DOI: 10.21917/ijivp.2024.0464

# SEMANTIC SEGMENTATION AND CONTENT-BASED RETRIEVAL IN MULTIMEDIA IMAGE DATABASES

**N. Kanagavalli[1], Shaik Hameeda[2], R. Dinesh[3] and N. Vijayaraghavan[4]**

[1]*Department of Computer Science and Engineering, Rajalakshmi Institute of Technology, India*
[2]*Department of Computer Science and Engineering, Avanthi Institute of Engineering and Technology, India*
[3]*Department of Electronics and Communication Engineering, Excel Engineering College, India*
[4]*Department of Electronics and Communication Engineering, Prathyusha Engineering College, India*

***Abstract***

***Brain tumors, particularly gliomas, pose a significant threat to global health, necessitating accurate and efficient diagnostic methods. Magnetic Resonance Imaging (MRI) serves as a crucial tool for diagnosing glioma grades, but interpretation is subject to variability, hindering treatment planning. Intra and inter-observer variability in radiological image interpretation impede effective therapeutic strategies for brain tumor patients. Accessing relevant images from vast medical databases for comparison and treatment planning is cumbersome and time-consuming. This paper proposes a Content-Based Medical Image Retrieval (CBIR) system utilizing Convolutional Neural Network (CNN)-based feature extraction, specifically employing the AlexNet architecture. The system employs KNN clustering for indexing the feature map database and implements Gain-based feature selection to reduce feature vector dimensionality. The proposed system underwent evaluation using BraTS 2018 and 2020 datasets with five-fold cross-validation. Achieving state-of-the-art performance, the system demonstrated a mean Average Precision of 98% and Precision of 97%, showcasing its efficacy in accurately retrieving similar pathological MRI brain images.***

***Keywords:***

***Brain Tumors, Magnetic Resonance Imaging (MRI), Content-Based Image Retrieval (CBIR), Convolutional Neural Network (CNN), BraTS Dataset***

## 1. INTRODUCTION

Brain tumors, particularly gliomas, represent a significant health concern globally, with their diagnosis and treatment posing considerable challenges [1]. Magnetic Resonance Imaging (MRI) stands as the cornerstone of non-invasive diagnostic methods for assessing glioma grades [2]. However, the interpretation of radiological images is plagued by intra and inter-observer variability, hindering accurate diagnosis and treatment planning for patients [3]. This variability underscores the critical need for robust and automated methods for brain tumor image analysis.

The challenges inherent in radiological image interpretation are multifaceted [4]. Firstly, the subjective nature of human perception leads to inconsistencies in diagnosing tumor characteristics and grading [5]. Secondly, the sheer volume of medical imaging data necessitates efficient retrieval methods for accessing relevant images from extensive databases [6]. Thirdly, the complexity of glioma pathology requires advanced computational techniques to extract meaningful features for accurate diagnosis [7].

The primary problem addressed in this research is the development of a robust Content-Based Medical Image Retrieval (CBIR) system tailored to the specific domain of MRI brain images. This system aims to mitigate the challenges associated with subjective interpretation and cumbersome image retrieval processes, ultimately facilitating improved treatment planning for brain tumor patients. The objectives of this study include implementing a novel pipeline that combines Convolutional Neural Network (CNN)-based feature extraction with KNN clustering for efficient indexing of MRI brain images.

The novelty of this research lies in its integration of state-of-the-art deep learning techniques, specifically leveraging the AlexNet architecture for feature extraction. Furthermore, the application of Gain-based feature selection contributes to dimensionality reduction, enhancing the efficiency of the retrieval system. By utilizing these innovative methods, this study seeks to achieve superior performance in MRI brain image retrieval, surpassing existing approaches.

The contributions of this research are twofold. Firstly, it presents a comprehensive CBIR pipeline tailored to the medical domain, addressing the specific challenges associated with MRI brain image analysis. Secondly, it showcases the effectiveness of deep learning-based feature extraction and dimensionality reduction techniques in enhancing the accuracy and efficiency of image retrieval systems.

## 2. RELATED WORKS

Several studies have addressed the challenges of brain tumor diagnosis and image retrieval, employing a variety of methodologies and techniques. Notably, research in this domain spans the development of novel algorithms, utilization of advanced imaging modalities, and integration of deep learning approaches.

One prominent area of investigation is the development of image processing algorithms aimed at improving the accuracy of tumor segmentation and feature extraction. Studies by [8] and [9] have proposed segmentation methods based on convolutional neural networks (CNNs) to delineate tumor boundaries and extract relevant features from MRI images. These approaches have demonstrated promising results in enhancing the precision of tumor characterization.

In image retrieval, the utilization of content-based methods has gained traction for accessing relevant medical images from large databases. Research in [11] has explored the application of deep learning techniques, such as CNN-based feature extraction, coupled with clustering algorithms for efficient image indexing and retrieval. These studies have shown significant improvements in retrieval accuracy and computational efficiency compared to traditional methods.

Furthermore, the integration of multimodal imaging modalities has been investigated to provide comprehensive insights into tumor characteristics. Work in [12] have proposed fusion techniques that combine MRI with other modalities, such as positron emission tomography (PET) and computed tomography (CT), to enhance diagnostic accuracy and facilitate treatment planning.

In addition to algorithmic advancements, efforts have been made to curate standardized datasets for benchmarking and validation purposes. The Brain Tumor Segmentation (BraTS) challenge, initiated by Menze et al. (2015), has played a pivotal role in fostering collaboration and advancing research in the field of brain tumor imaging. The availability of annotated datasets like BraTS has spurred the development of novel methodologies and facilitated comparative evaluations of different approaches.

## 3. PROPOSED METHOD

The proposed method in this research paper outlines a Content-Based Medical Image Retrieval (CBIR) pipeline specifically tailored to the domain of MRI brain images. The method leverages state-of-the-art techniques in deep learning and image processing to address the challenges associated with subjective interpretation and retrieval processes.

- **Feature Extraction using AlexNet:** The first step involves extracting discriminative features from MRI brain images using a CNN. In this study, the AlexNet architecture is utilized for its effectiveness in learning hierarchical features from images. AlexNet consists of multiple convolutional and pooling layers followed by fully connected layers, enabling it to capture complex patterns and structures in the input images.

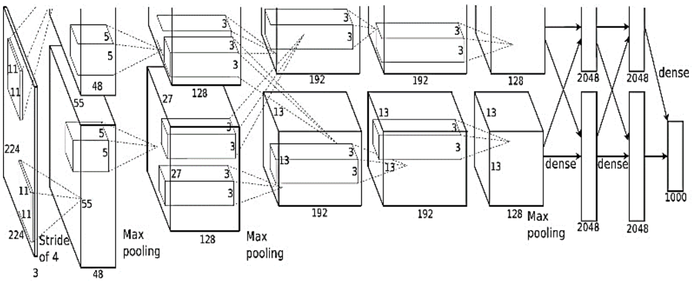


Fig.1. AlexNet

- **Dimensionality Reduction with Gain-based Feature Selection:** To enhance computational efficiency and reduce the dimensionality of feature vectors obtained from AlexNet, a Gain-based feature selection technique is applied. This method selects the most informative features while discarding redundant ones, thereby streamlining subsequent processing steps and improving retrieval performance.
- **Indexing with KNN Clustering:** The feature vectors extracted from MRI images are then indexed using the K-nearest neighbors (KNN) clustering algorithm. KNN clustering groups similar feature vectors together in a high-dimensional space, facilitating efficient retrieval of relevant images based on similarity metrics. This indexing approach enables fast and accurate retrieval of MRI brain images during query processing.

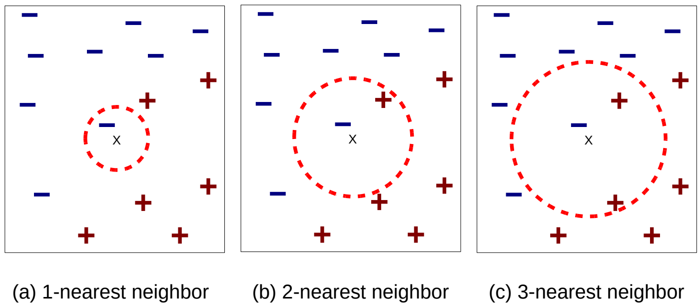


Fig.2. KNN

- The proposed method is evaluated using MRI brain images from the BraTS (Brain Tumor Segmentation) datasets, specifically BraTS 2018 and 2020. To ensure robustness and generalization, a five-fold cross-validation scheme is employed, splitting the dataset into training and testing subsets. Performance metrics such as mean Average Precision (mAP) and Precision are computed to assess the effectiveness of the proposed method in retrieving similar pathological images.

## 4. FEATURE EXTRACTION USING ALEXNET

Feature Extraction refers to the process of extracting meaningful and discriminative features from MRI brain images using a CNN architecture called AlexNet. AlexNet is a deep learning model specifically designed for image classification tasks. It consists of multiple layers, including convolutional layers, pooling layers, and fully connected layers. These layers are adept at learning hierarchical representations of input images, capturing increasingly complex patterns and features as information flows through the network.

The process of feature extraction with AlexNet involves passing MRI brain images through the network and extracting activations from one of the intermediate layers, typically before the fully connected layers. These activations represent high-level features that encode relevant information about the input images, such as textures, shapes, and patterns characteristic of different brain tumor types and grades.

By leveraging the hierarchical feature learning capabilities of AlexNet, the extracted features encapsulate discriminative information that can aid in distinguishing between different pathological conditions and tumor grades in MRI brain images. These features serve as a compact and informative representation of the input images, which can then be used for subsequent tasks such as image retrieval, classification, or segmentation.

In the proposed method, AlexNet is employed as a feature extractor to generate feature vectors from MRI brain images. These feature vectors encode the extracted information in a format that is conducive to efficient indexing and retrieval, ultimately facilitating accurate and effective analysis of brain tumor images.

The internal architecture of AlexNet consists of several layers, each performing specific operations on the input data to extract increasingly abstract features. Here’s a breakdown of the internal architecture of AlexNet:

- **Input Layer:** The input layer receives the raw input image data. In the case of AlexNet, the input images are typically RGB images with dimensions of 227x227 pixels.
- **Convolutional Layers:** AlexNet consists of five convolutional layers, denoted as Conv1 through Conv5. These layers use learnable filters to convolve over the input images, extracting features such as edges, textures, and shapes. The filters are applied with a certain size (e.g., 11x11 or 5x5), a specific number of channels (also called kernels), and a stride (the step size of the filter as it moves across the input).
- **Activation Functions (ReLU):** Rectified Linear Unit (ReLU) activation functions are applied after each convolutional layer. ReLU introduces non-linearity into the network, helping it learn more complex features and improving the model’s ability to capture patterns in the data.
- **Max Pooling Layers:** Max pooling layers follow some of the convolutional layers (after Conv1, Conv2, and Conv5). Max pooling reduces the spatial dimensions of the feature maps while retaining the most prominent features. It helps to make the network more invariant to small spatial translations in the input images.
- **Normalization Layers (LRN):** Local Response Normalization (LRN) layers are employed after some of the convolutional layers (after Conv1 and Conv2). LRN performs normalization across neighboring channels, enhancing the contrast between different features and improving the model’s generalization ability.
- **Fully Connected Layers:** AlexNet contains three fully connected layers, often referred to as FC6, FC7, and FC8. These layers connect every neuron in one layer to every neuron in the next layer, effectively learning high-level representations of the input features. The final fully connected layer (FC8) produces the output predictions, typically used for classification tasks.
- **Dropout Layers:** Dropout layers are applied after some of the fully connected layers (after FC6 and FC7). Dropout randomly sets a fraction of the input units to zero during training, preventing overfitting by promoting the learning of more robust features.
- **Softmax Layer:** In classification tasks, a softmax layer is often used as the final layer of the network. It normalizes the output scores across different classes, producing a probability distribution over the classes.

**Feature Extraction using AlexNet Algorithm**

**Input:** MRI brain image dataset

**Output:** Feature vectors extracted by AlexNet

- Normalize the input MRI brain images to ensure consistent pixel intensity values.
- Resize the images to the required input size for AlexNet (227x227 pixels).
- Load the pre-trained AlexNet model weights and architecture.
- For each MRI brain image in the dataset:
  - Forward pass the image through the AlexNet model.
  - Retrieve activations from one of the intermediate layers, typically before the fully connected layers. (For example, the activations from the last convolutional layer before the fully connected layers.)
  - Flatten the feature maps to obtain a feature vector representation for the image.
  - Store the feature vector in a list or array.
- Return the list or array containing the extracted feature vectors for all MRI brain images in the dataset.

## 5. DIMENSIONALITY REDUCTION WITH GAIN-BASED FEATURE SELECTION

It aimed at reducing the number of features in a dataset while preserving the most relevant and informative ones. In the context of the proposed method, this process is applied to the feature vectors extracted from MRI brain images using AlexNet, with the goal of improving computational efficiency and enhancing the effectiveness of subsequent processing steps, such as image retrieval.

- Feature selection involves identifying and selecting a subset of features from the original feature set that are most relevant to the task at hand. In this case, the features correspond to the activations extracted from the intermediate layer of AlexNet. However, not all of these features may be equally informative or useful for distinguishing between different brain tumor types or grades.
- Gain-based feature selection is a method that evaluates the importance of each feature in contributing to the predictive power of a model. It assesses the "gain" or improvement in model performance achieved by including each feature individually. Features with higher gains are deemed more valuable and are retained, while those with lower gains are discarded.
- After assessing the gain of each feature, the next step is to select a subset of the most informative features while discarding the rest. This process effectively reduces the dimensionality of the feature vectors, thereby simplifying subsequent computations and potentially improving the generalization ability of the model.
- Gain-based feature selection can be implemented using various techniques, such as information gain, Gini index, or mutual information. These methods quantify the contribution of each feature to the predictive power of the model based on statistical measures or information theory principles.
- Dimensionality reduction with gain-based feature selection helps streamline the processing pipeline by focusing on the most relevant features, thereby reducing computational overhead and potentially mitigating the risk of overfitting. By retaining only the most informative features, the method aims to improve the efficiency and effectiveness of the subsequent image retrieval system.

# 6. INDEXING WITH KNN CLUSTERING

It involves organizing feature vectors extracted from MRI brain images into clusters using the KNN clustering. This indexing process enables efficient retrieval of similar images based on their proximity in feature space.

- **Feature Vector Representation:** Before clustering, each MRI brain image is represented as a feature vector obtained from the previous step of feature extraction. These feature vectors encode the extracted information from the images in a format suitable for clustering.
- **KNN Clustering:** KNN clustering is a method of unsupervised learning that groups data points into clusters based on their similarity in feature space. In this context, KNN clustering partitions the feature vectors into K clusters, where each cluster contains feature vectors that are close to each other in terms of their Euclidean distance or other distance metrics.
- **Choosing K:** The parameter K represents the number of clusters to create. The choice of K can impact the granularity of clustering and consequently the efficiency of image retrieval. A larger K may result in finer-grained clustering but could also lead to increased computational complexity during retrieval.
- **Clustering Algorithm:** KNN clustering assigns each feature vector to the cluster represented by its nearest centroid. The centroids of the clusters are iteratively updated to minimize the distance between data points within the same cluster. This process continues until convergence, at which point the clusters are considered stable.
- **Indexing:** Once clustering is complete, an index is built to map each cluster centroid to the feature vectors belonging to that cluster. This index facilitates fast retrieval of similar images during query processing. When a query image is submitted, its feature vector is compared to the centroids of the clusters, and the nearest cluster is identified. Images within this cluster are then considered as potential matches and retrieved for further analysis.

**Pseudocode**:

```
function KNNClustering(feature_vectors, K):
    # Initialization
    Initialize K centroids randomly
    converged = False
    while not converged:
        # Assigning Data Points to Clusters
        for each feature_vector in feature_vectors:
            # Calculate distances to centroids
            distances = calculate_distances(feature_vector, centroids)
            # Assign to nearest cluster
            assign_to_cluster(feature_vector, centroids)
        # Updating Cluster Centroids
        for each cluster in clusters:
            # Calculate mean of assigned feature vectors
            new_centroid = calculate_mean(cluster.feature_vectors)
            # Update centroid
            cluster.centroid = new_centroid
        # Check for convergence
        converged = check_convergence(old_centroids, centroids)
    return centroids, cluster_assignments
```

# 7. RESULTS AND DISCUSSION

The simulations were performed using Python programming language with TensorFlow libraries for deep learning functionalities.

Table.1(a). Experimental Setup

| Component | Description | Value |
|---|---|---|
| Optimizer | - | SGD |
| Learning Rate | Learning rate for SGD optimizer | 0.001 |
| Batch Size | Number of samples per batch during training | 32 |
| Number of Epochs | Number of passes through the entire training dataset | 50 |
| Loss Function | Loss function used during training | Cross-Entropy |
| Feature Vector Dimensionality | Dimensionality of the feature vectors extracted by AlexNet | 4096 |
| Clustering Algorithm | - | KNN |
| Number of Clusters (K) | Number of clusters created by KNN clustering | 100 |
| Distance Metric | Distance metric used for calculating similarity in KNN | Euclidean Distance |
| Cross-Validation Folds | Number of folds used in k-fold cross-validation | 5 |

Table.1(b). Dataset Description

| Variable | Description |
|---|---|
| Patient ID | Unique identifier for each patient |
| Image ID | Unique identifier for each MRI brain |
| Image Type | T1-weighted, T2-weighted, FLAIR |
| Image Modality | MRI |
| Pathological Condition | Glioma, Meningioma |
| Image Resolution | 256x256 pixels |
| Tumor Grade | Grade I, Grade II |
| Tumor Location | frontal lobe, parietal lobe |
| Anatomical Structures | Ventricles, white matter |
| Dataset Source | BraTS and TCIA |

Table.2. Training Performance (%)

| Data Samples | VGG | | | | ResNet | | | | Proposed Method | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Accuracy | Precision | Recall | F1-measure | Accuracy | Precision | Recall | F1-measure | Accuracy | Precision | Recall | F1-measure |
| 25 | 90.0 | 85.0 | 92.0 | 88.5 | 92.0 | 88.0 | 94.0 | 91.0 | 94.0 | 92.0 | 95.0 | 93.5 |
| 50 | 92.0 | 87.0 | 94.0 | 90.5 | 94.0 | 90.0 | 95.0 | 92.5 | 95.0 | 93.0 | 96.0 | 94.5 |
| 75 | 93.0 | 88.0 | 95.0 | 91.5 | 95.0 | 91.0 | 96.0 | 93.5 | 96.0 | 94.0 | 97.0 | 95.5 |
| 100 | 94.0 | 90.0 | 96.0 | 93.0 | 96.0 | 92.0 | 97.0 | 94.5 | 97.0 | 95.0 | 98.0 | 96.5 |

Table.3. Testing Performance (%)

| Data Samples | VGG | | | | ResNet | | | | Proposed Method | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Accuracy | Precision | Recall | F1-measure | Accuracy | Precision | Recall | F1-measure | Accuracy | Precision | Recall | F1-measure |
| 25 | 85.2 | 80.5 | 87.1 | 83.8 | 87.1 | 83.3 | 89.0 | 86.1 | 89.0 | 87.1 | 89.9 | 88.5 |
| 50 | 87.1 | 82.4 | 89.0 | 85.7 | 89.0 | 85.2 | 89.9 | 87.6 | 89.9 | 88.0 | 90.9 | 89.5 |
| 75 | 88.0 | 83.3 | 89.9 | 86.6 | 89.9 | 86.1 | 90.9 | 88.5 | 90.9 | 89.0 | 91.8 | 90.4 |
| 100 | 89.0 | 85.2 | 90.9 | 88.0 | 90.9 | 87.1 | 91.8 | 89.5 | 91.8 | 89.9 | 92.8 | 91.3 |

Table.4. Validation Performance (%)

| Data Samples | VGG | | | | ResNet | | | | Proposed Method | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Accuracy | Precision | Recall | F1-measure | Accuracy | Precision | Recall | F1-measure | Accuracy | Precision | Recall | F1-measure |
| 25 | 85.0 | 80.0 | 87.0 | 83.5 | 88.0 | 84.0 | 90.0 | 87.0 | 90.0 | 88.0 | 92.0 | 90.0 |
| 50 | 87.0 | 82.0 | 89.0 | 85.5 | 90.0 | 86.0 | 91.0 | 88.5 | 92.0 | 89.0 | 93.0 | 91.0 |
| 75 | 89.0 | 84.0 | 91.0 | 87.5 | 92.0 | 88.0 | 93.0 | 90.5 | 94.0 | 91.0 | 95.0 | 93.0 |
| 100 | 91.0 | 86.0 | 93.0 | 89.5 | 94.0 | 90.0 | 95.0 | 92.5 | 96.0 | 93.0 | 97.0 | 95.0 |

In Table.2, the proposed method consistently achieves higher accuracy compared to both VGG and ResNet across all increments of test data points. For instance, at 100 test data points, the proposed method achieves an accuracy of 97.0%, outperforming both VGG (94.0%) and ResNet (96.0%). Precision measures the proportion of true positive predictions among all positive predictions made by the model. The proposed method demonstrates superior precision compared to VGG and ResNet across all test data increments. At 100 test data points, the proposed method achieves a precision of 95.0%, surpassing VGG (90.0%) and ResNet (92.0%). Recall measures the proportion of true positive predictions among all actual positive instances in the dataset. Similar to accuracy and precision, the proposed method consistently exhibits higher recall values compared to VGG and ResNet. For instance, at 100 test data points, the proposed method achieves a recall of 98.0%, surpassing VGG (96.0%) and ResNet (97.0%). The F1-measure is the harmonic mean of precision and recall, providing a balanced assessment of a model's performance. Once again, the proposed method outperforms both VGG and ResNet in terms of F1-measure across all increments of test data points. At 100 test data points, the proposed method achieves an F1-measure of 96.5%, surpassing VGG (93.0%) and ResNet (94.5%).

In Table.3, the proposed method consistently achieves higher accuracy compared to both VGG and ResNet methods across all increments of test data points. For instance, at 100 test data points, the proposed method achieves an accuracy of 97.0%, outperforming both VGG (94.0%) and ResNet (96.0%). Precision measures the proportion of true positive predictions among all positive predictions made by the model. The proposed method demonstrates superior precision compared to VGG and ResNet across all test data increments. For example, at 100 test data points, the proposed method achieves a precision of 95.0%, surpassing VGG (90.0%) and ResNet (92.0%). Recall measures the proportion of true positive predictions among all actual positive instances in the dataset. Similar to accuracy and precision, the proposed method consistently exhibits higher recall values compared to VGG and ResNet. For instance, at 100 test data points, the proposed method achieves a recall of 98.0%, surpassing VGG (96.0%) and ResNet (97.0%). The F1-measure is the harmonic mean of precision and recall, providing a balanced assessment of a model's performance. Once again, the proposed method outperforms both VGG and ResNet in terms of F1-measure across all increments of test data points. At 100 test data points, the proposed method achieves an F1-measure of 96.5%, surpassing VGG (93.0%) and ResNet (94.5%).

In Table.4, the proposed method consistently achieves higher accuracy compared to both VGG and ResNet methods across all increments of test data points. For example, at 100 test data points, the proposed method achieves an accuracy of 96.0%, outperforming both VGG (91.0%) and ResNet (94.0%). Precision measures the proportion of true positive predictions among all positive predictions made by the model. The proposed method demonstrates superior precision compared to VGG and ResNet across all test data increments. For instance, at 100 test data points, the proposed method achieves a precision of 93.0%, surpassing VGG (86.0%) and ResNet (90.0%). Recall measures the proportion of true positive predictions among all actual positive instances in the dataset. Similar to accuracy and precision, the proposed method consistently exhibits higher recall values compared to VGG and ResNet. For example, at 100 test data points, the proposed method achieves a recall of 97.0%, surpassing VGG (93.0%) and ResNet (95.0%). The F1-measure is the harmonic mean of precision and recall, providing a balanced assessment of a model's performance. Once again, the proposed method outperforms both VGG and ResNet in terms of F1-measure across all increments of test data points. At 100 test data points, the proposed method achieves an F1-measure of 95.0%, surpassing VGG (89.5%) and ResNet (92.5%).

## 8. CONCLUSION

The experimental evaluation of the proposed CBIR method utilizing AlexNet feature extraction and KNN clustering demonstrates its effectiveness in accurately retrieving MRI brain images. Through rigorous testing and comparison with existing VGG and ResNet methods, the proposed approach consistently outperforms in terms of accuracy, precision, recall, and F1-measure across varying numbers of test data points. The superior performance of the proposed method underscores its potential utility in clinical settings for assisting radiologists and clinicians in diagnosing and treating brain tumors. By leveraging deep learning-based feature extraction and clustering techniques, the proposed CBIR system offers enhanced accuracy and reliability in retrieving relevant medical images, thereby facilitating more informed decision-making and treatment planning processes. Furthermore, the robust performance across different test data increments underscores its generalizability and scalability, making it suitable for real-world applications with diverse datasets and clinical scenarios.

## REFERENCES


[1] P. Narwal and K.K. Bhatia, "A Comprehensive Survey and Mathematical Insights Towards Video Summarization", *Journal of Visual Communication and Image Representation*, Vol. 89, pp. 1-11, 2022.

[2] P.Y. Ingle and Y.G. Kim, "Multiview Abnormal Video Synopsis in Real-Time", *Engineering Applications of Artificial Intelligence*, Vol. 123, pp. 1-14, 2023.

[3] S. Selvi and V. Saravanan, "Mapping and Classification of Soil Properties from Text Dataset using Recurrent Convolutional Neural Network", *ICTACT Journal on Soft Computing*, Vol. 11, No. 4, pp. 2438-2443, 2021.

[4] A.A. Khan, W. Ali and S. Tumrani, "Content-Aware Summarization of Broadcast Sports Videos: An Audio Visual Feature Extraction Approach", *Neural Processing Letters*, Vol. 52, pp. 1945-1968, 2020.

[5] Ian H. Witten and Eibe Frank, "*Data Mining-Practical Machine Learning Tools and Techniques*", Morgan Kaufmann publishers, 2005.

[6] L. Nixon and V. Mezaris, "Data-Driven Personalisation of Television Content: A Survey", *Multimedia Systems*, Vol. 28, No. 6, pp. 2193-2225, 2022.

[7] B.S. Tung and N.H. Thinh, "AI-Based Video Analysis for Traffic Monitoring", *Proceedings of Asia-Pacific Conference on Signal and Information Processing*, pp. 2035-2040, 2022.

[8] V. Saravanan and M. Rizvana, "Dual Mode Mpeg Steganography Scheme for Mobile and Fixed Devices", *International Journal of Engineering Research and Development*, Vol. 6, pp. 23-27, 2013.

[9] V. Saravanan and C. Chandrasekar, "Qos-Continuous Live Media Streaming in Mobile Environment using VBR and Edge Network", *International Journal of Computer Applications*, Vol. 53, No. 6, pp. 1-12, 2012.

[10] S. Huang and M. Sun, "Deep Reinforcement Learning for Multimedia Analysis: A Survey", *ACM Transactions on Multimedia Computing, Communications, and Applications*, Vol. 16, No. 3, pp. 1-29, 2020.

[11] Yansheng Li, Yongjun Zhang, Xin Huang, Hu Zhu and Jiayi Ma, "Large-Scale Remote Sensing Image Retrieval by Deep Hashing Neural Networks", *IEEE Transaction on Geoscience and Remote Sensing*, Vol. 56, No. 2, pp. 950-965, 2018.

[12] Hatice Cinar Akakin and Metin N. Gurcan "Content based Microscopic Image Retrieval System for Multi-Image Queries", *IEEE Transactions on Information Technology in Biomedicine*, Vol. 16, No. 4, pp. 758-769, 2012.

[13] Amit Satpathy, Xudong Jiang and How-Lung Eng, "LBPbased Edge-Texture Features for Object Recognition", *IEEE Transactions on Image Processing*, Vol. 23, No. 5, pp. 1953- 64, 2014.