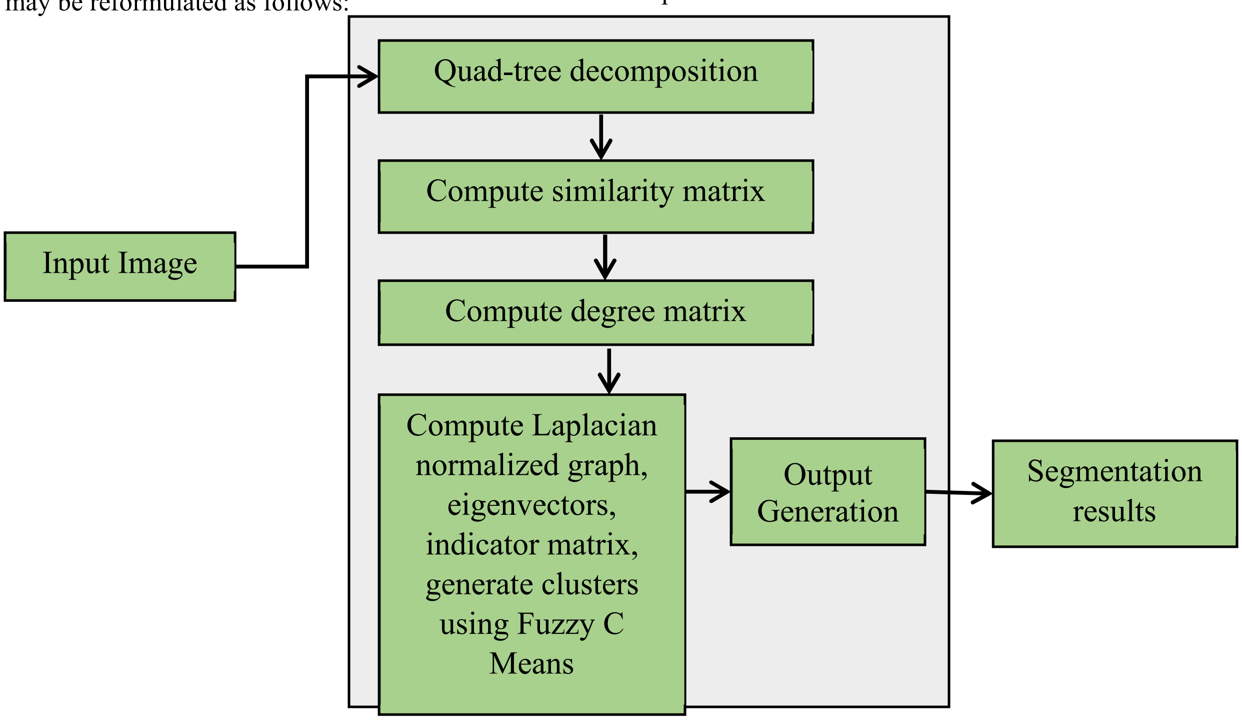


Figure 1: Illustrates Spectral Clustering-Based Image Segmentation

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 1 illustrates the spectral clustering-based image segmentation process. Input is the image that is to be segmented. Quad-tree decomposition divides the image into a hierarchical tree structure. The similarity matrix represents the similarity between different regions of the image. The degree matrix represents the number of connections between various image areas. Laplacian normalized graph, Eigenvectors, indicator matrix, and Generate Clustering: This step involves running a graph-based segmentation algorithm. The algorithm uses spectral clustering to choose the best way to partition the image. The output of the algorithm is a segmented image.

4.1.1 Proposed Algorithm

we proposed an algorithm known as Spectral Clustering-Based Image Segmentation (SCB-IS), which exploits the spectral clustering process for efficient image segmentation.

Algorithm: Spectral Clustering-Based Image Segmentation (SCB-IS)
Inputs:
Image I
Number of superpixels n
Number of clusters k
Output: Segmented image I'

1. Begin
2. Perform quad-tree decomposition to obtain superpixels $A = \{A_1, A_2, \dots, A_n\}$
3. Compute similarity matrix $\widetilde{W} \in \mathbb{R}^{m\times m}$ as in Eq. 11
4. Compute degree matrix $\widetilde{D}$ as in Eq. 12
5. Compute Laplacian normalized graph $\widetilde{L}_N = \widetilde{D}^{-\frac{1}{2}}(\widetilde{D} - \widetilde{W})\widetilde{D}^{-\frac{1}{2}}$.
6. Compute eigenvectors denoted as $t^1, t^2, \dots, t^k$
7. Compute indicator matrix $T = [t^1, t^2, \dots, t^k] \in R^{m\times k}$.
8. Compute $G = \widetilde{D}^{-\frac{1}{2}}T$
9. Transform G to G_p as in Eq. 17
10. Generate clusters using Fuzzy C-means
11. Output segmentation results
12. End

***Algorithm 1:** Spectral Clustering-Based Image Segmentation (Scb-Is)*

The SCB-IS algorithm is designed to segment images by clustering similar pixels. The process begins with a quad-tree decomposition to generate superpixels, which are then used to create a similarity matrix (as detailed in Equation 11) and a degree matrix (as per Equation 12). These matrices compute the normalized Laplacian graph, representing the image's structure. The algorithm then calculates the eigenvectors of the Laplacian matrix, denoted as V, and uses them to create an indicator matrix. This indicator matrix is transformed according to Equation 17, which likely involves a spectral embedding technique to convert the data into a space with fewer dimensions to make grouping more effective. Finally, the algorithm employs Fuzzy C-means clustering to generate the actual clusters, which are then used to output the segmented image I'. The SCB-IS algorithm is a sophisticated method that leverages spectral graph theory and fuzzy clustering to achieve precise segmentation of images; it finds utility in several applications, including satellite image processing, medical imaging, and object identification.

4.2 Deep Learning-Based Image Segmentation

The process for segmenting images using a deep learning-based algorithm is presented in this section. The SegNet model is recommended for picture segmentation despite several other deep learning models because of its efficient image segmentation method. A final layer for pixel-wise categorization, a related decoder network, and an encoder network make up SegNet, as shown in Figure 1. Analogous to the VGG16 network's initial thirteen convolutional layers, the encoder network consists of thirteen convolutional layers [1] for classifying objects. This allows us to start the training process with weights designed for classification on big datasets [58]. Alternatively, the higher-resolution feature maps at the deepest encoder output can be preserved while the completely connected layers are eliminated. Compared to other contemporary architectures, the SegNet encoder network has a notably smaller number of parameters [53], [54]. The decoder network has thirteen levels since every encoder layer has a corresponding decoder layer. The final decoder output is sent to a multi-class soft-max classifier, which generates class probabilities for every pixel.

A collection of feature maps is generated via convolution between each filter bank and encoder inside the encoder network. Batch normalization is applied after that [57], [58]). Next, an element-wise rectified linear non-linearity (ReLU) with a max(0, x) is implemented. Using a 2 × 2 window and stride 2 (non-overlapping window), a factor of two is applied to the max-pooling result. The minimal spatial changes in the input picture allow

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

for the translation invariance provided by max-pooling. An enormous input picture context is given to every subsampled feature map pixel. The feature maps still lose some spatial resolution, even though numerous layers of max-pooling and sub-sampling can provide more translation invariance for reliable classification.

The more lossy (border detail) visual depiction is not helpful in segmentation, where boundary delineation is essential. Therefore, boundary information must be captured and kept in the encoder feature maps before executing subsampling. After subsampling, all encoder feature maps can be stored if memory is not limited during inference. We suggest a more effective method of storing this data, as most real-world situations do not operate like this. For each encoder feature map, the maximum feature value locations are stored in memory for each pooling window; just the max-pooling indices must be retained. This is potentially significantly more affordable to store than remembering feature map(s) with float precision, as it only requires two bits for each 2 x 2 pooling window. Later in this study, as we will demonstrate, it is still suitable for real-world applications even if there is a slight loss of accuracy because of the smaller memory capacity.

The learned linked encoder feature map(s)' max-pooling indices are used by the corresponding decoder in the decoder network to upsample the input feature map(s). Step (s) of this procedure creates a sparse feature map. Dense feature maps are produced when a trainable decoder filter bank is applied to these feature maps. Subsequently, each of these maps undergoes a batch normalizing phase. The associated decoder produced a multi-channel feature map for the first encoder, which was closest to the input picture despite having three channels (RGB) in its encoder input. On the other hand, feature maps of the same size and quantity as the inputs from their matching encoders are produced by extra decoders in the network. The final decoder outputs the high-dimensional feature representation to a trainable soft-max classifier. With this soft-max technique, each pixel is categorized independently. When K is the number of classes, the soft-max classifier produces a channel picture of probabilities with K classes as its output. The expected segmentation for each pixel corresponds to the class with the highest probability. The final decoder outputs the high-dimensional feature representation to a trainable soft-max classifier. With this soft-max technique, each pixel is categorized independently. When K is the number of classes, the soft-max classifier produces a channel picture of probabilities with K classes as its output. The expected segmentation for each pixel corresponds to the class with the highest probability.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

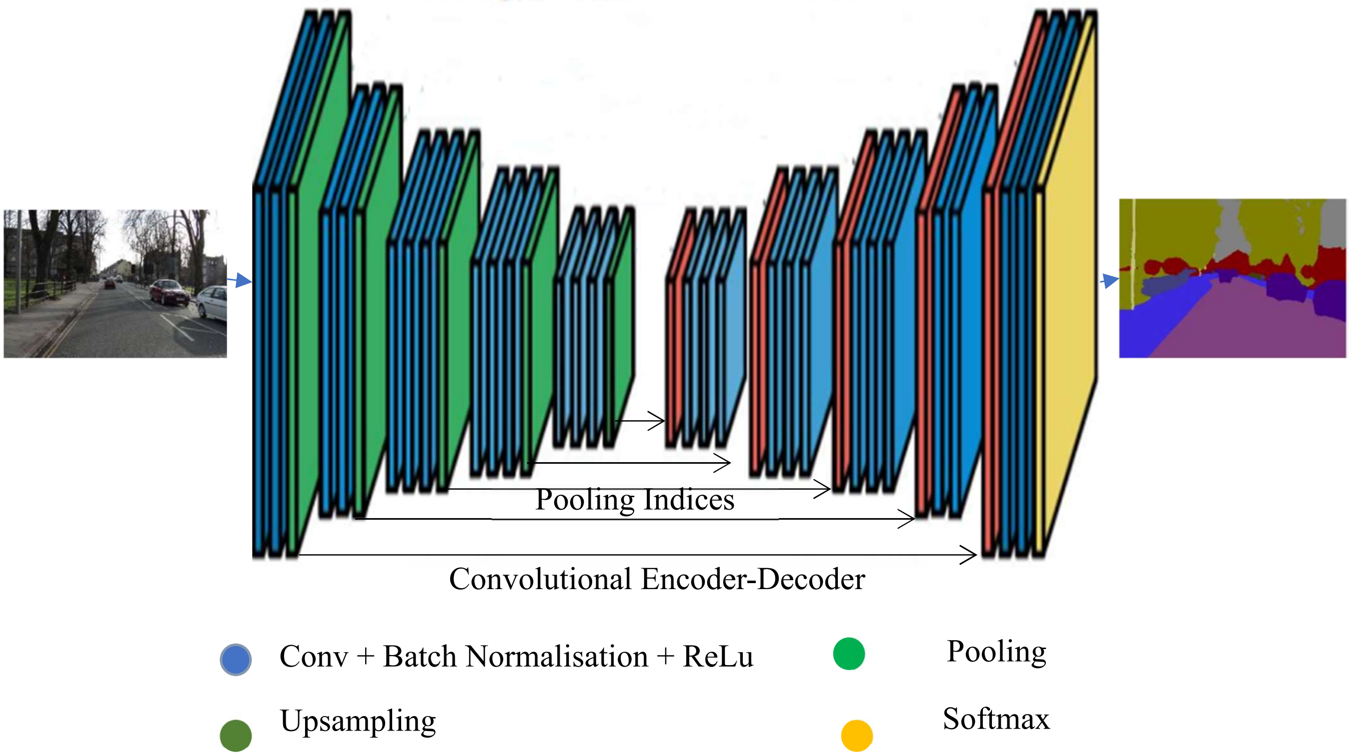


Figure 2: Overview Of Segnet Architecture Used For Image Segmentation

This RGB picture semantic segmentation technique uses a convolutional encoder-decoder architecture. The model's input is a typical RGB picture with three color channels (Red, Green, and Blue). Convolutional Encoder-Decoder Network has an encoder (on the left side of the network) and convolution + batch normalization + ReLU (blue layers). These layers perform convolution operations to extract features from the input image. Batch normalization helps stabilize and speed up the training, while the ReLU model becomes non-linear when the activation function is used. Pooling layers reduce the feature maps' spatial dimensions (width and height). This is typically done using max pooling; It increases the representations' sensitivity to tiny translations within the input picture. Upsampling layers expand the feature maps' spatial dimensionality. This process often uses transposed convolutions or interpolation methods to restore the original image size. Like the encoder, Convolution + Batch Normalization + ReLU layers further process the upsampled feature maps to improve the division. The last layer employs the softmax function to provide a probability distribution across the classes for every pixel in the input picture to aid the segmentation output. The decoder utilizes the encoder's pooling indices to guarantee that the upsampling procedure accurately conforms to the original spatial dimensions. The model's output is a segmented picture with a class name assigned to each pixel. Different colors in the segmented image represent different object classes (e.g., roads, cars, trees). The final output is a segmented image when each pixel is categorized into a predetermined class, enabling the identification of different objects and regions within the image. This Convolutional Encoder-Decoder architecture is widely used in various areas of semantic segmentation, including scene comprehension, medical picture analysis, and autonomous driving.

4.2.1 Training

Our baseline for the performance of the decoder versions is the CamVid road scenes dataset. This little collection of 360x480 RGB photos includes 367 training and 233 testing images of daytime and nighttime landscapes. Eleven classes—buildings, roads, vehicles, people, poles, signs, and sidewalks—must be broken apart to complete the work. Local contrast normalization [54] is used for the RGB input. The He et al. [57] method was initially used to set the encoder and decoder weights. We train all the versions using SGD using our Caffe implementation of SegNet-Basic [58], which has a fixed learning rate of 0.1 and momentum of 0.9 [59]. Up until the training loss converges, training is done on the variants.

The exercise package is randomized before each epoch to ensure that every image is utilized exactly once throughout every era. Each small batch of twelve photographs is then selected sequentially. By using a validation dataset, we choose the most effective model. For network training, As the goal function, we employ the cross-entropy loss [53]. The overall loss of a mini-batch is computed across all of its pixels. When the number of pixels in each class within the training set varies significantly, it becomes necessary to adjust the weighting of the loss according to the actual class (for instance, the CamVid dataset predominantly consists of pixels representing streets, sky, and buildings). The term "class balance" describes this. By dividing the class frequency by the ratio of the median of class frequencies computed across the whole training set, our approach—known as median frequency balancing [60]—determines the weight of a class in the loss function. This suggests the largest classes in the training set have weights more than 1, whereas the smaller classes have weights lower than 1. We also experimented with training the several versions with and without natural frequency balancing and without class balancing.

4.2.2 Proposed Algorithm

Our alternative algorithm, Deep Learning-Based Image Segmentation (DLB-IS), was suggested. This algorithm is based on the SegNet model, which is widely used for image segmentation in computer vision applications.

Algorithm: DLB-IS
Input: Image I
Output: Segmented image I'

1. Begin
2. Configure encoder
3. Configure decoder
4. Configure SegNet model m as in Figure 2
5. Compile m
6. features←EncoderExecution(I)
7. features'←EncoderOutcome(features) //compressed form of features
8. reconstructedFeatures←DecoderExecution(features')
9. I'←Segmentation(reconstructedFeatures, multi-class classifier)
10. Return I'
11. End

Algorithm 2: DLB-IS

Algorithm 2, DLB-IS, is a procedure used to divide a picture into smaller parts to simplify and transform the image's representation into something easier to understand and examine. The process generates a segmented image (I') as the output after receiving an input picture (I). The encoder is set up to process and extract features from the input picture. The encoder extracts characteristics, which the decoder reconstructs. As illustrated in Figure 2, a SegNet model is designed. This model is a DL architecture specifically tailored for picture division tasks. Model Compilation: The SegNet model (m) is compiled, typically defining the loss function, the optimizer, and other parameters necessary for the training process. The encoder processes the input image (I) to produce a set of features. After that, the encoder processes the extracted features to make a compressed version of the features. The decoder receives the compressed features and uses them to recreate the features. A multi-class classifier is then used to segment the rebuilt features. After the algorithm runs, the segmented picture (I') is returned. The algorithm is structured to leverage the power of deep learning to segment images accurately, an essential function in many applications, including autonomous driving, object identification, and medical imaging. This work is especially well-suited for the SegNet model, a convolutional neural network, since it can capture the spatial and contextual information of the image, Facilitating accurate separation of distinct items or areas within the picture.

5. EXPERIMENTAL RESULTS

This part shows the two techniques used in this paper's experiment findings. The results include observations of image segmentation using a semantic clustering approach and a deep learning approach based on the SegNet model.

5.1 Results of Fast Spectral Clustering

This subsection presents the results of the cluster spectral clustering approach.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

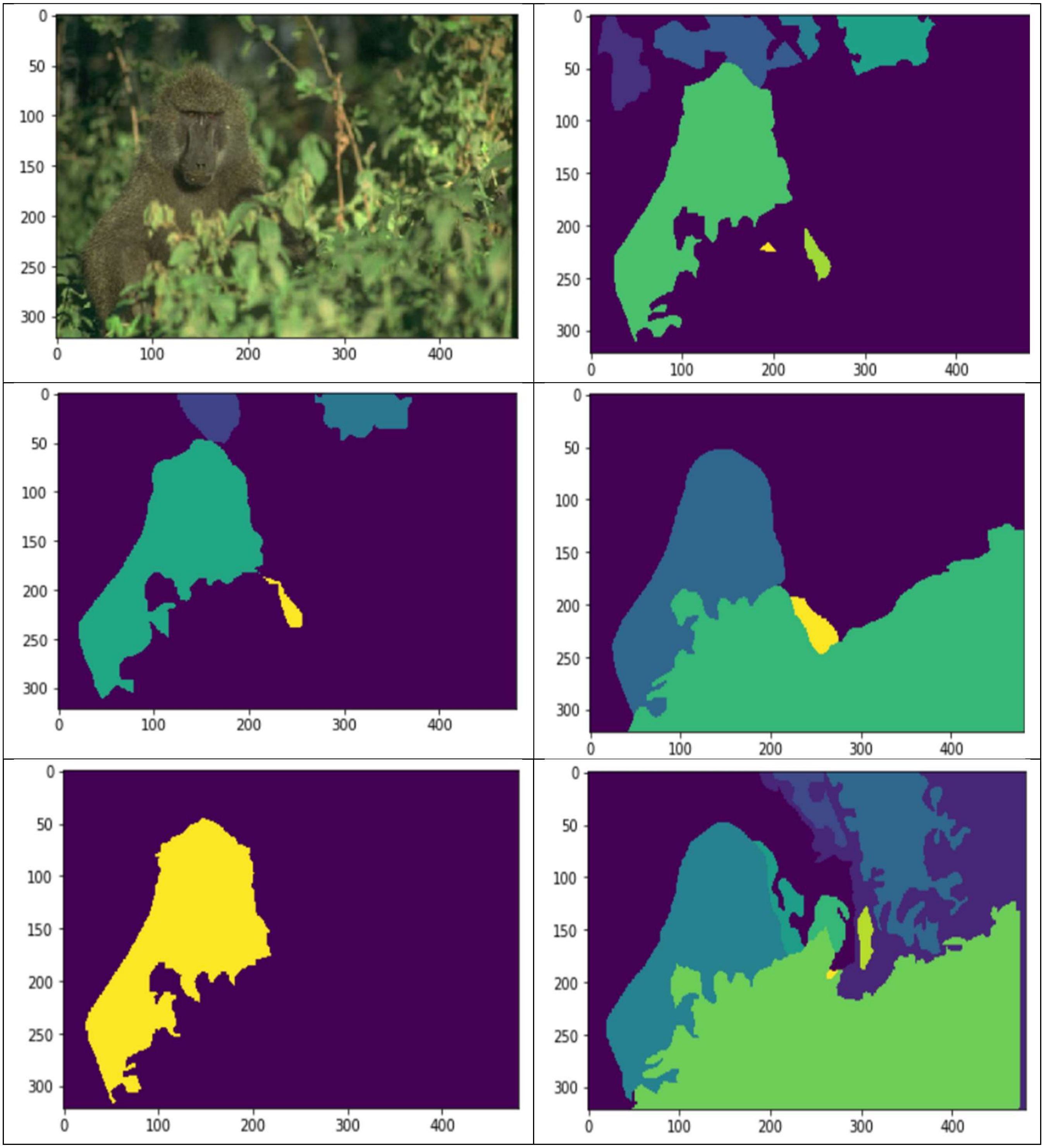


Figure 3: Results of ground truth segmentation

Journal of Theoretical and Applied Information Technology
15th March 2025. Vol.103. No.5

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

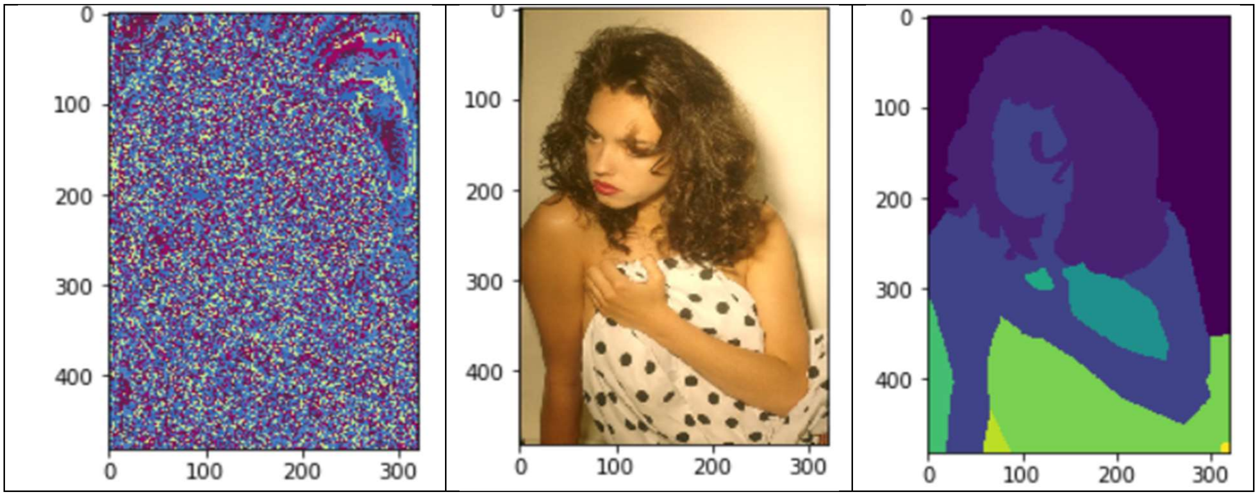


Figure 4: Results of spectral clustering with the full-size images (481x321) using the Annoy library

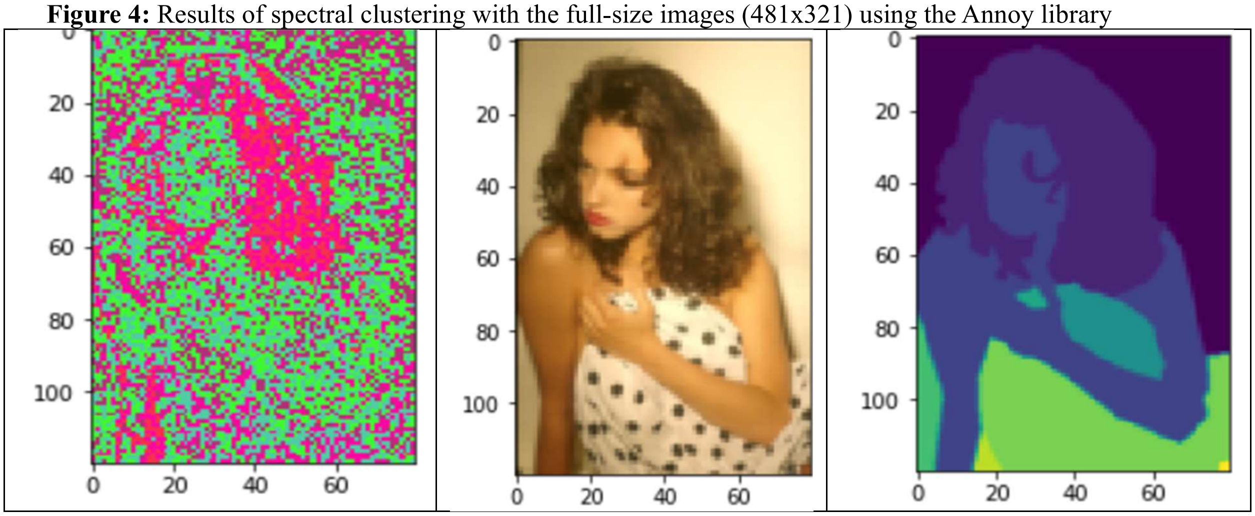


Figure 5: Results of spectral clustering with the resized images

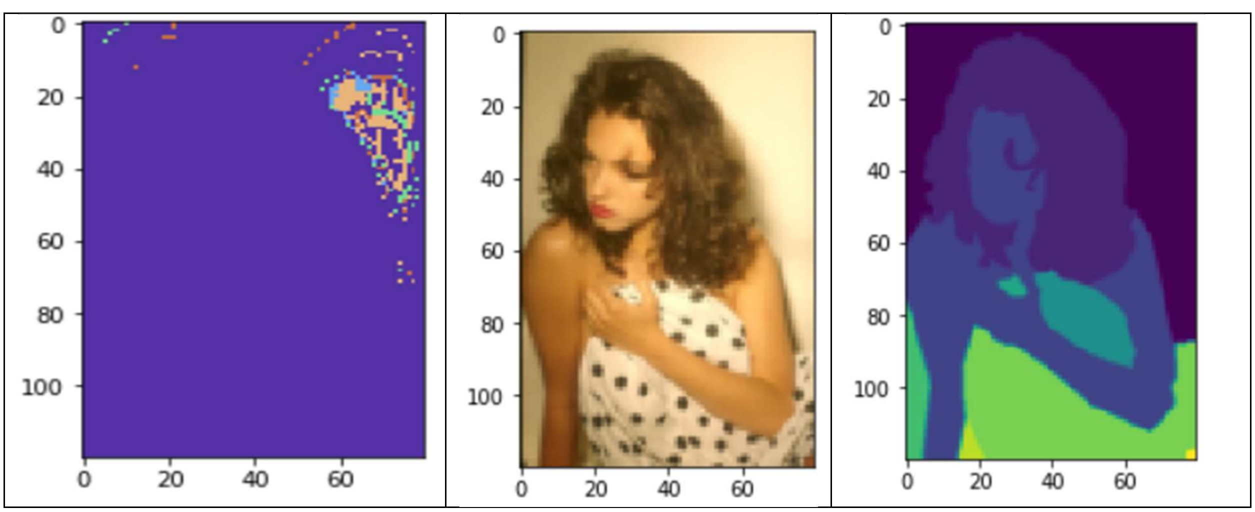


Figure 6: Results of spectral clustering with the resized images using SKlearn's implementation

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

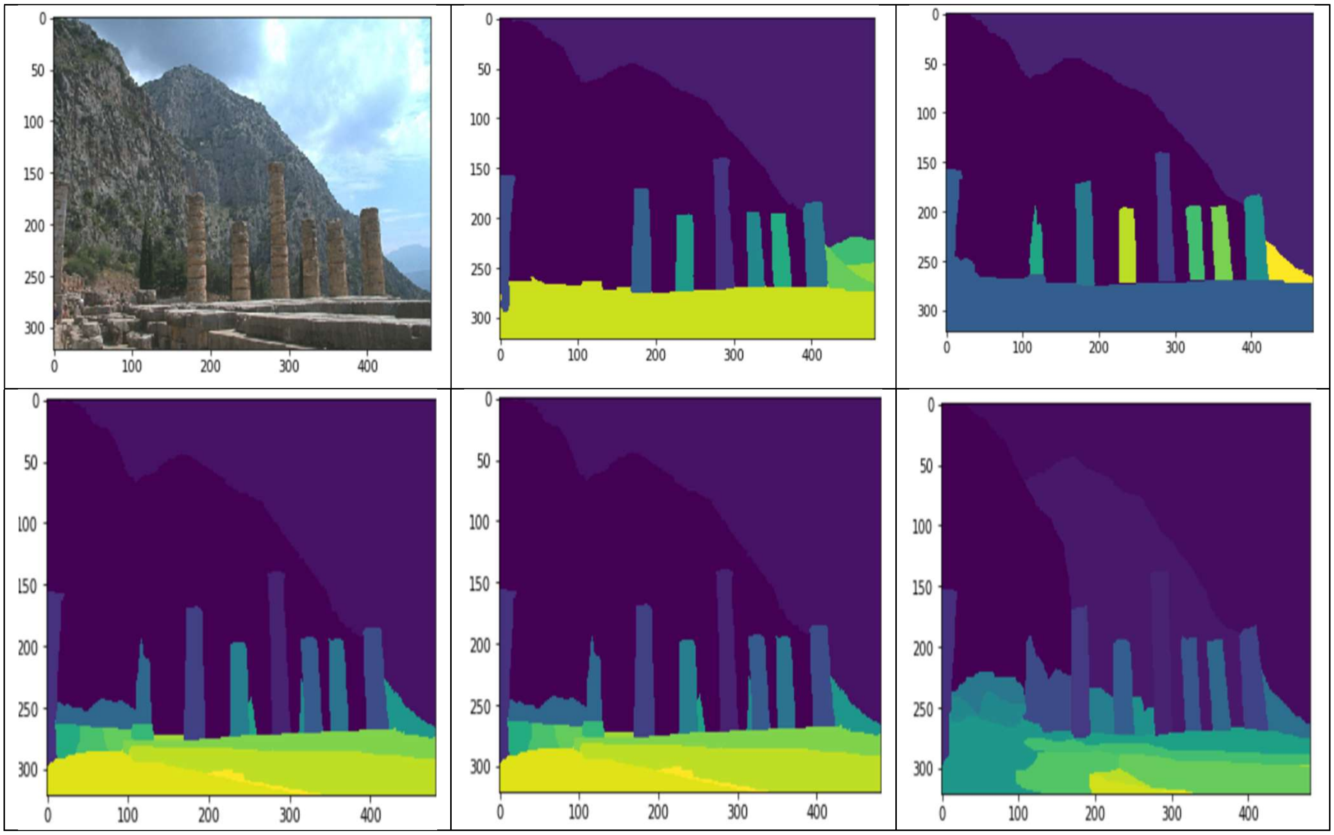


Figure 7: Image Visualization And Ground Truth Generation

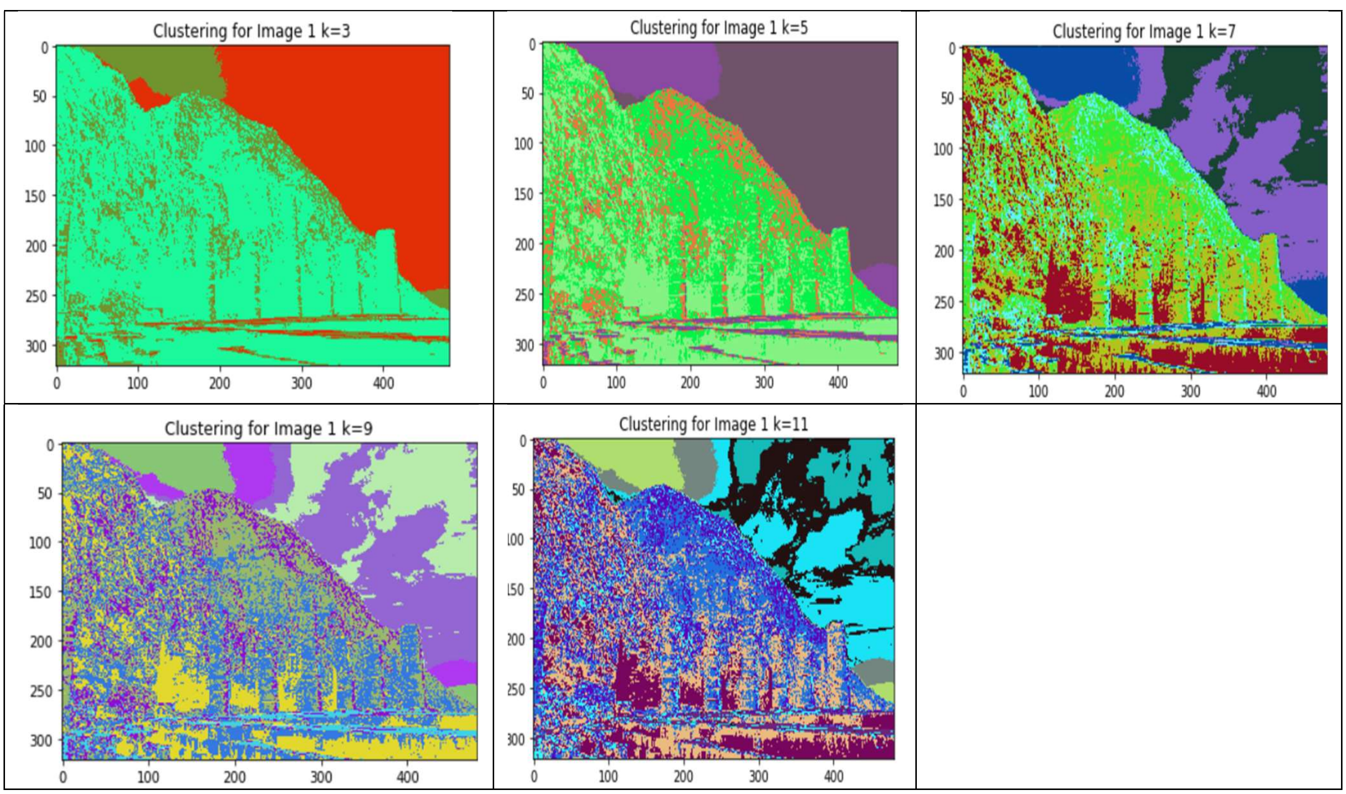


Figure 8: Results Of K-Means-Based Segmentation

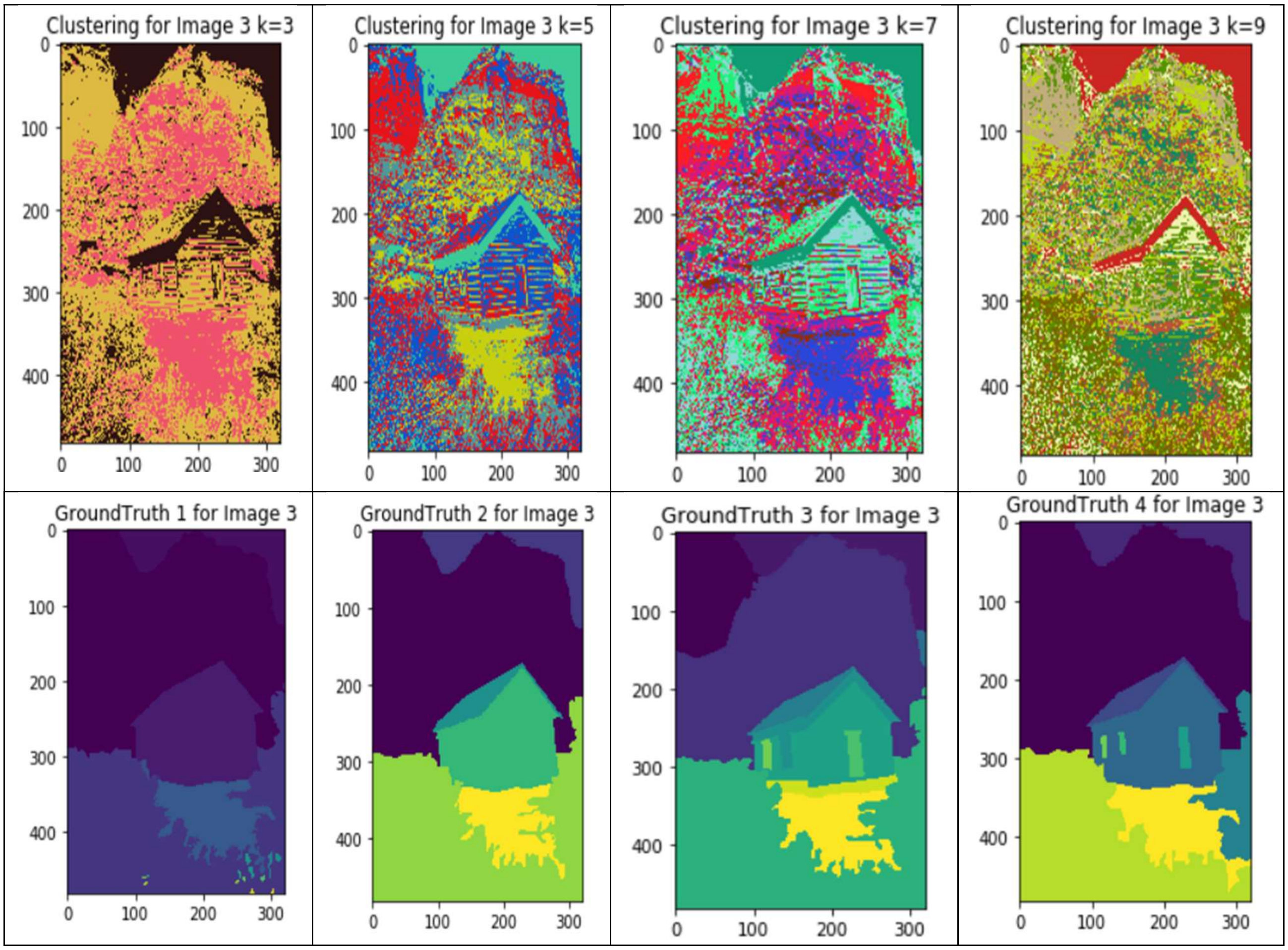
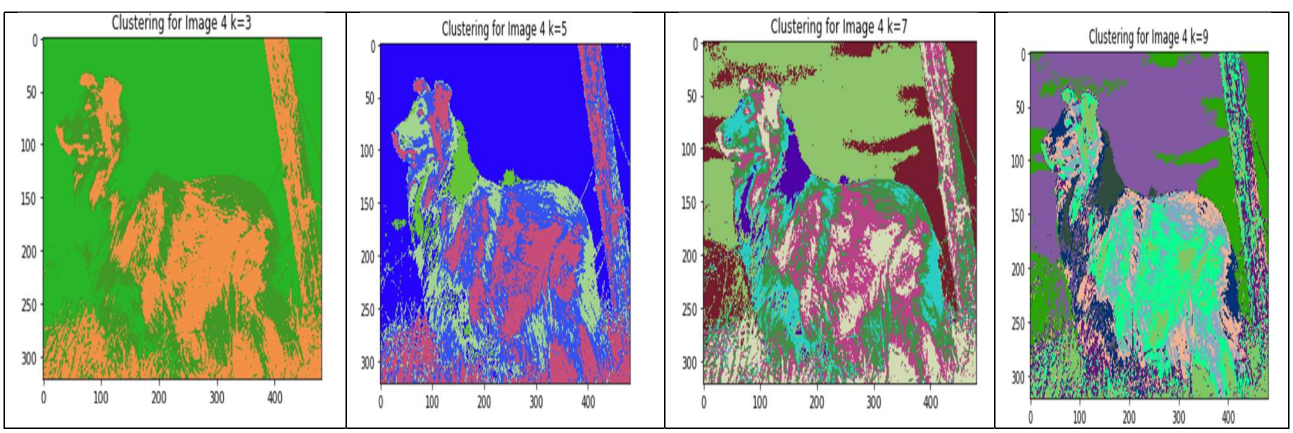


Figure 9: *Results of K-means-based segmentation*



ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

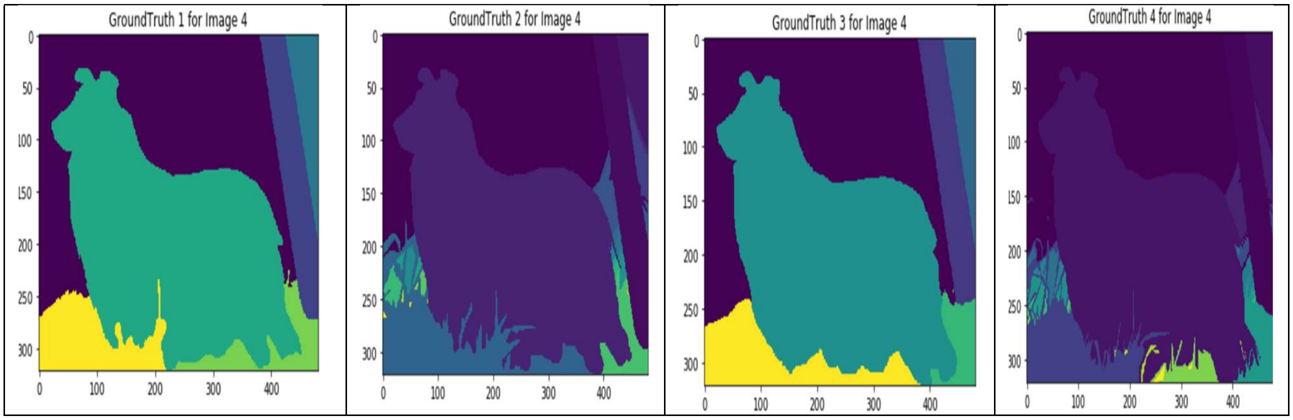


Figure 10: Results Of K-Means-Based Segmentation

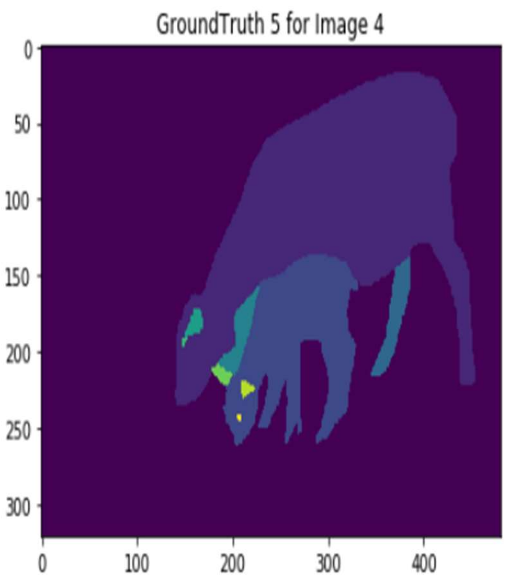


Figure 11: Ground Truth 5 For The Given Image

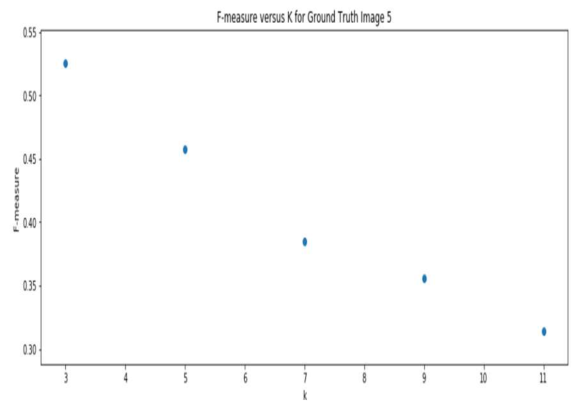


Figure 12: F-Measure And K-Value Analysis For Ground Truth Image 5

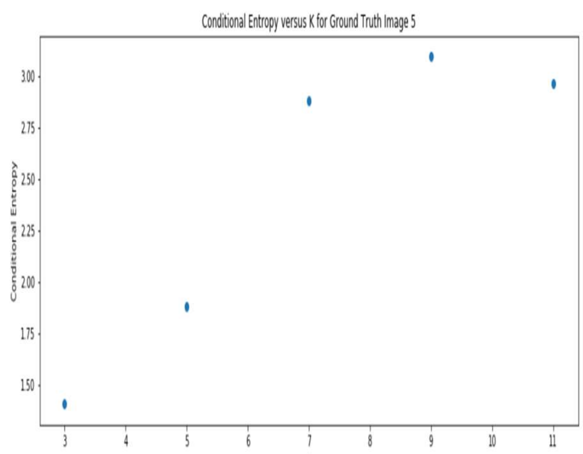


Figure 13: Conditional Entropy And K-Value Analysis For Ground Truth Image 5

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

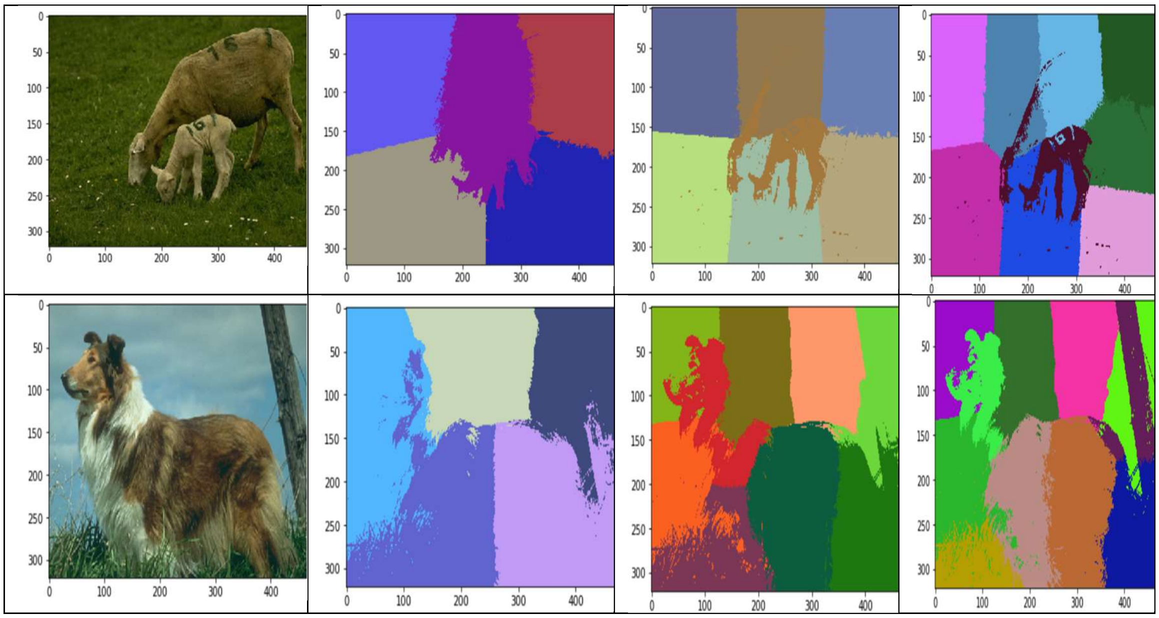


Figure 14: Results Of K-Means Clustering With Spatial Features

5.2 Results of SegNet-Based Segmentation

This segment shows the outcomes of our deep-learning method for picture segmentation. SegNet is the DL model used for picture segmentation.

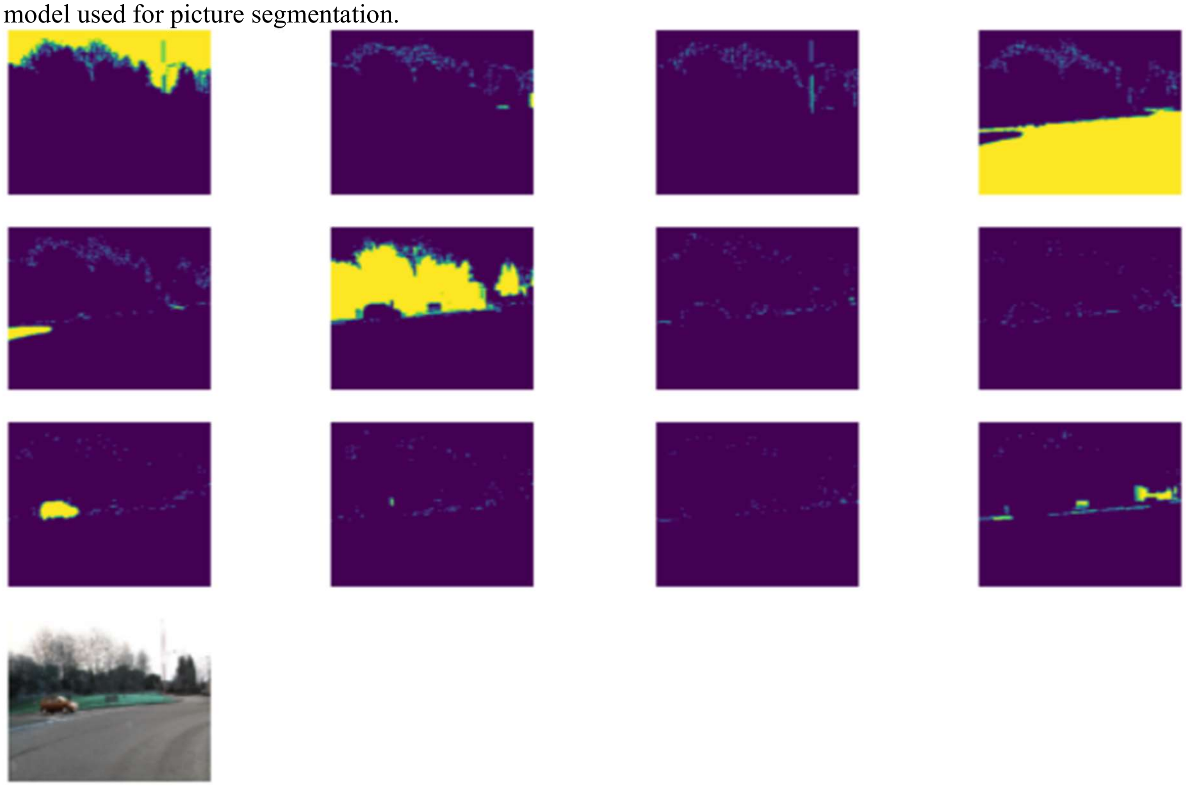


Figure 15: Results Of Segmentation For A Given Image

Journal of Theoretical and Applied Information Technology
15th March 2025. Vol.103. No.5

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

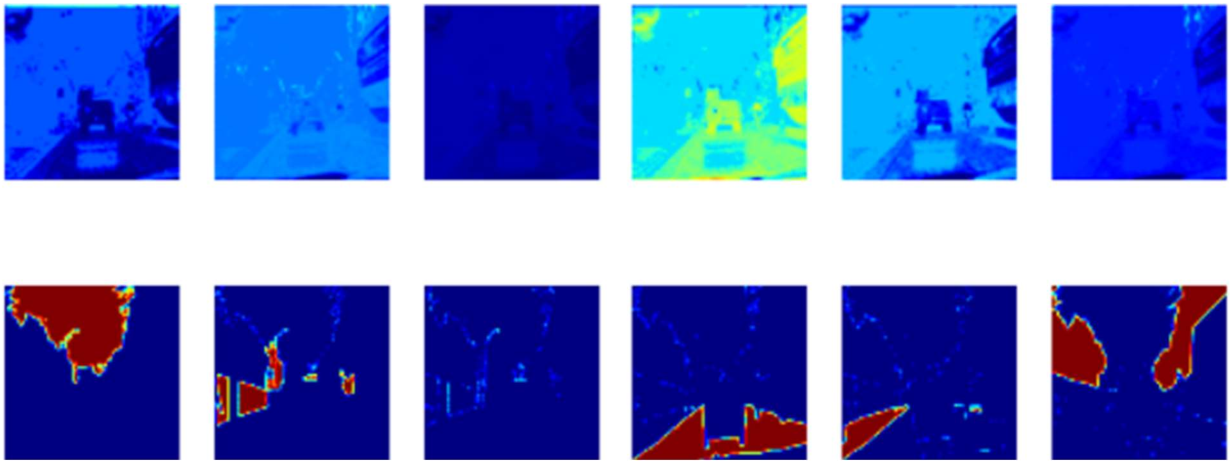


Figure 16: Intermediate Results Of Segmentation

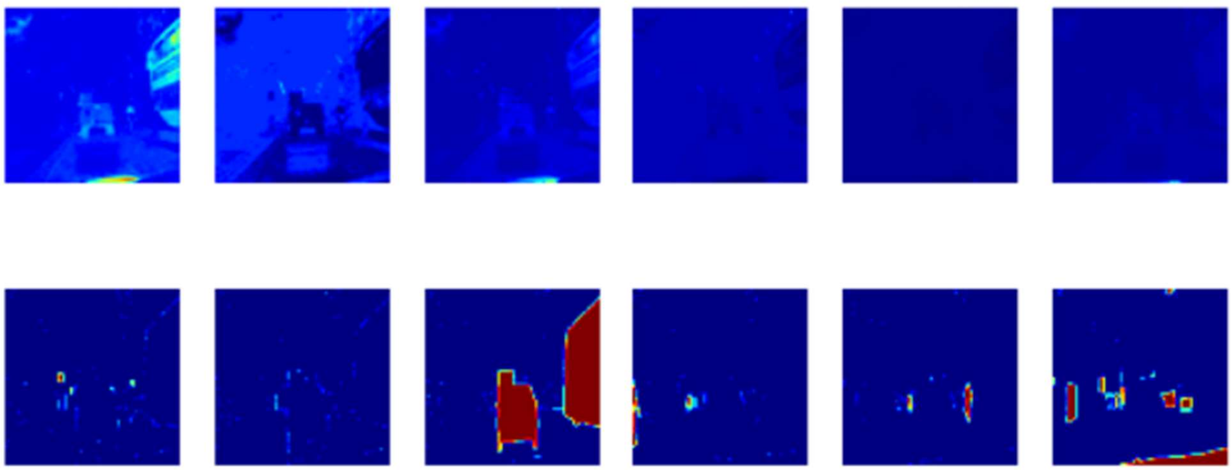


Figure 17: Intermediate Results Of Segmentation

Figure 18: Input Image For Segmentation

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 19: Random Sampling

Figure 20: Predicted Random Output

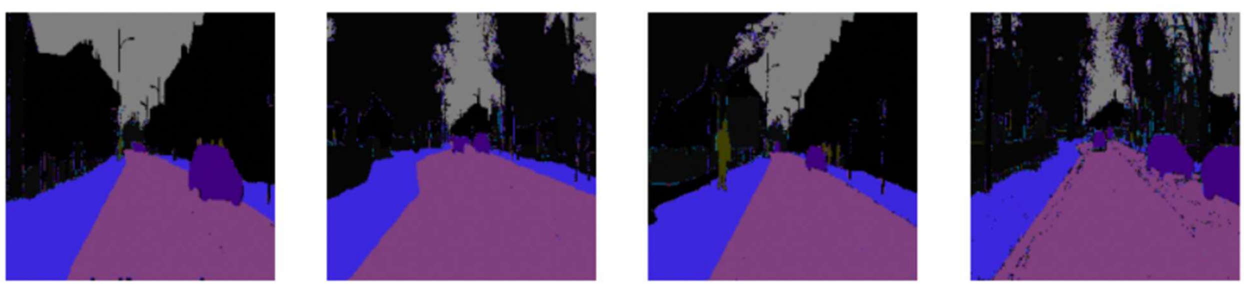


Figure 21: Results of segmentation

5.3 Comparative Analysis of Image Segmentation Models

This section compares different image segmentation models. The SegNet model's results in this paper's empirical study are compared with those of numerous other models.

Metric	Precision	Recall	F1 Score	IoU (Mean)	Accuracy
U-Net	81	78	79.4	63.5	85.3
FCN	83.5	80.8	82.1	66.2	87
DeepLab	80.2	77.5	78.8	62.3	84.8
SegNet	85.7	83	84.3	69.4	88.5

Table 1: Performance Comparison Of DL Models Used For Segmentation

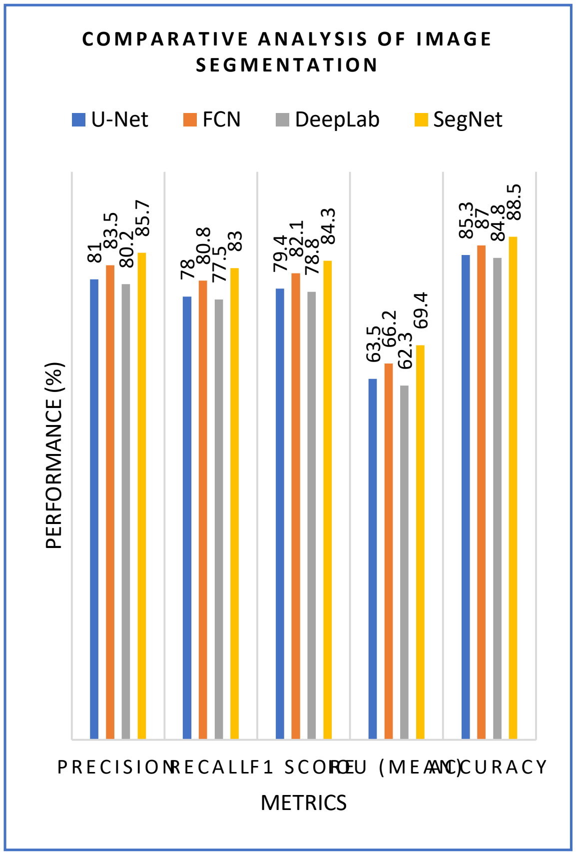


Figure 22: Performance Comparison Of DL Models

Figure 22 shows the output of four distinct picture segmentation models (U-Net, FCN, DeepLab, SegNet) across five measures of performance: Precision, Recall, F1 Score, IOU (Mean), and Accuracy. Precision performance includes U-Net: 81%, FCN: 83.5%, DeepLab: 83.2% and SegNet: 85.7%. Recall of the models includes U-Net: 78%, FCN: 80.8%, DeepLab: 77.5%, and SegNet: 83%. F1 scores of the models are U-Net: 79.4%, FCN: 82.1%, DeepLab: 78.8%, and SegNet: 84.3%. IOU performance of the models is U-Net: 63.5%, FCN: 66.2%, DeepLab: 62.3% and SegNet: 69.4%. The accuracy of the models is U-Net: 85.3%, FCN: 87%, DeepLab: 84.8%, and SegNet: 88.5%. SegNet performs the best across all metrics, achieving the highest precision (85.7%), recall (83%), F1 score (84.3%), IOU (69.4%), and accuracy (88.5%). FCN generally performs well, ranking second in most metrics. U-Net and DeepLab are competitive but do not consistently surpass SegNet and FCN. Results reveal that SegNet outperforms the other models across all evaluated metrics, making it the most reliable choice for image segmentation tasks among the compared models.

6. CONCLUSION AND FUTURE WORK

Our article presented the Spectral Clustering-Based Image Segmentation (SCB-IS) algorithm, which efficiently leverages spectral clustering to segment images. Additionally, we offered a deep learning-based segmentation technique using the widely employed SegNet. This method is known as Deep Learning-Based Image Segmentation (DLB-IS). Our empirical study using benchmark datasets found that many cutting-edge models are outperformed by our deep learning-based system, achieving the highest accuracy of 88.50%. The novelty of this work lies in integrating spectral clustering and deep learning approaches to achieve superior image segmentation accuracy. Our research demonstrates that the proposed Deep Learning-Based Image Segmentation (DLB-IS) method using SegNet outperforms existing models, making it a viable solution for medical imaging, object recognition, and autonomous systems, where precise segmentation is crucial. These learning-based approaches can be integrated into real-world computer vision applications to segment images efficiently. In our future work, features extracted from images aided by ML and DL approaches can be combined to make it a more efficient approach to image segmentation. Another important direction for our research work is to exploit such features in deep learning-based methodology towards multi-object detection and classification. In the future, it is also possible to exploit deep learning models and object detection approaches to recognize and describe a scene.

REFERENCES

[1] PRADEEP KUMAR DAS, DIYA V A, SUKADEV MEHER, RUTUPARNA PANDA AND AJITH ABRAHAM. (2022). A Systematic Review on Recent Advancements in Deep and Machine Learning Based Detection and Classification of Acute Lymphoblastic Leukemia. *IEEE*. 10, pp.81741 - 81763. http://DOI:10.1109/ACCESS.2022.3196037

[2] Valli Kumari Vatsavayi & Nagaraju Andavarapu. (2022). Identification and classification of wild animals from video sequences using hybrid deep residual convolutional neural network. *Springer*., pp.1-26. https://doi.org/10.1007/s11042-022-12852-w

[3] Mohammad Zolfaghari and Hedieh Sajedi. (2022). A survey on automated detection and classification of acute leukemia and WBCs in microscopic blood cells. *Springer*, pp.1-31.

[4] Shishir Kumar Shandilya, Aditya Srivastav, Kyrylo Yemets, Agni Datta, Atulya K. Nagar. (2023). YOLO-based segmented drone vs. bird detection dataset for deep and machine learning algorithms. *Elsevier*. 50, pp.1-10. https://doi.org/10.1016/j.dib.2023.109355

[5] SHAKHBOZ ABDIGAPPOROV, SHOKHRUKH MIRALIEV, VIJAY KAKANI and HAKIL KIM. (2023). Joint Multiclass Object Detection and Semantic Segmentation for Autonomous Driving. *IEEE*. 11, pp.37637 - 37649. http://DOI:10.1109/ACCESS.2023.3266284

[6] Nabin Kumar Naik, Prabira Kumar Sethy, Santi Kumari Behera and Rajat Amat. (2024). A methodical analysis of deep learning techniques for detecting Indian lentils. *Elsevier*. 15, pp.1-8. https://doi.org/10.1016/j.jafr.2023.100943

[7] Yichi Zhang, Zhenrong Shen and Rushi Jiao. (2024). Segment anything model for medical image segmentation: Current applications and future directions. *Elsevier*. 171, pp.1-15. https://doi.org/10.1016/j.compbiomed.2024.108238

[8] Hongxu Jiang, Muhammad Imran, Preethika Muralidharan, Anjali Patel, Jake Pensa, Muxuan Liang, Tarik Benidir, Joseph R. Grajo, Jason P. Joseph, Russell Terry, John Michael DiBianco, Li-Ming Su, Yuyin Zhou, Wayne G. Brisbane and Wei Shao. (2024). MicroSegNet: A deep learning approach for prostate segmentation on micro-ultrasound images. *Elsevier*. 112, pp.1-10. https://doi.org/10.1016/j.compmedimag.2024.102326

[9] Stefan Herdy, Emilio Rodríguez-Caballero, Thomas Pock and Bettina Weber. (2024). Utilization of deep learning tools to map and monitor biological soil crusts. *Elsevier*. 79, pp.1-16. https://doi.org/10.1016/j.ecoinf.2023.102417

[10] Jun Wei Roy Chong, Kuan Shiong Khoo, Kit Wayne Chew, Huong-Yong Ting, Iwamoto Koji, Roger Ruan, Zengling Ma and Pau Loke Show. (2024). Artificial intelligence-driven microalgae autotrophic batch cultivation: A comparative study of machine and deep learning based image classification models. *Elsevier*, pp.1-64. https://doi.org/10.1016/j.algal.2024.103400

[11] Sang-Yum Lee, Tri Ho Minh Le and Yeong-Min Kim. (2023). Prediction and detection of potholes in urban roads: Machine learning and deep learning based image segmentation approaches. *Elsevier*. 13, pp.1-13. https://doi.org/10.1016/j.dibe.2022.100109

[12] Zifei Luo, Wenzhu Yang, Yunfeng Yuan, Ruru Gou and Xiaonan Li. (2023). Semantic segmentation of agricultural images: A survey. *Elsevier*, pp.1-15. https://doi.org/10.1016/j.inpa.2023.02.001

[13] Md Hosne Mobarak, Mariam Akter Mimona, Md. Aminul Islam, Nayem Hossain, Fatema Tuz Zohura, Ibnul Imtiaz and Md Israfil Hossain Rimon. (2023). Scope of machine learning in materials research—A review. *Elsevier*. 18, pp.1-24. https://doi.org/10.1016/j.apsadv.2023.100523

[14] Mostafiz Ahammed, Md. Al Mamun and Mohammad Shorif Uddin. (2022). A machine learning approach for skin disease detection and classification using image segmentation. *Elsevier*, pp.1-15. https://doi.org/10.1016/j.health.2022.100122

[15] Tsimur Davydzenka, Daniel Sinclair, Nikhilesh Chawla and Pejman Tahmasebi. (2022). Deep-layers-assisted machine learning for accurate image segmentation of complex materials. *Elsevier*. 192, pp.1-33. https://doi.org/10.1016/j.matchar.2022.112175

[16] Dimitrios Loverdos and Vasilis Sarhosis. (2022). Automatic image-based brick segmentation and crack detection of masonry walls using machine learning. *Elsevier*. 140, pp.1-15. https://doi.org/10.1016/j.autcon.2022.104389

[17] Simrandeep Singh; Nitin Mittal; Diksha Thakur; Harbinder Singh; Diego Oliva and Anton Demin; (2021). *Nature and Biologically Inspired Image Segmentation Techniques . Archives of Computational Methods in Engineering.* http://doi:10.1007/s11831-021-09619-1

[18] Hamidreza Torbati-Sarraf; Sridhar Niverty; Rajhans Singh; Daniel Barboza; Vincent De Andrade; Pavan Turaga and Nikhilesh Chawla; (2021). *Machine-Learning-based Algorithms for Automated Image Segmentation Techniques of Transmission X-ray Microscopy (TXM) . JOM.* http://doi:10.1007/s11837-021-04706-x

[19] Yue Hou; Qiuhan Li; Chen Zhang; Guoyang Lu; Zhoujing Ye; Yihan Chen; Linbing Wang and Dandan Cao; (2021). The State-of-the-Art Review on Applications of Intrusive Sensing, Image Processing Techniques, and Machine Learning Methods in Pavement Monitoring and Analysis . Engineering. http://doi:10.1016/j.eng.2020.07.030

[20] Sanika Walvekar and Swati Shinde; (2021). Efficient Medical Image Segmentation Of COVID-19 Chest CT Images Based on Deep Learning Techniques . 2021 International Conference on Emerging Smart Computing and Informatics (ESCI). http://doi:10.1109/esci50559.2021.9397043

[21] Shervin Minaee; Yuri Y. Boykov; Fatih Porikli; Antonio J Plaza; Nasser Kehtarnavaz and Demetri Terzopoulos; (2021). Image Segmentation Using Deep Learning: A Survey . IEEE Transactions on Pattern Analysis and Machine Intelligence. http://doi:10.1109/tpami.2021.3059968

[22] Rizwan I Haque, Intisar and Neubert, Jeremiah (2020). Deep learning approaches to biomedical image segmentation. Informatics in Medicine Unlocked, 100297–. http://doi:10.1016/j.imu.2020.100297

[23] Hesamian, Mohammad Hesam; Jia, Wenjing; He, Xiangjian and Kennedy, Paul (2019). Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. Journal of Digital Imaging. http://doi:10.1007/s10278-019-00227-x

[24] Garcia-Garcia, Alberto; Orts-Escolano, Sergio; Oprea, Sergiu; Villena-Martinez, Victor; Martinez-Gonzalez, Pablo and Garcia-Rodriguez, Jose (2018). A Survey on Deep Learning Techniques for Image and Video Semantic Segmentation. Applied Soft Computing, S1568494618302813–. http://doi:10.1016/j.asoc.2018.05.018

[25] Guo, Zhe; Li, Xiang; Huang, Heng; Guo, Ning and Li, Quanzheng (2019). Deep Learning-based Image Segmentation on Multi-modal Medical Imaging. IEEE Transactions on Radiation and Plasma Medical Sciences, 1–1. http://doi:10.1109/TRPMS.2018.2890359

[26] Pop, Mihaela; Sermesant, Maxime; Jodoin, Pierre-Marc; Lalande, Alain; Zhuang, Xiahai; Yang, Guang; Young, Alistair and Bernard, Olivier (2018). Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges Volume 10663 || An Exploration of 2D and 3D Deep Learning Techniques for Cardiac MR Image Segmentation. , 10.1007/978-3-319-75541-0(Chapter 12), 111–119. http://doi:10.1007/978-3-319-75541-0_12

[27] Liu, Xiaolong; Deng, Zhidong and Yang, Yuhan (2018). Recent progress in semantic image segmentation. Artificial Intelligence Review. http://doi:10.1007/s10462-018-9641-3

[28] Lateef, Fahad and Ruichek, Yassine (2019). Survey on Semantic Segmentation using Deep Learning Techniques. Neurocomputing, S092523121930181X–. http://doi:10.1016/j.neucom.2019.02.003

[29] Wang, Guotai; Li, Wenqi; Zuluaga, Maria A.; Pratt, Rosalind; Patel, Premal A.; Aertsen, Michael; Doel, Tom; David, Anna L.; Deprest, Jan; Ourselin, Sebastien and Vercauteren, Tom (2018). Interactive Medical Image Segmentation using Deep Learning with Image-specific Fine-tuning. IEEE Transactions on Medical Imaging, 1–1. http://doi:10.1109/TMI.2018.2791721

[30] Tajbakhsh, Nima; Jeyaseelan, Laura; Li, Qian; Chiang, Jeffrey N.; Wu, Zhihao and Ding, Xiaowei (2020). Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. Medical Image Analysis, 63, 101693–. http://doi:10.1016/j.media.2020.101693

[31] Huang, Hong; Meng, Fanzhi; Zhou, Shaohua; Jiang, Feng and Manogaran, Gunasekaran (2019). Brain image segmentation based on FCM clustering algorithm and rough set. IEEE Access, 1–1. http://doi:10.1109/ACCESS.2019.2893063

[32] Kim, Wonjik; Kanezaki, Asako and Tanaka, Masayuki (2020). Unsupervised Learning of Image Segmentation Based on Differentiable Feature Clustering. IEEE Transactions on Image Processing, 29, 8055–8068. http://doi:10.1109/TIP.2020.3011269

[33] Cuong Manh Hoang and Byeongkeun Kang. (2023). Pixel-level clustering network for unsupervised image segmentation. *Elsevier*. 127, pp.1-13. https://doi.org/10.1016/j.engappai.2023.107327

[34] Branislav Panic´, Matej Borovinšek, Matej Vesenjak, Simon Oman and Marko Nagode. (2024). A guide to unsupervised image segmentation of mCT-scanned cellular metals with mixture modelling and Markov random fields. *Elsevier*. 239, pp.1-13. https://doi.org/10.1016/j.matdes.2024.112750

[35] Juan M. Muñoz-Ocaña, Justo Puerto and Antonio M. Rodríguez-Chía. (2024). Improved heuristics for solving large-scale Scanning Transmission Electron Microscopy image segmentation using the ordered median problem. *Elsevier*. 163, pp.1-17. https://doi.org/10.1016/j.cor.2023.106524

[36] Harsh Yadav, Jaspreeti Singh and Anjana Gosain. (2023). Experimental Analysis of Fuzzy Clustering Techniques for Outlier Detection. *Elsevier*. 218, pp.959-968. https://doi.org/10.1016/j.procs.2023.01.076

[37] Cheng-Jui Tseng and Changjiang Tang. (2023). An optimized XGBoost technique for accurate brain tumor detection using feature selection and image segmentation. *Elsevier*. 4, p. 1-9. https://doi.org/10.1016/j.health.2023.100217

[38] Amin Golzari Oskouei and Mahdi Hashemzadeh. (2022). CGFFCM: A color image segmentation method based on cluster-weight and feature-weight learning. *Elsevier*. 11, pp.1-5. https://doi.org/10.1016/j.simpa.2022.100228

[39] Himanshu Mittal; Avinash Chandra Pandey; Mukesh Saraswat; Sumit Kumar; Raju Pal and Garv Modwel; (2021). A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets. Multimedia Tools and Applications. http://doi:10.1007/s11042-021-10594-9

[40] Qureshi, Mohammad Naved, and Ahamad, Mohd Vasim (2018). An Improved Method for Image Segmentation Using K-Means Clustering with Neutrosophic Logic. Procedia Computer Science, 132, 534–540. http://doi:10.1016/j.procs.2018.05.006

[41] BSDS500 - The Berkeley Segmentation Dataset and Benchmark. Retrieved from https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/

[42] Saxena A, Prasad M, Gupta A, Bharill N, Patel OP, Tiwari A, Er MJ, Ding W, Lin CT (2017) A review of clustering techniques and developments. Neurocomputing 267:664–681.

[43] MacQueen J et al (1967) Some methods for classification and analysis of multivariate observations. In: Proc. of berkeley symposium on mathematical statistics and probability. USA pp 281–297.

[44] Langham AE, Grant P (1999) Using competing ant colonies to solve k-way partitioning problems with foraging and raiding strategies. In: Proc. of springer european conference on artificial life. Switzerland, pp 621–625.

[45] Steinbach M, Karypis G, Kumar V et al (2000) A comparison of document clustering techniques. In: Proc. of ACM international conference on knowledge discovery and data mining workshop on text mining. USA, pp 1–20.

[46] Phillips SJ (2002) Acceleration of k-means and related clustering algorithms. In: Lecture notes of springer workshop on algorithm engineering and experimentation. USA, pp 166–177.

[47] Zhang B, Hsu M, Dayal U (2001) K-harmonic means-a spatial clustering algorithm with boosting. In: Lecture notes of springer workshop on temporal, spatial, and spatio-temporal data mining. France, pp 31–45.

[48] Chaturvedi A, Green PE, Caroll JD (2001) K-modes clustering. J Classif 18:35–55.

[49] Park HS, Jun CH (2009) A simple and fast algorithm for k-medoids clustering. Expert Syst Appl 36:3336–3341.

[50] Ng RT, Han J (2002) Clarans: A method for clustering objects for spatial data mining. IEEE Trans Knowl Data Eng 5:1003–1016.

[51] Lutkepohl, Helmut."Handbook of matrices." Computational statistics and Data analysis, 2.25, pp. 243, 1997.

[52] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," International

Journal of Computer Vision (IJCV), pp. 1–42, April 2015.
[53] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in CVPR, pp. 3431–3440, 2015.
[54] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," in ICCV, pp. 1520–1528, 2015.
[55] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," CoRR, vol. abs/1502.03167, 2015.
[56] V. Badrinarayanan, B. Mishra, and R. Cipolla, "Understanding symmetries in deep networks,".
[57] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in ICCV, pp. 1026–1034, 2015.
[58] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in Proceedings of the 22nd ACM international conference on Multimedia, pp. 675–678, ACM, 2014.
[59] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in Proceedings of COMPSTAT'2010, pp. 177–186, Springer, 2010.
[60] D. Eigen and R. Fergus, "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture," in ICCV, pp. 2650–2658, 2015.
[61] Srikanth, G., Raghavendran, C.V., Prabhu, M.R., Radha, M., Kumari, N.V.S. & Francis, S.K., 2025. Climate Change Impact on Geographical Region and Healthcare Analysis Using Deep Learning Algorithms. Remote Sensing in Earth Systems Sciences, 130359.
[62] Ravikumar Ch1, Marepalli Radha2, Maragoni Mahendar3, Pinnapureddy Manasa"A comparative analysis for deep-learning-based approaches for image forgery detection International journalof Systematic Innovation https://doi.org/10.6977/IJoSI.202403_8(1).0001
[63] Anitha Patil. (2019). Distributed Programming Frameworks in Cloud Platforms. International Journal of Recent Technology and Engineering (IJRTE). 7(6), pp.611-619.
[64] A. Patil and S. Govindaraj, "An AI Enabled Framework for MRI-based Data Analytics for Efficient Brain Stroke Detection," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ACCAI58221.2023.10201136.
[65] Sreedhar Bhukya, A Novel Methodology for Secure De duplication of Imagedata in Cloud Computing using Compressive Sensing and Random Pixel Exchanging, Journal of Theoretical and Applied Information Technology (JATIT), Vol.102. No 4, ISSN: 1992-8645 (2024), SCOPUS.
[66] Sreedhar Bhukya, Multiclass Supervised Learning Approach for SAR-COV2 Severity and Scope Prediction: SC2SSP Framework, Volume (12) issue (1), 2025-01-31, SCOPUS.
[67] K. Swathi and Dr. G. Narsimha. (2024). Cyber Attacks and Anomaly Detection in Networking Based on Deep Learning- A survey. Springer., pp.11-20.
[68] K. Swathi, and G. Narsimha. (2023). Analysis of Deep Learning Approaches for Intelligent Intrusion Detection System. IEEE., pp.1-9. DOI:10.1109/ICAEECI58247.2023.10370800
[69] K. Swathi, and G. Narsimha. (2023). Robust Deep Learning Based Framework for Detecting Cyber Attacks from Abnormal Network Traffic. IJRITCC. 11(7), pp.341-353. https://doi.org/10.17762/ijritcc.v11i7.7958