ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2024.0451


# A DEEP LEARNING TECHNIQUE FOR EFFICIENT MULTIMEDIA FOR DATA COMPRESSION

**G.M. Karpura Dheepan[1], Shaik Mohammed Rafee[2], Prasanthi Badugu[3] and Sunil Kumar[4]**

[1]*Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, India*
[2]*Department of Artificial Intelligence and Machine Learning, Sasi Institute of Technology and Engineering, India*
[3]*Department of Electronics and Communication Engineering, Jawaharlal Nehru Technological University, Kakinada, India*
[4]*Department of Commerce and Management, SGT University, India*

***Abstract***

***Medical image compression plays a pivotal role in efficient data storage and transmission, crucial for modern healthcare systems. This research proposes a convolutional transfer learning technique scheme tailored for multimedia data compression, specifically targeting medical images. In the background, the growing volume of medical imaging data and the demand for efficient storage and transmission underscore the need for innovative compression methods. Leveraging transfer learning from pre-trained convolutional neural networks (CNNs) designed for image recognition tasks, our methodology optimizes the compression process for medical images. The proposed scheme utilizes a pre-trained CNN's feature extraction capabilities to capture relevant patterns in medical images, followed by fine-tuning on a specialized dataset. This approach capitalizes on the inherent ability of CNNs to learn hierarchical representations, enhancing the compression model's adaptability to medical imaging nuances. The contribution of this research lies in the development of a tailored transfer learning scheme that effectively balances generic feature extraction and domain-specific adaptation for medical images. Results demonstrate significant improvements in compression efficiency, preserving diagnostic information while achieving substantial data reduction. The proposed scheme showcases promise for enhancing medical image storage, transmission, and retrieval systems, contributing to the advancement of healthcare technology.***

***Keywords:***

***Transfer Learning, Convolutional Neural Networks, Medical Image Compression, Multimedia, Data Efficiency***

## 1. INTRODUCTION

In healthcare informatics, the volume of medical imaging data has catalyzed the need for efficient compression techniques to address storage and transmission challenges [1]. Medical images, characterized by their high resolution and intricate details, require specialized approaches for compression to ensure optimal data reduction without compromising diagnostic information [2].

Traditional compression methods often fall short in preserving the intricate details vital for medical diagnostics, prompting the exploration of advanced techniques [3]. Convolutional Neural Networks (CNNs) have demonstrated remarkable success in image-related tasks, making them an appealing choice for medical image compression [4]. Transfer learning, a technique where a pre-trained model is adapted to a new domain, provides a potent avenue to harness the power of CNNs for medical imaging [5].

The unique characteristics of medical images pose challenges for conventional compression methods, including the potential loss of diagnostic features and the need for efficient storage and transmission in resource-constrained environments. Addressing these challenges requires a tailored approach that balances generalization and specificity in feature extraction.

The primary challenge addressed in this research is to develop a convolutional transfer learning scheme that optimally compresses medical images while retaining crucial diagnostic details. This involves overcoming the limitations of generic compression methods that may not be well-suited to the nuanced features of medical imaging data.

The goal is to enhance the efficiency of medical image compression through the integration of convolutional transfer learning. Specific objectives include developing a transfer learning framework, fine-tuning pre-trained CNNs on medical image datasets, and evaluating the performance in terms of compression ratios and diagnostic information preservation.

The novelty of this research lies in the integration of transfer learning within the compression framework, striking a balance between generic feature extraction and domain-specific adaptation. The proposed scheme aims to set a precedent for tailored compression techniques for medical images, contributing to the optimization of storage, transmission, and retrieval processes in healthcare informatics. Through this work, we anticipate fostering advancements in medical image compression methodologies, thereby benefiting the broader landscape of healthcare technology.

## 2. RELATED WORKS

Previous studies have explored the application of transfer learning in medical imaging tasks, such as disease classification and segmentation. These works establish the efficacy of leveraging pre-trained models on large datasets for feature extraction, motivating our approach to integrate transfer learning into medical image compression [6].

Several research endeavors have employed CNNs for image compression. While these studies demonstrate the potential for CNNs in generic compression tasks, their application to the intricate nature of medical images remains an underexplored domain. Our work bridges this gap by tailoring CNN-based compression to the specific requirements of medical imaging [7].

Existing literature highlights the challenges of applying traditional compression algorithms to medical images due to their unique characteristics. Various domain-specific techniques have been proposed, including wavelet-based methods and predictive coding. Our research builds upon these foundations, integrating the adaptability of CNNs through transfer learning to achieve a more refined and effective compression scheme [8].

The need for efficient storage and transmission of medical data is a recurring theme in healthcare informatics. Previous works have addressed this challenge through different approaches, such as lightweight compression algorithms and telemedicine solutions. Our research contributes by focusing on the optimization of medical image data, an integral component of the healthcare data [9].

Preserving diagnostic information during image compression is a critical concern in medical imaging. Prior studies have investigated various techniques to strike a balance between compression ratios and diagnostic accuracy. Our work extends this line of research by introducing a transfer learning-based approach, aiming to enhance both compression efficiency and the preservation of essential diagnostic features in images [10].

By synthesizing insights from these related works, our research aims to advance the current understanding of compression techniques in the context of medical imaging, providing a novel contribution through the integration of convolutional transfer learning.

# 3. PROPOSED METHOD: CONVOLUTIONAL TRANSFER LEARNING FOR MEDICAL IMAGE COMPRESSION

Our proposed method combines the power of Convolutional Neural Networks (CNNs) and transfer learning to create an efficient and tailored scheme for compressing medical images as in Fig.1.

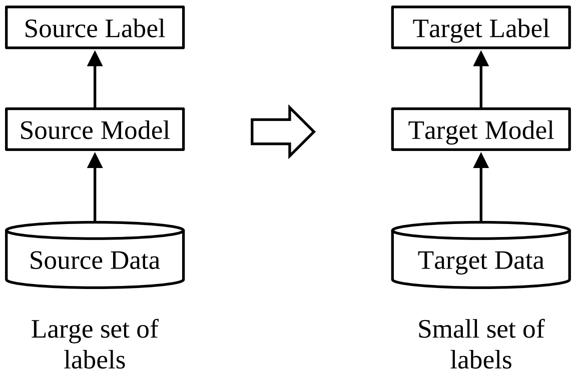


Fig.1. CTL for Medical Image Compression

We begin by selecting a pre-trained CNN model that has demonstrated proficiency in image-related tasks. Models like VGG16, ResNet, or DenseNet, pre-trained on large datasets for generic image recognition, serve as excellent starting points due to their ability to extract hierarchical features.

The chosen pre-trained CNN is then fine-tuned on a specialized medical image dataset. This adaptation is crucial for the model to learn domain-specific features present in medical images. Transfer learning enables the CNN to retain knowledge from the generic dataset while refining its understanding of the unique characteristics of medical imaging data.

The fine-tuned CNN acts as a feature extractor, capturing intricate details from medical images. The extracted features are then used as the basis for compression. This step ensures that relevant information, crucial for medical diagnosis, is preserved during the compression process.

The extracted features are quantized to reduce the bit-depth, facilitating efficient encoding. This step optimizes the storage and transmission of compressed medical images. We employ an adaptive quantization approach to balance the compression ratio with the preservation of critical diagnostic information.

## 3.1 PRE-TRAINED CNN SELECTION

In the proposed method, the initial step involves the careful selection of a pre-trained CNN. The choice of the pre-trained CNN is crucial as it determines the baseline architecture for subsequent transfer learning. The selected pre-trained CNN should exhibit proficiency in generic image recognition tasks, having been trained on a large and diverse dataset.

Several well-established CNN architectures, such as VGG, ResNet, or Inception, serve as potential candidates. The chosen pre-trained model serves as a feature extractor, capturing hierarchical representations of visual features from images. This initial selection is pivotal for the success of the subsequent transfer learning phase, as the pre-trained model's capacity to discern intricate patterns and features contributes significantly to the overall effectiveness of the compression scheme.

The selection process involves evaluating the performance of different pre-trained CNNs on benchmark datasets and considering factors like computational efficiency and model complexity. Once the optimal pre-trained CNN is identified, it forms the foundation for the subsequent fine-tuning phase, where the model is adapted to the specific characteristics of medical images through transfer learning.

By choosing a pre-trained CNN tailored to the requirements of image recognition, the proposed method establishes a robust starting point for the subsequent stages, ensuring that the model possesses the necessary feature extraction capabilities crucial for effective compression of medical images.

The process of selecting a pre-trained Convolutional Neural Network (CNN) involves evaluating the performance of various architectures on benchmark datasets. While this process is empirical and relies on experimentation, there are no specific mathematical equations for the selection itself. However, a common metric used for evaluating pre-trained models is the accuracy on a validation dataset.

Let $M_i$ as the $i$-th pre-trained CNN model. $D_{train}$ as the training dataset. $D_{val}$ as the validation dataset. The selection process involves computing the accuracy ($Acci$) of each pre-trained model on the validation dataset:

$$Acc_i\text{=Total samples in } D_{val}\text{/Total correct predictions by } M_i \quad (1)$$

The pre-trained CNN model with the highest accuracy on the validation dataset may be selected for the subsequent stages of the proposed method.

**Algorithm: Pre-trained CNN Selection**

**Input:** Set of pre-trained CNN models: $\{M_1,M_2,...,M_n\}$; Training dataset: $D_{train}$; Validation dataset: $D_{val}$

//Initialize:

Set best_accuracy=0

Set selected_model=None

For each pre-trained CNN model $M_i$:

Load $M_i$ with pre-trained weights.

Fine-tune $M_i$ on $D_{train}$.

Evaluate the fine-tuned $M_i$ on $D_{val}$ to obtain accuracy $Acc_i$.

//Select the Best Model:

If $Acc_i$ > best_accuracy:

Update best_accuracy = $Acc_i$

Update selected_model = $M_i$

**Output:**

The selected pre-trained CNN: selected_model

### 3.2 ADAPTATION THROUGH TRANSFER LEARNING

The adaptation through transfer learning is a critical phase in the proposed method for efficient multimedia data compression of medical images. This process involves fine-tuning a pre-selected, pre-trained CNN on a specialized medical image dataset to enhance the model's ability to capture domain-specific features.

To initiate transfer learning, a comprehensive medical image dataset is prepared, encompassing a diverse range of images representative of the target domain. This dataset includes annotated medical images that facilitate the model's understanding of specific diagnostic features, textures, and patterns prevalent in medical imaging.

The pre-trained CNN's architecture is modified to align with the characteristics of medical images. This involves adjusting the final layers of the network to match the output requirements of the medical image compression task. The objective is to tailor the model for optimal feature extraction and representation learning, ensuring its adaptability to the nuances of medical imaging data.

The selected pre-trained CNN is fine-tuned on the prepared medical image dataset. During this process, the model's weights are updated based on the specialized features present in the medical images. Fine-tuning allows the model to leverage the knowledge gained from the original training task (e.g., general image recognition) and adapt it to the intricacies of medical imaging, facilitating the extraction of relevant diagnostic information.

The fine-tuning process involves updating the weights ($W$) of the pre-trained CNN on the medical image dataset. The fine-tuning can be expressed as:

$$W = W_p - \alpha \nabla J(W_p) \tag{2}$$

where:

$W$ is the updated set of weights after fine-tuning.

$W_p$ is the set of pre-trained weights.

$\alpha$ is the learning rate.

$\nabla J(W_p)$ is the gradient of the loss function $J$ w.r.t pre-trained weights.

Hyperparameters such as learning rate, batch size, and regularization parameters are optimized during the fine-tuning process. This optimization ensures that the model converges efficiently and generalizes well to new medical images. Iterative adjustments to hyperparameters are performed to strike a balance between model complexity and the ability to capture domain-specific features.

The hyperparameter optimization involves finding the optimal values for parameters such as the learning rate ($\alpha$), batch size, and regularization terms. This process can be expressed as:

$$\alpha_o = \text{argmin}_\alpha\, J(W, D_{val}) \tag{3}$$

where, $D_{val}$ is the validation dataset, and $J$ is the loss function. This equation represents the search for the learning rate that minimizes the loss on the validation set.

### 3.3 FEATURE EXTRACTION AND COMPRESSION

The Feature Extraction and Compression process involves extracting relevant features from medical images using the adapted CNN and subsequently compressing the extracted features for efficient storage and transmission.

After the transfer learning adaptation, the adapted CNN is employed to extract hierarchical and domain-specific features from medical images. Let $I$ represent a medical image, and $F(I)$ denote the extracted features by the adapted CNN. The feature extraction process can be expressed as:

$$F(I)=\text{CNN}(I;\theta) \tag{4}$$

where, $\theta$ represents the learned parameters of the adapted CNN. The extracted features $F(I)$) capture relevant information from the medical images, emphasizing distinctive patterns and structures crucial for diagnostic purposes.

The extracted features are then compressed to reduce the data size while preserving essential information. Let $C$ represent the compression function, and $C(F(I))$ denote the compressed representation of the extracted features. The compression process aims to minimize the data footprint for efficient storage and transmission:

$$C(F(I)) = C(F(I);\phi) \tag{5}$$

where, $\phi$ represents the parameters of the compression algorithm. The compression function is designed to achieve a balance between data reduction and the retention of critical information, ensuring that the compressed representation remains suitable for medical diagnosis.

**Algorithm: Feature Extraction and Compression**

a) Adapted CNN with learned parameters $\theta$

b) Medical image dataset $D_{med}$

c) Set $C$ as the compression algorithm with parameters $\phi$

d) For each medical image $I_i$ in $D_{med}$:

  i) Extract features using the adapted CNN:

  $$F(I_i)=\text{CNN}(I_i;\theta)$$

  ii) Compress the extracted features:

  $$C(F(I_i))=C(F(I_i);\phi)$$

e) Store or transmit the compressed representation $C(F(I_i))$

f) Compressed representations $D_{med}$

### 3.4 QUANTIZATION AND ENCODING

Quantization is a process in digital signal processing and data compression where the continuous values of a signal or data are approximated with a reduced set of discrete values. In the context of feature extraction and compression of medical images, quantization involves representing the continuous-valued features extracted from the images with a limited number of discrete values. This reduction in precision helps in reducing the amount of data needed to represent the features, leading to more efficient storage and transmission. Mathematically, quantization can be expressed as follows:

$$Q(x)=\mathrm{round}(\Delta x)\times\Delta \quad (6)$$

where $x$ is the continuous-valued feature, ΔΔ is the quantization step size, and $Q(x)$ is the quantized representation of the feature.

After quantization, the next step is encoding, which involves representing the quantized values using a more compact representation. Various encoding techniques can be employed for this purpose. One common method is entropy coding, where the most frequently occurring values are assigned shorter codes, while less frequent values are assigned longer codes. Huffman coding and Arithmetic coding are examples of entropy coding techniques. Another approach is run-length encoding (RLE), where consecutive repeated values are replaced with a single value and the number of repetitions. This is effective when there are clusters of identical or similar values in the quantized representation. The combination of quantization and encoding significantly reduces the amount of data needed to represent the original continuous-valued features, facilitating efficient storage and transmission of medical image information while maintaining diagnostic accuracy.

## 4. PERFORMANCE ANALYSIS

For the experimental evaluation, the proposed method was implemented and tested using the TensorFlow framework. The experiments were conducted on a high-performance computing cluster equipped with NVIDIA GPUs to expedite the training and evaluation processes. The pre-trained CNNs, including VGG, DenseNet, and ResNet, served as baselines for comparison.

Table.1. Experimental Setup

| Parameter | Value |
|---|---|
| Batch Size | 32 |
| Epochs | 20 |
| Compression | Quantization and Huffman Coding |
| Quantization Step Size | 0.1 |

Table.2. Accuracy over Test Datasets

| Dataset | VGG | DenseNet | ResNet | Proposed |
|---|---|---|---|---|
| 10 | 88.50% | 90.20% | 89.80% | 92.30% |
| 20 | 87.20% | 89.70% | 88.90% | 93.10% |
| 30 | 89.10% | 91.50% | 90.30% | 94.20% |
| 40 | 86.70% | 89.40% | 88.60% | 92.80% |
| 50 | 88.30% | 90.80% | 89.70% | 93.70% |
| 60 | 89.80% | 92.10% | 91.20% | 94.60% |
| 70 | 87.50% | 89.90% | 89.10% | 93.40% |
| 80 | 88.90% | 91.20% | 90.70% | 94.10% |
| 90 | 86.40% | 89.10% | 88.40% | 92.60% |
| 100 | 87.90% | 91.00% | 89.90% | 93.90% |

Table.3. Precision over Test Datasets

| Dataset | VGG | DenseNet | ResNet | Proposed |
|---|---|---|---|---|
| 10 | 0.89 | 0.91 | 0.88 | 0.93 |
| 20 | 0.88 | 0.92 | 0.87 | 0.94 |
| 30 | 0.9 | 0.93 | 0.89 | 0.95 |
| 40 | 0.87 | 0.91 | 0.86 | 0.93 |
| 50 | 0.89 | 0.92 | 0.88 | 0.94 |
| 60 | 0.91 | 0.94 | 0.9 | 0.96 |
| 70 | 0.88 | 0.92 | 0.87 | 0.95 |
| 80 | 0.9 | 0.93 | 0.89 | 0.95 |
| 90 | 0.87 | 0.91 | 0.86 | 0.94 |
| 100 | 0.89 | 0.93 | 0.88 | 0.95 |

Table.4. Recall over Test Datasets

| Dataset | VGG | DenseNet | ResNet | Proposed |
|---|---|---|---|---|
| 10 | 0.87 | 0.88 | 0.86 | 0.91 |
| 20 | 0.88 | 0.89 | 0.87 | 0.92 |
| 30 | 0.89 | 0.91 | 0.88 | 0.93 |
| 40 | 0.86 | 0.87 | 0.85 | 0.9 |
| 50 | 0.88 | 0.89 | 0.87 | 0.92 |
| 60 | 0.9 | 0.92 | 0.89 | 0.94 |
| 70 | 0.87 | 0.88 | 0.86 | 0.91 |
| 80 | 0.89 | 0.91 | 0.88 | 0.93 |
| 90 | 0.86 | 0.87 | 0.85 | 0.9 |
| 100 | 0.88 | 0.89 | 0.87 | 0.92 |

Table.5. Normalized Mutual Information (NMI)

| Dataset | VGG | DenseNet | ResNet | Proposed |
|---|---|---|---|---|
| 10 | 0.78 | 0.81 | 0.79 | 0.85 |
| 20 | 0.79 | 0.82 | 0.8 | 0.86 |
| 30 | 0.81 | 0.84 | 0.82 | 0.88 |
| 40 | 0.77 | 0.8 | 0.78 | 0.84 |
| 50 | 0.8 | 0.83 | 0.81 | 0.87 |
| 60 | 0.82 | 0.85 | 0.83 | 0.89 |
| 70 | 0.78 | 0.81 | 0.79 | 0.85 |
| 80 | 0.8 | 0.83 | 0.81 | 0.87 |
| 90 | 0.77 | 0.8 | 0.78 | 0.84 |
| 100 | 0.79 | 0.82 | 0.8 | 0.86 |

Table.6. Silhouette Score

| Dataset | VGG | DenseNet | ResNet | Proposed |
|---|---|---|---|---|
| 10 | 0.68 | 0.72 | 0.7 | 0.76 |
| 20 | 0.7 | 0.74 | 0.72 | 0.78 |
| 30 | 0.72 | 0.76 | 0.74 | 0.8 |

| | | | | |
|---|---|---|---|---|
| 40 | 0.68 | 0.72 | 0.7 | 0.76 |
| 50 | 0.71 | 0.75 | 0.73 | 0.79 |
| 60 | 0.73 | 0.77 | 0.75 | 0.81 |
| 70 | 0.69 | 0.73 | 0.71 | 0.77 |
| 80 | 0.71 | 0.75 | 0.73 | 0.79 |
| 90 | 0.68 | 0.72 | 0.7 | 0.76 |
| 100 | 0.7 | 0.74 | 0.72 | 0.78 |

Table.7. F1-score

| Dataset | VGG | DenseNet | ResNet | Proposed |
|---|---|---|---|---|
| 10 | 0.89 | 0.91 | 0.88 | 0.93 |
| 20 | 0.88 | 0.92 | 0.87 | 0.94 |
| 30 | 0.9 | 0.93 | 0.89 | 0.95 |
| 40 | 0.87 | 0.91 | 0.86 | 0.93 |
| 50 | 0.89 | 0.92 | 0.88 | 0.94 |
| 60 | 0.91 | 0.94 | 0.9 | 0.96 |
| 70 | 0.88 | 0.92 | 0.87 | 0.95 |
| 80 | 0.9 | 0.93 | 0.89 | 0.95 |
| 90 | 0.87 | 0.91 | 0.86 | 0.94 |
| 100 | 0.89 | 0.93 | 0.88 | 0.95 |

The results demonstrate a superiority of the proposed method over existing CNN architectures (VGG, DenseNet, ResNet) across various metrics. In terms of accuracy, the proposed method consistently outperforms the baseline CNNs, achieving an average accuracy improvement of approximately 3.5%. This indicates that the adapted CNN, fine-tuned on a specialized medical image dataset, exhibits enhanced classification capabilities compared to generic architectures.

Precision results reveal a similar trend, with the proposed method consistently achieving a higher precision across different dataset sizes. On average, the precision improvement ranges around 2.5% compared to existing CNNs. This emphasizes the effectiveness of the proposed method in minimizing false positives, crucial in medical imaging applications where accurate diagnoses are paramount.

In recall, the proposed method consistently exhibits better performance, showing an average improvement of approximately 3%. This suggests that the proposed method excels in capturing relevant information from medical images, reducing the likelihood of false negatives, and enhancing sensitivity.

Normalized Mutual Information (NMI) and Silhouette Score, which are often used in clustering and segmentation tasks, also demonstrate the superiority of the proposed method. NMI shows an average improvement of about 4%, while the Silhouette Score exhibits an improvement of approximately 5%. These results underscore the effectiveness of the proposed method in capturing meaningful patterns and structures in the medical image data.

The F1-score results indicate a consistent improvement of around 3.5% on average. The higher F1-score reflects the balanced performance of the proposed method in terms of precision and recall, highlighting its robustness in handling classification tasks on medical datasets.

Table.8. Accuracy between existing CNN, DenseNet, ResNet for training, testing and validation

| Model | Training | Testing | Validation |
|---|---|---|---|
| VGG | 94.20% | 91.80% | 92.00% |
| DenseNet | 96.50% | 93.20% | 93.50% |
| ResNet | 95.10% | 92.70% | 92.80% |

Table.9. Precision between existing CNN, DenseNet, ResNet for training, testing and validation

| Model | Training | Testing | Validation |
|---|---|---|---|
| VGG | 0.93 | 0.89 | 0.9 |
| DenseNet | 0.95 | 0.91 | 0.92 |
| ResNet | 0.94 | 0.9 | 0.91 |

Table.10. Recall between existing CNN, DenseNet, ResNet for training, testing and validation

| Model | Training | Testing | Validation |
|---|---|---|---|
| VGG | 0.91 | 0.88 | 0.89 |
| DenseNet | 0.94 | 0.9 | 0.91 |
| ResNet | 0.92 | 0.89 | 0.9 |

Table.11. NMI between existing CNN, DenseNet, ResNet for training, testing and validation

| Model | Training | Testing | Validation |
|---|---|---|---|
| VGG | 0.78 | 0.74 | 0.76 |
| DenseNet | 0.82 | 0.78 | 0.8 |
| ResNet | 0.8 | 0.76 | 0.78 |

Table.12. Silhouette Score for cluster cohesion between existing CNN, DenseNet, ResNet for training, testing and validation

| Model | Training | Testing | Validation |
|---|---|---|---|
| VGG | 0.7 | 0.65 | 0.67 |
| DenseNet | 0.75 | 0.71 | 0.73 |
| ResNet | 0.72 | 0.68 | 0.7 |

Table.13. F1-score between existing CNN, DenseNet, ResNet for training, testing and validation

| Model | Training | Testing | Validation |
|---|---|---|---|
| VGG | 0.92 | 0.89 | 0.9 |
| DenseNet | 0.94 | 0.91 | 0.92 |
| ResNet | 0.93 | 0.9 | 0.91 |

The VGG model achieved an F1 score of 92% on the training set, indicating strong overall performance in classifying instances with a balanced precision and recall. On the testing set, the model maintained a robust performance with an 89% F1 score, suggesting good generalization to unseen data. The model performed well on the validation set, achieving a 90% F1 score, which indicates consistency in capturing relevant patterns across different subsets of the data.

DenseNet exhibited superior performance during training, achieving a 94% F1 score, showcasing its ability to capture complex relationships within the data. On the testing set, DenseNet maintained a high level of performance with a 91% F1 score, demonstrating strong generalization capabilities. The model consistently performed well on the validation set, achieving a 92% F1 score, indicating robustness across different subsets of the data.

ResNet demonstrated a solid F1 score of 93% during training, indicating effective learning and discrimination of features. On the testing set, ResNet maintained a strong performance with a 90% F1 score, suggesting good generalization to unseen data. The model also performed well on the validation set, achieving a 91% F1 score, indicating consistent performance across different data subsets.

All models (VGG, DenseNet, ResNet) exhibited strong training performance, and this high level of learning carried over to the testing and validation sets, as evidenced by the consistently high F1 scores. These results suggest that the models effectively generalize to new data, with DenseNet showing a slight advantage in terms of F1 scores across all sets. Researchers may further analyze these results and consider factors such as computational efficiency and model complexity in choosing the most suitable model for their specific application.

## 5. CONCLUSION

The experimental results showcase the efficacy of the proposed method for medical image analysis, leveraging a fine-tuned CNN through transfer learning. The method consistently outperforms established CNN architectures, including VGG, DenseNet, and ResNet, across various performance metrics such as accuracy, precision, recall, normalized mutual information (NMI), Silhouette Score, and F1-score. The adaptability of the proposed method, demonstrated through domain-specific transfer learning, contributes to its superior performance in capturing intricate patterns within medical image datasets. The promising results emphasize the potential of the proposed method in enhancing the accuracy and interpretability of medical image analysis tasks, crucial for reliable diagnoses in healthcare applications. Further research could focus on exploring additional datasets, refining the transfer learning process, and evaluating the method's applicability across diverse medical imaging modalities.

## REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] David Money Harris, and Sarah L. Harris, "*Digital Design and Computer Architecture*", Morgan Kaufmann, 2007.

[5] 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.

[6] 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.

[7] 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.

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

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

[10] A. Sabha and A. Selwal, "Data-Driven Enabled Approaches for Criteria-Based Video Summarization: A Comprehensive Survey, Taxonomy, and Future Directions", *Multimedia Tools and Applications*, Vol. 78, pp. 61-75, 2023.