RESEARCH METHODOLOGY ASSIGNMENT-1

VOORE MANASWINI

CB.SC.P2CSE24020

METAHEURISTICS HYBRID MODEL BASED BREAST CANCER CLASSIFICATION ON THERMOGRAM IMAGES

Defining, Redefining and Formalizing Problems

Breast cancer is a serious global health issue and hence detecting it in early stage is very crucial for an effective treatment. Thermography, which uses heat patterns to identify tumors, is really a non-invasive way to detect breast cancer. However, analysing these thermal images is a bit challenging in the process of diagnosis. Current methods are either invasive or rely on basic algorithms.

This research aims to develop a better approach for detecting breast cancer from thermal images by focusing and improving how we select and analyse features in the images and also improving the accuracy of diagnosis, thereby finally classifying them into cancerous and non-cancerous images.

Formulating Hypothesis

The hypothesis is a hybrid metaheuristic model, that combining two optimisation techniques-Genetic Algorithm (GA) and Ant Colony Optimisation (ACO) for feature selection, along with the MobileNetV3 for feature extraction, will improve the accuracy of breast cancer detection from thermogram images. Specifically, we hypothesize that this approach will perform better than the existing model that uses GA and Grey Wolf Optimisation (GWO) for feature selection, in terms of accuracy, for early diagnosis.

Suggesting Solutions or Solution Approaches

The proposed approach or solution involves a model with MobileNetV3 architecture for an efficient feature extraction from thermogram images. The MobileNetV3 model, with its depth-wise convolutions and squeeze-and-excitation blocks, is expected to capture critical features related to thermal breast images. Then, these extracted features are further used as input in developing a hybrid metaheuristic model that sequentially integrates GA and ACO picking out the most relevant features. After the feature extraction and selection, a Random Forest Classifier will be used to classify the images as benign or malignant. Besides this, a comparative analysis with an existing GA-GWO model will be conducted to evaluate the proposed system, GA-ACO.

Collecting, Experimenting and Analysing Data

In this research, we used a dataset of thermographic images from the Database for Mastology Research (DMR-IR), which includes 762 healthy and 760 cancerous samples. These images were pre-processed through resizing and normalisation, to improve their quality before

analysis. Then, feature extraction is performed using MobileNetV3. These extracted features are used as input to obtain more relevant features from them using GA followed by ACO optimisation techniques. Hence, GA-ACO hybrid model results in providing the most relevant possible features which are used for further data analysis.

Data analysis involves evaluating the performance of the model in detecting breast cancer, using various metrics, including accuracy, precision, recall, f1-score, and AUC-ROC across different training and testing split ratios (80-20,75-25 and 70-30). The proposed system involves training the model using Random Forest Classifier.

In the experiment, we found that using an 80-20 split ratio for training and testing data gave us the better accuracy of 99.6%. The proposed model has achieved its benchmark against the existing approach of GA-GWO model, in the aspect of accuracy across generations and the other performance metrics. The experiment showed that our approach of selecting features with GA-ACO improved the accuracy of the breast cancer detection process. Eventually, this refined feature selection also enhanced the classification of images into cancerous and non-cancerous images.

Conclusion

The experimental results validate the hypothesis. The results confirmed our hypothesis that the GA-ACO model is better than GA-GWO model.

The GA-ACO model achieved 99.6% accuracy while GA-GWO model reached 97.5% accuracy. This proves that using GA-ACO for selecting features along with MobileNetV3 for feature extraction is an effective and efficient way to detect breast cancer early from thermogram images and to classify them.

Future Scope

This research or study contributes new insights into breast cancer detection showing that hybrid model of GA-ACO with deep-learning models like MobileNetV3, can enhance the accuracy of breast cancer detection from thermal images, resulting in early diagnosis. This approach could possibly refine further and lead to a better diagnostic tool in future after tremendous research, helping doctors and patients in detecting breast cancer earlier and more accurately and treatment planning. Further research could be built on this work to develop advanced methods, like metaheuristics and deep-learning applications, for analysing medical images.