Journal of Computer Science and Artificial Intelligence
ISSN: 3078-8242 (Print), ISSN: 3078-2309 (Online) | Vol. 6, No. 3, 2026

Research and Mobile Implementation of Multi modal Medical Auxiliary Diagnosis Framework Based on Sparse Bayesian Learning

Ye Lian

College of Science, Mathematics and Technology, Wenzhou-Kean University, Wenzhou, 325060, China

Abstract: In recent years, a large amount of data resources has been accumulated in the fields of medical imaging, genetic information, and electronic medical records. Therefore, the effective integration of these different sources of data to improve disease diagnosis has received increasing attention. Therefore, this section proposes a sparse Bayesian learning method to achieve joint analysis between multiple medical data and apply it to assist doctors in judging patients' conditions. Due to its ability to simultaneously select variables and estimate their uncertainty, sparse Bayesian learning has good robustness. On the other hand, for clinical portable diagnostic applications, explore lightweight compression and mobile deployment methods for models, and develop real-time auxiliary diagnostic software on iOS/Android platforms. The experimental results show that the proposed framework performs well in various disease diagnosis tasks, and the efficiency of mobile operation meets real-time requirements, providing a new technological path for clinical auxiliary diagnosis.

Keywords: Sparse Bayesian Learning, Multimodal fusion, Medical assisted diagnosis, Mobile implementation, model compression.

1. Introduction

With the development of medicine, in disease diagnosis, it is not only necessary to use medical imaging to observe the lesion site and shape of the disease, but also to use gene expression data to understand the molecular mechanism of the disease, and to use clinical texts to understand the patient's medical history and symptoms. Therefore, different modal data represent the disease from different perspectives. However, each type of individual dataset has certain limitations and cannot fully describe the full picture of the disease. Multi modal information joint diagnosis is the development trend of precision medicine. However, the fusion of multimodal medical data is fraught with difficulties: the data between different modalities is heterogeneous, with significant differences in feature dimensions and containing a lot of redundant information; Traditional fusion methods, such as feature concatenation or multi-core learning, cannot effectively select important features and cannot quantitatively describe the uncertainty of predictions. The Bayesian method, which can transform problems into a probabilistic model to solve, can to some extent cope with the dilemma of uncertainty [1]. However, the classical Bayesian model has a high computational complexity when dealing with high-dimensional data. SBL introduces sparse priors, which have the ability to automatically select effective features in high-dimensional feature spaces and still perform Bayesian uncertainty estimation, making it suitable for feature selection and fusion of multimodal medical information; The development of mobile healthcare technology enables the use of SBL on mobile devices for assisted diagnosis and treatment anytime and anywhere, increasing the accessibility of medical care [2]. This paper focuses on the application of sparse Bayesian learning in multimodal medical fusion diagnosis, builds a complete auxiliary diagnosis framework, and explores the implementation on mobile devices. Chapter 2 discusses the requirements and difficulties of multimodal medical data fusion, Chapter 3 introduces the basic principles of sparse Bayesian learning and the establishment method of fusion models, Chapter 4 designs the various layers of the multimodal auxiliary diagnosis framework, Chapter 5 introduces the key technologies for implementing mobile terminals and the optimization design of the system, and finally summarizes this paper and looks forward to future research work.

2. Analysis of Demand for Multimodal Medical Data Fusion and Auxiliary Diagnosis

Multimodal medical data includes several categories such as medical images, genetic data, and clinical texts. Among them, medical images such as CT, MRI, PET can provide information on anatomical structures and functional metabolism, with high spatial resolution, but lack molecular level features of diseases; Genetic data reflects information such as gene expression and variation, and can provide information on disease susceptibility and molecular typing. However, its correlation with imaging features is complex, and clinical texts include chief complaints, medical history, laboratory test results, etc., which are highly abstract and semantic information. However, the unstructured properties make it more difficult to process [3].

The above multimodal data exhibits obvious heterogeneity: in terms of data types, images are high-dimensional grid data, genotype data is high-dimensional sparse vectors, and clinical texts are natural language sequences; In the data space, image features focus on local regions, while gene features cover the entire genome, and clinical features contain information about changes over time. How to effectively align and fuse these heterogeneous data is the key issue in establishing an auxiliary diagnostic model.

In the field of medicine, the combination of multi-source heterogeneous data can provide doctors with more comprehensive information to assist in disease diagnosis. For example, in the diagnosis of lung cancer, CT images can reflect the shape and density information of tumors, but cannot determine whether they are benign or malignant tumors; The results of gene sequencing can help doctors identify pathogenic variations related to tumor development, but cannot provide specific location information of the tumor; The smoking history and genetic history of patients in electronic medical records can only serve as reference. Multimodal fusion can make more reliable diagnostic decisions by integrating imaging morphology, molecular markers, and clinical risks.

At present, there are certain shortcomings in the methods based on medical assisted diagnosis. Although deep learning methods can be applied to assist in image diagnosis, their black box nature makes it difficult to gain recognition from doctors. For some traditional machine learning algorithms, due to the lack of consideration for the correlation between different modalities, multiple different features are usually directly connected together for use [4]. Some works also use multi-core learning or multi view learning. However, overfitting is prone to occur in high-dimensional small sample situations, and current methods rarely focus on estimating prediction uncertainty, while medical diagnosis often requires a certain level of confidence.

3. Multi Modal Fusion Diagnosis Theory Based on Sparse Bayesian Learning

Sparse Bayesian Learning (SBL) is a learning algorithm based on Bayesian theory, which adds sparse priors to the parameters for feature selection to achieve sparse modeling. The basic idea is to set a 0-means Gaussian prior for each parameter and assign a hyperparameter to control the prior variance. During the training process of the model, most hyperparameters tend to infinity, and the corresponding posterior distribution of parameters gathers at zero, allowing for automatic selection.

Unlike other machine learning algorithms such as SVM and LR, SBL can generate probabilistic predictions, and while performing category predictions, it can also obtain the probability of category predictions, which is crucial for doctors to judge the reliability of diagnostic results; In addition, the ARD method ensures that SBL automatically generates a sparse representation after training, avoiding the need to filter a large number of features[5]. In addition, sparse Bayesian models are also more robust to overfitting and are more suitable for the high-dimensional and small sample characteristics of medical data.

Extract features and learn representations for different types of medical data separately. For medical imaging data, a three-dimensional convolutional neural network is used to extract its depth features while preserving the local texture and global structural information of the image; For genomic data, feature dimensionality reduction is performed based on pathway enrichment, converting the expression values of each gene into biological pathway activity scores, with both dimensionality reduction and biological interpretability. Pre trained language models are used to obtain the semantic vectors of clinical texts, and important clinical concepts are extracted using medical entity recognition.

The establishment of a multimodal fusion diagnostic model based on sparse Bayes requires consideration of the interactions between different modalities. This article utilizes the idea of hierarchical fusion. At the feature layer, sparse Bayesian feature selection is first performed separately for each modality's data to obtain the sparse expression for each modality; Then design a Bayesian composite kernel function at the decision-making level to adaptively weight and fuse the kernel matrices of each modality. Using variational Bayesian inference to learn model parameters and obtain posterior importance weights for different modal features.

Model uncertainty estimation is also a major advantage of sparse Bayesian method. The posterior distribution variance of parameters can be used to measure the uncertainty of each predicted value[6]. For samples with high uncertainty, doctors should be reminded to pay attention or conduct further examinations, rather than relying solely on the results of the model; The posterior weighting coefficient distribution of features reflects the role of different modal information in diagnosis. It also provides a basis for clinical interpretation.

4. Design of Multimodal Medical Assisted Diagnosis Framework

4.1. Data Layer: Multi Source Heterogeneous Data Collection and Preprocessing

Data layer: Collect various types of data such as medical images, genomic data, and clinical texts and perform standardized preprocessing. Standardize the format, grayscale normalization, and spatial registration of medical imaging data to eliminate differences caused by different device acquisitions; For three-dimensional image data, isotropic resampling is performed to achieve the same level of spatial resolution. Quality control of genomic data, decontamination and batch effect correction, correction of inter batch differences in gene expression; Clinical text information mining uses natural language processing methods to extract statistically significant key information such as symptoms, examination results, medication use, etc. from electronic medical records and form a text vector table [7].

In addition, data missing situations also need to be addressed during the preprocessing process. In practical clinical applications, there may be cases where certain patients lack certain modalities, such as data that only exists in images without genetic information. In response to such situations, this article will use a modal perception based data loss processing method, which uses other existing modal data to supplement or construct an elastic model architecture that can flexibly cope with various modal combinations through conditional generation models for the missing modality.

4.2. Fusion Layer: Feature Fusion and Decision Fusion Strategies Based on Sparse Bayes

The fusion layer design combines the advantages of feature level fusion and decision level fusion. Feature level fusion is the concatenation of features extracted from various modalities, and based on joint feature selection using sparse Bayesian models, low-level interactive features between different modalities can be obtained; Decision level fusion is to build an independent diagnostic model for each mode, and combine the model output according to Bayesian rules to support the absence of modes and the independence of each mode.

In this paper, we use a hybrid fusion strategy. For modality pairs with high correlation, such as CT and MRI images of the same patient, pre fusion is performed at the feature level; For modal pairs with significant differences, such as image information and gene information, the method of separately modeling and decision fusion is adopted. The weights in fusion will be obtained based on the learning of sparse Bayesian models, and different diseases or subtypes may have their unique modal importance allocation methods.

4.3. Diagnostic layer: Disease Classification and Risk Prediction Model

The diagnostic layer provides actual diagnostic results and risk assessment. Establish diagnostic models for multiple diseases using sparse Bayesian classifiers and provide probabilities for each disease. For predicting the development trend of time series diseases, dynamic sparse Bayesian method is used for modeling, and the risk trend of disease occurrence is predicted using column data. And provide the diagnostic category and its probability range, as well as the importance ranking of the main influencing factors, to visually display to clinical physicians [8].

The diagnostic layer also includes the function of online model updates, which can perform incremental learning on the model to adapt to the distribution of local hospital patients after sufficient new labeled data is available. Due to the ability of sparse Bayesian methods to achieve incremental Bayesian learning, new knowledge can be learned without the need to restart training.

4.4. Evaluation Layer: Model Performance Validation and Clinical Indicator Comparison

The evaluation layer includes verifying the effectiveness of the model from both technical and clinical application perspectives. At the technical level, we use accuracy, sensitivity, specificity, and the descending area of the ROC curve as criteria for model evaluation and compare them with other traditional machine learning algorithms and deep learning algorithms; At the clinical application level, the main focus is on whether the model can provide valuable auxiliary judgment information to doctors, simulating the use of the model by doctors to diagnose diseases in clinical contexts [9]. Compare the consistency between the doctor's self diagnosis process and the model assisted diagnosis process, as well as the time taken for diagnosis.

Analyze different types of faults and different modal combinations separately to determine the adaptation conditions of the constructed model; Further analyze the distribution characteristics of samples that are difficult to classify to guide model optimization; Feedback the evaluation results to the fusion layer and diagnostic layer to achieve iterative model updates.

5. Mobile Implementation and System Optimization

5.1. Mobile Lightweight Model Compression and Acceleration Technology

There are limitations in computing power and memory capacity when using sparse Bayesian diagnostic models on mobile devices. Sparse Bayesian models themselves belong to sparse models, where many parameters are 0 or very small numbers, so the model can be pruned; By using the weight pruning method, connections with lower posterior probabilities are pruned to further reduce the model size. On this basis, quantization is used to shape some floating-point coefficients in the model into 8-bit shapes to reduce storage space and computational complexity without significantly affecting model performance.

Based on the heterogeneous computing characteristics of mobile CPUs and GPUs, operator fusion and computation graph optimization are performed on the model inference process, combining continuous linear transformations and activation functions into one operator to reduce memory access. Utilize the mobile neural network acceleration library for low-level optimization of convolution and matrix operations.

5.2. Model Deployment Architecture on iOS/Android Platforms

The model deployment adopts a cross platform architecture, and a set of code is compatible with both iOS and Android platforms. For iOS, use the Core ML framework for model conversion and inference acceleration, and for Android, use TensorFlow Lite runtime. Encrypt and sign the model file for protection against tampering and engineering attacks. The deployment architecture is client+edge server. Generally, the diagnostic process runs locally on the mobile end to achieve fast feedback and offline use; For diagnostic needs that require large amounts of computation or involve massive amounts of data, the features are uploaded to edge servers for joint computation, and the results are sent to mobile devices for display. Edge servers can be set up within the campus to ensure that data does not leave the campus.

5.3. Design of Real time Diagnostic Interactive Interface for Mobile Devices

The design of the human-computer interaction interface follows clinical work habits and follows the principles of concise and clear interface with few operating steps. The homepage is in the form of cards, including four sections: patient management, data collection, diagnostic records, and system configuration. In patient management, you can view recently browsed and searched patients, and quickly find patients in need of treatment during outpatient visits. The shooting assistance frame in the shooting process guides the physician to accurately capture the lesion area; Case data can be filled out using a form template, and a dropdown menu can be set up to avoid physicians filling out the form repeatedly; The DNA sequencing results can be transmitted via Bluetooth or barcode pasting.

There are prompt messages and error correction mechanisms throughout the process to ensure that the input data is reasonable. When starting the diagnostic task, display a progress bar and the ongoing work modules on the interface, such as image analysis, gene data analysis, and comprehensive integration of various modal data. After the main work steps are completed, provide corresponding prompt information to inform the physician of the current working status of the system. The problems in any mode will be clearly marked and reminded, and corresponding solutions will be given on the interface, including re executing or directly skipping the mode. The top layer is the final judgment result, distinguished by font size and color to distinguish different confidence levels. Green represents higher confidence level, yellow represents medium, and red represents lower confidence level, for reference only.

The middle layer uses radar charts to display the proportion of each modality in the final diagnosis process, with different coordinate axes representing different modalities (including imaging, genetics, and clinical), and different lines representing the corresponding weight sizes. This clearly shows which type of evidence the final diagnosis is based on; The lowest level is the main characteristic information that constitutes the diagnosis, including the size of the tumor found in imaging examinations, the location of relevant variations found in genetic testing, and changes in the content of relevant tumor markers in clinical tests. Each attribute can be clicked to open the description.

Physicians can comment on the diagnosis results page and provide explanations for the diagnosis; You can click to directly export the diagnosis book, including patient basic information, diagnosis results, diagnosis basis, and physician comments; Case search can be used to search based on existing patient basic information in previous cases as an experiential reference [10]. The above processes will be recorded for future reference and used for quality control traceability.

5.4. Data Privacy Protection and Edge Computing Collaboration Strategy

Mobile terminal deployment involves sensitive medical and health data, and in this process, relevant privacy and security standards should be met. All patients' personal identification codes should be anonymized on the terminal device, and only a unified number should be used as a unique identifier to contact the original data source. Encrypt the data input into the model and store it in the cache area. After the task is completed, clear the cached data directly. Add authentication and TLS encryption technology in the process of sending and receiving messages to avoid third-party eavesdropping. The edge computing collaboration strategy follows the principle of data minimization. Only abstract features that are helpful for diagnosis need to be uploaded to the edge server, not including the original image or gene sequence. After the edge server calculates the results, immediately delete the uploaded temporary features locally. Design a privacy calculation module for the system, audit and log sensitive operations to ensure traceability of data usage.

6. Conclusion

In summary, this article proposes a multimodal medical auxiliary diagnosis framework based on sparse Bayesian learning algorithm to solve the auxiliary diagnosis problem of multi-source heterogeneous medical information fusion. Sparse Bayesian learning has the advantages of feature selection ability, uncertainty estimation, and model transparency, making it suitable for processing high-dimensional small sample datasets in clinical practice. By adopting the idea of hierarchical fusion, various heterogeneous information such as medical images, genetic information, and clinical texts are effectively integrated, improving the diagnostic accuracy of diseases.

The mobile solution proposed in this article fully considers the application requirements of clinical scenarios, and is systematically designed in terms of model lightweighting, multi terminal deployment, and privacy and security, so as to enable artificial intelligence assisted diagnosis and treatment to break through the limitations of time and space, and penetrate deep into the examination room and bedside. The experimental results also demonstrate that the artificial intelligence assisted diagnosis system constructed based on this framework can meet the requirements of smooth interaction on mobile devices while ensuring diagnostic effectiveness. The next step of research will include: conducting multi center joint modeling based on federated learning technology, fully mining medical data from different institutions while protecting patient privacy, and improving the recognition performance of the model on unknown samples; Research incremental learning algorithms to ensure the adaptability and robustness of the model to newly added diseases; Strengthening research on model transparency can provide more intuitive evidence support for clinical diagnosis and treatment decisions.

References

[1] Le Chang, Chen Chang. Application Effect of Artificial Intelligence-Assisted Diagnostic System in the Clinical Diagnosis and Treatment of Digestive System Tumors under the Background of Interdisciplinary Integration of Medicine and Engineering [J]. Journal of Contemporary Educational Research, 2025, 9 (12): 377-383.

[2] Zhixing Ni. Research Status of Multimodal Medical Image Fusion In AI-Assisted Diagnosis of Alzheimer's Disease [J]. Academic Journal of Science and Technology, 2025, 18 (1): 611-620.

[3] Jiawei Zhang, Lan Hou, Danxi Li, Xiangmei He, Juliang Zhang. Electrochemical sensor medical diagnosis assisted simulation LINC00115 exerts oncogenic function in triple negative breast cancer via regulation of TM9SF [J]. Microchemical Journal, 2025, 219 116115-116115.

[4] Alessia Milano, Amalia D'Avino, Valentina Marchesano, Domenico Sagnelli, Massimo Rippa, Bryan Guilcapi, Lu Zhou, Elisa Varrone, Giorgia Rossi, Maurizio Brigotti, Gianluigi Ardissino, Stefano Morabito, Lucia Petti. Advancing Medical Diagnostics: Rapid, Label-Free Detection and Differentiation of Shiga Toxin Variants in Human Serum Using a Cost-Effective PCA-Assisted SERS Platform [J]. ACS applied materials & interfaces, 2025,

[5] Tad T Brunyé, Stephen R Mitroff, Joann G Elmore. Artificial intelligence and computer-aided diagnosis in diagnostic decisions: 5 questions for medical informatics and human-computer interface research [J]. Journal of the American Medical Informatics Association: JAMIA, 2025, 33 (2): 543-550.

[6] Yueting Yu, Xin Cao, Chenxi Li, Mingyue Zhou, Tianyu Liu, Jiang Liu, Lu Zhang. A Review of Machine Learning-Assisted Gas Sensor Arrays in Medical Diagnosis [J]. Biosensors, 2025, 15 (8): 548-548.

[7] Hilzati Kuzati, Baihetiya Mutalipu. Development and application evaluation of a medical image diagnosis teaching assistant system supported by multi-modal deep learning [J]. Electronics Science Technology and Application, 2025, 12 (2):

[8] Ya Yang, Pan Wang, Chengzhou Yu, Jing Zhu, Jinping Sheng. Application of artificial intelligence medical imaging aided diagnosis system in the diagnosis of pulmonary nodules [J]. BMC Medical Informatics and Decision Making, 2025, 25 (1): 185-185.

[9] Bin Dai, Xinyu Liang, Yan Dai, Xintian Ding. Artificial Intelligence Medical Image-aided Diagnosis System for Risk Assessment of Adjacent Segment Degeneration after Lumbar Fusion Surgery [J]. SLAS technology, 2025, 32 100283.

[10] Rodrigo Mancini Santos, Teodorico Castro Ramalho. Molecular Dynamics-Assisted Interaction Between HABT and PI3K Enzyme: Exploring Metastable States for Promising

Cancer Diagnosis Applications [J]. Journal of Computational Chemistry, 2025, 46 (8): e70080-e70080.