Mousumi Gupta
Sujata Ghatak
Amlan Gupta
Abir Lal Mukherjee   *Editors*

Artificial Intelligence on Medical Data

Proceedings of International Symposium, ISCMM 2021

Springer

Lecture Notes in Computational Vision and Biomechanics

Volume 37

Research related to the analysis of living structures (Biomechanics) has been carried out extensively in several distinct areas of science, such as, for example, mathematics, mechanical, physics, informatics, medicine and sports. However, for its successful achievement, numerous research topics should be considered, such as image processing and analysis, geometric and numerical modelling, biomechanics, experimental analysis, mechanobiology and enhanced visualization, and their application on real cases must be developed and more investigation is needed. Additionally, enhanced hardware solutions and less invasive devices are demanded. On the other hand, Image Analysis (Computational Vision) aims to extract a high level of information from static images or dynamical image sequences. An example of applications involving Image Analysis can be found in the study of the motion of structures from image sequences, shape reconstruction from images and medical diagnosis. As a multidisciplinary area, Computational Vision considers techniques and methods from other disciplines, like from Artificial Intelligence, Signal Processing, mathematics, physics and informatics. Despite the work that has been done in this area, more robust and efficient methods of Computational Imaging are still demanded in many application domains, such as in medicine, and their validation in real scenarios needs to be examined urgently. Recently, these two branches of science have been increasingly seen as being strongly connected and related, but no book series or journal has contemplated this increasingly strong association. Hence, the main goal of this book series in Computational Vision and Biomechanics (LNCV&B) consists in the provision of a comprehensive forum for discussion on the current state-of-the-art in these fields by emphasizing their connection. The book series covers (but is not limited to):

- Applications of Computational Vision and Biomechanics
- Biometrics and Biomedical Pattern Analysis
- Cellular Imaging and Cellular Mechanics
- Clinical Biomechanics
- Computational Bioimaging and Visualization
- Computational Biology in Biomedical Imaging
- Development of Biomechanical Devices
- Device and Technique Development for Biomedical Imaging
- Experimental Biomechanics
- Gait & Posture Mechanics
- Grid and High Performance Computing on Computational Vision and Biomechanics
- Image Processing and Analysis
- Image processing and visualization in Biofluids
- Image Understanding
- Material Models
- Mechanobiology
- Medical Image Analysis
- Molecular Mechanics
- Multi-modal Image Systems
- Multiscale Biosensors in Biomedical Imaging
- Multiscale Devices and BioMEMS for Biomedical Imaging
- Musculoskeletal Biomechanics
- Multiscale Analysis in Biomechanics
- Neuromuscular Biomechanics
- Numerical Methods for Living Tissues
- Numerical Simulation
- Software Development on Computational Vision and Biomechanics
- Sport Biomechanics
- Virtual Reality in Biomechanics
- Vision Systems
- Image-based Geometric Modeling and Mesh Generation
- Digital Geometry Algorithms for Computational Vision and Visualization

In order to match the scope of the Book Series, each book has to include contents relating, or combining both Image Analysis and mechanics.

Indexed by Scopus.

Mousumi Gupta · Sujata Ghatak · Amlan Gupta ·
Abir Lal Mukherjee

Editors

Artificial Intelligence on Medical Data

Proceedings of International Symposium,
ISCMM 2021

Editors
Mousumi Gupta
Sikkim Manipal Institute of Technology
Majitar, India

Amlan Gupta
Sikkim Manipal Institute of Medical
Sciences
Sikkim, India

Sujata Ghatak
Institute of Engineering and Management
Kolkata, India

Abir Lal Mukherjee
Henry Ford Hospital
Detroit, MI, USA

Preface

Medical information is generated ubiquitously in many settings ranging from personal physician offices and hospitals to government and non-government health agencies. Some of these data are individual data points of single patients, while others are conglomeration of health demographic data in large server databases. Then, there are data that are generated in clinical laboratories testing different healthy persons for screening for disease along with basic science research that are being translated to clinical trials. The analysis of such data may supply insights for decision making from the level of individual patient/health worker to large-scale policy making for the entire population of country, region or racial groups. Such diverse generation of data has given rise to many developments in the method to analyse data. One of the primary concerns of these platforms has been sieving and sifting proper analysable data.

Many computer scientists and health professionals have been working in tandem, but there are also many individual efforts that go undocumented. Thus, this second symposium of ISCMM 2021 was conducted to encourage such individuals and small groups to present their findings and observations. Many of them have innovative ideas or findings that get drowned by a cacophony of larger groups pushing their agendas.

The symposium gave emphasis to use of artificial intelligence in medical images. This could be microscopic or radiologic images and image obtained by signal processing of individual patient to images generated when bid population data visualization is done.

Presentations having application of artificial intelligence in some side of medical information or those pointing to areas of signal processing to create or analyse medical images were included have wider scope for interaction. Papers related to medical image processing and artificial intelligence include neuroimaging for Alzheimer's disease and brain tumour detection, mammography, cancer and tuberculosis detection in chest X-ray images, artificial ophthalmoscopic and retinoscopy images of eye, microscopic images of blood smear, melanoma cells, viral detection in electron microscopy CT images in COVID-19 patients a very contemporary phenomenon. Some applications of artificial intelligence in signal processing of medical or health-related issues have been covered in chapters dealing with wearable

tachycardia monitoring system, walking assistant for vision-impaired. Some authors have looked at artificial intelligence in medical databases, especially those generated by COVID-19 pandemic. Some images not usually disease but particularly important in health care is wearing mask during the pandemic. There are two chapters on use of artificial intelligence to detect mask. It forms a very major issue in Preventive and Community Medicine Departments of Academic Hospitals. One major problem faced by medical professionals is transmission of images without compromising on confidentiality of patients, and there are chapters dedicated to use of encryption and decryption of medical image data.

The chapters are segregated into distinct aspects of intelligent computing in health care, biomedical signal processing; health monitoring; computational image processing; image informatics. This will enable the reader to develop a directed structured approach to research in widely different applications of artificial intelligence on medical science and medical data.

The book aims to create a bridge the postgraduate student and research communities working in the disciplines of Computer Science and Engineering with those medical science. Thus, some of the independent problem-solving exercises are converted to cross-disciplinary activity. However, the researchers working in the allied domain of pattern recognition and others will also find the book useful as it will have some newer approaches, computational intelligence methodologies and applications in medical science.

The editors believe that the selected papers will encourage and give opportunity to both academic researchers and industrial workers to project their research contribution to wider community. Each chapter inspires young scholars to learn newly created avenues of research at an international academic forum. Researchers and research organizations can refer the book for developmental works and incremental researches.

Majitar, India   Mousumi Gupta
Kolkata, India   Sujata Ghatak
Sikkim, India   Amlan Gupta
Detroit, USA   Abir Lal Mukherjee

Contents

Health Monitoring

About the Editors

Dr. Mousumi Gupta is an Associate professor of Computer Application at SMIT with more than a decade of teaching and research experience in the field of pattern recognition. She graduated with majoring in mathematics and did her masters in computer application. She further specialized in Information Technology with an M.Tech. from Sikkim Manipal Institute of Technology (SMIT) in 2007. She completed her Ph.D. from Sikkim Manipal University in 2015 during which she worked in the area of pattern recognition using digital image processing. She started her professional career after joining as a lecturer in the Department of Computer Science and Engineering in 2007 at SMIT. Her present research area include development of pattern recognition algorithm for digital pathology images. Presently, she has two research grants from ICMR and 26 peer reviewed publications.

Dr. Sujata Ghatak is an Assistant Professor of Computer Applications at Institute of Engineering and Management, India. She graduated from Calcutta University majoring in Computer Science following which she received Masters in Computer Science. She completed Master of Technology in Computer science from Sikkim Manipal Institute of Technology (SMIT), Majitar, Sikkim. She is one of the editor for Springer LNNS conference proceedings for IEMIS 2018. Her research interests include Image processing, Pattern Recognition, Applied Dynamical System in image processing, Natural Language Processing and Data Science. She is the member of several technical functional bodies such as The Society of Digital Information and Wireless Communications (SDIWC), Internet Society as a Global Member (ISOC), International Computer Science and Engineering Society (ICSES). She has chaired several sessions in various conferences in India. She has published several papers in reputed journals and conferences.

Dr. Amlan Gupta graduated in medical sciences from NMCH Patna in 1989 and Post-Graduated in Pathology from Calcutta University in 1995. After a brief interlude of private practice, he started his academic career in MCOMS Pokhara as an Assistant professor in 1997. He moved from there to SMIMS Gangtok and joined as an Associate professor in March of 2002. He became a professor in 2007. As a faculty

he has been teaching, researching and taking care hospital diagnostics including surgical pathology and blood banking. His present area of research is wound healing and neural tissue healing. During his tenure he has 32 publications with 4 extramural research grants including one on quantitative histopathology.

Dr. Abir Lal Mukherjee is a board certified anatomic and clinical patholo- gist, working at Henry Ford Hospital, Detroit, Michigan as Chief of Clinical Neuropathology. He is a medical graduate from the oldest medical college in India, Calcutta Medical College. He post-graduated in pathology from JIPMER Pondicherry. He had further specialization in pathology at MD Anderson Cancer Center Houston Texas. He began his academic career in MCOMS Pokhara as assistant professor in 1996. He rose to the position of Associate Director, Pathology Residency Program and Director of Autopsy and Neuropathology, Department of Pathology and Laboratory Medicine, Temple University, School of Medicine, Philadelphia, PA. He is member of major scientific and educational committees, e.g., Graduate Medical Education Committee. He has 19 high impact publications. His primary research interests are in the area of Morphological spectrum of central nervous system malig- nancy and post mortem appearances of CNS in encephalitis and general incidence of undiagnosed malignancies in autopsy.

Biomedical Signal Processing

Comprehensive Study—A Deep Learning and Machine Learning Classification Methods for Cardiogram Images

D. Damodharan and Amit Kumar Goel

Abstract An echocardiogram is one of the heart testing techniques which is widely used in medical field for test and detecting the human heart illness, at present stage of heart functions, diseases, blood flow identification and any other issues to predict. In medicine, the resolution and quality of the pictures are most important. The problem of the medical images is not able to find issues as sooner and also various kinds noises mixed with it. This paper completely analysis the cardiovascular issue in all the multiple fields of technical views surveyed, with basic related issues to cardio diseases pattern identification and its identify the category of the diseases. At most challenging tasks are removing the noises, classifying and recovering the images in advanced technology with less cost. The machine learning techniques and models will help to improve the quality of the pictures and classify the images to practice for clinical and research purpose.

Keywords Classification · Cardiogram · Heart diseases and issues · Medical image

1 Introduction

Due to environmental changes, we see a lot of diseases are facing in the world. Due changing the environment and people food and habits are creating various new type of diseases are entering, so for in this situation only solution is identified early stage and control the diseases in earlier stage. One of the main things that human being needs to know is that cardiovascular disease is one of the foremost cause of deaths among men and women, not breast cancer and so on. Unhealthy food habits, genetic disorders, and age factors like old age people are affected more by these diseases. For Following some basic principle, we got good health and maintain every one body and heath, that is ABCD principles A-active, B-Blood pressure C-Cholesterol and D-Good balanced diet. It will care of diabetes and no smoke, working environment, adequate sleep daily.

D. Damodharan (✉) · A. K. Goel
School of Computing Science and Engineering, Galgotias University, Greater Noida, India
e-mail: itechdamu@gmail.com

In this part delivers images of the correlated similar areas of this broadside such as machine learning, its methods with transitory descriptions, information pre-processing, concert evaluation system of measurement and a brief description of the utmost used sentiment illness dataset.

A. Supervised learning: Supervised learning is an approach to creating AI artificial intelligence model in which are trained using labeled training data and based on the output, it provided input data well correct output data to ML Machine learning model [1, 2].
B. Unsupervised learning: The approach of this method is not included in the dataset. Hear, there is no structure to give a proper format. So as a result, the algorithm attempts to recognize comparisons between input variables besides too classifies them accordingly. The clustering approach is part of the unsupervised learning strategy [3].
C. Reinforcement learning: This method is in managed and unaided realizing, where the model works on its exhibition as it collaborates with the climate. Consequently, it figures out how to address its slip-ups. It should get the right outcome through assessment and evaluate various conceivable outcomes [4].

The limit of the information are had structure and different Patten and different class of in the fields so for the identify impression method to work and practice strategies; especially the whole body various technique that is typically castoff the assessment. In this paper, we generally exertion on the idea that pre-owned association cycles to investigate cardiovascular sickness.

2 Classification ML Techniques

This section provides a brief overview of the furthermost extensively used organization methods used for cardio disease prediction. The methods described here are all examples of supervised mathematics techniques.

2.1 Naive Bayes (NB)

Classifiers based on naive Bayes theorem belong to the probabilistic classifier family. These classifiers assume that all features are independent, and this is essential in making predictions. They are easy to construct and perform well regardless of sample size. Hence, it is suitable for use in the medical science field as well as in diagnosing diseases [5, 6].

As far as naive Bayes (NB) classifiers go, the difference between them is mainly based on their assumptions about how they will model $P(x_i|y)$, where x is the posterior likelihood of class (x, target) given predictor (y, attributes). However oversimplified their assumptions may be, naive Bayes classifiers have proved useful in a number

of real-world applications, including document classification and spam filtering. Detailed estimation of parameters is possible by using a small set of training data.

A classifier is based on naive Bayes method; it can be terribly fast, related to other methods based on additional complex statistical analysis. An extrication of the class restricted features leads to an estimation of each distribution as an independent one-dimensional distribution. Through improving the image dimensional related issues, the curse of dimensionality becomes less severe.

This calculation includes tracking down the *k*-nearest information focuses in the preparation set to the information point for which an objective worth is inaccessible and appointing the normal worth of the discovered information focuses to it.

2.2 Artificial Neural Network (ANN)

Algorithms have been established in order to mimic like the nerve cell in the brain. Each node or neuron is connected to another, and its output is the input of the next node. Despite the fact that each node receives multiple inputs, it only produces one output. The layer involves as an input beside concealed layer and a production layer. Multi-layer perceptions (MLP) are one of the most widely used types of ANN. Under different circumstances, each layer contains a different number of neurons.

A network consists of several components and various layers. Artificial neurons are arranged in large number layers consisting of neural network. Each layer involves an established of artificial nerve cell. Analyzing layers cutting-edge a neural network we will explore various types of layers. Programmers can input data in a variety of formats as indicated by the name. Input and output layers are separated by a hidden layer. By calculating all the hidden features and patterns, it finds hidden patterns and features.

They are equipped with numerical values that can accomplish more than one task simultaneously. Unlike traditional programming, which is stored on one machine, traditional programming uses the network as a storage medium. One or two data pieces disappearing from one place does not affect the network's ability to function. From Fig. 1, AI neural network is training is conducted on inadequate data, and most output may still be produced. It is the importance of missing data that determines the performance loss here. For an ANN to be able to adapt, examples must be determined and the network must be encouraged through these examples in order to achieve the desired output. When the network does not see the event in entirely of its pieces, it will produce false results, because sequence of the network is straight comparative to selected artworks. ANN can still generate output even when some cells have been removed, making it fault-tolerant.

According to Das et al. [7] an artificial neural network (ANN) ensemble technique was designed for the system in question. Through the combination of multiple models trained on the same task, we were able to achieve greater generalization. 89.01% accuracy was observed when model predictions were made using SAS enterprise miner 5.2 software.
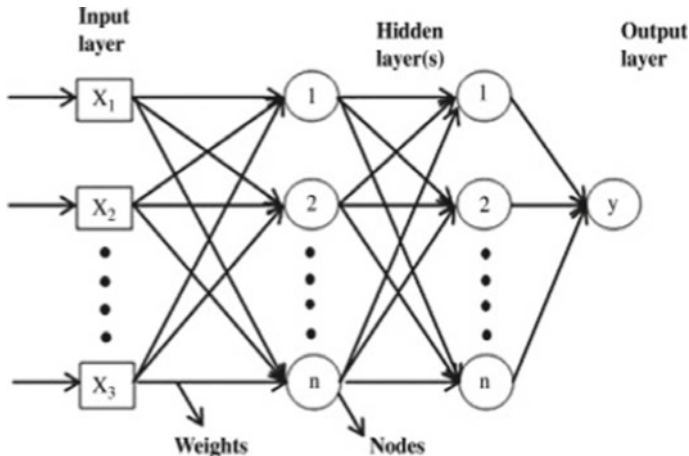
Fig. 1 AI neural networks (ANN)

Chen et al. [8] devised an algorithm that predicts heart disease using artificial neural networks (HDPS). Research performed in this study utilized learning vector quantization (LVQ), a type of ANN. Users can make predictions using the developed system by filling in the thirteen medical attributes. As a result, it displays the prediction result as either healthy or unhealthy, along with the ROC curve and various measurements of accuracy, sensitivity, specificity, and the time spent displaying the result. During development of the system, C programming language was used, as well as C# to implement the user interface. Eighty percentage, 85%, and 70%, respectively, were revealed to be the accuracy, sensitivity, and specificity of the model.

Dagare and Apte in [6] developed a predetermination system (HDPS) to predict whether patients would suffer from heart problems. Both datasets were attained from UCI's source and contain thirteen medicinal characteristics, including the Cleveland heart disease dataset. The precision of this model was improved by adding two additional attributes, obesity, and smoking, making fifteen in all. During the experiment, we used WEKA tool. When using the thirteen attributes, the prediction accuracy was 99.25%, while when using the fifteen attributes, the prediction accuracy was nearly 100%.

2.3 Decision Tree (DT)

This calculation has a DT technique likes flowchart and tree design construction. It comprises of twigs, plant leaves like, hubs besides a cause hub. The inside hubs hold the qualities though the divisions address the aftereffect of separately trial on every hub. DT is broadly utilized meant for grouping resolves since it need not bother with much information in the field or situation the boundaries for it to effort [9].

Fig. 2 Decision tree classification to predict
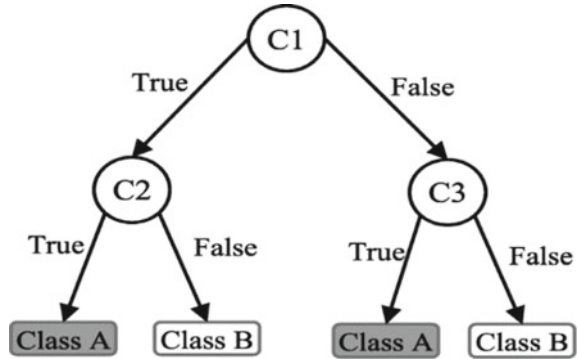
Among supervised machine learning procedures, the decision tree procedure is one of the greatest commonly methods to use for both classification and regression problems can be solved it. The decision tree conducts the tree illustration of the problem to solve the problem of predicting a target variable whose leaf node corresponds to a class label and whose attributes are placed on the internal node of the tree [3] (Fig. 2).

Our training set starts with the root, which we refer to as the entire set. Feature values with categorical attributes are preferred; however, continuous values will be converted to discrete values before a model can be built. Defining Entropy and Information Gain: Before we jump on to decision tree algorithms, we need to understand what is Entropy.

As a measure of disorder, impurity, or uncertainty, entropy measures the amount of impurity. Decision trees have their own way of handling data split according to entropy. As a result, decision trees have different boundary parameters. In a system with entropy values ranging from 0 to 1, less entropy is guaranteed to be trusted.

Equation for Entropy:

$$H(S) = -\text{probability of } \log_2(p+) - \\ - \text{probability of } \log_2(p-) \qquad (1)$$

$(p+)$ % of positive class.
$(p-)$ % of negative class.

Formulating the values:

$$H(S) = -x/y \, \log_2(x/y) - x'/y' \log_2(x'/y') \qquad (2)$$

S refers to the number of positives in the calculation as form this Eq. 1 we are calculated probability xy represents the number of rows.

The same applies if we perform the entropy function, we get a bit, which is an attribute as there are 50% positives and 50% negatives. This splitting will not end, so we will not reach a pure subset. We will have either all positives or all negatives

in this situation since we have a pure subset. In a tree, the steps to split based on which features to focus on at each level are guided by information gain. Simplicity is best when it comes to our tree, so we want it to be small. It is ideal to have the fewest daughter nodes at each step in this process. It is commonly understood that information is a measure of purity in business.

We can determine what information values tell us about a class based on its representation at a particular time point on the tree. First, the divided with the maximum information expansion will be engaged and a subsequent divided will be made till all of the derived nodes have a 0-information gain, or when all the nodes are pure.

Information gain:

$$\text{Gain}(S, A) = H(s) \cdot (|Sv|/|s|) * H(Sv) \tag{3}$$

Based on the algorithm, each split is analyzed for information gain; the split that results in the highest information gain is selected. The information gain corresponds to calculating the average value of the entropy following the split based on all the inputs.

Sv Total of sample after the split.
S Total of Sample.

Information gain plays an important role in determining which split receives selection.

Based on the distribution of classes in a dataset, Gini Impurity of new, random data indicates the likelihood of the class being incorrectly classified if it were assigned a random class label. Considering entropy has a maximum value of 1 and Gini impurity has a maximum value of 0.5, the Gini impurity requires less computational time, as it does not contain a logarithmic function.

A decision tree J48 algorithm was used Prerana and Shivaprakash at in [10] to select features and predict heart disease. For training purposes, 240 records were used, while for testing, 120 records were used. The dataset contained thirteen medical attributes/features. By selecting only certain features, the accuracy increased to 76.67%, which was better than the accuracy achieved by using all features. Additionally, improving accuracy to 85% can be achieved by restricting irrelevant features. The J48 algorithm is said to enable high prediction accuracy by selecting the minimum features.

Patel et al. [11] used several decision tree algorithms to assess the incidence or the presence of cardio disease using the WEKA tool on the UCI dataset. In addition to J48, logistic model trees and random forests were also tested. An accuracy of 56.76% was achieved by the J48 algorithm when compared to the rest.

2.4 K-Nearest Neighbor (KNN)

There are several kinds of machine learning algorithms, but the *K*-nearest neighbors (KNN) algorithm is capable of both regression and classification predictions. But this method is mostly used by the industrial sector to solve problems of classification and prediction. It is a good candidate for defining KNN because it is characterized by two characteristics.

KNN is a lazy learning method that routines all the data for physical activity training and classification, allowing lazy learning. The neural network takes into account no assumptions about what is underlying the data, which makes it a nonparametric learning algorithm (Fig. 3).

In the *K*-nearest neighbors (KNN) algorithm, the values of new data points are predicted by analyzing the characteristics of their features, which is in turn raised on how closely they match the parameters in working the traditional. Here are some steps that will help us understand how it works. The first step to implementing any algorithm is to acquire data. So, we need to load training and test data in the first phase of KNN.

In the next step, we must choose the nearest data points, where *K* is the value. *K* may be an integer of any size. The third step is to apply the following formula to each of the test points:

- Analyze the distance among each row of verification for the test information and every row of working out facts using any of the following methods: Manhattan/Hamming, Euclidean distance. The most commonly used technique to calculate distance is Euclidean.
- Categories the data to ascending order founded on the detachment value.
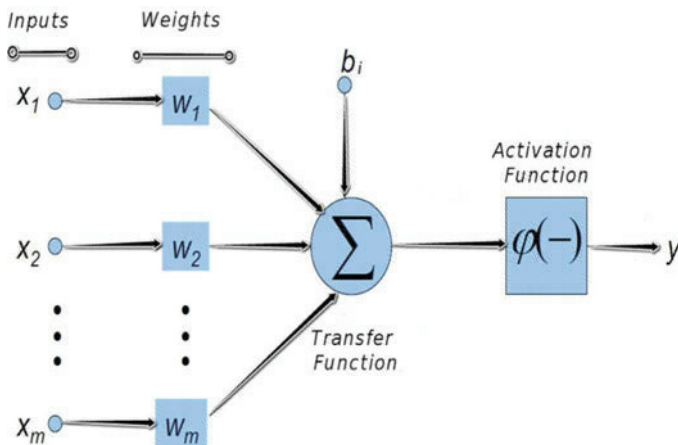
Fig. 3 KNN algorithm architecture diagram

- The following step will select the topmost K rows in the range based on the sort order.
- The class for the test point will now be determined by the most frequent class for the rows.

Based on the Cleveland dataset, Samu et al. in [12] practical K-Nearest Neighbor (KNN) to expect cardio disease. In the study, KNN applications alone and KNN applications with voting were compared. Data are divided into subsets by voting, and each subset applies the classifier. A cross-validation of tenfold is used for the evaluation. Various values for K showed an accuracy range of 94–97.4% without voting. A 97.4% accuracy was attained when $K = 7$. In spite of this, the accuracy did not improve using the voting technique. In the results, we noted a decrease in accuracy at $K = 7$–93%.

2.5 Support Vector Machine (SVM)

This calculation takes a valuable characterization precision. It is characterized by way of limited dimensional direction space which comprises of a measurement for each element/quality of an item.

A.   Hereditary Algorithm

It is a developmental calculation that is constructed dependent on Darwin's hypothesis of advancement. It impersonates techniques in nature like transformation, hybrid, and normal determination. One of the for the most part utilized benefits of the hereditary calculation is its use to introduce loads of the neural organization standard [13]. That is the reason its utilization close by ANN is seen in numerous investigates to create a half and half forecast mode.

B.   Troupe Learning

This technique joins various classifiers into one typical to build the precision. There are n kinds of collaborative learning technique. The main sort is deceiving, which is conglomerating classifiers of the comparative kind by casting a ballot method. Boosting is the subsequent kind, which resembles sacking, so far the new prototypical is influenced by past models' outcomes. Loading is the 3rd sorting which resources collecting AI classifiers for different sorts to deliver one model [14].

Among many supervised learning model, SVMs are used for equally classification and addition to regression. Primarily, it is second hand in machine learning to resolve classification glitches. This algorithm seeks to determine the best decision edge which, along with m-dimensional categories, will allow us to categorize new data points in the future easily. A hyperplane is a boundary that represents the best decision. This helps set up the hyperplane by selecting extreme points and vectors. Thus, support vector machine is called a method that considers these extreme cases as
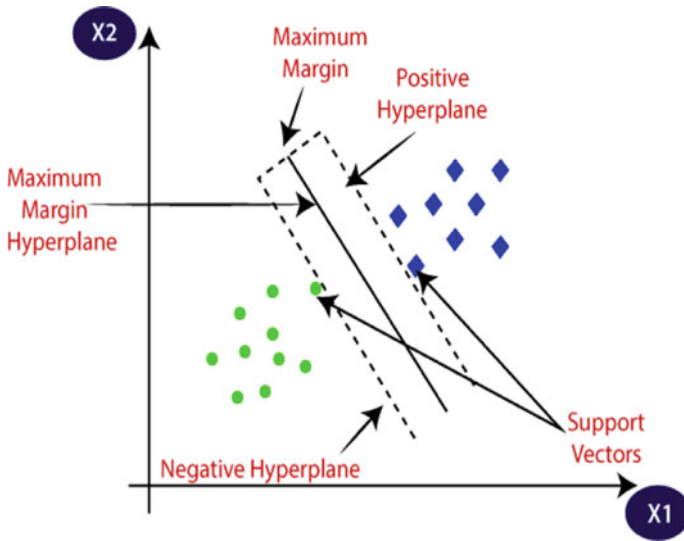
Fig. 4 SVM model architecture with maximum and vector with $x1, x2$

support vectors. As shown in the following diagram, there are two different types of categorizations which are classified by defining a decision boundary or a hyperplane:

By using an example, one may understand how SVM algorithms work. Consider a dataset that has two tags (green and blue). Figure 4: The dataset also has two features $x1$ and $x2$, and these can be found here. A classifier should be able to distinguish green from blue coordinate combinations $(x1, x2)$.

Amelio et al. [15] examined SVM method types on the UCI database data to assess their accuracy in diagnosing heart disease. There were several types of SVMs studied various algorithms. An algorithm called Min–Max scaler was initially applied to the dataset. Using the above SVM algorithms, we then trained the algorithm on the dataset. With a 61.86% overall accuracy, BTSVM had improved performance as compared to the other algorithms.

RNNs are typically regarded as deep learning networks that are used for an unpredictability of tasks related to US image analysis in medicine field [16]. With RNNs, the complexity of the network is resolute by the quantity of sequences of input sample data. Xt and $ht1$, the last two states of a plain RNN, are mapped in a nonlinear fashion to give ht at time t. In this case, W-Weight and Weight-R are both collective over time, with b representing the bias parameter.

3   Conclusion

In our study, analysis reviewed machine learning approaches used in three popular tomography modes in ultrasonography, including an overview of standing machine learning techniques typically convolution neural networks founded that are premeditated for classification and segmenting diverse cardiac image patterns and structures. Furthermore, we discussed the fresh improvement of deep learning model-based segmentation approaches in each of the three segments, charted the possible of the deep learning-based segmentation procedures, and outlined any remaining limitations that could make them unsuitable for wide clinical application. With this review, we intend to provide a deeper understanding of the deep learning algorithms that have contributed many advanced techniques to cardiac image segmentation as well as highlight some of the common challenges that this field faces that will require future research.

References

1. Gárate-Escamila AK, El Hassani AH, Andrès E (2020) Classification models for heart disease prediction using feature selection and PCA. Inform Med Unlock 19:100330. ISSN 2352-9148. https://doi.org/10.1016/j.imu.2020.100330
2. Liu S, Wang Y, Yang X, Lei B, Liu L, Xiang Li S, Dong N, Wang T (2019) Deep learning in medical ultrasound analysis: a review. Engineering.https://doi.org/10.1016/j.eng.2018.11.020
3. Gavali P, Banu JS (2019) Deep convolutional neural network for image classification on CUDA platform. In: Deep learning and parallel computing environment for bioengineering systems, pp 99–122
4. Mohanasundaram R, Malhotra AS, Arun R, Periasamy PS (2019) Deep learning and semi-supervised and transfer learning algorithms for medical imaging. In: Deep learning and parallel computing environment for bioengineering systems, pp 139–151
5. Balaji K, Lavanya K (2019) Medical image analysis with deep neural networks. In: Deep learning and parallel computing environment for bioengineering systems, pp 75–97.https://doi.org/10.1016/b978-0-12-816718-2.00012-9
6. Dangare C, Apte S (2012) A data mining approach for prediction of heart disease using neural networks. Int J Comput Eng Technol 3(3):30–40
7. Das R, Turkoglu I, Sengur A (2009) Effective diagnosis of heart disease through neural networks ensembles. Expert Syst Appl 36(4):7675–7680
8. Chen C, Qin C, Qiu H, Tarroni G, Duan J, Bai W, Rueckert D (2020) Deep learning for cardiac image segmentation: a review. Front Cardiovasc Med 7:25. https://doi.org/10.3389/fcvm.2020.00025
9. Arivazhagan S, Deivalakshmi S, Kannan K, Gajbhiye BN, Muralidhar C, Lukose SN, Subramanian MP (2007) Performance analysis of wavelet filters for image denoising. Adv Comput Sci Technol 1(1):1–10. ISSN 0973-6107
10. Prerana THM, Shivaprakash NC et al. Prediction of heart disease using machine learning algorithms—Naïve Bayes, introduction to PAC algorithm. Compar Algorithms HDPS 3:90–99
11. Zhenya Q, Zhang Z (2021) A hybrid cost-sensitive ensemble for heart disease prediction. In: BMC medical informatics and decision making
12. Samundeeswari ES, Saranya PK, Manavalan R (2016) Segmentation of breast ultrasound image using regularized K-means (ReKM) clustering. In: International conference on wireless communications, signal processing and networking. IEEE, pp 1379–1383

13. Gonzalez RC, Woods RE (2002) Digital image processing, 2nd edn. Prentice-Hall, Englewood Cliffs, NJ
14. Damodharan D, Kumar Goel, A (2021) Novel approach on classification in machine learning model-based USG cardiogram images. In; 2021 international conference on advance computing and innovative technologies in engineering (ICACITE), pp 947–950. https://doi.org/10.1109/ICACITE51222.2021.9404752
15. Amelio L, Amelio A (2018) Classification methods in image analysis with a special focus on medical analytics. In: Intelligent systems reference library, pp 31–69.https://doi.org/10.1007/978-3-319-94030-4_3
16. Damodharan D, Suresh KN, Anandhan K, Jerald Nirmal KS (2019) J Innov Comput Sci Eng 9(1):29–33. Print ISSN: 2278-0947. Online ISSN: 2455-3506

Face Mask Recognition Using Modular, Wavelet and Modular-Wavelet PCA

Kavya Duvvuri, Kanisettypalli Harshitha, Vemula Srihitha, and Sarada Jayan

Abstract The world is currently under the grip of the COVID-19 pandemic. The only possible escape from this pandemic is wearing a mask. Mask checks are being done in most public places. This project brings out a good technique for recognizing a mask on a face. Principal component analysis (PCA) is a dimensionality reduction method which is used in image and signal processing. It uses only the principal components of a dataset and ignores rest of the components. In this paper, three different pre-processing methods (modular, wavelet and a combination of both) have been performed on an image dataset. The resultant data has been processed through PCA. Through comparison of the processed data vectors, the similarity between images has been established. Also, a comparison between different PCA techniques is developed through recognition of distorted faces. Initially, a comparison between different PCA techniques is developed through recognition of distorted faces. Then, a comparison is done between the Modular PCA, Wavelet PCA and Modular-Wavelet PCA techniques by observing the performance for mask recognition. Finally, a conclusion is drawn as to which method is the best in terms of mask recognition.

Keywords PCA · Modular PCA · Wavelet PCA · Modular-Wavelet PCA

K. Duvvuri (✉) · K. Harshitha · V. Srihitha
Department of Computer Science Engineering, Amrita School of Engineering, Bengaluru, Amrita Vishwa Vidyapeetham, Bengaluru, India
e-mail: krishnakavya.duvvuri@gmail.com

S. Jayan
Department of Mathematics, Amrita School of Engineering, Bengaluru, Amrita Vishwa Vidyapeetham, Bengaluru, India
e-mail: j_sarada@blr.amrita.edu

1 Introduction

Most of the popular applications in image processing and signal processing have been possible due to various results from linear algebra [1–6].

Principal component analysis (PCA), that is developed using linear algebra, is a data processing method which ensures dimensionality reduction. It facilitates data storage and analysis. In this paper, PCA in combination with a few pre-processing techniques, is used for face mask recognition. The first section of this paper gives an introduction to the paper along with a brief literature survey. Section 2 describes the methods of image representation, and Sect. 3 details PCA and its three types that are used in this work—Modular PCA, Wavelet PCA and Modular-Wavelet PCA. Section 4 gives the experimental results and analysis, and the paper is concluded in Sect. 5.

According to the eigenface algorithm [2], any grayscale image with dimension $m \times n$ can be represented as a $1 \times (m \times n)$ row vector. A matrix can be formed with its rows as the row vector representations of every image of an image dataset. This matrix is processed using PCA to reduce its dimensions. A query image given by the user must be processed in the same way and compared with the row vectors of the matrix in order to select the most similar image from the given dataset.

The same method can be applied to pre-processed images. There are three different methods used in this paper: modular PCA, Wavelet PCA and Modular-Wavelet PCA [1].

In Modular PCA [3], the image that is to be processed is split into k equal sub-images. These sub-images are separately processed and converted into row vectors. The sub-image matrices are compared with the corresponding sub-image vector of the query image.

In Wavelet PCA [4], the image is first processed through discrete wavelet transformation (DWT). PCA is further applied to the LL sub-band of the obtained transformation. The same method is applied to the query image before comparison. DWT technique is useful in facial expression identification as demonstrated in [7]. In [8] Wavelet PCA has been combined with neuro-fuzzy fusion rule for medical image fusion.

In Modular-Wavelet PCA, a combination of both these methods is used to have the advantages of both. For comparing the similarity of the given query image with the images of the input dataset, Euclidean distance has been used. Interestingly, Euclidean distance can also be used in recognizing facial expressions as shown in [9].

In [10], a comparative study of several face recognition methods like PCA, convolutional neural network (CNN), linear discriminant analysis (LDA) and discrete cosine transformation (DCT) have been illustrated.

2 Image Representation

Image pixels can generally be represented in two different ways: RGB representation and grayscale representation. In RGB representation, the color of a pixel can be represented by a vector of three components which represent the ratio of red, green and blue colors that forms the color of the pixel. In the grayscale representation, the color of a pixel is represented by a single numeric value that represents the grayscale intensity.

In order to be able to represent a pixel using a single numeric value in the row vector, all the images in the image dataset are converted to grayscale before any further processing. After converting the images to grayscale, they are resized to ensure uniform image dimensions throughout the process.

3 PCA and Its Types

3.1 Principal Component Analysis for Image Recognition

An image of $m \times n$ dimension is converted into a row vector of length $m \times n$. All the k images of the given image dataset are converted into row vectors. These row vectors are populated as the rows of a $k \times (m \times n)$ matrix. Mean of each column of this matrix is found, and the means are subtracted from all the respective column elements. The matrix thus obtained is called the mean centralized matrix (X).

The covariance matrix of the mean centralized matrix is found using Eq. 1:

$$\text{Cov}(X) = \frac{X^T X}{k - 1} \tag{1}$$

Now, the eigenvalues and eigenvectors of the covariance matrix are found, and the eigenvalues are arranged in decreasing order. The first i eigenvalues are considered as the most important, and the corresponding eigenvectors are arranged as the column vectors of a matrix. The matrix thus obtained is called the PCA matrix. The reduced matrix is obtained by multiplying the mean centralized matrix with the PCA matrix (P).

$$Y = PX \tag{2}$$

Equation 2 shows the formula to obtain the final reduced matrix using the mean centralized matrix and the PCA matrix (P). The query image given by the user is represented as a row vector after converting into grayscale. This row vector is centralized using the mean vector of the image dataset matrix. This centralized vector is multiplied with the PCA matrix obtained from the dataset matrix. The resultant vector is compared with each row vector of the reduced matrix to obtain the Euclidean
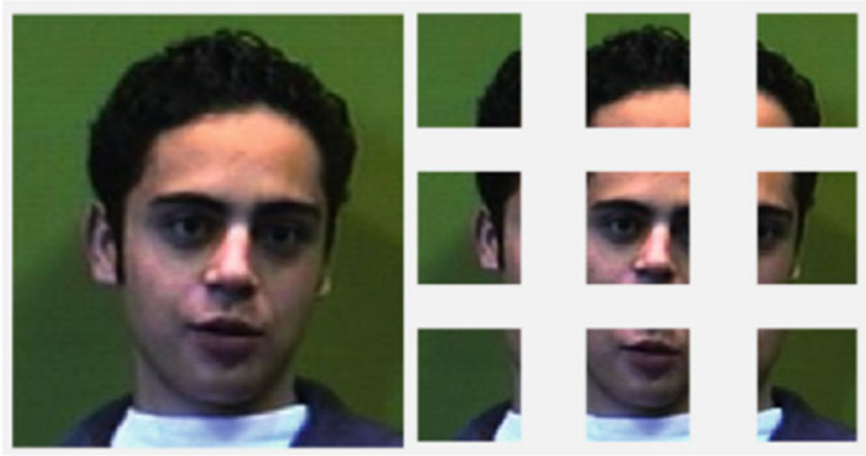
Fig. 1 Image before and after cropping into nine sub-images

distance vector. The image corresponding with the row vector with minimum distance from the centralized vector is said to be the recognized image.

3.2 Modular PCA

In Modular PCA all the images in the given image dataset are divided into k equal sub-images. Figure 1 gives an image before and after cropping into nine sub-images.

Corresponding PCA matrices and reduced matrices are obtained for each of these sub-images. The query image given by the user is also similarly divided into sub-images. A row vector is obtained for each of the sub-query image. The row vectors are compared with the rows of the matrices of the corresponding sub-images. Thus, each sub-query-image is matched with a recognized sub-image.

3.3 Wavelet PCA

In Wavelet PCA, each image of the given image dataset is processed through wavelet transformation. Wavelet transformation is a process of image compression. The first level Haar [11] decomposition of an image gives four sub-bands as the output: LL, LH, HL and HH. Figure 2 gives the images before and after Haar decomposition. LL sub-band [1] gives the approximation coefficients of the image and is approximately at half resolution of the original image. The LH sub-band gives the vertical detailed coefficients of the original image, while the HL sub-band gives the horizontal detailed
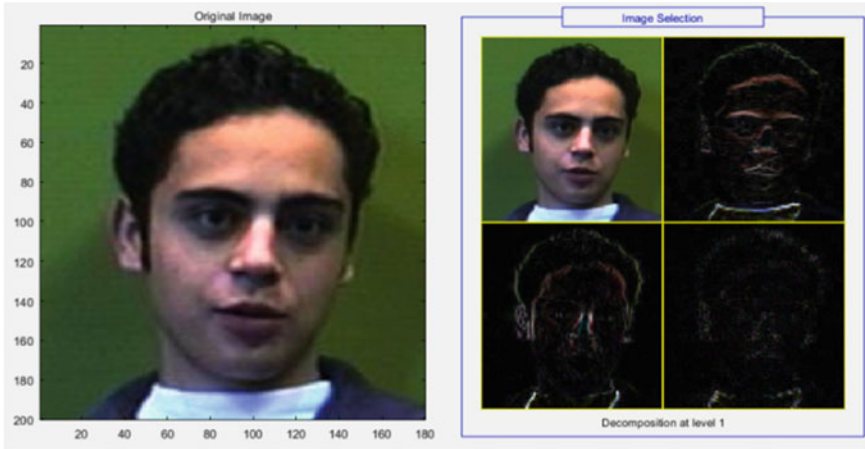
Fig. 2 Image before and after Haar decomposition

coefficients of the original image. The HH sub-band gives the diagonal detailed coefficients of the original image.

In this method, all the images in the given image dataset are decomposed using first level Haar transformation and each LL sub-band is processed through PCA. The given query image is also decomposed into the sub-bands, and the LL sub-band is processed through PCA. Finally, the query image vector is compared with the row vectors of the dataset matrix. The most similar image is identified and displayed.

3.4 Modular-Wavelet PCA

Modular-Wavelet PCA is a combination of Modular and Wavelet PCA. The images are first divided into k equal sub-images. These sub-images are decomposed into the four sub-bands using the first level Haar transformation. The LL sub-band of each sub-image is thus obtained and processed through PCA. The given query image is also similarly processed. The image vector is again compared with the row vectors of the reduced matrix, and then, the image with minimum distance is selected.

4 Experiments, Results and Analysis

In this work, two datasets have been used: Faces94 database [12] and Kaggle face mask dataset [13]. The Faces94 database has been used to compare plain PCA, Modular PCA, Wavelet PCA and Modular-Wavelet PCA. A set of 90 faces (three per individual) have been given as the image dataset. This dataset is loaded into the