IRAQI
Academic Scientific Journals

JPS NTU

ISSN: 2789-1089 EISSN: 2789-1097

NTU Journal of Pure Sciences

Available online at: https://journals.ntu.edu.iq/index.php/NTU-JPS/index

Review of Classification Medical Images Using Ensemble Learning

Fatima Maan Abdulatif [1], Aseel Waleed Ali[2]

1. Department of Software , College of Computer Sciences & Mathematics , University of Mosul , Mosul , Iraq,
2. Department of Software , College of Computer Sciences & Mathematics , University of Mosul , Mosul , Iraq.

A B S T R A C T

Machine learning and deep learning play an important role today in the field of classification and prediction of diseases, particularly through computer-assisted medical imaging and modern devices, which aid in early disease detection, lowering the actual risk, and assisting doctors in making a final diagnosis decision. Disease classification using ensemble learning techniques was developed to overcome the classification problem using a convolutional neural network (CNN). Although CNN produces good results, it requires a large dataset, limiting its use in classifying medical images due to its limited data availability and privacy. The research and literature have revealed that traditional machine learning algorithms perform poorly when trained using unbalanced datasets. This research aims to shed light on the use of ensemble learning in medical diagnosis, presents ensemble learning techniques used in classification, reviews previous research and work in which ensemble techniques were used, and proposes ensemble learning as a way to increase the accuracy and efficiency of classification systems by collecting the results of multiple classifiers and outputting the most voted results. These techniques improve the performance of a single model by using multiple models and combining their predictions.

Introduction

The development of technology has led to its use in many areas in our daily lives, as the development of technology has left no without bringing about a major change , a qualitative leap , and a reduction in time, effort, and financial cost . One of its most important uses is in medical diagnosis , which has developed at a rapid pace and it has become possible to communicate with doctor more easily by the internet or to use diagnostic applications directly without medical intervention if it is no possible to visit a doctor, especially in light of the spread of epidemics , and this is what prompted us to highlight in this research one of modern techniques, namely ensemble learning, have recently been widely used in medical diagnosis through images medical [1]. In our lives, we use an approach similar to ensemble by asking a question to elicit several people's opinions different ways to make a final decision, for example when taking the opinion of more than one doctor to classify a disease better than the opinion of one doctor by combining the individual decision of many opinion . This applies to the use of ensemble , which combines the results of several classifiers to produce the final classification, which helps to obtain better classification accuracy. Combining several models will combine the strength of each model and produce the best accuracy , because it is possible that using classification models individually will not be able to give good results. This part contains an explanation ensemble learning and its types.

MACHINE LEARNING AND DEEP LEARNING

Today many applications and smart systems that use artificial intelligence capabilities rely on machine learning , machine learning describes the ability of these systems to extract knowledge and learn from training data , built models and find results automatically[2] .

Deep learning is a type of machine learning based on the use of artificial neural networks , it is called a deep neural network , as the network contains a set of neurons with several parameters and layers between the input and output layers , and the learning takes place in a hierarchical manner and at different levels , which makes deep learning more powerful than traditional machine learning methods and is suitable for training larger and more complex data . Due to the lack of training data for medical images , another type of machine learning that is more capable of dealing with the problem of the lack of data will be discussed ,which is ensemble learning [3].

ENSEMBLE LEARNING

Research on ensemble learning is beginning to spread large since the 1990s , ensemble is considered a model of machine learning , where many classifiers called basic learners, are trained on the same data , unlike traditional machine learning algorithms that train a single model on the data. Ensemble learning is a wonderful thing because it strengthens the ability of basic (weak) learners by collect their predictions, as shown in Figure 1 .
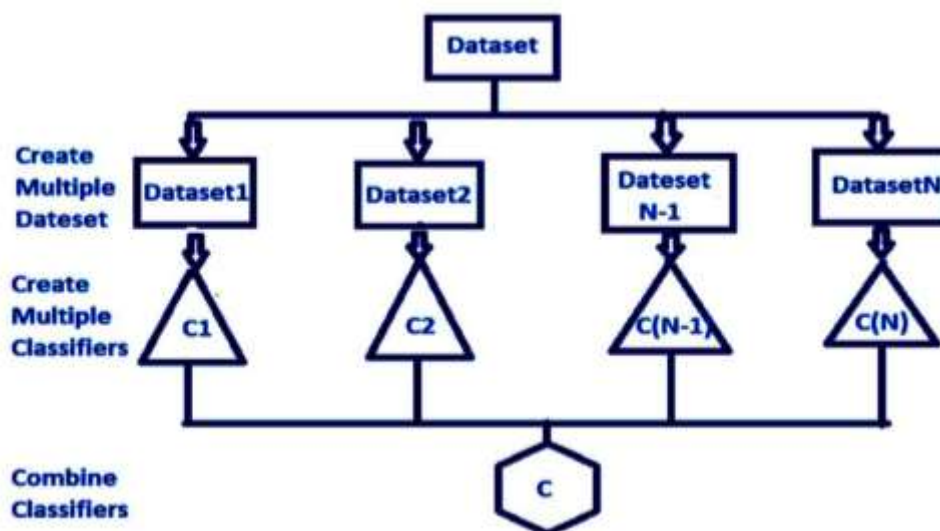

Figure 1: Ensemble learning

ENSEMBLE BUILDING

The ensemble is created by selecting the training data that is used by the basic classifiers, which can be configured in either a parallel structure such as bagging where the classifier is not affected by other classifiers, or in a way sequential or iterative such as boosting , where the classifier has an effect on the next classifier. So that the model works in each iteration to correct errors in the previous classifier , and then the classification models are combined in the most common ways. Such as the majority voting and the weighted average. The ensemble can also be classified into a homogeneous model, which uses classifiers of the same type, and heterogeneous model , which uses several types of classifiers. It is considered stronger than the previous type , as its strength lies in the diversity of the classifiers [4].

It is observed that the total error of the classification model decreases continuously until it can disappear and become an error-free model .In contrast , we have complexity in the way of understanding a group of classification models working on the same data [5]. Ensemble learning applications have become widely used recently, examples of its uses include remote sensing application, change detection , malware detection, intrusion detection, face recognition, emotion recognition and fraud detection.

Types of ensemble learning are generally classified into three types: bagging, boosting and stacking.

BAGGING

It is an abbreviation for bootstrap aggregating, and it is one of the ensemble methods that combines many basic learners trained on samples of training data (subsamples of the dataset). To build bagging , we select a set of training data by bootstrapping, then create classification models based on the number of subsets of the training data, and then training the model, and each model works independently of the other models, where the classification is done in parallel , and then all the results of the classification models are combined by majority vote, average or average weight according to the type of training models[6] . Bagging reduces the variance [1] , which makes the results stable and improved better than individually used classification models that are often exposed to noise. Random forest is a type bagging method where a set of decision trees is selected to build a classification models and samples of training data are randomly selected .

BOOSTING

It is similar to bagging, but the basic idea of boosting is to iterative and sequentially classify the classification models that were chosen to train the data. This technique involves classifying the data using a classifier (the weak learner) and calculating the classification result. Then the incorrectly classified samples are selected in the next model until good classification results are obtained or for a pre-determined number of classifiers used. Using boosting reduces bias and is the most common example of boosting is an adaboost where samples that were misclassified acquire more weight until they are selected and focused on in the next classifier . This type is used in binary classification [7].

3.1.3 STACKING

Stacking is one of the meta-learning method that it achieves excellent results, as the basic idea is to use basic learners (classifiers) to train data samples. Then another basic learner (classifier) called meta-learner to combine the results of the classifiers at the first level , and meta-learners at the second level . Stacking is a heterogeneous ensemble technique where several types of classification algorithm can be used on the same dataset. the diversity in the use of classification algorithms helps improve classification , and stacking is considered the most powerful type because it uses the strengths of each classifier that is used [8].

PREVIOUS WORKS

This section of the research focuses on studies and research in which researchers use ensemble learning technique to classification diseases . Researchers have used several types and methods of ensemble learning techniques depending on the type of disease that was classified and the type of dataset used.

- (Khuriwal & Mishra ,2018) classified breast cancer using the ensemble technique , where they used ANN and logistic algorithm individually and calculating the accuracy of each model . Then they used the voting method to combine their results. The open UCI dataset was used , and the best accuracy for classifying breast cancer diseases was obtained , reaching 98.50%[9].

- (Kassani et al.,2019) proposed an ensemble approach based on the use of deep learning algorithms to classification breast tissue images. The system consists of three networks VGG19, Mobilenet and Denesnet , and four datasets were used ICIAR, BreakHis, PatchCamelyon, and Bioimaging. they obtained a better classification than using classification models individually with an accuracy of 98.13%, 95.00%, 94.64% and 83.10% for the datasets above[10].

- (Abdar & Makarenkov,2019) presented this research to diagnose breast cancer tumors using ensemble learning technique . They used the UCI dataset available in the WBCD repository , and training was done using a boosting artificial neural network (BANN) model with two SVMs . Confidence weighted voting (CWV) was used to collect the final classification results, get 100% accuracy[11].

- (Assiri et al. ,2020) proposed an ensemble classification method for classifying breast tumors . They used the WBCD dataset , and the classification algorithms simple logistic regression , stochastic gradient descent optimization with SVM ,multilayer perceptron network , random decision , random decision tree, SVM with sequential minimal optimization ,naïve bayes and k-nearest neighbor the results of these classification were combined using majority voting . the classification model obtained an accuracy of 99.42%[12].

- An ensemble learning system for early prediction of breast cancer was developed by (Kumar et al. ,2022) using a dataset from the (UCI) Wisconsin Breast Cancer Data Repository machine learning containing 569 samples and each sample containing 32 features. They created a model using stacking ensemble learning on a set of (k-nearest –neighbor , random forest , logistic regression , SVM , decision tree , adaboost m1 , gradient boosting , stochastic gradient boosting , catboost and xgboosat classifier) and the model achieved excellent accuracy of up to 99.45 % compared to other classifiers [13].

- (Naga & Anitha ,2022) proposed the ensemble bagging weighted voting classification (EBWVC) technique for breast cancer classification .The (UCI) Wisconsin diagnosis breast cancer dataset was used, which consists of 699 samples . Several classifiers were used in the bagging technique and the classification results were collected by majority voting . They obtained a classification accuracy of 95% , which is much higher than other techniques SVM , random forest and adaboost [14].

- This study ,by (Shafi et al. ,2021) proposes an ensemble learning method to classify brain tumors and auto-immune diseases . It used 2,399 images taken from multiple sources . The data have been classified using various forms of SVM models, and their results have been combined with a majority voting method for the types used , with an accuracy of 97.,957%[15].

- (Archana & Komarasamy,2023) proposed a method to classification brain malignant tumor using bagging k-nearest neighbor (BKNN) using a dataset containing 3064 images divided into three categories . The proposed model obtained a classification accuracy of 97.7% and 96.9% when using KNN individually [16].

- (Ghalejoogh et al. ,2020) presented a system that combined classifiers, a stacking ensemble method , to improve automated diagnostic systems for skin cancer. Skin diseases were classified into (skin cancer , dysplasia and benign) . They used two datasets, PH2 (mendonca , celebl , mendonca marques , 2015) contains 200 samples and Ganster et al (2001) contains 270 samples . and two hybrid approaches were used Structure Based On Stacking (SBS) and Hierarchical Structure Based On Stacking (HSBS) . HSBS is better than SBS on the same set of data, with 96% accuracy [17].

- Anew approach was proposed by (Pratiwi et al.,2021) to classification skin diseases using ensemble learning . The HAM10000 dataset was used , it contains 10015 samples . They proposed a three network convolutional neural network (cnn) approach Resnet V2 , Inception V3 and Densenet 201 , that combines them to improve performance and obtain better accuracy than using a single CNN , where the accuracy was 97.7390% , specificity 97.73% and sensitivity 90.12%[18].

- (Sharon et al. , 2019) classified rheumatoid arthritis using three ensemble learning techniques random subspace , bagging and boosting . They used two datasets: the first represents lymphocytes in the images , blood cells obtained from microscopic imaging and contains 40 images , and the second is from the UCI bone diseases warehouse contains 310 images, Obtaining the highest accuracy of 97.50% when applying random subspace to the first dataset , 94.849% when applying bagging –RF to the second dataset , and 94.84 when applying adaboost to the second dataset[19].

- Health disorders of rheumatoid arthritis have been classified using ensemble learning by (Sundaramurthy et al. ,2020). They collected data from the sakthi rheumatology center , which contains digital data for patients. Classification was performed using three classification models random subspace , ada boosting and SVM. KNN and random forest to make basic accuracy measurements. The results were compared, but the SVM model was observed to be more accurate when used with KNN and random forest[20].

- (Kandel et al. ,2021) used ensemble learning to classification musculoskeletal fracture , and combined the results of individual CNN networks. They used a MURA datasets that includes seven

types of musculoskeletal and contains 40,005 images with size of 512*512 pixels to 512*97 pixels with the PNG extension , and the image size was changed to 96*96 pixels . More than one stacking technique structure was used and the results of the individual CNN network were combined using the average , the weighted averaged and majority vote for its outputs . it was found that stacking is superior 10 % to used individually CNN[21].

- (An et al. , 2020) used a three- layer approach to classification by deep ensemble learning at each layer to combine data from multiple sources . They used data from the National Alzheimer's Diseases Coordinating Center (NACC UDS) . Which maintains cumulative data consisting of different types of clinical data . Deep ensemble learning was used versus six other ensemble methods stacking , logiboost , bagging , vote , random forest and adaboost M1 . Results showed that deep ensemble learning outperformed the other methods, as the increase in accuracy was 4% and recall was 3% , reaching accuracy 86%[22].

- (Khan et al. , 2022) have proposed a more effective ensemble learning model for classifying Alzheimer's diseases . A dataset from Alzheimer's diseases neuroimaging initiative (ADNI) was used, which contains 2125 samples . The proposed system consists of a set of extreme gradient boosting , decision tree and SVM-polynomial kernel (XGB + DT + SVM) . The accuracy of this model reached 95.75%[23].

- (AlJame et al. ,2020) diagnosed Covid-19 using stacking ensemble learning .They used a dataset from Albert Einstein Hospital , Which contains 5644 images. They were trained using three types of classifiers :logistic regression , extra trees, and random forest . Then ,results of their classification were collected using extreme gradient boosting (XGBOOST), and the model's accuracy reached 99.38%[24].

- (Siswantining & Parlindungan , 2021) also used the stacking ensemble method to classification Covid-19 diseases . They used X-Ray data for the diseases , which consists of 170 images and applying the stacking technique to the neural network , SVM and CNN , and excellent classification accuracy of up to 95% was obtained compared to using the models individually , as they the accuracy is 92.5% when using a neural network , 87.5% when using CNN , and 87.5% when using SVM [25].

- Also , a solution was proposed by (Das et al. ,2021) to detect the Covid-19 diseases using ensemble learning with modern CNN models :InceptionV3, Denset201 and ResnetV2, they used these models individually and then combined them using the ensemble approach, where Weighted average ensemble techniques were used for the final output and X-ray images were used from various open sources of infected people contain 538 non-infected images and 468 infected images. The proposed system obtained a classification accuracy of 91.62% which is higher compared to using CNN individually[26].

Table 1: Summary of previous works

Researcher	Year	Diseases	Algorithm	Dataset	Size of dataset	Accuracy
Khuriwal & Mishra [9]	2018	breast cancer	ANN and logistic algorithm used the voting method to combine their results	UCI	-	98.50%.
Kassani et al [10]	2019	breast tissue	VGG19, Mobilenet and Denesnet	ICIAR, BreakHis, PatchCamelyon, and Bioimaging	-	98.13%, 95.00%, 94.64% 83.10%
Abdar & Makarenkov [11]	2019	breast cancer tumors	boosting artificial neural network (BANN) model with two SVM.Confidence weighted voting (CWV) was used to collect the final classification results	UCI	-	100%
Assiri et al [12]	2020	breast tumors	simple logistic regression , stochastic gradient descent optimization with SVM ,multilayer perceptron network, random decision , random decision tree, SVM with sequential minimal optimization ,naïve bayes and k-nearest neighbor the	WBCD	-	99.42% .

Researcher	Year	Diseases	Algorithm	Dataset	Size of dataset	Accuracy
Kumar et al [13]	2022	breast cancer	results of these classification were combined using majority voting stacking ensemblelearning on a set of (k-nearest – neighbor , random forest , logistic regression , SVM , decision tree , adaboost m1 , gradient boosting , stochastic gradient boosting , catboost and xgboosat classifier)	UCI	569 samples	99.45 %
Naga & Anitha [14]	2022	breast cancer	ensemble bagging weighted voting classification (EBWVC) technique	UCI	699 samples	95%
Shafi et al [15]	2021	brain tumors	using various forms of SVM models, and their results have been combined with a majority voting method	multiple sources	2,399 images	97.,957%.
Archana & Komarasamy [16]	2023	brain malignant tumor	bagging k-nearest neighbor (BKNN)	-	3064 images	97.7%
Ghalejoogh et al [17]	2020	skin cancer	two hybrid approaches were used Structure Based On Stacking (SBS) and Hierarchical Structure Based On Stacking (HSBS)	two datasets, PH2 (mendonca , celebl , mendonca marques , 2015) and Ganster et al (2001)	200 sample 270 sample	HSBS is better than SBS on the same set of data, with 96% accuracy
Pratiwi et al [18]	2021	skin diseases	a three network (cnn) approach Resnet V2 , Inception V3 and Densenet 201 , that combines them to improve performance	The HAM10000 dataset	10015 sample	97.7390%
Sharon et al [19]	2019	rheumatoid arthritis	three ensemble learning techniques random subspace , bagging and boosting	two datasets: the first represents lymphocytes in the images , blood cells obtained from microscopic imaging , and the second is from the UCI bone diseases,	40 images 310 images	97.50% when applying random subspace to the first dataset , 94.849% when applying bagging –RF to the second dataset , and 94.84 when applying adaboost to the second dataset
Sundaramurthy et al [20]	2020	rheumatoid arthritis	using three classification models random subspace , ada boosting and SVM. KNN and random forest to make basic accuracy measurements	sakthi rheumatology center	-	The results were compared, but the SVM model was observed to be more accurate when used with KNN and random forest.

Researcher	Year	Diseases	Algorithm	Dataset	Size of dataset	Accuracy
Kandel et al [21]	2021	musculoskeletal fracture	More than one stacking technique structure was used, combined the results of individual CNN networks	MURA datasets	40,005 images	stacking is superior 10 % to used individually CNN
An et al [22]	2020	Alzheimer's Diseases	. Deep ensemble learning was used versus six other ensemble methods stacking , logiboost , bagging , vote , random forest and adaboost M1	the National Alzheimer's Diseases Coordinating Center (NACC UDS)	-	86.8%
Khan et al [23]	2022	Alzheimer's diseases	set of extreme gradient boosting , decision tree and SVM-polynomial kernel (XGB + DT + SVM)	Alzheimer's diseases neuroimaging initiative (ADNI	2125 samples	95.75%.
AlJame et al [24]	2020	Covid-19 diseases	stacking ensemble learning by using three types of classifiers :logistic regression , extra trees, and random forest . Then results of their classification were collected using extreme gradient boosting	dataset from Albert Einstein Hospital	5644 images	99.38%
Siswantining & Parlindungan [25]	2021	Covid-19 diseases	the stacking technique to the neural network , SVM and CNN.	X-Ray data for the diseases	170 images	95%
Das et al [26]	2021	Covid-19 diseases	InceptionV3, Denset201 and ResnetV2, they used these models individually and then combined them using the ensemble approach, where Weighted average ensemble techniques were used for the final output	X-ray images were used from various open sources	538 non-infected images and 468 infected images	91.62%

In brief , the table above compares several ensemble learning methods that have been used in medical image classification research on a different datasets and to classification different diseases .As noted ,all ensemble classifiers produce good for results, as shown above from the accuracy each model , and of the results reached an accuracy of 100% when using boosting artificial neural network (BANN) model with two SVM. Confidence weighted voting (CWV) was used to collect the final classification results to classify breast cancer tumors.

DISCUSSION

This research proposes an ensemble learning method to classify medical images for diagnosing diseases using computers and smartphones and highlights the types of ensemble . It focuses on comparing the use of diverse ensemble methods for classification .

CONCLUSION

The use of traditional machine learning algorithms to medical images classification tends to perform poorly, especially when the dataset is unbalanced . To overcome this problem , ensemble learning algorithms were used, it helps with diversity in choosing training models .When a models is used individually it may be inappropriate for the training data or the classification results are unsatisfactory . Using different samples in each model, which leads to reducing variance and solves the problem of data size, whether small data or large data by taking a certain number of samples that match the classification model, which leads to obtaining better accuracy than using classifiers individually . This research was created to illustrate the effect of evolution technology on medical diagnosis and demonstrating the possibility of developing classification algorithms.

Ensemble learning, its types ,and some research reviews and studies related to its use in medical classification have been studied.

ACKNOWLEDGMENTS

COMPETING INTERESTS

We confirm that we have no conflicts of interest with regard to this research.

References

[1] Akinbo, R. S., & Daramola, O. A. (2021). Ensemble machine learning algorithms for prediction and classification of medical images. Machine Learning-Algorithms, Models and Applications.

[2] Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. Electronic Markets, 31(3), 685-695.

[3] Sharma, N., Sharma, R., & Jindal, N. (2021). Machine learning and deep learning applications-a vision. Global Transitions Proceedings, 2(1), 24-28.

[4] Zhou, Z. H., & Zhou, Z. H. (2021). Ensemble learning (pp. 181-210). Springer Singapore.

[5] Dong, X., Yu, Z., Cao, W., Shi, Y., & Ma, Q. (2020). A survey on ensemble learning. Frontiers of Computer Science, 14, 241-258.

[6] Zhang, Y., Liu, J., & Shen, W. (2022). A review of ensemble learning algorithms used in remote sensing applications. Applied Sciences, 12(17), 8654.

[7] Mienye, I. D., & Sun, Y. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. IEEE Access, 10, 99129-99149.

[8] Liang, M., Chang, T., An, B., Duan, X., Du, L., Wang, X., ... & Gao, H. (2021). A stacking ensemble learning framework for genomic prediction. Frontiers in genetics, 12, 600040.

[9] Khuriwal, N., & Mishra, N. (2018, March). Breast cancer diagnosis using adaptive voting ensemble machine learning algorithm. In 2018 IEEMA engineer infinite conference (eTechNxT) (pp. 1-5). IEEE.

[10] Kassani, S. H., Kassani, P. H., Wesolowski, M. J., Schneider, K. A., & Deters, R. (2019). Classification of histopathological biopsy images using ensemble of deep learning networks. arXiv preprint arXiv:1909.11870.

[11] (Abdar, & Makarenkov, 2019). CWV-BANN-SVM ensemble learning classifier for an accurate diagnosis of breast cancer. Measurement, 146, 557-570.

[12] Assiri, A. S., Nazir, S., & Velastin, S. A. (2020). Breast tumor classification using an ensemble machine learning method. Journal of Imaging, 6(6), 39.

[13] Kumar, M., Singhal, S., Shekhar, S., Sharma, B., & Srivastava, G. (2022). Optimized stacking ensemble learning model for breast cancer detection and classification using machine learning. Sustainability, 14(21), 13998.

[14] Naga, D. P., & Anitha, R. (2022). A novel ensemble bagging classification method for breast cancer classification using machine learning techniques. Traitement du Signal, 39(1), 229.

[15] Shafi, A. S. M., Rahman, M. B., Anwar, T., Halder, R. S., & Kays, H. E. (2021). Classification of brain tumors and auto-immune disease using ensemble learning. Informatics in Medicine Unlocked, 24, 100608.

[16] Archana, K. V., & Komarasamy, G. (2023). A novel deep learning-based brain tumor detection using the Bagging ensemble with K-nearest neighbor. Journal of Intelligent Systems, 32(1), 20220206.

[17] Ghalejoogh, G. S., Kordy, H. M., & Ebrahimi, F. (2020). A hierarchical structure based on stacking approach for skin lesion classification. Expert Systems with Applications, 145, 113127.

[18] Pratiwi, R. A., Nurmaini, S., Rini, D. P., Rachmatullah, M. N., & Darmawahyuni, A. (2021). Deep ensemble learning for skin lesions classification with convolutional neural network. IAES International Journal of Artificial Intelligence, 10(3), 563.

[19] Sharon, H., Elamvazuthi, I., Lu, C. K., Parasuraman, S., & Natarajan, E. (2019). Development of rheumatoid arthritis classification from electronic image sensor using ensemble method. Sensors, 20(1), 167.

[20] Sundaramurthy, S., Saravanabhavan, C., & Kshirsagar, P. (2020, November). Prediction and classification of rheumatoid arthritis using ensemble machine learning approaches. In 2020 International Conference on Decision Aid Sciences and Application (DASA) (pp. 17-21). IEEE.

[21] Kandel, I., Castelli, M., & Popovič, A. (2021). Comparing stacking ensemble techniques to improve musculoskeletal fracture image classification. Journal of Imaging, 7(6), 100.

[22] An, N., Ding, H., Yang, J., Au, R., & Ang, T. F. (2020). Deep ensemble learning for Alzheimer's disease classification. Journal of biomedical informatics, 105, 103411.

[23] Khan, Y. F., Kaushik, B., Chowdhary, C. L., & Srivastava, G. (2022). Ensemble model for diagnostic classification of Alzheimer's disease based on brain anatomical magnetic resonance imaging. Diagnostics, 12(12), 3193.

[24] AlJame, M., Ahmad, I., Imtiaz, A., & Mohammed, A. (2020). Ensemble learning model for diagnosing COVID-19 from routine blood tests. Informatics in Medicine Unlocked, 21, 100449.

[25] Siswantining, T., & Parlindungan, R. (2021, January). Covid-19 classification using X-ray imaging with ensemble learning. In Journal of Physics: Conference Series (Vol. 1722, p. 012072). IOP Publishing.

[26] Das, A. K., Ghosh, S., Thunder, S., Dutta, R., Agarwal, S., & Chakrabarti, A. (2021). Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network. Pattern Analysis and Applications, 24, 1111-1124.