International Journal of Health Sciences and Research
Vol.14; Issue: 1; January 2024
Website: www.ijhsr.org
ISSN: 2249-9571

Review Article

A Review on Ensemble Machine and Deep Learning Techniques Used in the Classification of Computed Tomography Medical Images

Sheik Imran[1], Pradeep N[2]

[1] Department of Information Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India.
[2] Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India.

Corresponding Author: Sheik Imran

DOI: https://doi.org/10.52403/ijhsr.20240124

ABSTRACT

Ensemble learning combines multiple base models to enhance predictive performance and generalize better on unseen data. In the context of Computed Tomography (CT) image processing, ensemble techniques often leverage diverse machine learning or deep learning architectures to achieve the best results. Ensemble machine learning and deep learning techniques have revolutionized the field of CT image processing by significantly improving accuracy, robustness, and efficiency in various medical imaging tasks. These methods have been instrumental in tasks such as image reconstruction, segmentation, classification, and disease diagnosis. The ensemble models can be divided into those based on decision fusion strategies, bagging, boosting, stacking, negative correlation, explicit/implicit ensembles, homogeneous/heterogeneous ensembles, and explicit/implicit ensembles. In comparison to shallow or traditional, machine learning models and deep learning architectures are currently performing better. Also, a brief discussion of the various ensemble models used in CT images is provided. We wrap up this work by outlining a few possible avenues for further investigation.

Keywords: *Computed Tomography, Ensemble, Deep learning, Machine Learning.*

1. INTRODUCTION

The use of technology in the health sector is expanding quickly, and one aspect of this growth is the development of medical imaging tools, which help to streamline medical imaging procedures. It alludes to a facet of medical procedures in the modern era where technology has supplanted conventional methods. The development of technology has played a significant role in improving imaging, which has improved medicine. Traditional approaches to diagnosing and clarifying the results of image processes take a long time to process, are prone to human error, and result in a general outcome that is not well aligned with history because the earlier viewpoints are hard to compare. These restrictions drove our study's efforts to shed light on the potential uses of Ensemble Machine Learning and Deep Learning Algorithms for the Prediction and Classification of Medical Images. [1]

Medical image analysis has been completely redesigned using deep learning, which produces outstanding results for image processing tasks like feature extraction, registration, segmentation, and classification. [2] The branch of artificial intelligence that is developing the fastest is called deep learning, and it has been applied successfully recently in many fields, including medicine. An overview of the research conducted in the following areas of application is provided: computerized pathology, brain, retina, neuro, bosom, heart, breast, bone, stomach, and musculoskeletal. Deep learning networks can be effectively used with large data for knowledge deployment, information exploration, and knowledge-based prediction. [4]

1.1 Medical imaging

Medical imaging pertains to the utilization of diverse methodologies to get diverse picture modalities from the human body, specifically the impacted region, for subsequent processing and to support patient diagnosis and treatment [15]. To keep up with the increasing population and the present trend of fewer medical experts per capita, medical image analysis is crucial in today's society. Numerous technical advancements that help humanity have been experienced by the healthcare sector, and more development is being made. A precise medical picture analysis will improve diagnostic times and allow for the prediction of treatment plans. This technique speeds up recovery times and saves many lives.

Radiology is a subspecialty of medicine that uses imaging technology for disease diagnosis and treatment. The various forms of diagnostic radiology exams include nuclear medicine, Positron Emission Tomography, Computed Tomography (CT), including CT angiography, Magnetic Resonance Imaging (MRI),[5] and Magnetic Resonance Angiography (MRA), Fluoroscopy, Mammography, and Plain X-Rays. [16]

1.2 Challenges in Medical Image Classifications

The conventional approach to interpreting results involves challenges with data bottlenecks, accuracy, speed, reliability, and precision. [11] The majority of these problems are being solved by machine learning algorithms and approaches. The demand for improved analysis and clarity in medical imaging prediction and classification grows as technology advances. The availability of data, technique validation, and the development of quicker and more accurate algorithms for patient-specific models are the problems that come with using machine learning applications [17].

1.3 Benefits of Medical Image Processing

The ability to investigate the anatomy and the internal workings of a human organ is the primary benefit of medical image processing. It is fascinating to be able to see inside and observe how things function.

- Image processing enables quick intervention and rapid treatment of the illness, and the fatality rate is significantly decreased.
- A further advantage is the in-depth understanding of interior anatomy, which improves the results of diagnosis and treatment.

2. LITERATURE SURVEY

Ammar Mohammed et al. in 2023 [1] proposed a system, The main problem with deep learning is that optimizing the hyper-parameters is a laborious and time-consuming task that requires a great deal of expertise and experience. This is true despite the diversity of deep learning architectures, their capacity to handle complex problems, and their ability to automatically extract features. The majority of these endeavors center on basic ensemble techniques that possess certain constraints. As a result, this review paper offers thorough analyses of the different ensemble learning strategies, particularly as they relate to deep learning. It also provides a detailed explanation of the different aspects or components that affect ensemble methods' effectiveness. Furthermore, it describes and correctly classifies several studies that make use of ensemble learning across a broad spectrum of fields.

Bagher Sistaninejhad et al. in 2023 [2] have presented a detailed overview of newly published deep learning-based methods from 2019 to 2022 in medical imaging. Recent developments in deep learning architectures have the potential to improve medical imaging diagnostic accuracy. However, for deep learning to surpass conventional machine learning models, a substantial amount of data is required. In reality, though, it can be challenging to find such databases with medical image content. This issue can be resolved with the use of pre-trained models and transfer learning. Pretrained models are frequently modified to better suit a given task. Pretrained models are popular because they ensure strong classification accuracy and speed up training. The techniques based on GANs have demonstrated efficacy in reconciling disparities between segmentation masks generated by models and ground truth.

M.A. Ganaiea et al. in 2021 [3] proposed evaluations on the most recent deep ensemble models, providing scholars with a thorough synopsis. The ensemble models can be broadly classified as explicit/implicit, homogeneous/heterogeneous, decision fusion techniques-based deep ensemble models, bagging, boosting, stacking, and negative correlation-based deep ensemble models.

Suganyadevi, S et al. in 2022 [4] proposed a system where medical imaging is a vital tool that connects the gaps between societal and scientific demands and can offer a significant synergy that could lead to advancements in each field. Based on the most recent scientific literature from 120 medical imaging research papers, our survey has shed light on the state of the art, which could be helpful to radiologists everywhere. We not only discovered that the ResNet architecture generally performs best, but we also discussed the main problems, obstacles, and future approaches.

Liu, Y., Ma, J., et al. in 2022 [5] proposed a system where the second special issue of IEEE transactions on Medical Imaging (TMI) is compiled to reflect the state of this quickly developing subject and serves as a sequel to the first special issue on the topic of deep tomographic reconstruction. This editorial summarizes the papers featured, reports our verification of the shared deep learning algorithms, and gives a brief historical overview that highlights the impetus behind the development of network-based, data-driven, and learning-oriented reconstruction methods.

B. Zhang, et al. in 2019 [6] proposed a method in which First, eight deep CNN learners with different architectures are trained and evaluated by 10-fold cross-validation. Each nodule has eight predictions from the eight primary learners. Second, fuse these eight predictions with the strategies of majority voting (VOT), averaging (AVE), or machine learning. Moreover, the correlation coefficients between the predictions of 10 ensemble learners are calculated, and the hierarchical clustering dendrogram is drawn. It is found that the ensemble learners achieve higher prediction accuracy (84.0% vs 81.7%) than single CNN learner.

Sarraf S. and Tofighi G. in 2016 [7] proposed a convolutional neural network to classify Alzheimer's brain from a normal healthy brain. The importance of classifying this kind of medical data is to potentially develop a prediction model or system to recognize the type of disease from normal subjects or to estimate the stage of the disease. Using Convolutional Neural Network (CNN) and the famous architecture LeNet-5, we successfully classified functional MRI data of Alzheimer's subjects from normal controls where the accuracy of test data on trained data reached 96.85%.

Alena-Kathrin Golla, et al. in 2021 [8] proposed a brand-new ratio-based sampling technique to train the DeepVesselNet, V-Net, and U-Net in two and three dimensions. Networks were trained using a cross-entropy loss and the Dice method. Twenty IRCAD participants were used to evaluate performance. The ensemble of top-performing networks was created. Seven distinct weighting algorithms have been investigated.

3. ENSEMBLE LEARNING

The general framework of any ensemble learning system is to use an aggregation function *G* to combine a set *h* of baseline classifiers, c_1, c_2, c_3, … c_h towards predicting a single output. [3]

Given a dataset of size *n* and features of dimension m, $D=\{(x_i, y_i)\}$, $1 \leq i \leq n$, $x_i \in R^m$ the prediction of the output based on this ensemble method is given by Eq. 1.

$$G\ (c_1, c_2, c_3 \ldots c_k)$$

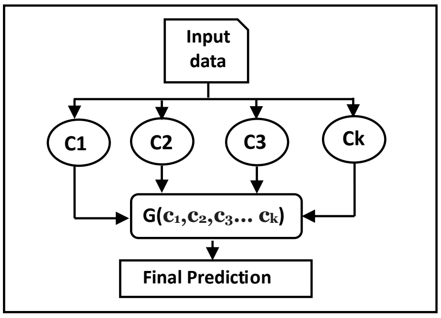


Fig. 1: General Framework for Ensemble Learning

Above Fig.1 depicts the overall abstract framework for group education. Every ensemble consists of a set of baseline classifiers (also known as classifiers ensemble) that have been trained on input data and provide predictions that are then aggregated to get a final prediction. By merging the predictions from several models, ensemble learning is a generic meta-approach to machine learning that aims to improve predictive performance. The way that ensemble techniques choose which baseline classifiers to train varies. Depending on whether the basic classifiers are homogeneous or heterogeneous ensembles, two methodologies are used to create variety among them.

3.1 Types of Ensembles: There are two types of ensembles as shown in the figures below.

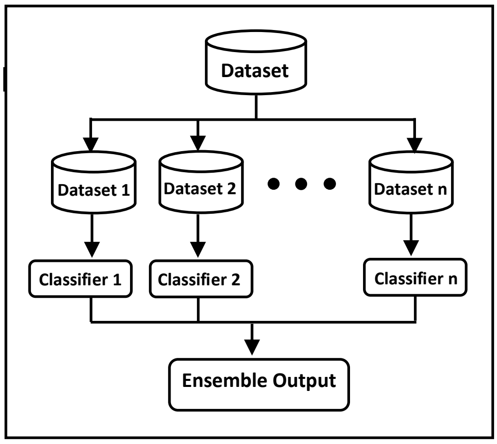


Fig. 2: Homogenous Ensemble

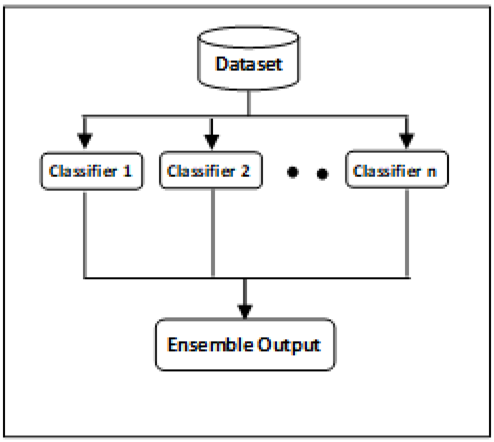


Fig. 3: Heterogenous Ensemble

A homogeneous ensemble is made up of baseline classifiers of the same kind, each with a distinct set of data as its basis. This strategy uses the same feature selection process regardless of the type of training data. While each classifier in a heterogeneous ensemble is based on the same data, there are varying numbers of baseline classifiers in a heterogeneous ensemble. Different feature selection techniques are used in heterogeneous classifiers for the same training set of data. Finally, because homogeneous ensemble approaches are simpler to comprehend and implement, researchers find them more enticing.

3.2 Categorization of Ensemble methods: Generally, three factors that influence an ensemble framework's performance can be used to view and define it. The first one is the reliance on the sequential or parallel-trained baseline models. The second feature is the use of fusion methods, which entail selecting an appropriate procedure for merging baseline classifier outputs via weight voting or meta-learning techniques. The third feature is the heterogeneity whether homogeneous or heterogeneous of the baseline classifiers that are involved. The popular ensemble methods' features are compiled in Table 1. We will go into more depth about those qualities in the sections that follow.

Method	Dependent	Fusion method	Heterogeneity
Bagging	Parallel	Weight Voting	Homogenous
Random Forest	Parallel	Weight Voting	Homogenous
Boosting	Sequential	Weight Voting	Homogenous
AdaBoost	Sequential	Weight Voting	Homogenous
Gradient Boosting	Sequential	Weight Voting	Homogenous
Extreme Gradient Boosting	Sequential	Weight Voting	Homogenous
Stacking	Parallel	Meta-Learning	Heterogeneous

Table 1: Categorization of several ensemble methods.

3.3 Training baseline classifiers: The second significant aspect of the ensemble system is the variety of the baseline classifiers. The sequential ensemble technique and the parallel ensemble technique are the two training methods for individual ensemble members at the heart of every ensemble-based system. Due to data dependency, distinct learners learn in succession using the sequential ensemble technique. [20] As a result, as seen in Fig. 4, the second model successively corrects the faults produced by the first model. Therefore, taking advantage of the reliance among base learners [18] is the primary benefit of sequential techniques. On the other hand, because there is no data dependency in the parallel ensemble technique, [19] base learners are generated simultaneously. Thus, as Fig. 5 illustrates, every piece of data in the base learner is produced independently. The fundamental benefit of this method is taking advantage of the independence among base learners. Because one model's errors are different from those of another independent model, the ensemble model may compute the average of the errors.

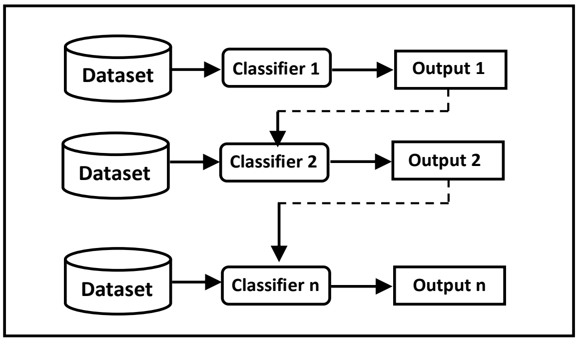


Fig. 4: Parallel Processing

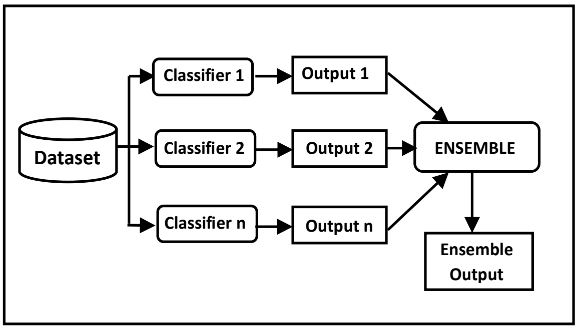


Fig. 5: Sequential Processing

3.4 Fusion method: Combining the outputs of the baseline classifiers into a single output is known as output fusion. The voting method and the meta-learning approach are the two fusion techniques. model is to the prediction.

3.4.1 Voting method: These weights can be used to determine the weighted average for either class 0 or class 1 by multiplying each prediction by the classifiers' weights to create a weighted sum, which is then divided by the sum of the classifiers' weights. [23]

- **Max Voting:** The first and most widely used voting technique is maximum voting, also referred to as hard voting or majority voting. The concept of maximum voting is gathering guesses for every class label and projecting which class label will receive the greatest number of votes, as indicated by the function.[21]

- **Averaging Voting:** Averaging votes is the second voting technique. By extracting predictions from several models and averaging them, a final forecast can be made. This is the concept behind average voting. The arithmetic mean, which is the sum of the forecasts divided by the total number of guesses, is used to get the average prediction. Voting techniques are typically applied to regression or classification issues to enhance predictive performance. Furthermore, bagging and boosting techniques are best integrated with voting techniques.[22] A voting ensemble, comprising three techniques: maximum voting, averaging voting, and weighted average voting, is the first fusion approach.
- **A weighted Average Voting:** A significantly modified form of averaging voting, weighted average voting is the third voting method. The concept of weighted average voting involves assigning varying weights to the baseline learners based on how important each model is to the prediction. These weights can be used to determine the weighted average for either class 0 or class 1 by multiplying each prediction by the classifiers' weights to create a weighted sum, which is then divided by the sum of the classifiers' weights.[23]

3.4.2 Meta-learning method

The second fusion method is called "learning to learn," or meta-learning. It is the process of picking up knowledge from other learners. Learning that is dependent on prior experience with other tasks is referred to as "meta-learning." As a result, by altering some parts of the learning algorithm in response to experiment findings, it is utilized to enhance the performance and outcomes of a learning algorithm.[24]

4. ENSEMBLE METHODS

Stacking, boosting, and bagging are the three most widely used methods for group learning. Every one of these methods provides a different way to increase the accuracy of predictions. Every technique has a distinct application, and its choice is contingent upon several variables.

4.1 Bagging: By modifying a stochastic distribution of the training datasets, bagging aims to produce more diversified predictive models by ensuring that even minor modifications to the training data set result in notable variations in the model predictions. The terms "bagging" and "aggregating" refer to the same thing. When using bootstrapping, the training dataset is replicated by training the ensemble models on Bootstrap.[25] When determining the final forecast, the model's predictions are voted on by the majority of participants to arrive at the outcome. Bagging has the benefit of lowering variance and overcoming the problem of overfitting.

4.2 Boosting: Boosting is a process that goes step by step, with each new model trying to fix the mistakes made by the preceding one. Boosting involves fitting successively multiple weak learners in an extremely adaptable manner. Each model in the sequence is fitted, and observations in the dataset that the preceding models in the sequences handled poorly are given additional weight. The concept of correcting prediction errors is central to boosting ensembles. The models are fitted and added to the ensemble sequentially, with the third model correcting the second, the second model attempting to correct the predictions of the first, and so on. Boost algorithms include three types, namely, Adaptive Boosting (AdaBoost), Stochastic Gradient Boosting (SGB), and Extreme Gradient Boosting (XGB), also known as XGBoost.[25]

4.3 Stacking: The stacking method, also known as Stacked Generalization, is a model ensembling technique used to combine information from multiple predictive models to generate a new model (meta-model). Stacking aims to create a single robust model from multiple heterogeneous strong learners. Stacking differs from bagging and boosting in that [25]

- It combines strong learners.
- It combines heterogeneous models.
- It consists of creating a Metamodel.

A metamodel is a model created using a new dataset.

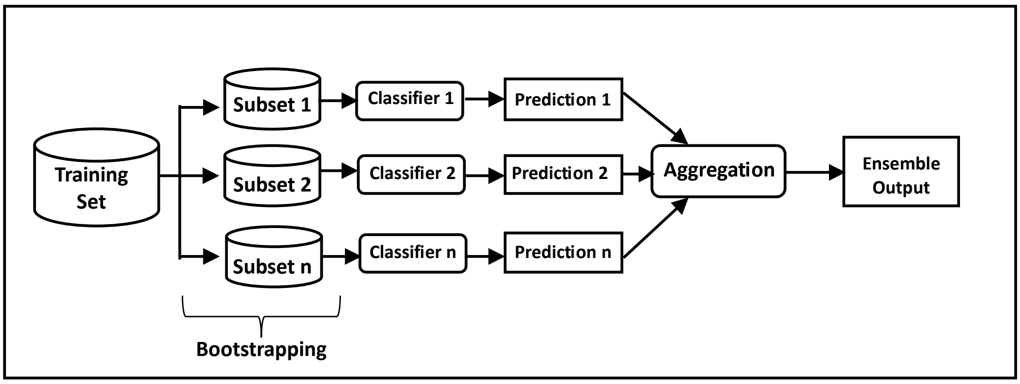


Fig. 6: Bagging

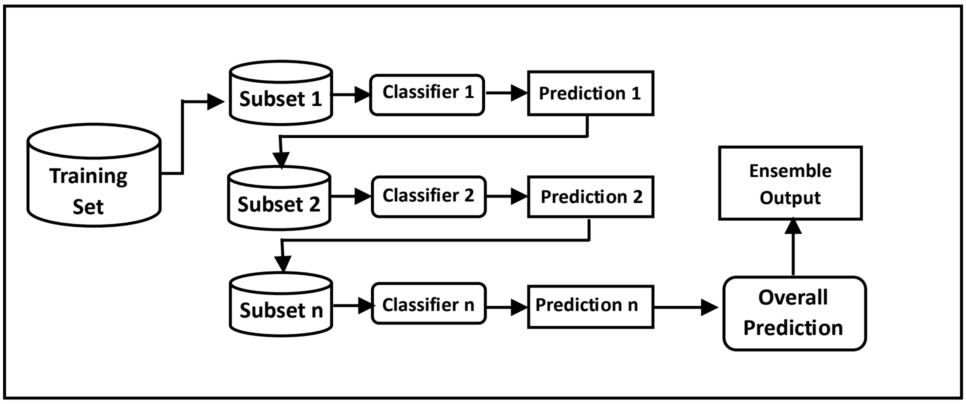


Fig. 7: Boosting.

Individual heterogeneous models are trained using an initial dataset. These models make predictions and form a single new dataset using those predictions. This new data set is used to train the metamodel, which makes the final prediction. The prediction is combined using weighted averaging. Because stacking combines strong learners, it can combine bagged or boosted models. A learner is instructed to combine the individual learners through a generic technique called stacking. In this case, the combiner is referred to as a second-level learner, or meta-learner, while the individual learners are referred to as first-level learners.

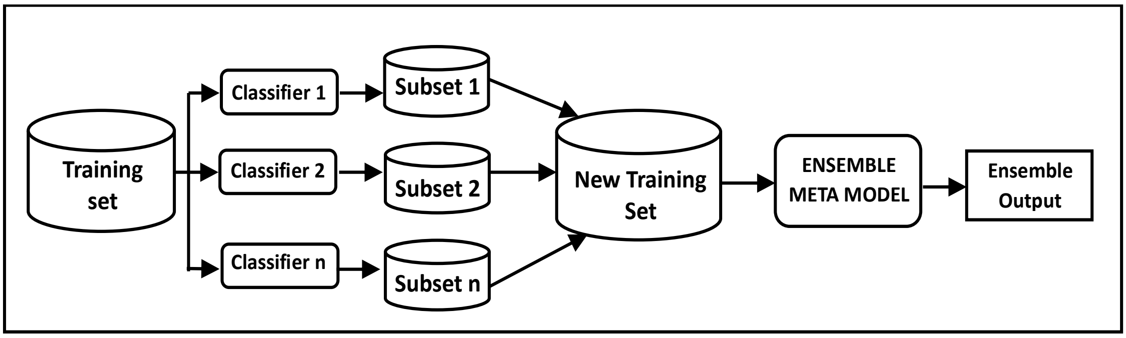


Fig 8: Stacking

4.4 Comparison analysis of the Ensemble Models: All three of these methods can work with either classification or regression problems.

Conversely, yet, this is also true. Using bagging to lessen bias or underfitting is not recommended.

Techniques	Purpose	Base Learner Types	Base Learner Training	Aggregation
Bagging	Reduce Variance	Homogeneous	Parallel	Max voting, Averaging
Boosting	Reduce Bias	Homogeneous	Sequential	Weighted Average
Stacking	Improve Accuracy	Heterogeneous	Meta Model	Weighted Average

Table 2: Comparison analysis of Bagging, Boosting, and Stacking

The vulnerability of boosting to variation or overfitting is one of its drawbacks. Therefore, using boosting to reduce variance is not advised. When it comes to lowering variation, bagging will be more effective than boosting.

This is because bagging does not aid in the reduction of bias and is more biased. The prediction accuracy of stacked models is higher than that of bagging or boosting. However, the drawback of using bagged or boosted models is that they require a lot more time and processing resources.

4.5 Ensemble deep learning: Stochastic training algorithms are utilized in deep neural network models, which are nonlinear learning techniques. This indicates that it is extremely adaptable, capable of approximating any mapping function and learning intricate correlations between variables. This flexibility has the drawback of increasing the variation required by the models. By training many deep models for the problem and pooling their predictions, ensemble deep learning technique opportunities can overcome the large variance of the deep model. Thus, training multiple baseline deep models and combining a few rules to generate predictions is referred to as ensemble deep learning approaches.[26] The goal of ensemble deep learning is to efficiently integrate the main advantages of multiple deep learning models using an ensemble learning system. Because deep neural networks include millions to billions of hyper-parameters and require a lot of time and space to train several base deep learners, ensemble learning using deep learning models is more challenging. Hyper-parameters are thus a problem when using ensemble deep learning methods.[9] When adjusting the data level or the baseline model level, ensemble learning procedures are developed. When manipulating data, fresh training sets are created to train various base learners using sampling or cross-validation data (re-sampling).

Four methods for conducting deep learning based on the ensemble represented by are shown in Fig. 9. (A) Applying numerous fundamental models to the same set of data. (B) Applying several fundamental model architectures with identical data. (C) Applying a wide range of fundamental models to a wide range of data samples. (D) Applying several fundamental model structures to a wide range of data samples.[26] By contrasting these approaches, it can be seen that both strategy A and strategy C work well with both conventional learning methods and deep learning models.

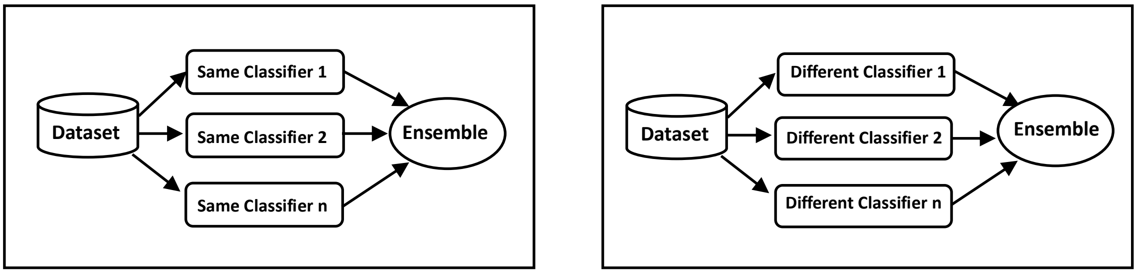
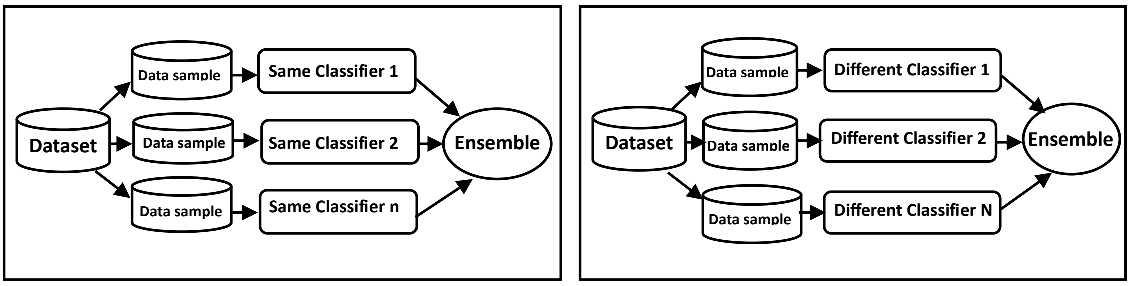


Fig. 9: (A) Same Data and Similar Neural Network **(B): Same Data and different Neural Network**



(C): Data sample and Similar Neural Network **(D): Data sample different Neural Network**

4.6 List of Ensemble learning Machine and Deep learning approaches used in medical image processing. [27]

Table 3: List of ensemble methods used in medical image processing from 2000.

Study	Authors	Baseline Classifiers	Fusion Method	Domain
Zhang et al. (2020)	Zhang, Lee, & Chen	VGG16 (CNN)	Stacking	MRI Brain Tumor Segmentation
Patel and Wu (2021)	Patel and Wu	ResNet50 (CNN) [7]	Boosting	X-ray Image Classification
Kim et al. (2022)	Kim, Garcia, & Wang	InceptionV3 (CNN)	Bagging	Histopathology Image Analysis
Chen and Singh (2023)	Chen and Singh	DenseNet (CNN)	Voting	Ultrasound Image Reconstruction
Rahman et al. (2020)	Rahman, Li, & Park	SVM (Support Vector Machine)	Voting	Mammogram Classification
Garcia and Patel (2021)	Garcia and Patel	Random Forest	Stacking	PET Scan Image Segmentation
Wang et al. (2022)	Wang, Chen, & Gupta	U-Net (CNN)	Bagging	Retinal OCT Image Analysis
Lee and Kim (2023)	Lee and Kim	MobileNetV2 (CNN)	Boosting	Endoscopy Image Classification
Singh et al. (2022)	Singh, Zhang, & Martinez	ResNeXt (CNN)[9]	Stacking	CT Scan Reconstruction
Huang and Wang (2021)	Huang and Wang	Decision Trees	Voting	Brain MRI Lesion Detection
Patel et al. (2022)	Patel, Nguyen, & Kim	AlexNet (CNN)	Bagging	Cardiac MRI Segmentation
Kim and Lee (2023)	Kim and Lee	SVM (Support Vector Machine)	Boosting	Histology Slide Classification
Chen et al. (2022)	Chen, Zhang, & Garcia	InceptionResNetV2 (CNN)	Stacking	Chest X-ray Abnormality Detection [5]

5. PERFORMANCE EVALUATION METRICS

1. **Sensitivity, Specificity, and Accuracy:** These metrics are utilized to measure the ability of a machine learning and deep learning model to correctly classify lung tumors, distinguishing between malignant and benign cases. They assess the model's ability to minimize false negatives (missed malignancies) and false positives (misdiagnosing benign tumors as malignant).

2. **Receiver Operating Characteristic Curve (ROC):** To balance sensitivity and specificity, the ROC curve, and its associated area under the curve (AUC) offers a thorough assessment of the model's performance across a range of thresholds.

3. **Precision, Recall, and F1-Score:** Precision is the percentage of accurately identified malignant cases. Recall, also known as sensitivity, quantifies the model's potential to detect all malignant tumors correctly. The F1-score combines precision and recall, providing a balanced evaluation metric.

4. **Other Metrics:** Additional metrics such as the Dice coefficient, Jaccard index, and Cohen's kappa coefficient are discussed, highlighting their relevance to lung tumor classification tasks.

Table 3. Metrics used in machine learning and deep learning literature.

Metric	Definition	Note
Sensitivity	SE = TP/(TP + FN)	True positive rate (TPR) or recall
Specificity	SP = TN/(TN + FP)	True negative rate (TNR)
Accuracy	ACC = (TP + TN)/(TP + TN + FP + FN)	Total true results
Precision	PPV = TP/(TP + FP)	Positive predicted value (PPV)
F1-Score	F1 = 2TP/(2TP + FP + FN)	Relates the sensitivity and precision measures
ROC	The X-axis represents the false positive rate, and the Y-axis represents the true positive rate, in a curve that illustrates the relationship between sensitivity and specificity.	Receiver Operating Characteristic (ROC) curve

TP = True Positives; TN = True Negatives; FP = False Positives; FN = False Negatives.

6. Advantages of Ensemble Machine Learning Techniques in CT Image Processing:

Some of the notable advantages of using ensemble machine learning models are:

- Improved accuracy and robustness by combining diverse models.
- Reduction in overfitting by leveraging the wisdom of multiple models.
- Effective handling of complex relationships and patterns in CT images.

6.1 Challenges of Ensemble Machine Learning Techniques:

Using ensemble machine learning models has several noteworthy challenges, including:

- Computational resources are required for training and maintaining multiple models.
- Ensemble selection and tuning parameters demand expertise and computational time.
- Potential difficulties in interpreting and explaining results due to increased complexity.

6.2 Future Directions of Ensemble Machine Learning Techniques:

- **Ensemble Diversity:** Exploring and developing more diverse and complementary models for ensemble learning to enhance performance in CT image processing tasks.
- **Interpretable Ensembles:** Integrating interpretability techniques into ensemble models to improve transparency and trust in medical decision-making.
- **Efficiency and Scalability:** Develop computationally efficient ensemble

methods that can scale to handle larger CT datasets and real-time applications.

In summary, ensemble machine learning techniques in CT image processing offer a powerful framework for improving accuracy, robustness, and generalization, making them integral in advancing diagnostic capabilities and clinical decision support systems in medical imaging. Continued research and innovation in ensemble methodologies hold promise for further enhancing the analysis and interpretation of CT images for healthcare applications.

7. Advantages of Ensemble Deep Learning Techniques in CT Image Processing:

Some of the notable advantages of using ensemble Deep learning models are:

- Improved accuracy and robustness by combining diverse models.
- Effective handling of complex relationships embedded in CT images.
- Enhanced generalization and reduced overfitting.
- Better feature representation for classification or segmentation tasks.

7.1 Challenges Ensemble Deep Learning Techniques:

Using ensemble Deep learning models has several noteworthy challenges, including:

- Computational complexity due to training multiple models.
- Ensemble design and selection of diverse architectures require expertise.
- Potential increased memory and resource requirements.

7.2 Future Directions Ensemble Deep Learning Techniques:

The future of ensemble techniques in CT image processing is promising in the Integration of explainable AI methods for better interpretability in medical decision-making. It also focuses on developing lightweight and efficient ensembles for real-time applications and Exploration of novel architectures and fusion strategies for improved performance.

CONCLUSION

In conclusion, ensemble machine learning and deep learning techniques have significantly advanced CT image processing, offering improved accuracy, robustness, and generalization across various medical imaging tasks. As technology advances and new methods emerge, these ensemble approaches will continue to play a crucial role in enhancing diagnostic and clinical decision-support systems in healthcare. Finding the ideal hyperparameters, however, necessitates a tiring approach in the search space, making it a laborious and time-consuming process. Deep ensemble learning has therefore been used in several research projects across a wide range of domains, most centered around straightforward ensemble techniques. This research included a thorough analysis of the different ensemble learning strategies, with a focus on deep learning. It is important to remember that utilizing a collection of deep learning models with basic averaging techniques is not a wise decision and is highly susceptible to biased baseline models. However, adding diversity to an ensemble's deep learning model can make it more resilient against biased baseline models. Training various baseline deep learning architectures over multiple data samples can yield the desired diversity.

Declaration by Authors

Acknowledgement: None

Source of Funding: None

Conflict of Interest: The authors declare no conflict of interest.

REFERENCES

1. Ammar Mohammed, Rania Kora. A comprehensive review on ensemble deep learning: Opportunities and challenges. *Journal of King Saud University* - Computer and Information Sciences. Volume 35. Issue 2, 2023; ISSN 1319-1578.

https://doi.org/10.1016/j.jksuci.2023.01.014. Pages 757-774
2. Bagher Sistaninejhad, Habib Rasi, Parisa Nayeri. A Review Paper about Deep Learning for Medical Image Analysis. Computational and Mathematical Methods in Medicine. vol. 2023. Article ID 7091301. 10 pages. 2023; https://doi.org/10.1155/2023/7091301.
3. M.A. Ganaiea, Minghui Hub, A.K. Malika, M. Tanveera, P.N. Suganthanb, Ensemble deep learning: A review. Elsevier. 2022;
4. Suganyadevi, S., Seethalakshmi, V. & Balasamy. K. A review on deep learning in medical image analysis. Int J Multimed Info Retr 11, 2022; https://doi.org/10.1007/s13735-021-00218-1. pages 19–38
5. Liu, Y., Ma, J., et al. Deep Ensemble Learning for CT Image Denoising. *IEEE Transactions on Medical Imaging.* 2021;
6. Saman Sarraf, Ghassem Tofighi. Classification of Alzheimer's Disease Using fMRI Data and Deep Learning Convolutional Neural Networks. arXiv:1603.08631v1. March 2016;
7. Alena-Kathrin Golla et al. Convolutional Neural Network Ensemble Segmentation with Ratio-Based Sampling for the Arteries and Veins in Abdominal CT scans, IEEE transactions on biomedical engineering. May 2021;
8. Moeskops, P, Viergever, M. A., Mendrik, A. M, & de Vries, L. S. Non-Local Patch-Based CNN for Whole Slide Tissue Image Classification. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2016;
9. Li, C., Liu, J., et al. Ensemble Learning-Based Method for Anomaly Detection in CT Images. IEEE Access. 2021;
10. Liu, S., Gao, Y., et al.Deep Learning-Based Classification of Liver Cancer Histopathology Images Using Only Global Labels. IEEE Transactions on Medical Imaging. 2019;
11. Yasaka, K., Akai, H., & Abe, O. Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study. RadioloOpen 2018;
12. Tang, A., Tam, R., et al. Association of Convolutional Neural Network-Derived MRI Visceral Adipose Tissue with Risk of Coronary Artery Disease. JAMA Network Open. 2019;
13. Zhu, X., et al. Ensemble Learning-Based System for Pneumonia Detection and Classification on Chest CT Images. IEEE Access. 2020;
14. Suetens P. Fundamentals of Medical Imaging, Cambridge University Press, ISBN:9781316671849. 2017; DOI:10.1017/9781316671849.
15. MedlinePlus. Imaging and radiology. 2021; Available from: https://medlineplus.gov/ency/article/007451.html.
16. Jürgen W, Lorenz C. Four challenges in medical image analysis from an industrial perspective. Editorial Med Image Anal. 2016; 33:44-49. DOI: 10.1016/j.media.2016.06.023.
17. R.M. Saeed, S. Rady, T.F. Gharib An ensemble approach for spam detection in Arabic opinion texts J. King Saud Univ. - Comput. Information. Sci., 34 (1). 2022; pp. 1407-1416.
18. J. Tang, Q. Su, B. Su, S. Fong, W. Cao, X. Gong Parallel ensemble learning of convolutional neural networks and local binary patterns for face recognition Comput. Methods Programs Biomed., 197. 2020; p. 105622.
19. N. Sultana, N. Sharma, K.P. Sharma, S. Verma A sequential ensemble model for communicable disease forecasting Curr. Bioinform. 15 (4). 2020; pp. 309-317.
20. R. Delgado, A semi-hard voting combiner scheme to ensemble multi-class probabilistic classifiers Appl. Intell. 52 (4). 2022; pp. 3653-3677.
21. B.M. Hopkinson, A.C. King, D.P. Owen, M. Johnson-Roberson, M.H. Long, S.M. Bhandarkar Automated classification of three-dimensional reconstructions of coral reefs using convolutional neural networks PloS One. 15 (3). 2020; p. e0230671.
22. Mohan, M.A. Ghazanfar, M.A. Azam, A. Karami, K.H. Alyoubi, A.S. Alfakeeh Stock market prediction using machine learning classifiers and social media. news J. Ambient Intell. Humanized Comput. 2020; pp.
23. S. Kuruvayil, S. Palaniswamy Emotion recognition from facial images with simultaneous occlusion, pose, and illumination variations using meta-learning

J. King Saud Univ.-Comput. Information. Sci. 34 (9). 2022; pp. 7271-7282.

24. https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/.
25. H. Saleh, S. Mostafa, A. Alharbi, S. Elsentiment, T. Alkhalifah. Heterogeneous ensemble deep learning model for enhanced Arabic sentiment analysis Sensors. 22 (10).2022; p.3707.
26. I.E. Livieris, A. Kanavos, V. Tampakas, P. Pintelas A weighted voting ensemble self-labeled algorithm for the detection of lung abnormalities from x-rays Algorithms. 12 (3). 2019; p. 64.

How to cite this article: Sheik Imran, Pradeep N. A review on ensemble machine and deep learning techniques used in the classification of computed tomography medical images. *Int J Health Sci Res.* 2024; 14(1):201-213. DOI: *https://doi.org/10.52403/ijhsr.20240124*

******