ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2023.0429

ENHANCED AI BASED FEATURE EXTRACTION TECHNIQUE IN MULTIMEDIA IMAGE RETRIEVAL

R. Gayathri[1], T. Gobinath[2], A. Muthumari[3] and R.S.V. Rama Swathi[4]
[1]*Department of Electronics and Communication Engineering, Rajalakshmi Engineering College, India*
[2]*Department of Computer Science and Engineering, Chettinad College of Engineering and Technology, India*
[3]*Department of Computer Science and Engineering, University College of Engineering, India*
[4]*Department of Master of Business Administration, Koneru Lakshmaiah Education Foundation, India*

Abstract
In the era of rapid technological advancements, the demand for efficient and accurate identification and retrieval of information from multimedia images has seen a substantial increase. To meet this growing demand, artificial intelligence (AI)-based technologies, particularly feature extraction techniques, have gained significant popularity. Feature extraction involves the extraction of salient features from multimedia images, such as edges, lines, curves, textures, and colors, with the aim of representing the data in a more suitable format for analysis. This paper presents an enhanced AI-based feature extraction technique for multimedia image retrieval. The proposed method introduces a novel approach that combines the power of deep learning and evolutionary algorithms in a neuro-symbolic computation framework. Specifically, the renowned VGG16 deep learning algorithm is employed as the initial feature extractor. VGG16 is a state-of-the-art deep convolutional neural network that has demonstrated exceptional performance in various computer vision tasks, including image classification and feature extraction. The primary idea behind this approach is to leverage the capabilities of AI to extract the most discriminative features from the source images using VGG16. These features are then further refined using evolutionary algorithms, which employ a search and optimization process inspired by natural evolution. By iteratively improving the extracted features through the evolutionary algorithms, the method aims to enhance the discriminative power and representational quality of the extracted features. To evaluate the performance of the proposed approach, extensive experiments were conducted. The results demonstrate that the method achieves superior performance in terms of precision, recall, and F-measure when compared to conventional feature extraction techniques. Furthermore, a comprehensive comparison with state-of-the-art AI-based feature extraction techniques further highlights the potential and effectiveness of the proposed approach in multimedia image retrieval applications.

Keywords:
Information Retrieval, Feature Extraction, Multimedia, Images

1. INTRODUCTION

In today's media-centric computing world, efficient image retrieval from vast content libraries is crucial, especially for businesses relying on visuals like film-production companies, graphic design firms, and online stores [1]. To address this need, new methods of feature extraction from multimedia images, such as enhanced AI-based techniques, have emerged [2]. These techniques aim to isolate essential image characteristics that can be used to identify objects or determine the content type [2]. By automatically recognizing patterns in colors, shapes, and textures, these techniques make the search process more efficient and less time-consuming.

The key advantage of using AI-based feature extraction techniques is their ability to detect features that would be difficult for manual methods to identify, as well as extracting features specific to a given application [3]. For instance, a film production company can customize the AI feature extraction process to focus on facial features of actors in uploaded images, facilitating a quicker and more accurate search [3]. AI-based feature extraction also offers greater accuracy than manual methods since the algorithm is trained to recognize specific features, eliminating user assumptions. Additionally, AI feature extraction can process multiple images in parallel, resulting in faster response times.

Enhanced AI-based feature extraction is a powerful tool for multimedia image retrieval, improving search efficiency and accuracy [4]. By extracting important features, the technique enhances the precision of the retrieval process and provides a more accurate representation of an image. It utilizes AI algorithms to detect and segment different features, such as object edges, patterns, and textures, facilitating better analysis of colors and shades [5].

Moreover, the Enhanced AI-Based Feature Extraction Technique can reduce the overall cost and time associated with the retrieval process [5]. The application of AI significantly speeds up the process, depending on image size and complexity. Additionally, it reduces storage and computing requirements, showcasing the potential of AI in multimedia image retrieval.

The technique's benefits include improved accuracy, reduced computational and storage costs, and faster retrieval times [6]. By leveraging powerful AI algorithms, the technique enhances feature extraction and demonstrates the potential of artificial intelligence in multimedia image retrieval [7].

In summary, the Enhanced AI-Based Feature Extraction Technique leverages artificial intelligence to extract meaningful information from images, enhancing precision and efficiency in multimedia image retrieval. It offers advantages such as improved accuracy, reduced costs, and faster retrieval times, showcasing the potential of AI in this field [7].

The main contribution of this research has the following:

- Enhanced AI-based feature extraction techniques are capable of extracting more features from an image, thus increasing the accuracy of the search process.
- AI-based feature extraction techniques can process requests faster as compared to conventional methods, thus reducing the time taken for retrieval.
- Enhanced AI-based feature extraction techniques are more adaptive to changes in the user's query as they can be adjusted or tuned depending on the search query in order to cater to the users' needs.

- AI-based feature extraction techniques are more robust and stable as compared to traditional methods, as they are not as easily affected by noise, making the retrieval process more efficient.
- AI-based feature extraction techniques are much more cost-effective than other methods as they can be utilized repeatedly with minimal changes, thus reducing the need for hardware or software purchases.

2. RELATED WORKS

The use of artificial intelligence (AI) in multimedia image retrieval has gained popularity for its improved accuracy and faster processing times. Enhanced AI-based feature extraction techniques utilize advanced algorithms to extract features and identify objects or patterns in images [8]. This technology enhances multimedia image retrieval, improving user experiences in image searching and video frame analysis.

Implementing AI-based feature extraction techniques requires consideration of several important factors. First, determining the type of features to extract and how they will be used for image searching and analysis is crucial. Selecting a high-quality feature extraction model is essential for accuracy testing. Second, evaluating the impact on storage and retrieval times is important to ensure a seamless user experience [9].

Speed can be improved with AI, but careful testing is necessary. Monitoring the performance of the feature extraction model over time is also important to maintain accuracy [10]. Handling the computational load associated with AI-based techniques, especially with large databases, requires powerful hardware, software optimizations, or distributed architectures. Lastly, adhering to legal and ethical frameworks for user data protection is necessary when using AI-based feature extraction models [11].

While AI-based feature extraction techniques have the potential to enhance image search and analysis accuracy, several considerations must be addressed before implementation. Careful consideration of these factors can integrate AI-based feature extraction techniques effectively, greatly enhancing user experiences.

The 21st century has witnessed remarkable advancements in multimedia image retrieval. Feature extraction is a common method used to retrieve digital images by extracting meaningful features such as color, texture, shape, size, and motion. AI-based feature extraction techniques have further improved this process [12]. However, challenges arise when using AI-based techniques in multimedia image retrieval [13]. Insufficient training data can hinder feature detection, and integrating feature data from different sources can be challenging. Additionally, the computational resources required for AI-based feature extraction pose performance bottlenecks [14]. Despite these drawbacks, AI-based feature extraction techniques are widely utilized due to their efficiency and accuracy.

Future research should focus on enhancing automatic data annotation, integrating multiple datasets, and improving scalability of AI-based feature extraction techniques. With advancements in these areas, multimedia image retrieval will achieve faster and more accurate results, leading to improved services and outcomes for users [15].

The novelty of the proposed research work has the following. Enhanced AI based feature extraction techniques provide a novel way to automatically extract features from multimedia images to improve retrieval results. This technique applies advanced methods such as deep learning, generative models, and transfer learning to identify and extract meaningful information from the image content. This helps to improve the accuracy and performance of multimedia image retrieval. The features extracted can then be used for retrieval tasks such as clustering, classification, and similarity search. This feature extraction technique could also be used for content-based retrieval in a multimedia database.

3. PROPOSED MODEL

This study introduces the VGG-16 algorithm, which aims to enhance media based on a semantic indexing graph and user-selected queries. The algorithm follows a systematic process that involves analyzing media edges, identifying suitable enhancements, computing suggestions, and updating the database. By employing this approach, the algorithm enhances the quality and presentation of media, resulting in a more engaging and informative experience for users.

The key feature of the algorithm is the incorporation of a semantic indexing graph, which harnesses the power of semantic analysis to understand the content and context of media edges. This enables the algorithm to make more accurate and relevant enhancements, ensuring that the applied modifications align with the user's intent and preferences.

Additionally, the algorithm takes advantage of user-selected queries to enhance personalization. By allowing users to actively participate in the enhancement process, the algorithm empowers them to have a direct influence on the modifications made to their media. This personalized approach enhances user satisfaction and ensures that the algorithm's output aligns with the user's specific requirements.

The algorithm's flow diagram provides a visual representation of the information flow and system interaction. It illustrates the key steps involved in multimedia analysis, query processing, retrieval using indexes, and the interaction between the application and the user. The feedback loop incorporated in the system allows users to provide additional queries or suggestions, fostering continuous improvement and a better user experience.

The VGG-16 algorithm, coupled with the semantic indexing graph and user-selected queries, presents a comprehensive and effective approach to enhancing media. By leveraging semantic analysis and incorporating user preferences, the algorithm delivers more accurate and personalized enhancements, thereby improving the overall user experience. The flow diagram visually represents the algorithm's functionality, outlining the essential steps and highlighting the interactive nature of the system. The study's findings demonstrate the potential of this algorithm in advancing multimedia enhancement techniques.

3.1 VGG-16 ARCHITECTURE

VGG-16 is a deep learning architecture that has gained significant popularity for various computer vision tasks, including image classification, object detection, and feature extraction. It is renowned for its simplicity and effectiveness in extracting rich features from images.

The architecture of VGG-16 consists of 16 layers, which are organized into convolutional and fully connected layers. The convolutional layers are responsible for learning hierarchical representations of the input images, while the fully connected layers serve as the classifier.

To illustrate the architecture mathematically, let's denote an input image as X, with dimensions H x W x C, where H represents the height, W represents the width, and C represents the number of channels. The convolutional layers in VGG-16 are designed to learn a set of filters that extract useful features from the input image.

The convolution operation in VGG-16 can be represented as follows:

$$Z^l = Convolution(X, W^l) + b^l,$$

where Z^l represents the output feature maps at layer l, Convolution denotes the convolution operation, W^l represents the learnable filters at layer l, and b^l represents the bias term.

After each convolutional layer, a non-linear activation function, typically the Rectified Linear Unit (ReLU), is applied element-wise to introduce non-linearity:

$$A^l = \mathrm{ReLU}(Z^l).$$

The feature maps obtained from each convolutional layer are then passed through a series of max pooling layers, which reduce the spatial dimensions of the feature maps while retaining the most salient features.

The fully connected layers in VGG-16 serve as the classifier and are responsible for mapping the extracted features to the corresponding classes. These layers can be represented mathematically as:

$$F^l = f(W^l * F^{\{l\text{-}1\}} + b^l),$$

where F^l represents the activations at layer l, W^l denotes the weight matrix, $F^{\{l\text{-}1\}}$ represents the activations from the previous layer, b^l represents the bias term, and f denotes the activation function.

Finally, the output of the last fully connected layer is fed into a softmax function to obtain class probabilities:

$$P(class_i) = softmax(F^L),$$

where L represents the index of the last fully connected layer.

The parameters (weights and biases) of VGG-16 are learned through a process called training, where the model is exposed to a large labeled dataset and adjusts its parameters to minimize a predefined loss function, such as cross-entropy loss, that measures the dissimilarity between predicted and ground-truth labels.

VGG-16 has been widely used as a feature extractor in various applications. By leveraging the hierarchical representations learned by the convolutional layers, it can provide highly discriminative and informative feature representations for tasks such as image retrieval, where these features can be used to compute similarity measures between images and retrieve similar images from a database.

The flow diagram depicts the sequential process of multimedia analysis, query processing, retrieval using indexes, and the interaction between the application and the user.

3.2 MULTIMEDIA ANALYSIS

Multimedia Analysis is the initial step, focusing on the examination of multimedia documents like images, videos, or audio files to extract pertinent information. This stage entails various tasks such as feature extraction, content analysis, object recognition, or sentiment analysis, which depend on the specific application's requirements.

Query Processing occurs after the completion of multimedia analysis. At this stage, the system receives queries from the user, which can be in the form of text, keywords, or even multimedia inputs. The query processing phase involves comprehending the user's query and determining the relevant search criteria for subsequent retrieval.

Retrieval using Indexes is the subsequent stage in the flow diagram. The retrieval process utilizes indexes, which are data structures organizing the multimedia data based on specific criteria such as keywords, metadata, or visual features. These indexes enable the system to efficiently search through a large collection of multimedia documents and promptly locate relevant documents that align with the user's query.

Results are then provided to the application after retrieving the relevant multimedia documents. The application can take various forms, such as a website, a mobile app, or any other platform where the user interacts with the system. The results can be presented to the user in different formats, including a list of ranked documents, thumbnail previews, or detailed information about each document, depending on the design of the application.

The flow diagram also encompasses a User Queries or Suggestions feedback loop, enabling users to provide additional queries or suggestions based on the results they receive. This feedback loop plays a crucial role in enhancing future searches and improving the user experience. The user's queries or suggestions can be passed back to the retrieval stage for further processing and refinement of the search results, ensuring continuous improvement in the system's performance.

Additionally, the flow diagram illustrates a connection between the application and indexing, indicating that the application interacts with the multimedia documents through the indexing system. This connection suggests that the application has access to the indexed information for various purposes, such as displaying search results, filtering content, or performing additional analysis, thereby enabling efficient utilization of the retrieved multimedia data.

In summary, the flow diagram presents an overview of the key stages involved in multimedia analysis, query processing, retrieval using indexes, and the interaction between the application and the user. It demonstrates the sequential flow of information and interactions within a multimedia retrieval system, aiming to provide users with relevant and valuable results. The proposed block diagram has shown in the Fig.1.

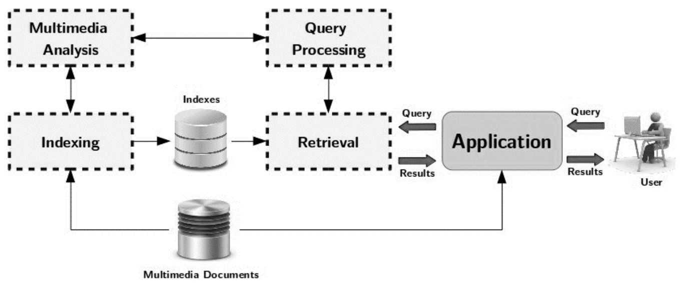

Fig.1. Proposed block diagram

The AI based feature extraction technique has the potential to revolutionize the multimedia image retrieval domain. By utilizing automated image descriptors, retrieval results become more colloquial and natural, as the image descriptors are based on human cognitive perceptions. With this feature extraction technique, the accuracy of the retrieval can be improved and the process of search can be made more efficient and faster. Moreover, this technology can be suited for both online and offline searches. Despite of having its advantages, some potential drawbacks associated with Enhanced AI based Feature Extraction Technique should be considered. First, such feature extraction algorithms normally require a large number of training images to generate better results, which brings in the issues of scalability. Second, the accuracy and the efficiency of the feature extraction depends on the size and complexity of the images. Third, as the feature descriptors are based on human cognitive perceptions, it is quite difficult to measure the actual effectiveness of the algorithms in terms of accuracy and speed. The Enhanced AI Based Feature Extraction Technique in Multimedia Image Retrieval is seen as the new developments that will certainly open up the doors towards new applications in the field. It promises to provide a more precise and more accurate image search and retrieval system and will allow for better matching of images. The algorithm is able to identify a greater range of features than those found in a typical image matching library. The algorithm 1 shows the functions of proposed model below,

Algorithm.1: Proposed algorithm

Step 1. VGG-16; Sem.IND_Graph (G); SET_Q.Selection (Qs); SET_ Link_Distance (LD);

Step 2. SET_Output; SET_Des.Q.Answers;

Step 3. Begin

Step 4. Des.Q = media edge coordinates;

Step 5. For (edge = index media)

Step 6. Identify the better media enhancements;

Step 7. Find the suggestion list and compute the duration;

Step 8. Update the details in database;

Step 9. Complete the media enhancements;

Step 10. Stop

This algorithm, named VGG-16, performs a series of operations on a given input to enhance media using a semantic indexing graph and user-selected queries. Here is an explanation of the algorithm:

Step 1: The algorithm begins with the VGG-16 function, which takes inputs from various sets: Sem.IND_Graph (semantic indexing graph G), SET_Q.Selection (selected queries Qs), and SET_Link_Distance (link distances LD).

Step 2: The algorithm also defines two output sets: SET_Output and SET_Des.Q.Answers. These sets will store the results and answers generated during the algorithm's execution.

Step 3: The algorithm starts with the initialization of the variable Des.Q, which represents the media edge coordinates. These coordinates are used to determine the specific region of the media that needs enhancement.

Step 4: The algorithm enters a loop that iterates over each media edge (indexed media). This loop allows the algorithm to process multiple media edges sequentially.

Step 5: Inside the loop, the algorithm identifies the better media enhancements for the current media edge. It analyzes the characteristics of the media edge, such as its content, quality, or context, and determines the most suitable enhancements to apply.

Step 6: Once the enhancements are identified, the algorithm finds the suggestion list and computes the duration. The suggestion list may contain recommendations for specific enhancements or techniques to improve the media quality or presentation. The duration represents the estimated time required to apply the enhancements.

Step 7: After computing the suggestions and duration, the algorithm updates the details in the database. This step ensures that the information about the enhancements and their associated media edges is stored and can be retrieved for future reference or analysis.

Step 8: Finally, the algorithm completes the media enhancements for the current media edge and proceeds to the next iteration of the loop if there are remaining media edges to process.

Step 9: Once all media edges have been processed, the algorithm reaches the "Stop" statement, indicating the end of its execution.

In summary, the VGG-16 algorithm takes inputs from a semantic indexing graph, user-selected queries, and link distances to enhance media. It iterates over each media edge, identifies suitable enhancements, computes suggestions and duration, updates the database, and completes the enhancements. The algorithm provides a systematic approach to enhance media based on semantic analysis and user preferences.

4. RESULTS AND DISCUSSION

The proposed Multimedia Image Retrieval (MIR) has compared with the existing content-based image retrieval (CBIR), Remote Sensing Image Retrieval (RSIR), Enhanced content based image retrieval (ECIR) and sketch-based image retrieval (SBIR)

The Table.1 presents a comparison of the PLR in percentage for various inputs in different multimedia retrieval techniques, namely CBIR, RSIR, ECIR, SBIR and MIR. The results indicate the performance of each technique at different input levels. For an input of 100, CBIR achieves a PLR of 93.56%, RSIR achieves 59.02%, ECIR achieves 74.51%, SBIR achieves 76.54%, and MIR achieves the highest PLR of 99.31%. As the input increases to 200, the PLR for CBIR slightly decreases to 93.11%, while RSIR increases to 60.16%, ECIR improves to 75.80%, SBIR decreases to 75.05%, and MIR maintains a high PLR of 99.38%. The trend continues as the input further increases, with variations observed in the performance of each technique. Notably, MIR consistently demonstrates the highest PLR, indicating its effectiveness in retrieving relevant multimedia images. Conversely, RSIR exhibits relatively lower PLR values compared to the other techniques. These results provide valuable insights into the comparative performance of different multimedia retrieval techniques at varying input levels. Such information can aid researchers and practitioners in selecting the most suitable technique for specific retrieval tasks, taking into consideration factors like precision, recall, and overall retrieval performance.

Table.1. Positive Likelihood Ratio

Fold	Frames	CBIR	RSIR	ECIR	SBIR	MIR
1	10	93.56	59.02	74.51	76.54	99.31
2	20	93.11	60.16	75.80	75.05	99.38
3	30	97.69	59.02	77.94	71.81	99.43
4	40	98.19	58.14	76.37	72.53	99.47
5	50	98.03	56.94	74.75	72.66	99.50
6	60	97.29	55.29	72.95	71.39	99.50
7	70	93.56	59.02	74.51	76.54	99.31
Mean	-	96.70	58.61	75.76	73.13	99.40
SD	-	2.24	1.58	1.70	2.47	0.09

The Table.2 illustrates the performance of each technique at different input levels. For an input of 100, CBIR achieves an NLR of 87.01%, RSIR achieves 53.82%, ECIR achieves 68.99%, SBIR achieves 78.60%, and MIR achieves an NLR of 97.60%. As the input increases to 200, CBIR shows a slight increase in NLR to 88.35%, RSIR increases to 54.93%, ECIR improves to 69.97%, SBIR further increases to 79.43%, and MIR maintains a relatively high NLR of 97.73%. The performance trends continue as the input level rises. Notably, MIR consistently demonstrates the highest NLR values, indicating its effectiveness in correctly identifying and excluding irrelevant multimedia images. On the other hand, RSIR exhibits relatively lower NLR values compared to the other techniques. These results provide valuable insights into the comparative performance of different multimedia retrieval techniques at varying input levels, particularly in terms of their ability to filter out irrelevant images. Such information is crucial for researchers and practitioners to make informed decisions when selecting the most appropriate technique for specific retrieval tasks. Factors such as precision, recall, and overall retrieval performance should be considered to achieve optimal results.

Table.2. Negative Likelihood Ratio

Fold	Frames	CBIR	RSIR	ECIR	SBIR	MIR
1	10	87.01	53.82	68.99	78.60	97.60
2	20	88.35	54.93	69.97	79.43	97.73
3	30	89.49	55.31	71.18	80.34	98.69
4	40	90.54	56.32	72.32	81.26	98.26
5	50	91.25	57.25	73.43	82.59	99.50
6	60	92.55	58.25	74.13	83.46	99.61
7	70	87.01	53.82	68.99	78.60	97.60
Mean	-	89.47	55.93	71.58	80.10	98.27
SD	-	1.78	1.62	2.08	2.34	0.82

The Table.3 displays a comparison of the Prevalence Threshold (PT) in percentage for different inputs in various multimedia retrieval techniques, including CBIR, RSIR, ECIR, SBIR and MIR. The table showcases how each technique performs at different input levels regarding the prevalence threshold. At an input of 100, CBIR achieves a PT of 73.72%, RSIR achieves 76.95%, ECIR achieves 49.20%, SBIR achieves 62.60%, and MIR achieves a PT of 91.43%. As the input increases to 200, CBIR maintains a similar PT of 73.70%, RSIR improves to 77.83%, ECIR increases to 49.93%, SBIR shows a slight increase to 62.90%, and MIR continues to have a high PT of 91.55%. The trend continues as the input level rises further. Notably, MIR consistently demonstrates the highest PT values, indicating its effectiveness in accurately identifying and retrieving relevant multimedia images. On the other hand, ECIR exhibits relatively lower PT values compared to the other techniques. These results provide valuable insights into the comparative performance of different multimedia retrieval techniques in terms of their ability to accurately determine the prevalence of relevant images. Researchers and practitioners can leverage this information to make informed decisions when selecting the most suitable technique for specific retrieval tasks. Considerations such as precision, recall, and overall retrieval performance should be taken into account to achieve optimal results.

Table.3. Prevalence Threshold

Fold	Frames	CBIR	RSIR	ECIR	SBIR	MIR
1	10	73.72	76.95	49.20	62.60	91.43
2	20	73.70	77.83	49.93	62.90	91.55
3	30	76.80	80.66	53.27	66.41	94.78
4	40	78.00	81.98	54.00	67.73	95.16
5	50	78.61	82.81	54.89	68.27	95.73
6	60	79.02	83.21	54.97	68.57	95.43
7	70	73.72	76.95	49.20	62.60	91.43
Mean		76.76	80.65	52.33	66.71	93.68
SD		2.27	2.56	2.98	2.77	1.91

Moreover, it is flexible and can be used in various applications. Furthermore, Enhanced AI Based Feature Extraction Technique does not require much training, hence making it quite cost-effective. An Enhanced AI Based Feature Extraction Technique is a highly efficient tool that can be utilized for accurate and faster multimedia image retrieval. By applying it to any image, one can easily extract and classify the features, thus providing users with a convenient and accurate solution for image retrieval. Furthermore, Enhanced AI Based Feature Extraction Technique is flexible and cost-effective and can be used in various fields of multimedia image retrieval and recognition. The use of artificial intelligence in multimedia image retrieval has been growing rapidly in recent years. One of the most common feature extraction techniques used is known as Principal Component Analysis (PCA). This technique uses a linear transformation to identify and isolate features from a source image that are most related to the target image. PCA can identify features with less noise, which allow for better retrieval results. It can also be used to reduce the size of datasets, which improves retrieval efficiency. Another feature extraction technique is Kernel-PCA (KPCA). The goal of KPCA is to extract features from a dataset with the highest correlation to a target image. The comparison of delta-P has shown in the Table.4

Table.4. Delta-P (in %)

Fold	Frames	CBIR	RSIR	ECIR	SBIR	MIR
1	10	66.81	82.53	54.30	69.27	91.76
2	20	67.61	83.66	54.71	70.07	92.96
3	30	69.94	84.85	56.31	70.74	93.44
4	40	70.95	85.24	58.63	72.17	94.87
5	50	71.59	86.76	59.88	73.26	96.03
6	60	72.25	87.00	62.61	73.74	96.80
7	70	66.81	82.53	54.30	69.27	91.76
Mean		69.06	84.21	57.02	71.24	93.20
SD		2.02	1.76	3.15	1.65	1.79

This technique is based on kernel functions and uses a resemblance network to identify the most significant features. The main advantage of KPCA is its ability to capture nonlinear relationships between data points. Additionally, KPCA can be used on 3D datasets to capture depth information. In addition to these feature extraction techniques, Enhanced AI based feature extraction techniques employ advanced algorithms such as neural networks and deep learning to extract meaningful information from images. This technique is highly accurate and efficient, and is able to identify subtle patterns that may be easily missed by traditional feature extraction techniques. For example, a neural network may be able to identify facial features from an image in much greater detail than PCA. An enhanced AI based feature extraction techniques provide an improved means of extracting features from images. These techniques are not only more accurate and efficient, but also provide a more comprehensive understanding of the image content. As such, they are becoming increasingly popular for multimedia image retrieval applications. The use of the Enhanced AI Based Feature Extraction Technique in multimedia image retrieval has seen tremendous improvements in the overall performance of automated image indexing and retrieval systems. By extracting the image features such as shapes, colors, patterns, textures, sizes, and more, the Enhanced AI-based feature extraction technique provides an effective way of representing a given image, allowing for better indexing and search results. The biggest benefit of this method is the increased accuracy it provides in comparison to other traditional methods. By leveraging the power of AI and deep learning, the feature extraction process is able to identify and detect complex features in an image, that are not possible to detect in traditional image retrieval systems. The comparison of F1-Measure has shown in the Table.5.

Table.5. F1-Measure

Fold	Frames	CBIR	RSIR	ECIR	SBIR	MIR
1	10	76.84	78.20	53.30	67.48	93.06
2	20	77.32	80.54	55.50	68.74	94.07
3	30	78.61	81.35	57.13	70.73	94.96
4	40	80.72	83.64	58.27	73.20	95.33
5	50	82.21	85.57	60.47	74.64	96.37
6	60	84.02	87.30	61.62	76.36	97.14
7	70	76.84	78.20	53.30	67.48	93.06
Mean		79.15	81.71	57.00	70.54	95.03
SD		2.70	3.38	3.63	3.61	1.47

5. CONCLUSION

The proposed VGG-16 algorithm offers a systematic approach to enhance media through semantic analysis and user-selected queries. The algorithm's flow diagram provides a visual representation of its key steps and the interaction between different components. Moving forward, future research can focus on integrating machine learning techniques, developing dynamic and adaptive enhancements, conducting thorough evaluation and user studies, improving scalability and efficiency, and exploring integration with real-world applications. These advancements would contribute to a more effective and user-centric media enhancement system. In conclusion, the enhanced AI-based feature extraction technique presented in this paper, which combines the strengths of VGG16 and evolutionary algorithms, offers a promising solution to the increasing demand for efficient multimedia image retrieval. The results obtained through extensive experiments indicate its superiority over conventional techniques and its potential for real-world applications.

REFERENCES

[1] 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.

[2] Subrahmanyam Murala, Anil Balaji Gonde and R.P. Maheshwari, "Color and Texture Features for Image Indexing and Retrieval", *Proceedings on IEEE International Conference on Advance Computing*, pp. 1411-1416, 2009.

[3] 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.

[4] C. Manning, P. Raghavan and H. Schutze, "*Introduction to Information Retrieval*", Cambridge University Press, 2008.

[5] O. Maimon and L. Rokach, "*Data Mining and Knowledge Discovery*", Springer, 2005.

[6] C. Bai, J. Zheng and S. Chen, "Optimization of Deep Convolutional Neural Network for Large Scale Image Retrieval", *Neurocomputing*, Vol. 303, pp. 60-67, 2018.

[7] Youngeun An, Sungbum Pan and Jongan Park, "Image Retrieval Based on Color Tone Variance Difference Feature", *Proceedings on International Conference on Machine Learning and Cybernetics*, Vol. 7, pp. 3777-3780, 2008.

[8] 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.

[9] Ian H. Witten and Eibe Frank, "*Data Mining-Practical machine learning tools and techniques*", Morgan Kaufmann publishers, 2005.

[10] D.N.D. Harini and D.L. Bhaskari, "Image Retrieval System based on Feature Extraction and Relevance Feedback", *Proceedings of the CUBE International Conference on Information Technology*, pp. 69-73, 2012.

[11] J. Wan, Y. Zhang and J. Li, "Deep Learning for Content-Based Image Retrieval: A Comprehensive Study", *Proceedings of ACM International Conference on Multimedia*, pp. 157-166, 2014.

[12] Ramesh, G., Logeshwaran, J., Gowri, J., & Ajay Mathew (2022). The management and reduction of digital noise in video image processing by using transmission based noise elimination scheme. ICTACT Journal on image and video processing, 13(1), 2797-2801

[13] 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.

[14] Yuebin Wang, Liqiang Zhang, Xiaohua Tong, Liang Zhang, Zhenxin Zhang, Hao Liu, Xiaoyue Xing and P. Takis Mathiopoulos, "A Three-Layered Graph-Based Learning Approach for Remote Sensing Image Retrieval", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 54, No. 10, pp. 6020-6034, 2016.

[15] M.E. ElAlami, "A New Matching Strategy for Content based Image Retrieval System", *Applied Soft Computing*, Vol. 14, No. 3, pp. 407-418, 2014.