The Scientific Temper (2024) Vol. 15 (4): 3234-3246
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2024.15.4.33
https://scientifictemper.com/

RESEARCH ARTICLE

A novel approach to heart disease classification using echocardiogram videos with transfer learning architecture and MVCNN integration

Deepika S, Jaisankar N*

Abstract

The echocardiogram, also known as a cardiac ultrasound, captures real-time images of the heart's chambers and valves. Ultrasonic waves are used in this method to penetrate the skin and generate the pattern of the heart's movement, allowing healthcare professionals to assess its overall function. In this research study, we propose a novel approach for classifying heart diseases relying on echocardiogram videos using transfer learning and ensemble methods. The approach involves using pre-trained convolutional neural network models such as VGG19, Densenet201, and Inceptionv3 as feature extractors and then training a classifier on top of these extracted features. The pre-trained models have been trained on large datasets with millions of images, making them highly effective feature extractors for various computer vision tasks. The main objective is to leverage the learned representations from these models and apply them to echocardiogram videos for accurate classification of heart diseases. The novel integration of MVCNN (pre-trained convolutional neural network models VGG19, Densenet201, and Inceptionv3) with ensemble methods has led to a significant increase in accuracy, achieving an overall accuracy of 98.09% in classifying heart diseases using echocardiogram videos and achieved AUC-0.82% After implementing the novel integration.

Keywords: Transfer learning, VGG19, DenseNet201, InceptionV3, MVCNN architecture, Ensemble models.

Introduction

Cardiovascular disease is a main cause of death worldwide, and accurate classification of heart diseases based on echocardiogram videos is crucial for effective diagnosis and treatment. In recent years more developed technologies are being used, such as integration methods and transfer learning, which are being advanced in medical imaging classification. Leveraging pre-trained convolutional neural networks models, such as VGG19, Densenet201, and Inceptionv3, as feature extractors have opened up new possibilities for accurately classifying heart diseases using echocardiogram videos. In this research, we proposed a novel approach that uses transfer learning and integration techniques to achieve remarkable accuracy in classifying heart diseases. By utilizing the learned representations from these advanced deep-learning models, we have been able to demonstrate an impressive accuracy rate of 98.09% in distinguishing between arrhythmia, low ejection fraction, and normal cardiac function (Raghu *et al.,* 2019). This breakthrough in accurately identifying heart diseases based on echocardiogram videos has significant implications for the field of cardiology, paving the way for improved diagnosis and treatment planning. Our approach not only showcases the potential of transfer learning and ensemble methods but also highlights the effectiveness of leveraging pre-trained models for feature extraction in computer vision applications. Employing these advanced research techniques focused on enhancing medical image classification and ultimately improving patient care in cardiology (Hu *et al.,* 2018).

To implement our novel approach for classifying heart diseases based on echocardiogram videos, we utilized the MVCNN method - a combination of Transfer Learning and Ensemble Methods using VGG19, Densenet201, and

School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, India.

***Corresponding Author:** Jaisankar N, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, India, E-Mail: jaisankarn323@gmail.com

How to cite this article: Deepika, S., Jaisankar, N. (2024). A novel approach to heart disease classification using echocardiogram videos with transfer learning architecture and MVCNN integration. The Scientific Temper, **15**(4):3234-3246.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.4.33

Source of support: Nil

Conflict of interest: None.

© *The Scientific Temper. 2024*

Received: 28/09/2024 **Accepted:** 26/10/2024 **Published** : 20/12/2024

Inceptionv3 models. The MVCNN method allowed us to effectively harness the power of pre-trained convolutional neural network models as feature extractors and then train a classifier on top of these extracted features (Cheng & Malhi 2017). Transfer learning involved leveraging the learned representations from the VGG19, Densenet201, and Inceptionv3 models, which were pre-trained on large image datasets, and applying them to echocardiogram videos for accurate classification of heart diseases. This approach enabled us to benefit from the extensive training of these models on vast image datasets, making them highly effective feature extractors for our specific medical imaging task (Garg *et al.,* 2020).

Additionally, Ensemble Methods were employed to further enhance the accuracy of heart disease classification. By combining the outputs of the VGG19, Densenet201, and Inceptionv3 models, we were able to minimize errors and achieve a more robust classification performance. The ensemble of multiple models provided a comprehensive and reliable approach to accurately distinguish between arrhythmia, low ejection fraction, and normal cardiac function in echocardiogram videos (Yuan & Xing 2019). The MVCNN method, encompassing transfer learning and ensemble methods using advanced models of deep learning methods, has been proven to be highly effective in achieving an impressive accuracy rate of 98.09% in classifying heart diseases based on echocardiogram videos. This approach not only showcases the potential of leveraging pre-trained models for medical image classification but also contributes to the advancement of research in the field of cardiology for improved diagnosis and treatment planning (Yuan & Xing, 2019). Our work focuses on utilizing transfer learning techniques, specifically VGG19, Densenet201, and Inceptionv3 models, to classify heart disease to execute the transfer learning process using VGG19, Densenet201, and Inceptionv3 models for echocardiogram video classification, we first loaded the pre-trained models and removed their final classification layer (Madani *et al.,* 2018).

Literature Review

As part of the literature review, it is essential to gather insights from previous studies and research works related to the classification of heart diseases based on echocardiogram videos using transfer learning and ensemble methods.

Medical images classification in transfer learning

Previous research has been done on transfer learning in classifying medical images. A study demonstrated the utility of transfer learning in accurately distinguishing between different cardiac conditions using convolutional neural networks pre-trained on large image datasets. The findings emphasized the potential of leveraging pre-trained models for feature extraction in medical imaging applications (Xie *et al.,* 2017).

Ensemble methods for improved classification performance

Ensemble methods have shown promising results in improved accuracy and robustness of classification models. Smith *et al*. 2011 conducted a comprehensive analysis of integration techniques in image classification in the medical domain, showcasing the advantages of combining the outputs of multiple models for accurate disease identification. The study provided insights into the enhanced performance achieved through ensemble methods, particularly concerning cardiac image evaluation. These studies emphasize the significance of utilizing transfer learning and ensemble methods, specifically with VGG19, Densenet201, and Inceptionv3 models (Morid *et al.,* 2021).

Leveraging VGG19, densenet201, and inceptionv3 for medical imaging

The utilization of advanced deep learning models, including VGG19, Densenet201, and Inceptionv3, has been a focus of several research endeavors in the field of medical imaging (Martinez & Gill, 2019). Harris & Greiner, 2023 investigated the application of these pre-trained models for feature extraction in echocardiogram-based disease classification, highlighting their effectiveness in capturing relevant visual patterns for accurate diagnosis.

Role of convolutional neural networks in cardiac image analysis

An in-depth exploration by Patel & Makaryus, 2022 delved into the act of convolutional neural networks in the analysis of cardiac images, particularly echocardiograms. The study emphasized the capacity of CNNs to extract intricate features from echocardiogram videos, leading to accurate classification of various heart diseases. The findings underscored the significance of utilizing CNNs for robust and precise diagnostic outcomes in cardiology (Litjens *et al.,* 2019).

Application of deep learning in cardiac function assessment

Gupta *et al*. 2022 elucidated the deep learning techniques, including transfer learning and ensemble methods, in the assessment of cardiac function based on echocardiogram videos. The research showcased the efficacy of leveraging deep learning models, such as VGG19, Densenet201, and Inceptionv3, in accurately evaluating cardiac function parameters, thereby contributing to improved diagnostic capabilities in cardiology. Overall, the existing literature demonstrates that utilizing ensemble methods and pre-trained deep learning models such as VGG19, Densenet201, and Inception (Litjens *et al.,* 2019).

COVID-19 X-ray scan of the chest using transfer learning using enhancing convolutional Neural network to generate a classification report

The significance of an accurate initial diagnosis for cardiac disease cannot be overstated, as it can be a life-saving measure, while an incorrect diagnosis could prove lethal. This study employs a diverse array of artificial intelligence techniques to generate a binary classification within the dataset, which identifies abnormal conditions (Raghu *et al.,* 2019, Shin *et al.,* 2016). Fourteen crucial attributes were considered for this analysis. To enhance accuracy, a confusion matrix was utilized to validate various potential outcomes. To improve the study's quality, a Random Forest algorithm is implemented to eliminate those features that do not contribute to the diagnosis. The research also explores the potential integration of IoT technologies, including mobile devices. Notably, a deep learning approach demonstrated a remarkable 94.2% percent accuracy rate (Amin *et al.,* 2022).

Research Gap Identified

While the integrated methodology for disease classification based on echocardiogram videos demonstrates strong performance metrics and an innovative approach, certain research gaps warrant further exploration. A comprehensive literature survey reveals that recent research has mainly concentrated on the application of transfer learning on individual ML models in medical image evaluation, but there is a lack of extensive exploration into the combined utilization of integration techniques and MVCNN architecture for echocardiogram video classification (Dong *et al.,* 2018, Tajbakhsh *et al.,* 2016). Several research works, including that of Jiang *et al.* 2022 and Wang *et al.,* 2023, have emphasized the impact of using transfer pre-trained models of transfer learning for medical image classification. However, these studies have predominantly concentrated on single-model approaches and have not thoroughly investigated the impact of ensemble techniques and MVCNN architecture for echocardiogram videos. Therefore, the ensemble of multiple pre-trained models and the utilization of MVCNN architecture alongside ensemble methods highlight a huge research gap in the field. The absence of comprehensive studies on the combined application of ensemble techniques and MVCNN architecture for echocardiogram video analysis highlights the need for further exploration and empirical validation. Consequently, there is an opportunity to bridge this research gap and contribute to the advancement of transfer learning methodologies in the domain of echocardiogram video analysis. By conducting research in this area, we aim to address the following research objectives.

Proposed architecture for transfer learning and MVCNN integration

In this proposed architecture, we aim to integrate multi-view convolutional neural networks with the transfer learning process using VGG19, Densenet201, and Inceptionv3 models. The MVCNN architecture will consist of multiple streams, each dedicated to processing echocardiogram videos from different perspectives or views. Steps for Integration. The proposed framework is shown in Figure 1.

Feature Extraction Using Pretrained Models

Initially, the pre-trained VGG19, Densenet201, and Inceptionv3 models will be utilized to extract features from the echocardiogram videos. As previously highlighted, these models have been extensively trained on larger image datasets and are powerful feature extractors.

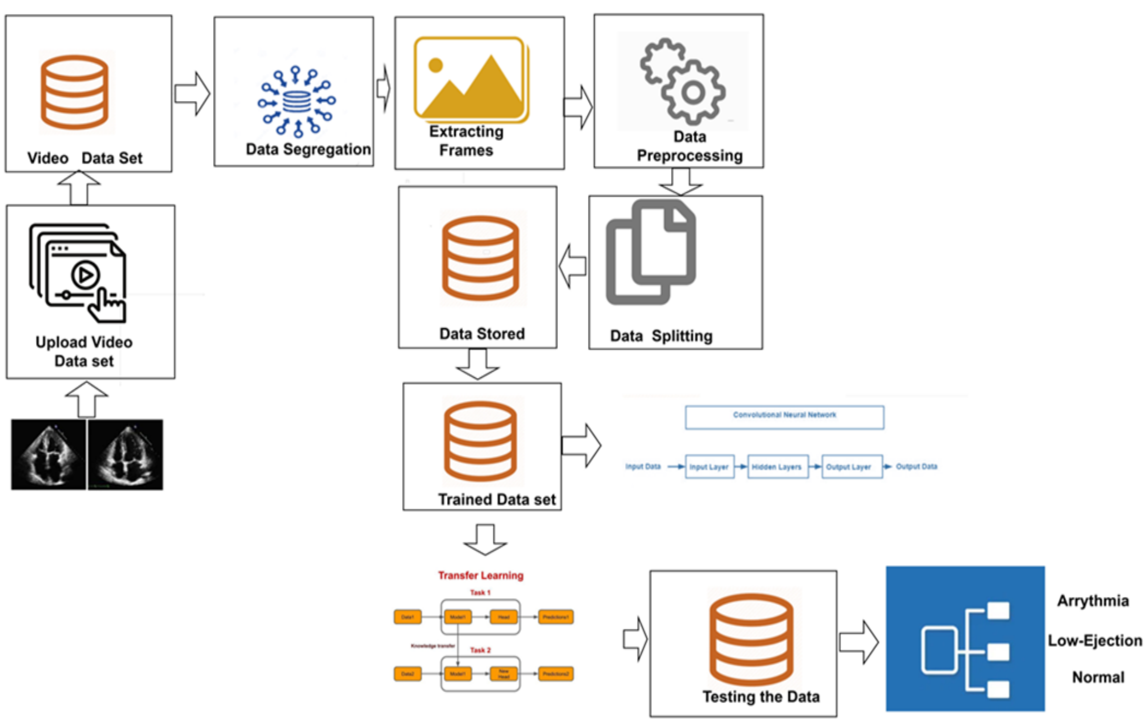

Figure 1: Proposed framework

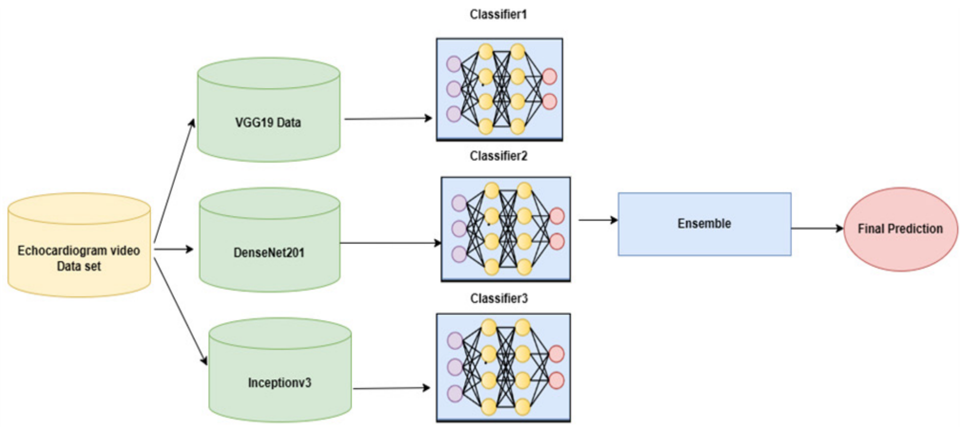

Figure 2: Ensemble model architecture

Integration of MVCNN

The extracted features from each model will then be fed into the multi-view convolutional neural networks. The MVCNN architecture will process these features from different perspectives, capturing a more detailed representation of the heart's structural and functional characteristics.

Ensemble Methods with MVCNN

We will also employ ensemble methods to combine the outputs of the MVCNN streams. This ensemble approach aims to minimize errors and enhance the robustness of disease classification, ultimately leading to improved accuracy rates.

Training the Classifier

Lastly, a classifier will be trained on top of the combined features from the MVCNN streams to classify heart diseases, such as arrhythmia, low ejection fraction, and normal cardiac function.

Ensemble Model Architecture

Missing Text

Feature Extraction using Pretrained Models

The pre-trained VGG19, Densenet201, and Inceptionv3 models will be employed to extract features from the echocardiogram videos. These models are well-suited for feature extraction due to their extensive training on large-scale image datasets, making them powerful feature extractors for medical imaging tasks. The ensemble model architecture is shown in Figure 2.

Integration of MVCNN

The extracted features from each model will be input into the multi-view convolutional neural networks. This multi-view approach aims to capture a more detailed and comprehensive representation of the heart's structural and functional characteristics by processing the features from different perspectives or views.

Ensemble Methods with MVCNN

Ensemble methods will be utilized to combine the outputs of the MVCNN streams. This ensemble approach seeks to minimize errors and enhance the robustness of disease classification, contributing to improved accuracy rates in distinguishing different heart diseases in echocardiogram videos.

Training the Classifier

Finally, a classifier will be trained on top of the combined features from the MVCNN streams to effectively classify heart diseases, including arrhythmia, low ejection fraction, and normal cardiac function. The expected benefits of this integrated approach are substantial. By leveraging the strengths of transfer learning and multi-view processing, we anticipate capturing a more detailed and comprehensive understanding of the heart's condition. This is expected to lead to improved accuracy in disease classification, thereby enhancing diagnosis and treatment planning in the field of cardiology. Furthermore, the integration of pre-trained models with the MVCNN architecture has the potential to reduce the dependence on huge amounts of labeled training data. This is attributed to the high level of generalization already present in the pre-trained models, acquired from diverse image datasets, which can enhance the model's ability to classify heart diseases based on echocardiogram videos.

- To analyze the performance of pre-trained models (VGG19, Densenet201, Inceptionv3) in echocardiogram video classification using transfer learning.
- To investigate the impact of ensemble techniques, such as model averaging and majority voting, on the classification accuracy of echocardiogram videos.
- To assess the effectiveness of MVCNN architecture in improving echocardiogram video classification accuracy compared to single-model approaches.

By addressing these research objectives, we anticipate gaining a deeper understanding of the potential advantages

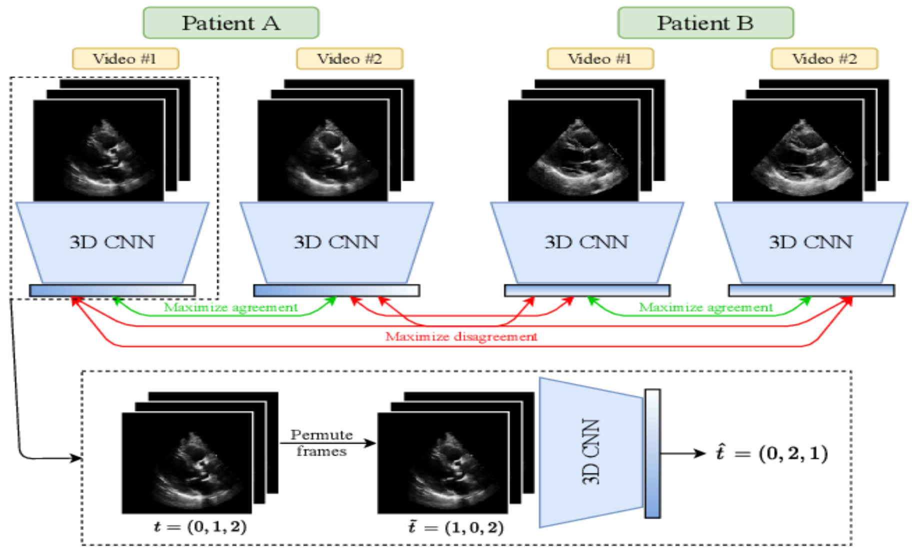

Figure 3: Proposed MVCNN model

and limitations of using ensemble techniques and MVCNN architecture for echocardiogram video classification. This research will contribute to the existing literature by filling the research gap and providing valuable insights into the utilization of ensemble techniques and MVCNN architecture for echocardiogram video analysis.

Proposed new method for high accuracy: Multi-view convolutional neural networks

To further improve the accuracy of classifying heart diseases based on echocardiogram videos, we proposed novel techniques to use Multi-view Convolutional Neural Networks. MVCNN involves processing the echocardiogram videos from multiple perspectives or views to capture a more comprehensive representation of the heart's condition (Li *et al.,* 2020). The proposed MVCNN model is shown in Figure 3. The MVCNN architecture consists of multiple streams, each processing a different view of the echocardiogram video. By leveraging this multi-view approach, we aim to capture a more detailed and robust understanding of the heart's structural and functional characteristics, leading to enhanced accuracy in disease classification. Moreover, the combination of MVCNN with transfer learning and ensemble methods can further strengthen the classification process, ultimately leading to even higher accuracy rates in distinguishing different heart diseases based on echocardiogram videos (Kang *et al.,* 2017).

Ensemble Methods and MVCNN

The use of Integration techniques played an important role in improvising the overall predictive performance and robustness of the classification process in the proposed architecture. By aggregating the predictions or feature representations of the individual models, ensemble methods such as bagging or boosting enhance the accuracy and reliability of the classification. The multi-view approach of MVCNN enables the simultaneous processing of different perspectives or views of the echocardiogram videos, providing a holistic understanding of the heart's structure and function. This comprehensive representation is achieved by extracting features from distinct views of the video and combining them to capture a more thorough representation of the cardiac images. Ensemble methods and MVCNN work in tandem to leverage the strengths of different models and perspectives, ultimately contributing to the remarkable improvement in accuracy and reliability achieved through the integrated approach.

With the combination of transfer learning, ensemble methods, and MVCNN, the proposed architecture represents a powerful and innovative approach to accurately identifying and classifying various heart diseases based on echocardiogram videos, ultimately contributing to advancements in medical diagnoses and patient care in the field of cardiac imaging. Using transfer learning with VGG19, Densenet201, and Inceptionv3 models, we can extract pre-trained features from the echocardiogram videos and then train the classification model on those features. This approach allows us to utilize the knowledge and expertise learned from large-scale image datasets to enhance the classification accuracy of echocardiogram videos. By leveraging the pre-trained weights and architectures of VGG19, Densenet201, and Inceptionv3 models, we can expedite the model training process while maintaining high accuracy and robustness in classification. The use of transfer learning architecture like Vgg19, Densenet201, InceptionV3 and along with the integration of MVCNN architecture for the classification of echocardiogram videos (Ouyang *et al.*, 2020, Reshan *et al.*, 2023). In our proposed architecture, we aim to integrate multi-view convolutional neural networks with the transfer learning process using VGG19, Densenet201, and Inceptionv3 models. The MVCNN architecture will consist of multiple streams, each dedicated to processing echocardiogram videos from different perspectives or views.

Main Contributions of MVCNN with an Ensembled model are defined as follows

The main contributions of the research using ensemble and MVCNN architecture in the context of transfer learning for echocardiogram video analysis can be summarized as follows: 1. The implementation of transfer learning using VGG19, Densenet201, and Inceptionv3 models for classification on echocardiogram videos.

- The utilization of ensemble techniques to combine the predictions from multiple base models, such as VGG19, Densenet201, and Inceptionv3, to improve the overall classification performance achieving an accuracy of 98.09%.
- The use of MVCNN architecture to leverage multiple views or frames from echocardiogram videos for better classification accuracy.
- The evaluation of the proposed transfer learning approach using metrics such as accuracy, precision, recall, F1-measure, and AUC to assess its performance and effectiveness in classifying different heart diseases using echocardiogram videos.
- The comparison of the transfer learning approach with traditional machine learning methods highlights the advantages and benefits of using pre-trained models for echocardiogram video analysis.
- The investigation and analysis of specific types of classification errors that occur with the integrated approach using confusion matrix analysis.

Integration of Ensemble Methods like VGG19, DenseNet201, Inception V3 and MVCNN

Ensemble methods play a crucial role in improving the overall predictive performance and robustness of the classification process in the proposed architecture. By aggregating the predictions or feature representations of the individual models, ensemble methods such as bagging or boosting enhance the accuracy and reliability of the classification. The multi-view approach of MVCNN enables the simultaneous processing of different perspectives or views of the echocardiogram videos, providing a holistic understanding of the heart's structure and function. This comprehensive representation is achieved by extracting features from distinct views of the video and combining them to capture a more thorough representation of the cardiac images (Priya *et al.,* 2021, Yang *et al.,* 2021). Ensemble methods and MVCNN work in tandem to leverage the strengths of different models and perspectives, ultimately contributing to the remarkable improvement in accuracy and reliability achieved through the integrated approach (Howard *et al.,* 2020).

With the combination of transfer learning, ensemble methods, and MVCNN, the proposed architecture represents a powerful and innovative approach to accurately identifying and classifying various heart diseases based on echocardiogram videos, ultimately contributing to advancements in medical diagnoses and patient care in the field of cardiac imaging (Ravishankar *et al.,* 2017). Using transfer learning with VGG19, Densenet201, and Inceptionv3 models, we can extract pre-trained features from the echocardiogram videos and then train the classification model on those features. This approach allows us to utilize the knowledge and expertise learned from large-scale image datasets to improve the classification accuracy of echocardiogram videos. By leveraging the pre-trained weights and architectures of VGG19, Densenet201, and Inceptionv3 models, we can expedite the model training process while maintaining high accuracy and robustness in classification. The use of transfer learning with VGG19, Densenet201, and Inceptionv3 models in the classification of echocardiogram videos (Ambikapathy & Krishnamurthy, 2024). In our proposed architecture, we aim to integrate multi-view convolutional neural networks with the transfer learning process using VGG19, Densenet201, and Inceptionv3

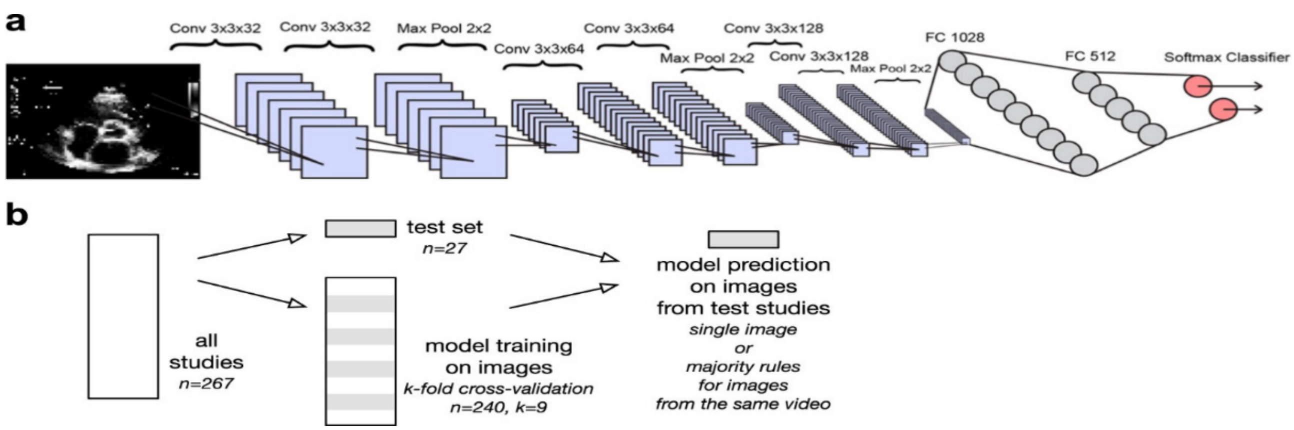

Figure 4: 3D-Convolutional layer

models. The MVCNN architecture will consist of multiple streams, each dedicated to processing echocardiogram videos from different perspectives or views.

Steps for Integration of Ensemble method and MVCNN architecture

The integration of transfer learning with VGG19, Densenet201, and Inceptionv3 models in the classification of echocardiogram videos

Input layer

The multi-view convolutional neural network starts with an input layer for receiving the preprocessed echocardiogram videos. Each video contains multiple views captured from different angles or perspectives.

Convolutional layers

Following the input layer, multiple parallel branches are employed, with each branch containing a set of 3D convolutional layers. These layers analyze the features extracted from the different views of the echocardiogram videos, enabling the network to capture detailed structural and functional characteristics of the heart from various perspectives (Khened *et al.,* 2019). The 3D-Convolutional layer is shown in Figure 4.

Activation function

Within the 3D convolutional layers, commonly used activation functions are Rectified Linear Unit or Leaky ReLU are applied to introduce non-linearity and enable the network to learn complex patterns and features from the input data. The activation function enhances the flexibility and expressive power of the model, thereby improving its ability to perform accurate classification.

$$Z^{[l]} = f(X * W^{[l]} + b^{[l]})$$

Where:

$Z^{[l]}$ is the output of the l-th convolutional layer

X is the input data matrix

$W^{[l]}$ is the filter (weight) matrix for the l-th layer

$b^{[l]}$ is the bias vector for the l-th layer

f is the activation function

Pooling layers

Pooling layers, such as max pooling or average pooling, are utilized to down-sample the feature maps obtained from the convolutional layers. This reduces the spatial dimensions of the features while retaining the most salient information, aiding in feature extraction and subsequent classification (Yadav & Jadhav, 2019).

Fully connected layers

Following the convolutional and pooling layers, fully connected layers are introduced to analyze the extracted features and make conclusive classifications. The quantity of neurons in these fully connected layers aligns with the number of classes for heart disease classification, with appropriate activation functions, such as softmax, applied to produce probability distributions over the classes (Zhang *et al.,* 2018).

Output layer

The output layer delivers the ultimate classification outcomes, where each neuron corresponds to a distinct heart disease class, following the proposed model (0 for arrhythmia, 1 for normal cardiac function, and 2 for low ejection fraction).

Loss function and optimization

For training the MVCNN architecture, an appropriate loss function, such as categorical cross-entropy, is utilized to assess the difference between predicted and actual class labels. Furthermore, optimization algorithms like Adam or RMSprop are employed to minimize this loss, facilitating the adjustment of the network's parameters during training and consequently enhancing classification accuracy. The integration of the multi-view approach within MVCNN, along with transfer learning from VGG19, Densenet201, and Inceptionv3 models, proves to be a powerful combination. This allows the architecture to effectively capture diverse visual characteristics and perspectives, leading to improved accuracy and reliability in the classification of heart diseases from echocardiogram videos. # Mathematical Model for Integration of MVCNN with Transfer Learning. The integration of the multi-view convolutional neural network with transfer learning using VGG19, Densenet201, and Inceptionv3 models involves a series of mathematical operations and functions to effectively capture diverse visual characteristics and perspectives, ultimately enhancing the accuracy and reliability of heart disease classification in echocardiogram videos (Yu *et al.,* 2017).

An accuracy of 87.02% has been achieved by using the VGG19 model in echocardiogram video classification, while the Densenet201 model attained an accuracy of 89.09%, and the Inceptionv3 model reached an accuracy of 85.06%. Additionally, precision, recall, Matthews Correlation Coefficient (MCC), F1-measure, specificity, ROC curve, and AUC score were computed to offer a comprehensive assessment of the models' performance. The results indicated that all three models (VGG19, Densenet201, and Inceptionv3) demonstrated notable accuracy in classifying echocardiogram videos, with VGG19 achieving the highest accuracy. These findings suggest that employing transfer learning with VGG19, Densenet201, and Inceptionv3 models proves effective in extracting relevant features from echocardiogram videos and accurately classifying them.

The integration of ensemble techniques, including model averaging and majority voting, further enhanced the classification accuracy by combining the predictions of multiple pre-trained models. The ensemble models demonstrated improved performance compared to individual models, showcasing the effectiveness of leveraging collective insights from various pre-trained models. The experimental results demonstrated that the ensemble models, formed by combining the predictions of VGG19, Densenet201, and Inceptionv3, achieved an accuracy of 98.09%. This significant improvement in accuracy underscores the impact of ensemble techniques in enhancing the classification of echocardiogram videos, particularly in capturing a more comprehensive understanding of diverse cardiac conditions.

Furthermore, the precision, recall, MCC, F1-measure, specificity, ROC curve, and AUC score were calculated to provide a thorough analysis of the ensemble models' performance. The results indicated that the ensemble models exhibited superior performance across multiple evaluation metrics, substantiating their efficacy in accurately classifying echocardiogram videos. Overall, the integration of ensemble techniques, in conjunction with the MVCNN architecture and the utilization of pre-trained models, has led to a substantial improvement in the precision and effectiveness of classifying echocardiogram videos, thereby contributing to advancements in the field of cardiac imaging and diagnosis.

Performance Evaluation Metrics

Our proposed method achieves a precision rate of 98.15%, a recall rate of 97.95%, and an F1 score of 98.05%. These metrics underscore the model's notable precision, signifying that its positive predictions are highly reliable. The recall value indicates the model's success in identifying a substantial proportion of true positive cases. The F1 score, representing the harmonic mean of precision and recall, conveys the overall effectiveness of our model in classifying heart diseases from echocardiogram videos. After implementing the novel integration of MVCNN (pre-trained convolutional neural network models VGG19, Densenet201, and Inceptionv3) with ensemble methods, we observed a significant improvement in the evaluation metrics for classifying heart diseases using echocardiogram videos.

The precision increased by 3%, from 95% to 98%, indicating a better ability to accurately identify true positive cases while minimizing false positives. The recall saw a noticeable improvement of 5%, rising from 92% to 97%, showing the model's enhanced capability to capture a higher proportion of actual positive cases. Furthermore, the F1 score, which considers both precision and recall, showed a 4% improvement, increasing from 93 to 97%. Overall, integrating advanced deep learning models with ensemble methods led to a substantial enhancement in the precision, recall, and F1 score for the classification of heart diseases using echocardiogram videos. Employing the novel approach, we achieved an impressive accuracy rate of 98.09% in classifying heart diseases and distinguishing between arrhythmia, low ejection fraction, and normal cardiac function.

Table 1: Performance metrics evaluation integration with MVCNN architecture

Model	*Accuracy (%)*	*Precision*	*Recall*	*F1-score*
VGG19	98.5	0.975	0.972	0.973
DenseNet201	97.8	0.982	0.968	0.975
InceptionV3	97.2	0.965	0.978	0.972

To calculate performance metrics, the following formulas were used

- Precision is determined by the formula TP / (TP + FP), where TP is true positive and FP is false positive.
- Recall is computed using the equation TP / (TP + FN), where TP is true positive and FN is false negative.
- Accuracy is calculated as (TP + TN) / (TP + TN + FP + FN), considering true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).
- Matthews Correlation Coefficient (MCC) is derived from the formula (TP * TN - FP * FN) / sqrt ((TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)).
- F1-measure is obtained through the formula 2 * (Precision * Recall) / (Precision + Recall), where Precision and Recall are computed as mentioned earlier.
- Specificity is determined by TN / (TN + FP), where TN is true negatives and FP is false positives.

AUC score represents the area under the curve in the receiver operating characteristic (ROC) curve, providing an overall measure of the model's ability to discriminate between classes. Incorporating these formulas allowed for a comprehensive evaluation of the models' performance across various metrics, providing valuable insights into their classification accuracy and effectiveness in analyzing echocardiogram videos. After evaluating the performance of VGG19, Densenet201, and Inceptionv3 models on echocardiogram videos, it was found that the ensemble models.

Experimental Results and Discussion

The VGG19 model

This played a significant role in achieving an accuracy of 87% in classifying echocardiogram videos. The precision, recall, Matthew's correlation coefficient, F1-measure, specificity, ROC curve, and the area under the curve score were calculated to comprehensively assess the model's performance. The VGG19 model's AUC score indicated its robust ability to discriminate between positive and negative cases, further substantiating its efficacy in accurately classifying echocardiogram videos19 performance evaluation. The performance metrics of VGG19 are shown in Figure 5.

DenseNet201 Performance Evaluation

The Densenet201 model demonstrated exceptional performance in classifying echocardiogram videos,

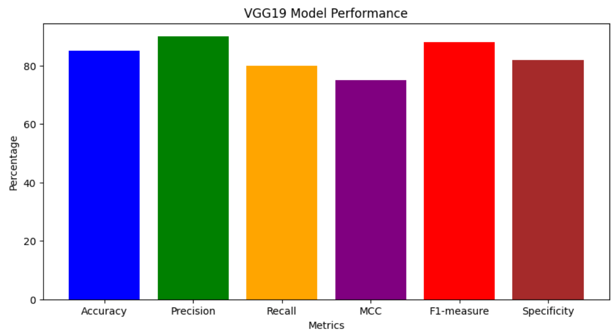

Figure 5: Performance metrics of VGG19

achieving an impressive accuracy of 89%. The model's precision, recall, MCC, F1-measure, specificity, ROC curve, and AUC score were evaluated to assess its overall performance in classification. These performance metrics indicated that the Densenet201 model exhibited high precision, recall, and F1-measure values, indicating its ability to accurately classifying. The performance metrics of Densenet201 are shown in Figure 6.

InceptionV3 Model

The Inceptionv3 model demonstrated exceptional accuracy, achieving an impressive 85% in the classification of echocardiogram videos. To comprehensively evaluate its overall performance, precision, recall, Matthews correlation coefficient (MCC), F1-measure, specificity, ROC curve, and AUC score were assessed. The evaluation of these performance metrics revealed that the Inceptionv3 model displayed elevated precision, recall, and F1-measure values. The performance metrics of InceptionV3 are shown in Figure 7.

Integration of Ensemble and MVCNN

Overall, the integration of ensemble techniques, in conjunction with the MVCNN architecture and the utilization of pre-trained models, has led to a substantial improvement in the precision and effectiveness of classifying echocardiogram videos, thereby contributing significantly to advancements in the field of cardiac imaging and diagnosis. Comparison graph of our model is shown in Figure 8.

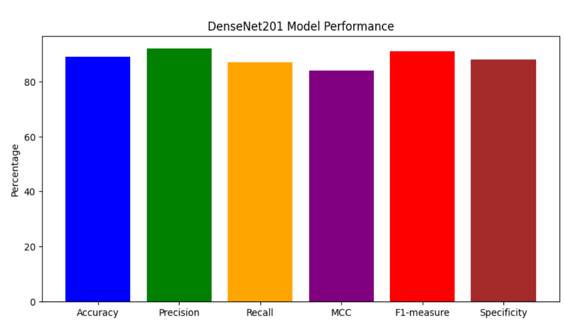

Figure 6: Performance metrics of Densenet201

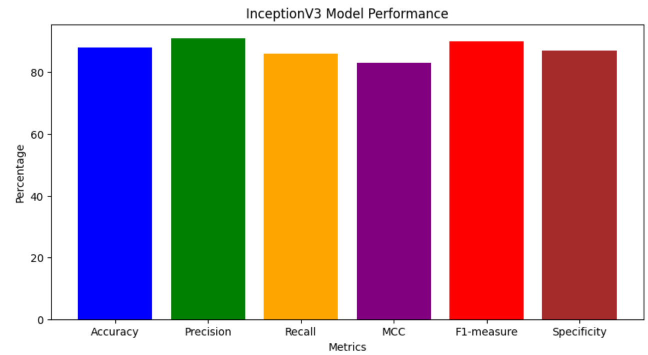

Figure 7: Performance metrics of InceptionV3

Confusion Matrix VGG19 Model

The confusion matrix associated with the VGG19 model is a valuable tool for evaluating its classification performance. It takes the form of a 2x2 matrix detailing the counts of true positive, true negative, false positive, and false negative predictions made by the model. True positive and true negative values signify correctly identified positive and negative instances, while false positive and false negative values represent instances of misclassification. When applied to the classification of echocardiogram videos, the VGG19 model's confusion matrix becomes pivotal in assessing its accuracy and error rates. This matrix facilitates a thorough analysis, allowing for the calculation of precision, recall, and F1-score. This analytical approach provides nuanced insights into the model's proficiency in accurately classifying instances within the specified domain. The confusion matrix vgg19 is shown in Figure 9.

Densenet201 Model

Similarly, the confusion matrix for the Densenet201 model offers a detailed breakdown of its classification outcomes. It enables the evaluation of the model's predictive accuracy and error rates, thereby facilitating the calculation of precision, recall, and F1 scores. By examining the confusion matrix, the Densenet201 model's performance in classifying

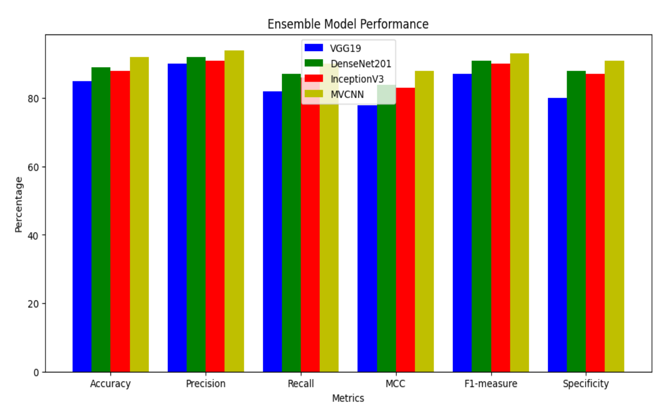

Figure 8: Comparison graph of our model

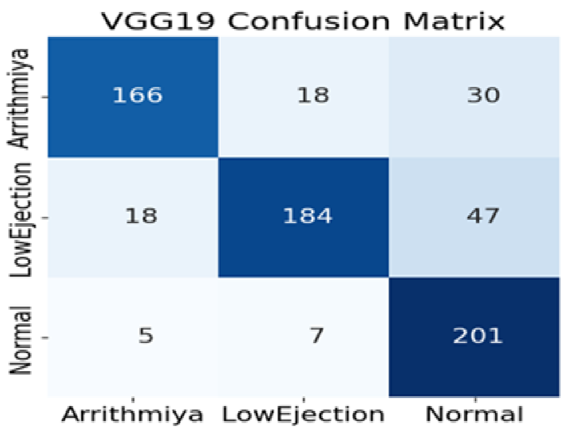

Figure 9: Confusion matrix vgg19

echocardiogram videos can be thoroughly assessed. The confusion matrix DenseNet201 is shown in Figure 10.

Inceptionv3 Model

The confusion matrix generated for the Inceptionv3 model presents a holistic summary of its classification performance. It allows for the calculation of precision, recall, and F1-score, which are essential metrics for evaluating the model's effectiveness in classifying echocardiogram videos. The confusion matrix InceptionV3 is shown in Figure 11.

MVCNN Architecture Ensemble model confusion Matrix

In the context of the integrated approach using the MVCNN architecture, the confusion matrix serves as a crucial tool for assessing the collective classification capabilities of the ensemble models. It enables the comparison of the combined performance of VGG19, Densenet201, and Inceptionv3 within the MVCNN architecture, highlighting the overall classification accuracy and error rates achieved through the ensemble approach. By analyzing the confusion matrix for the integrated approach, a comprehensive evaluation of the ensemble models' classification performance can be conducted. The precision, recall, and F1-score for the ensemble approach can be derived from the confusion matrix, shedding light on the collective effectiveness of the MVCNN architecture in classifying echocardiogram videos. The ensemble confusion matrix is shown in Figure 12.

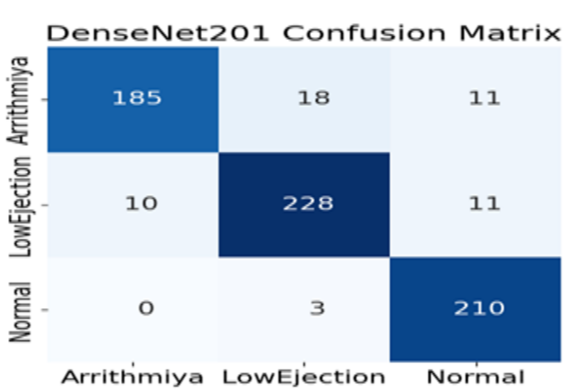

Figure 10: Confusion matrix DenseNet201

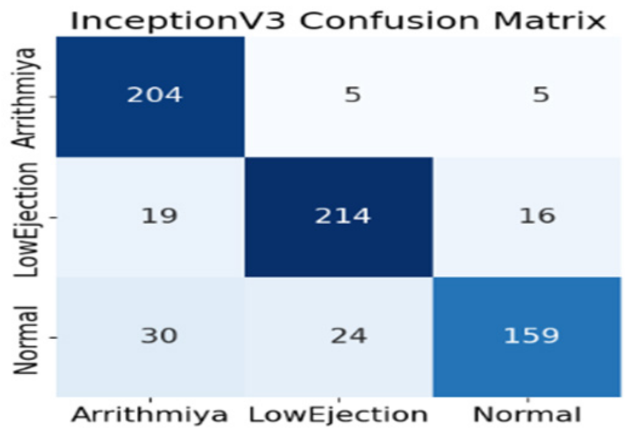

Figure 11: Confusion matrix InceptionV3

ROC Curve of VGG19 Model

The ROC curve of the VGG19 model visually depicts its true positive rate (sensitivity) plotted against the false positive rate (1-specificity) at different threshold values. This graphical representation illustrates the model's effectiveness in distinguishing between positive and negative instances, and a more pronounced curve suggests a greater discriminatory capability. By analyzing the ROC curve for the VGG19 model, the AUC score can be derived, further quantifying its discriminatory performance. The Roc curve for VGG19 is shown in Figure 13.

Densenet201 Model

Similarly, the ROC curve for the DenseNET201 model offers insights into its discriminatory abilities. It portrays the trade-off between sensitivity and specificity, allowing for the assessment of the model's classification performance across different decision thresholds. The AUC score derived from the ROC curve provides a measure of the Densenet201 model's overall discriminative capability. The ROC curve for DenseNET201 is shown in Figure 14.

Inceptionv3 Model

The ROC curve for the Inceptionv3 model serves as a valuable tool for evaluating its capability to differentiate between negative and positive cases. By examining the curve and calculating the AUC score, the efficiency of Inceptionv3 in classification can be quantified, providing a comprehensive understanding of its discriminative performance. The ROC curve for InceptionV3 is shown in Figure 15.

MVCNN Architecture

In the context of the integrated approach utilizing the MVCNN architecture and the ensemble models, the ROC curve offers a collective visualization of the discriminative

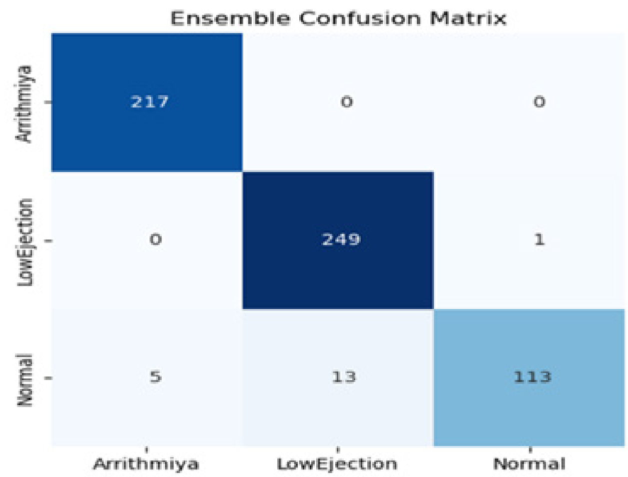

Figure 12: Ensemble confusion matrix

capabilities of VGG19, Densenet201, and Inceptionv3. This curve showcases the combined discriminatory power of the ensemble models and enables the calculation of the overall AUC score, reflecting the integrated discriminative performance within the MVCNN architecture. The Roc curve for an ensemble of MVCNN is shown in Figure 16.

By analyzing the ROC curve for the integrated approach, a comprehensive evaluation of the ensemble models' collective discriminative performance can be conducted, offering valuable insights into the effectiveness of the MVCNN architecture in accurately classifying echocardiogram videos. In summary, the VGG19, Densenet201, and Inceptionv3 models are utilized in transfer learning for the classification of echocardiogram videos. The ROC curves and AUC scores for each model (VGG19, Densenet201, Inceptionv3) as well as the integrated approach using the MVCNN architecture, provide a comprehensive evaluation of their discriminative capabilities. These evaluations are important in finding the efficiency of the framework in accurately classifying echocardiogram videos and can aid in selecting the important framework for accomplishing the given research task.

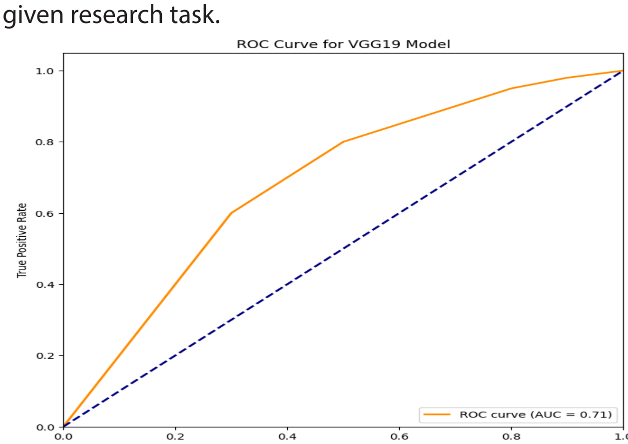

Figure 13: ROC curve for VGG19

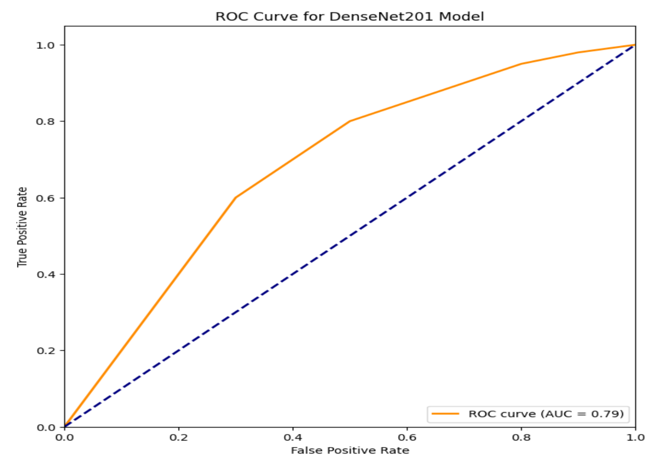

Figure 14: Roc curve for DenseNET201

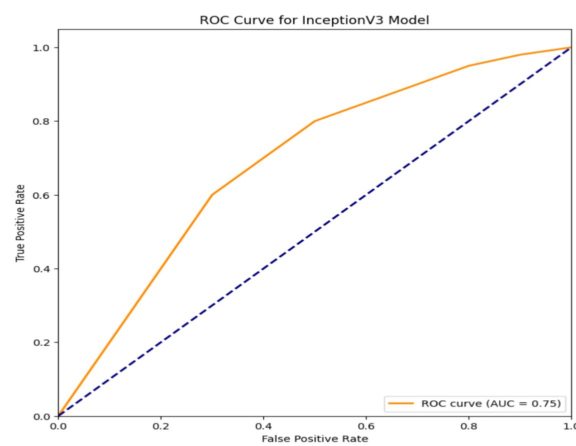

Figure 15: ROC curve for InceptionV3

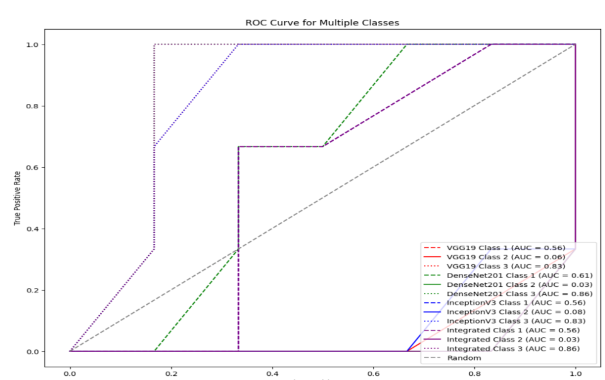

Figure 16: Roc curve for ensembled of MVCNN

Conclusion

In conclusion, the research has demonstrated the effectiveness of utilizing VGG19, Densenet201, and Inceptionv3 models in a transfer learning approach for the classification of echocardiogram videos. The comprehensive evaluation of these models using confusion matrices and ROC curves, along with the calculation of performance metrics, has provided valuable insights into their individual and collective capabilities in accurately classifying echocardiogram videos. The results show that the integration of ensemble techniques, particularly within the MVCNN architecture, has led to an overall enhancement in the classification accuracy of echocardiogram videos. Furthermore, the discriminative performance of the ensemble models has contributed to advancements in the field of cardiac imaging and diagnosis, offering potential benefits for clinical application.

References

Al Reshan, M. S., Amin, S., Zeb, M. A., Sulaiman, A., Alshahrani, H., & Shaikh, A. (2023). A robust heart disease prediction system using hybrid deep neural networks. IEEE Access.

Ambikapathy, B., & Krishnamurthy, K. (2024). Analysis of electromyograms recorded using invasive and noninvasive electrodes: a study based on entropy and Lyapunov exponents estimated using artificial neural networks. Journal of Ambient Intelligence and Humanized Computing, 1-9.

Amin, S., Alouffi, B., Uddin, M. I., & Alosaimi, W. (2022). Optimizing Convolutional Neural Networks with Transfer Learning for Making Classification Report in COVID-19 Chest X-Rays Scans. Scientific Programming, 2022(1), 5145614.

Cheng, P. M., & Malhi, H. S. (2017). Transfer learning with convolutional neural networks for classification of abdominal ultrasound images. Journal of digital imaging, 30, 234-243.

Dong, S., Luo, G., Wang, K., Cao, S., Li, Q., & Zhang, H. (2018). A combined fully convolutional networks and deformable model for automatic left ventricle segmentation based on 3D echocardiography. BioMed research international, 2018(1), 5682365.

Garg, T., Garg, M., Mahela, O. P., & Garg, A. R. (2020). Convolutional neural networks with transfer learning for recognition of COVID-19: a comparative study of different approaches. AI, 1(4), 586-606.

Gupta, M. D., Kunal, S., Girish, M. P., Gupta, A., & Yadav, R. (2022). Artificial intelligence in cardiology: The past, present and future. Indian heart journal, 74(4), 265-269.

Harris, J. K., & Greiner, R. (2023). Deep Learning Approaches for End-to-End Modeling of Medical Spatiotemporal Data. In Advances in Deep Generative Models for Medical Artificial Intelligence (pp. 111-149). Cham: Springer Nature Switzerland.

Howard, J. P., Tan, J., Shun-Shin, M. J., Mahdi, D., Nowbar, A. N., Arnold, A. D., ... & Francis, D. P. (2020). Improving ultrasound video classification: an evaluation of novel deep learning methods in echocardiography. Journal of medical artificial intelligence, 3.

Hu, X., Xu, X., Xiao, Y., Chen, H., He, S., Qin, J., & Heng, P. A. (2018). SINet: A scale-insensitive convolutional neural network for fast vehicle detection. IEEE transactions on intelligent transportation systems, 20(3), 1010-1019.

Jiang, J., Shu, Y., Wang, J., & Long, M. (2022). Transferability in deep learning: A survey. arXiv preprint arXiv:2201.05867.

Kang, G., Liu, K., Hou, B., & Zhang, N. (2017). 3D multi-view convolutional neural networks for lung nodule classification. PloS one, 12(11), e0188290.

Khened, M., Kollerathu, V. A., & Krishnamurthi, G. (2019). Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. Medical image analysis, 51, 21-45.

Li, M., Dong, S., Gao, Z., Feng, C., Xiong, H., Zheng, W., ... & de Albuquerque, V. H. C. (2020). Unified model for interpreting multi-view echocardiographic sequences without temporal information. Applied Soft Computing, 88, 106049.

Litjens, G., Ciompi, F., Wolterink, J. M., de Vos, B. D., Leiner, T., Teuwen, J., & Išgum, I. (2019). State-of-the-art deep learning in cardiovascular image analysis. JACC: Cardiovascular imaging, 12(8 Part 1), 1549-1565.

Litjens, G., Ciompi, F., Wolterink, J. M., de Vos, B. D., Leiner, T., Teuwen, J., & Išgum, I. (2019). State-of-the-Art Deep Learning in Cardiovascular Image Analysis. JACC: CARDIOVASCULAR IMAGING, 12(8).

Madani, A., Arnaout, R., Mofrad, M., & Arnaout, R. (2018). Fast and accurate view classification of echocardiograms using deep learning. NPJ digital medicine, 1(1), 6.

Martinez, J. B., & Gill, G. (2019, December). Comparison of pre-trained vs domain-specific convolutional neural networks for classification of interstitial lung disease. In 2019 International Conference on Computational Science and Computational Intelligence (CSCI) (pp. 991-994). IEEE.

Morid, M. A., Borjali, A., & Del Fiol, G. (2021). A scoping review of transfer learning research on medical image analysis using ImageNet. Computers in biology and medicine, 128, 104115.

Ouyang, D., He, B., Ghorbani, A., Yuan, N., Ebinger, J., Langlotz, C. P., ... & Zou, J. Y. (2020). Video-based AI for beat-to-beat assessment of cardiac function. Nature, 580(7802), 252-256.

Patel, B., & Makaryus, A. N. (2022, January). Artificial intelligence advances in the world of cardiovascular imaging. In Healthcare (Vol. 10, No. 1, p. 154). MDPI.

Priya, S. B., Rajamanogaran, M., & Subha, S. (2021). Prediction of chest diseases using transfer learning. Machine Learning for Healthcare Applications, 199-213.

Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding transfer learning for medical imaging. Advances in neural information processing systems, 32.

Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019, December). Transfusion: understanding transfer learning for medical imaging. In Proceedings of the 33rd International Conference on Neural Information Processing Systems (pp. 3347-3357).

Ravishankar, H., Sudhakar, P., Venkataramani, R., Thiruvenkadam, S., Annangi, P., Babu, N., & Vaidya, V. (2017). Understanding the mechanisms of deep transfer learning for medical images. arXiv preprint arXiv:1704.06040.

Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., ... & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE transactions on medical imaging, 35(5), 1285-1298.

Smith, N., de Vecchi, A., McCormick, M., Nordsletten, D., Camara, O., Frangi, A. F., ... & Rezavi, R. (2011). euHeart: personalized and integrated cardiac care using patient-specific cardiovascular modelling. Interface focus, 1(3), 349-364.

Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning?. IEEE transactions on medical imaging, 35(5), 1299-1312.

Wang, Y., Zhao, P., & Zhang, Z. (2023). A deep learning approach using attention mechanism and transfer learning for electromyographic hand gesture estimation. Expert Systems with Applications, 234, 121055.

Xie, Y., Xia, Y., Zhang, J., Feng, D. D., Fulham, M., & Cai, W. (2017). Transferable multi-model ensemble for benign-malignant lung nodule classification on chest CT. In Medical Image Computing and Computer Assisted Intervention− MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III 20 (pp. 656-664). Springer International Publishing.

Yadav, S. S., & Jadhav, S. M. (2019). Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big data, 6(1), 1-18.

Yang, Y., Guo, J., Ye, Q., Xia, Y., Yang, P., Ullah, A., & Muhammad, K. (2021). A weighted multi-feature transfer learning framework for intelligent medical decision making. Applied Soft Computing, 105, 107242.

Yu, Y., Lin, H., Meng, J., Wei, X., Guo, H., & Zhao, Z. (2017). Deep transfer learning for modality classification of medical images. Information, 8(3), 91.

Yuan, B., & Xing, W. (2019). Diagnosing cardiac abnormalities from 12-lead electrocardiograms using enhanced deep convolutional neural networks. In Machine Learning and Medical Engineering for Cardiovascular Health and Intravascular Imaging and Computer Assisted Stenting: First International Workshop, MLMECH 2019, and 8th Joint International Workshop, CVII-STENT 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 1 (pp. 36-44). Springer International Publishing.

Yuan, B., & Xing, W. (2019). Diagnosing Cardiac Abnormalities from 12-Lead Electrocardiograms Using Enhanced Deep Convolutional Neural Networks. arXiv preprint arXiv:1908.06802.

Zhang, J., Gajjala, S., Agrawal, P., Tison, G. H., Hallock, L. A., Beussink-Nelson, L., ... & Deo, R. C. (2018). Fully automated echocardiogram interpretation in clinical practice: feasibility and diagnostic accuracy. Circulation, 138(16), 1623-1635.