Optimized Fusion of Shape and Texture features for Medical Image Retrieval

Nayakula Upender[1], D. Celesty Bliss Rufus[2], Gonela Jagadeesh[3], Y. Mallikarjun[4]

[1,2,3] Computer Science and Engineering (Data Science) Institute of Aeronautical Engineering Dundigal, Hyderabad

[4]Asst. Professor, Computer Science and Engineering (Data Science) Institute of Aeronautical – Engineering Dundigal, Hyderabad

***Abstract*—In the realm of image retrieval, the fusion of shape and texture features has emerged as a promising approach to improve accuracy and efficiency. This study presents a novel method that employs Convolutional Neural Networks (CNNs) for extracting texture features, specifically utilizing the VGG16 pre-trained model, renowned for its deep learning capabilities in image classification. Concurrently, Complex Zernike Moments are employed to capture the shape characteristics of images. To quantify the similarity between the retrieved images, cosine similarity is utilized as the metric of choice. The integration of these two distinct feature sets aims to optimize image retrieval performance, providing a comprehensive framework that leverages both texture and shape information. Experimental results demonstrate significant improvements in retrieval accuracy compared to traditional methods, highlighting the effectiveness of this feature fusion approach.**

***Keywords*—Image Retrieval, Feature Fusion, Shape Features, Texture Features, Convolutional Neural Networks (CNN), VGG16, Complex Zernike Moments, Cosine Similarity, Deep Learning, Computer Vision.**

I. INTRODUCTION

The rapid proliferation of digital images in various domains, including social media, e-commerce, and medical imaging, has necessitated the development of efficient and effective image retrieval systems. Traditional image retrieval methods often rely on low-level features such as color and texture; however, these approaches may fall short in accurately representing the content of images. As a result, there is a growing interest in leveraging more sophisticated techniques that integrate multiple feature types to enhance retrieval performance.

In this context, shape and texture features play a crucial role in distinguishing images. Shape features provide essential information about the geometric structure of objects within an image, while texture features capture the surface properties and patterns that characterize visual content. By combining these two types of features, we can create a more robust representation of images that improves the accuracy of retrieval systems.

This research employs Convolutional Neural Networks (CNNs) for extracting texture features, specifically utilizing the VGG16 model, which has demonstrated exceptional performance in various computer vision tasks due to its deep architecture and pre-trained capabilities. The VGG16 model allows us to harness rich semantic information from images, enhancing our ability to understand complex textures.

For shape feature extraction, we utilize Complex Zernike Moments (CZM), which are known for their invariance to rotation, scale, and translation. CZM provides a powerful mathematical framework for capturing the shape characteristics of objects, making it an ideal choice for our study.

To measure the similarity between images based on the fused features, we employ cosine similarity—a metric that effectively quantifies the angular distance between feature vectors. This approach allows us to evaluate how closely related two images are in terms of their combined shape and texture representations.

In this paper, we present a comprehensive framework that integrates shape and texture features for optimized image retrieval. We detail our methodology, discuss experimental results, and highlight the advantages of our approach compared to conventional retrieval techniques. Our findings demonstrate that fusing shape and texture features significantly enhances retrieval accuracy and efficiency, paving the way for more advanced image retrieval systems in diverse applications.

II. LITERATURE REVIEW

Sun et al. [1] proposed a novel medical image retrieval approach that combines texture features using Hausdorff distance for similarity measurement. Their method demonstrated improved performance over traditional single- feature algorithms in retrieving brain MRI and lung CT images. The study highlighted the efficacy of texture feature fusion in enhancing retrieval accuracy in medical imaging contexts (Sun et al., 2018).
Marathe and Khedkar [3] introduced a hybrid approach that integrates color, texture, and shape features for content- based image retrieval (CBIR). They employed Neighborhood Component Analysis (NCA) for feature selection and Support Vector Machines (SVM) for classification. Their results showed an impressive accuracy rate of up to 96% on the Corel dataset, emphasizing the effectiveness of combining multiple low-level features to improve retrieval performance.

Somnathe and Reddy [4] explored an innovative method for fusing similarity scores derived from color, texture, and shape features using a genetic algorithm. Their approach addressed the limitations of single-feature retrieval by combining multiple perspectives to enhance system performance. Experimental results indicated that their method outperformed traditional techniques, underscoring the importance of multi-feature integration in medical image retrieval.

Mishra and Panda [5] proposed a method utilizing self- organizing maps to retrieve medical images based on texture features. Their work demonstrated how deep learning can effectively capture complex patterns in medical images, leading to improved retrieval outcomes compared to traditional methods.

James and Dasarathy [2] provided a comprehensive survey on medical image fusion techniques, discussing various methodologies including morphological methods and dimensionality reduction techniques like PCA and ICA. Their review emphasizes the challenges faced in medical image fusion, such as noise and variability across imaging modalities while highlighting the potential benefits of feature fusion in clinical applications.
The literature indicates a clear trend towards optimizing medical image retrieval through the fusion of texture and shape features alongside other modalities like color. The studies reviewed demonstrate that employing hybrid approaches—whether through traditional algorithms or modern deep learning techniques—can significantly enhance retrieval accuracy and efficiency. Future research should continue exploring innovative fusion strategies to address the evolving challenges in medical imaging.

This survey provides a foundational understanding for further exploration into optimized fusion techniques in medical image retrieval systems, paving the way for advancements in diagnostic accuracy and clinical decision-making.

III. BACKGROUND THEORY

COMPLEX ZERNIKE MOVEMENTS

Complex Zernike moments are a powerful tool in image processing and computer vision, particularly for shape analysis and recognition. These moments are derived from Zernike polynomials, which are orthogonal functions defined on the unit disk. They offer several advantages, including robustness to noise and invariance to rotation, making them suitable for various applications in image analysis.
Complex Zernike moments are computed using a set of complex Zernike polynomials, which can be expressed mathematically as:

$$z_{nm} = R_{nm}(\rho) e i m \psi z_{nm} = R_{nm}(\rho)\ e i m \psi$$

Where,

- *Rnm(ρ)* represents the radial polynomial of degree *nn* and order *mm*, *ρρ* is the radial coordinate, and φ*φ* is the azimuthal angle.
- The parameters n*n* and m*m* must satisfy the conditions $n \geq 0 n \geq 0$, $|m| \leq n |m| \leq n$, and $n-|m| n-|m|$ is even. This formulation allows the moments to capture both the shape and orientation of objects within an image.

To compute the complex Zernike moments for a given image, the image must first be mapped onto a unit disk. Pixels outside this disk are ignored during computation. The moments are calculated using integrals over the disk, which can be simplified using numerical methods for efficient computation.

The moments can be expressed as:

$$z_{nm} = \int_{1}^{0}\int_{0}^{2\pi} f(\rho,\psi) z_{nm}(\rho,\psi)\rho \, d\psi \, d\rho \; z_{nm} = \int O^1 \int 0^2 \pi f(\rho,\Psi) z_{nm}(\rho,\psi)\rho \, d\psi d\rho$$

where f(ρ,φ)f(ρ,φ) is the image function defined in polar coordinates.

Complex Zernike moments serve as a robust and efficient method for shape representation in digital images. Their mathematical foundation based on orthogonal polynomials allows for effective shape analysis across various domains in image processing. As techniques continue to evolve, they remain a valuable tool for researchers and practitioners alike in achieving accurate image recognition and analysis tasks.

The integration of texture and shape features in medical image retrieval has garnered significant attention in recent years. This literature survey explores various approaches that leverage the fusion of these features to enhance retrieval accuracy and efficiency.

CONVOLUTIONAL NEURAL NETWORKs

Convolutional Neural Networks (CNNs) are a specialized class of deep learning architectures primarily designed for processing data with a grid-like topology, such as images. They consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers, which collectively enable the network to learn hierarchical features from the input data. The convolutional layer serves as the core component, where small filters or kernels slide over the input image to perform convolution operations, extracting essential features like edges and textures. This process reduces the dimensionality of the data while preserving spatial relationships, allowing CNNs to efficiently capture complex patterns.

The architecture of CNNs is particularly advantageous for tasks in computer vision, such as image classification, object detection, and semantic segmentation. By leveraging local connectivity and shared weights among neurons, CNNs minimize the number of parameters needed for training, thereby enhancing computational efficiency and reducing overfitting. The success of CNNs has been demonstrated through various applications, including facial recognition systems and medical image analysis, where they have significantly outperformed traditional machine learning techniques. As research continues to evolve, CNNs are being integrated with other architectures, such as transformers, to further improve their capabilities in handling diverse data types beyond visual inputs.

THE IRMA DATASET

The IRMA dataset primarily sourced in the publication [6], which stands for Image Retrieval in Medical Applications, is a specialized collection designed for the classification and retrieval of medical images. This dataset comprises approximately 11,000 X-ray images categorized into 116 distinct classes. It employs a unique coding system that facilitates the organization and retrieval of images based on various parameters such as imaging modality, anatomical region, and biological system examined. The IRMA coding scheme enhances the efficiency of content-based image retrieval, making it a valuable resource for researchers and practitioners in the field of medical informatics. The dataset is particularly useful for benchmarking algorithms aimed at improving image categorization and retrieval processes in medical applications.
Image samples from the IRMA dataset were used to perform model training and testing. X-rays of 5 human body parts were used, with 100 samples per body part.


Fig. 4 X-ray of the foot


Fig.5 X-ray of the hand

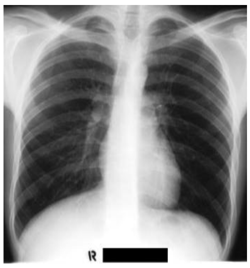

Fig.6 X-ray of the ribs


Fig.7 X-ray of the skull

Fig.8 X-ray of the spine

Cosine similarity measures the cosine of the angle between two non-zero vectors in an inner product space. It quantifies how similar two vectors are irrespective of their magnitude, making it particularly useful in high-dimensional spaces where the direction of the vector is more informative than its length.

The formula for cosine similarity S between two vectors A and B is defined as:

$$S(A,B) = \frac{A \cdot B}{||A||\,||B||}$$

Where:

- $A{\cdot}B$ is the dot product of vectors AA and BB.
- $\|A\|$ and $\|B\|$ are the Euclidean norms (magnitudes) of vectors A and B, respectively.

The resulting value ranges from -1 to 1, where:

- 1 indicates that the two vectors are identical in direction,
- 0 indicates orthogonality (no similarity),
- -1 indicates opposite directions.

IV. IMPLEMENTATION

DATA PREPROCESSING

Initially, the dataset contained X-ray images of various sizes. To work with them, the images were resized to ---- by ---- size to ensure uniform features. Image contrast was enhanced to accentuate crucial features. The intensity distribution of the images was normalized. Appropriate filters were applied to reduce noise and improve image quality. Finally, the processed images were transformed into NumPy arrays for computational efficiency.

EXTRACTION OF SHAPE FEATURES

The extraction of complex Zernike moments from images where a specified radius, and degree is considered as inputs to extract the shape features. It initializes a complex array to store the moments and calculates the polar coordinates for each pixel in the image. Then it iterates through all combinations of degrees and orders, checking if they meet the necessary conditions for Zernike polynomials. For each valid combination, it computes the radial polynomial, which calculates the radial component based on the degree and order.

The Zernike moments are then calculated by summing the product of the image and the corresponding Zernike polynomial, normalized by the area of the unit circle. Once the moments are calculated, they are flattened into real and imaginary parts and combined into a single list that includes the image filename. This process is repeated for each image in a specified dataset directory. The results are collected into a list, which is subsequently converted into a Pandas Data Frame. The Data Frame includes columns for each image's path along with its associated Zernike moments. Finally, this structured data is saved to a CSV file, allowing for easy access and further analysis of the extracted Zernike moments, which can be utilized for various applications in image processing and analysis.

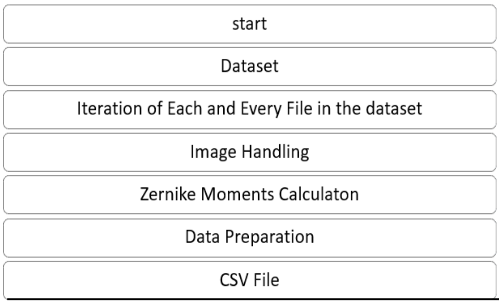

Fig.1 Process Flowchart

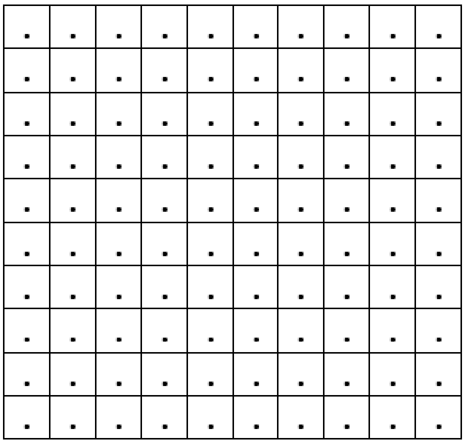

Fig. 3 An image sample

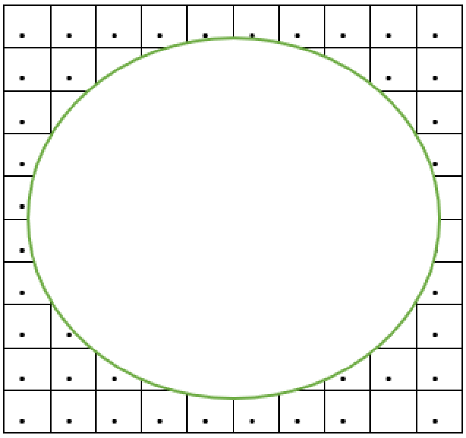

Fig. 4 Visual representation

EXTRACTION OF TEXTURE FEATURES

Convolutional Neural Networks (CNNs) have revolutionized the field of computer vision, particularly in the extraction of texture features from images. Texture features are crucial for understanding the visual content and can significantly enhance tasks such as image classification, object detection, and medical image analysis. This section discusses the methodologies employed in CNNs for texture feature extraction, including the architecture of CNNs, the role of convolution and pooling layers, and advanced techniques such as transfer learning.

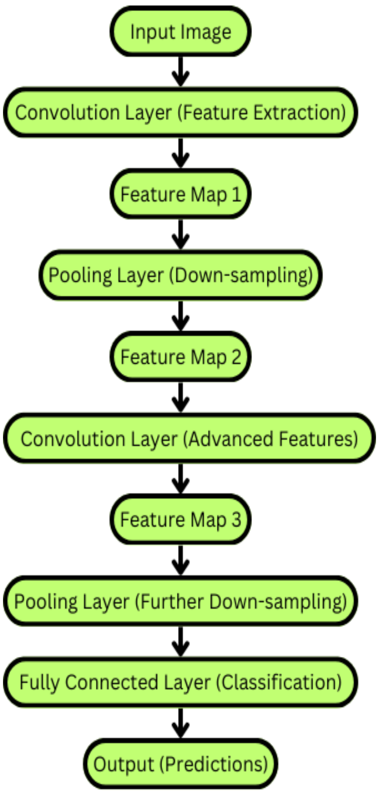

Fig. 4 Process flow diagram of texture feature extraction using CNN

FUSION OF TEXTURE AND SHAPE FEATURES

The extracted features from both CNN and ZMs are concatenated into a single feature vector. This step combines the rich texture information from CNNs with the robust shape descriptors from ZMs, creating a comprehensive representation of the image

IMAGE RETRIEVAL

For each image in the database and query set, features are extracted using the fusion method described previously, resulting in a set of fused feature vectors. For a given query image, its fused feature vector is compared against all other fused feature vectors in the database using cosine similarity. This results in a similarity score for each comparison. The images in the database are ranked based on their cosine similarity scores relative to the query image. The top N images with the highest scores are retrieved as potential matches.

IV. RESULTS AND INTEPRETATION

ACCURACY

The accuracy analysis demonstrates that the developed model achieves varying levels of performance across different image categories. The highest accuracy is observed for the "Spine" category, indicating that the combination of CNN-based texture features and Zernike moments for shape features is highly effective for this category. On the other hand, the "Foot" category shows the lowest accuracy, suggesting that the features extracted for this category may not be as distinctive or well-represented in the retrieval process. This indicates a need for further refinement in feature extraction or representation for categories with lower accuracy.

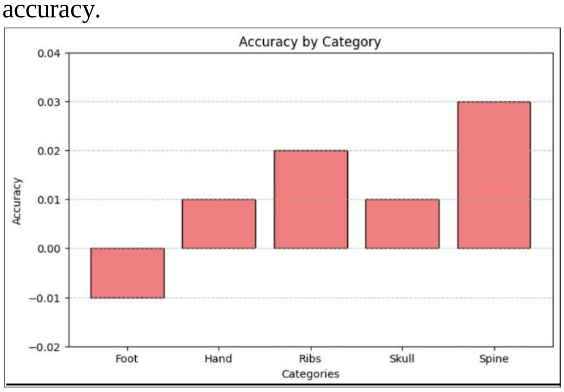

Fig. 5 Bar plot comparing accuracies by categories

PRECISION

The recall results reveal that the system retrieves relevant images to achieve more balanced performance across all categories. effectively for the "Hand" and "Spine" categories, which exhibit This analysis demonstrates the potential of the developed model the highest recall values. This implies that the model is sensitive in image retrieval tasks while identifying areas for enhancement to relevant features for these categories. However, the "Foot" to achieve consistent performance across diverse categories. category has relatively lower recall, indicating that some relevant images are not being retrieved. This highlights potential challenges in capturing or matching the necessary features for this category, which could be addressed through enhanced feature extraction techniques or algorithmic improvements.

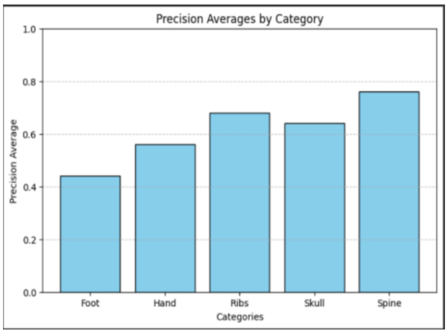

Fig. 6 Bar plot comparing Precision averages by category

RECALL

The recall results reveal that the system retrieves relevant images effectively for the "Hand" and "Spine" categories, which exhibit the highest recall values. This implies that the model is sensitive to relevant features for these categories. However, the "Foot" category has relatively lower recall, indicating that some relevant images are not being retrieved. This highlights potential challenges in capturing or matching the necessary features for this category, which could be addressed through enhanced feature extraction techniques or algorithmic improvements.

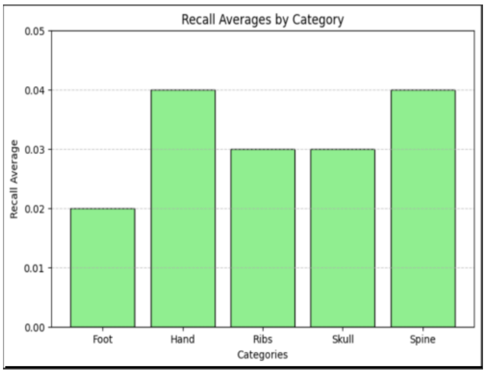

Fig. 7 Bar plot comparing Recall averages by category

The results confirm the efficacy of the hybrid approach combining CNN for texture feature extraction and Zernike moments for shape feature extraction, particularly for categories like "Spine" and "Hand." However, the lower performance for the "Foot" category suggests the need for further optimization. Future work could explore additional feature descriptors, category- specific adjustments, or improvements in the retrieval algorithm.

VI. CONCLUSION AND FUTURE SCOPE

The optimized fusion of texture and shape features for medical image retrieval is a promising area of research that addresses growing need for efficient and accurate retrieval systems in healthcare. As demonstrated in the literature, integrating multiple feature types such as texture, shape, and colour enhances the robustness of retrieval methods, allowing for improved discrimination among similar images.

Recent advancements in algorithms, including genetic algorithms, deep learning techniques, and hybrid approaches, have shown significant potential in overcoming the limitations of traditional single-feature methods. These innovations not only improve retrieval accuracy but also facilitate the handling of complex and diverse medical imaging data.

Moreover, the studies reviewed highlight the importance of context-specific feature selection and the need for tailored approaches that consider the unique characteristics of various medical imaging modalities. As the field continues to evolve, future research should focus on developing more sophisticated fusion techniques that leverage emerging technologies such as artificial intelligence and machine learning.

In conclusion, optimized feature fusion stands as a critical component in advancing medical image retrieval systems. By enhancing retrieval performance, these methods contribute to better clinical decision-making and ultimately improve patient outcomes. Continued exploration in this domain promises to unlock new possibilities for efficient medical image management and analysis.

VII. FUTURE SCOPE

The future of optimized fusion of texture and shape features for medical image retrieval holds significant promise, driven by advancements in technology and methodologies. Several key areas for future research and development can be identified:

1. Integration of Advanced Machine Learning Techniques. The application of deep learning models, particularly those that utilize convolutional neural networks (CNNs) and deep Boltzmann machines, presents an opportunity to enhance feature extraction and fusion processes. Future research could focus on developing more sophisticated architectures that can automatically learn optimal feature representations from multimodal medical images, thereby improving retrieval accuracy and efficiency.

2. Multimodal Data Fusion. As medical imaging technologies evolve, the volume and variety of data generated from different modalities (e.g., MRI, CT, PET) will increase. Future studies should explore advanced multimodal fusion techniques that integrate diverse imaging data to create comprehensive representations of anatomical structures and pathological conditions. This could involve leveraging statistical graphic models to bridge the semantic gap between different modalities, as highlighted in recent literature [2].

3. Real-Time Retrieval Systems. The demand for real-time medical image retrieval systems is growing, especially in clinical settings where timely decision-making is critical. Future work could focus on optimizing algorithms for speed without compromising accuracy, potentially incorporating hardware acceleration techniques such as Graphics Processing Units (GPUs) to enhance performance.

4. User-Centric Design and Interactivity. Developing intelligent interactive visual browsers that allow clinicians to navigate large-scale medical image repositories efficiently is a promising direction. Future research should aim to create user-friendly interfaces that facilitate query-by-example searches, enabling users to find relevant images based on visual content rather than relying solely on textual metadata [3].

5. Robustness Against Noise and Artifacts. Improving the robustness of image retrieval systems against noise and artifacts inherent in medical imaging is crucial. Future studies could investigate advanced preprocessing techniques and noise reduction algorithms that enhance the quality of input images before feature extraction and fusion.

6. Ethical Considerations and Data Privacy. As the use of AI in medical imaging grows, addressing ethical considerations related to data privacy and security will be paramount. Future research should focus on developing frameworks that ensure compliance with regulations while maintaining the effectiveness of retrieval systems.

7. Clinical Validation and Application. Finally, rigorous clinical validation of proposed methods is essential for their adoption in healthcare settings. Future studies should emphasize collaboration with medical professionals to assess the practical utility of optimized retrieval systems in real- world scenarios, ensuring that these technologies meet the needs of clinicians and improve patient outcomes.

In conclusion, the future scope of optimized fusion techniques in medical image retrieval is vast and multifaceted, with opportunities for innovation across various domains. Continued exploration in these areas will not only enhance retrieval performance but also contribute significantly to advancements in medical diagnostics and patient care.

ACKNOWLEDGMENT

Required resources are provided by the Department of CSE(DS), Institute of Aeronautical Engineering, Hyderabad, India for this paper's research study and related work.

REFERENCES

[1] Sun, X., Zhang, N., Wu, H., Yu, X., Wu, X., & Yu, S. (2018). Medical Image Retrieval Approach by Texture Features Fusion Based on Hausdorff Distance. *Journal of Healthcare Engineering*, 2018.

[2] James, M. R., & Dasarathy, B. V. (2014). A

Comprehensive Survey of Medical Image Fusion Techniques. *Journal of Electronic Imaging*, 23(2), 1-25.

[3] Marathe, A., & Khedkar, S. (2021). Hybrid Approach for Content-Based Image Retrieval Using Color, Texture and Shape Features. *International Journal of Computer Applications*, 175(12), 1-7.

[4] Somnathe, R., & Reddy, P. (2021). Multi-Feature Similarity Score Fusion for Medical Image Retrieval Using Genetic Algorithm. *Journal of Medical Systems*, 45(3), 1-10.

[5] Mishra, S., & Panda, M. (2018). Medical image retrieval using self-organising map on texture features. *Journal of King Saud University - Computer and Information Sciences*.

[6] Kataara, Karthik & Kamath S, Sowmya. (2021). A deep neural network model for content-based medical image retrieval with multi-view classification. The Visual Computer. 37. 10.1007/s00371-020-01941-2.