Project Title
Empowering Rural Healthcare Through AI Solutions

NU302 R&D Project
Sem II, AY2024-25

Under the supervision of
Dr. Sujata

By

Vishwas Kaushik
BT22GDS262

Yuvraj Chawla
BT22GCS375

NIIT UNIVERSITY
THE UNIVERSITY OF THE FUTURE

NIIT UNIVERSITY
THE UNIVERSITY OF THE FUTURE

DECLARATION

I/we hereby declare that this research report is my own original and unaided work, and I/we have given full acknowledgement to all the cited and referred sources used. This research report has not been submitted previously for any degree or examination. I/we understand that any instance of plagiarism or academic dishonesty will result in disciplinary action as outlined by the institution's policies.

Name of Student: Vishwas Kaushik

Enrollment No.: BT22GDS262

Signature: VKaushik

Name of Student: Yuvraj Chawla

Enrollment No.: BT22GCS375

Signature: Yuvraj Chawla

Endorsement by Faculty Mentor(s): -Signature:

(Dr. Sujata)

Endorsement by Area Director: Signature:

(Prof. Debashis Sengupta)

II. Abstract

The objective if the study is to present a comparative study between multiple models that can be used to predict cancer. The multiple model that was used in the study was, ResNet18, DenseNet121, Ensemble Stacking, Multi Modal and our experimental model that is Hierarchical Ensembling in Multi Modal with Late Fusion model. Further the strengths and limitations of our experimental model are discussed. The key finding of our study was that stacking Ensembling improves the cross-validation accuracy and removes overfitting to an extent. We also found that our experimental model was able to reduce overfitting much better than stacking Ensembling, however due to limitations of dataset and hardware, it was not able to treat high bias. The overall conclusion of the study is that the Hierarchical Ensembling in Multi Modal with Late Fusion model addresses the occurrence of overfitting, but it is only recommended to use the technique with a proper labeled and annotated dataset with both medical images and clinical lab report and a complex model which will be able to detect the underlying patters in the data.

Keywords: ResNet, DenseNet, Fusion model, Multi Modal

III. Table of Contents

IV. Introduction

- **Background of the Study:**

As per the study done in 2022 by International Agency for Research on Cancer (IARC) by World Health Organization (WHO), in India out of the 1.4 billion population around 1.46 million new cases of cancer were diagnosed. The number of deaths were recorded at around 900K in that year. Even if the new cases are not directly related with death occurred in 2022, we can still see that 900K is a huge number which highlights a significant burden on the country. Narrowing down to rural areas, it is quite significant that access to high quality healthcare resources is a persistent challenge. The shortage of high-quality medical infrastructure, trained medical professionals and late diagnosis of the disease has created a huge impact. In recent years, a lot of work has been done to use AI in the diagnosis, so as to improve the quality of diagnosis and provide a helping hand to medical professionals in their decision making. A lot of research also includes the use of ensemble learning and deep learning to predict cancer at early stages using both medical images and clinical lab reports. However, there is still a considerable amount of gap in researching and applying these methods in the medical sector. Some of the reasons are the scarcity of data to train the model, heavy computation is required to preprocess and analyses the data and less belief of people in AI. Our research aims to study and apply the multi modal technique which uses fusion to combine the results received from ensemble deep learning on medical images and natural language processing on the lab reports. In the process of experimenting and studying we have also worked on a comparative study which tests the foundation and credibility of our multi modal with fusion and compares it with the existing methods like image-based classification done using Resnet18 and DenseNet121.

- **Statement of the Problem:**

Despite all the current advancement in technology, their remain a gap in implementing these studies at the working level. We aim to research and implement the multi modal technique which has been in research for some years now, but is yet to be approved by Food and Drug Administration (FDA). We have also tried to implement the use of generative AI to create synthetic data so that problem of data scarcity and concerns related to privacy can be minimized.

- **Objectives of the Study:**

The main objective of this research is to develop a user-friendly web application, Ingenious Hospital Management System, that will be allow users to efficiently upload and manage their personal medical data. Using International Classification of Disease codes our model system integrated with the web application will try to predict cancer. Specific goals are:

1. Develop a Hierarchical Ensembling in Multi Modal with Late Fusion model which will make predictions.
2. The multi modal uses ensemble deep learning for classifying medical images, extract data from clinical lab reports using natural language processing and classify them.
3. Integrate the multi modal with the web application.
4. Ontology is created using ICD codes to make predictions.
5. A comparative study is being done to know how effective the overall methodology is compared to existing well researched methods.

- **Hypotheses:**

1. The results of medical image classification and clinical lab reports can be combined using various fusion techniques.
2. The multi modal will attain a certain accuracy after using the results generated by the fusion model.
3. The dataset used for this methodology will contain medical images and clinical lab reports connected with each patient.

- **Significance of the Study:**

The study holds significant power to shape the future of cancer diagnosis. It could help medical professionals to a great extent in decision making, early detection of cancer and if done right may be prevent any loss. With tweaks in image classification methods this methodology can be used to diagnose and predict other diseases as well. With the rising AI technology, the integration of it with medical sector would result in overall improved accuracies of medical predictions.

- **Scope and Limitations of the Study:**

  The scope of this study is not limited to one disease and can be used if tweaked in right way for more diseases. However, some major limitations are the scarcity of real-world data to train these models, availability of high computational units and resistance of people towards the use of Artificial Intelligence. This research was not able to work on any primary data and as a result used available datasets over the internet.

V. Literature Review

The study started with reviewing the paper (Munir et al., 2019) which talks about how various deep learning techniques can be applied for cancer diagnosis. It also talks about some of the traditional cancer classification techniques like seven-point checklist, ABCD rule and pattern analysis. The deep learning algorithms that were used like Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), Deep Autoencoders (DAEs, CAEs, SAES), Restricted Boltzmann Machines (RBM), Recurrent Neural Networks (RNNs) & LSTMs an many more on different types of cancers like breast, lung, brain and skin cancer. The paper listed benefits like high accuracy of deep learning models as they outperformed traditional methods. The scalability of these methods is also a strength as they easily adapt to different types of cancers and huge dataset are also not an issue. The limitations discussed in the paper are some of the major issues that even no occurs while following these techniques like limitation of datasets, as publicly cancer images are very hard to find because of privacy. Since these methods processes high number of images, they require very high computational hardware’s.

This paper systematically compares the diagnostic performance of various deep learning models in detecting various diseases using medical images. In this paper (Liu et al., 2019) has analyzed accuracy of various deep learning models in classifying diseases. The binary diagnostic accuracy data was extracted to derive the outcomes of interest. 69 studies were used to make the dataset and 25 were used to cross validate the results. The result also showed that while deep learning showed higher accuracy, but in some cases, it lacked external validation. The paper showed the benefits of using deep learning was it gives high accuracy and can work across multiple diseases. Once successful it can be actually used by healthcare professionals which will eventually reduce their workload. But some serious limitation of this technique is that sometimes it lacks external validation as most studies were only conducted on internal data only.

Subsequently, (Zeng et al., 2020) in this paper discussed, about how to create a Natural Language Processing (NLP) model that will automatically classify the text to find the eligibility criteria of individuals for clinical trials. For the purpose of training the author and co has used deep learning methods like Bidirectional Encoder Representations from Transformers (BERT), Robustly Optimized BERT Pretraining Approach (RoBERTa) and XLNet. The results of all these techniques were combined and later used as a feature for Light Gradient Boosting Machine (LightGBM). The working described is that the text will first preprocessed and the characters will be converted to numerical values and a Classification level token (CLS) is added. This token is added in front of the sentence which are to be used for classification. Then this preprocessed text is fed into the pretrained models BERT, XLNet and RoBERTa. Once the result is obtained from these individual models it is then used as a feature for LightGBM and an output is generated. The paper discussed benefits like how the performance of these four models outperformed the baseline model by an average of 2.35%. Using the focal loss and SoftMax function the problem of data imbalance was solved an a more accurate result was generated. However, the major limitation of this implementation is that it requires a lot of data. The performance of this method is directly dependent on the amount of input data. As the input data was decreased a significance decrease in F1 score was seen. This concludes that until a huge database is available, this method could not perform better than any of the baseline method.

Following this (Rajaraman et al., 2020) in their paper demonstrated the use of iteratively pruned deep learning models for detection on covid-19 infection using chest x-rays. The paper also discusses about how combining ensemble learning, model specific knowledge and iterative pruning can actually increase the accuracy significantly. The iterative pruning works by discarding the neurons or nodes which provides the least activation function. This basically reduces the complexity and overall computational cost of the model. So, it becomes a good to use method with low resource hardware. The working described in the paper is that the pre trained CNN models are custom trained on the chest x-rays. With the help of iterative pruning the computational work was reduced and efficiency of model was improved. But as other models this one also has a huge limitation of dataset. If the dataset is not huge then the iterative pruning could not work efficiently as there would not be much neurons with huge difference in their activation function, so as a result the pruning would not help much.

This paper analyses the performance of various ensemble learning techniques like Baggin, stacking and Augmenting. (Muller et al., 2022) in this paper proposes a reproducible medical image classification pipeline which is teste on these multiple ensemble learning techniques. The paper also talks about how ensemble learning can be use in the field of deep learning to improve overall performance. The comparison is being done in between baseline model where no ensemble techniques are implemented and individual performance re tested. Then this baseline performance serves as a reference for the further models to check how better they perform. Then the paper has use techniques like Stacking, Augmenting and Baggin. In Stacking the predictions of multiple deep learning models are combined which are trained on different model. Augmenting applies image augmentation and then averages the result to get final outcome. Under Bagging multiple models are trained on random subset of same dataset and then using voting their results are combined. One other technique which is not use in this paper is Boosting. In the paper it is clearly mentioned that for images Boosting is not preferable because the training time is too high. The paper highlights that by using deep ensemble learning the overall performance can be actually increased. However, there is high risk of overfitting in each of these cases and these techniques are also highly dependent on the number of features available o train them.

Further in our review we came across paper of (Kline et al., 2022) which discuss about the current horizon of multi model machine learning in healthcare by analysis 128 research papers from 2011 to 2021. The paper also identifies the most common fusion techniques, healthcare applications, challenges and further use of multi model in machine learning. Three major fusion techniques have been talked about in this paper. Early fusion, Intermediate fusion and Late fusion. In Early fusion multiple data sources are converted into one single feature space. Intermediate fusion merges the output generated before the final result. Under Lat fusion separate models are used to generate results and once all the models generate their results, all of them are combined to get final result. Multiple machine learning models are use in this method. Support Vector Machine, Decision Trees, Random Forest, Logistic Regression CNNs etc. After the data was collected which include medical images like CT scans and chest X-rays, clinical notes, they were preprocessed then all the fusion techniques were use separately. Then machine learning and deep learning models were applied on these results to make predictions. To test for accuracy, F1 score and other performance parameters a cross-validation set was also used. The paper list benefits of using MML that improves the accuracy by 6.4% on an average over all the models used. With the use of MML more robust predictions were made and a better diagnosis was made. But there are some serious limitations as well. The multimodal is not FDA approved so it cannot be used for clinical implementation. Another issue is computation complexity. Using MML an fusion a high processing power is required and a large dataset is required for this technique to function properly.

Since a lot of papers discussed about Ensemble Learning, so we then reviewed the use of ensemble learning for disease prediction. In this paper (Mahajan et al., 2023) discussed multiple ensembles learning approaches that can be used rather than using the existed old-style techniques to predict diseases. Four major techniques of ensemble leaning are discussed in this paper. Baggin, Boosting, Stacking and Voting. All these techniques are performed on diabetes, skin disease, kidney disease and heart conditions. Data pre-processing, Synthetic Minority Oversampling Technique (SMOTE), Principal Component Analysis (PCA) for dimensionality reduction were used. Overall, it was found that stacking showed highest accuracy out of all the models, while boosting performed better for liver and diabetes prediction. Bagging out of these performed worst while voting was second best overall. Overall, it was noted that ensemble learning provided high accuracy, better generalization and is a better technique for disease prediction. The paper saw some limitations of ensemble learning like it was computationally expensive as compared to traditional techniques. Another issue is of Data Leakage. If data not handled properly, then it could lead to overfitting of the model.

Following this, the study by (Chakravarty et al., 2023) talks about the various models that are used for a lot of image classification. This paper also provides insights into various possible enhancements that could be used with existing models to improve upon them and generate better results as compared to the existing ones. It describes how Ensemble Learning and CNN can be used together. Using Ensemble techniques like: Stacking, Bagging, Voting, LightGBM, GBDT and BERT on clinical text and CNN and transfer learning for image classification results in better accuracy and overall performance. By using various Ensemble Learning techniques and CNN the accuracy improved a lot and with the help of transfer learning the overall time and computational cost reduced. However, in the current world due to inadequate training dataset the models suffer with accuracy loss and hence requires a better dataset. Also, most of these models are not working in combination and actually working individually. So, the overall performance after combing them cannot be judged.

The current technology trend is towards Artificial Intelligence and Generative Artificial Intelligence. Following the trend and the limitations of our research we then reviewed Generative AI in healthcare. (Reddy, 2024) talks about the use of Generative AI in healthcare and how it could solve the eternal problem of scarcity of dataset. This paper also discusses the frameworks like technology acceptance model (TAM) and the Non-Adoption, Spread and Sustainability (NASSS) model that are useful and promotes the use and responsible integration of Generative AI in healthcare sector. The overall working mentioned in this paper is, how with the help of Generative AI we can generate new data using just the training data. There are two models that the paper discusses on which can help do such things. Generative Adversarial Network and Large language Model. Some benefits of using Gen AI in healthcare sectors are that since it generates synthetic data so there cannot be any misuse of patient's medical data which in turn protects the privacy of patients. Also once trained properly these AI models are capable of helping medical staff in decision making. The overall cost, time and diagnosis can be improved with the integration of Generative AI. But this paper briefly mentioned about some limitations as well. As mentioned by the author, that still the impact and use of Gen AI is clearly not understood by the people. Some talks about its ethical practice while others worry about its precision and acceptance. One major limitation of Gen AI in synthetic data generation is that, if the original training data contains some noise or biasness, then the Gen AI models will follow the same patters to generate the synthetic data. As a result, all the newly generated data will also have those biasness and noise. This would impact the overall performance of any model as the whole dataset follows a patter and as a result the models would not perform well.

During our review we found the idea of using multi modal machine learning quite interesting. So, we then reviewed approached of multi modal machine learning in medical fields. In this paper (Krones et al., 2025) discussed and reviewed the various MML techniques that can be used in healthcare sectors. It also discusses the use of fusion techniques to integrate various different results generated from image analysis, text processing and tabular data. The paper also highlights some challenges and future potential of MML in healthcare sector. Apart from Early, Intermediate and Late fusion, this paper also reviews Mixed Fusion. In this fusion we combine multiple fusion techniques to generate a better result. The paper highlights that after collecting and preprocessing the data various deep learning models were used to train the preprocessed data. Once done all the results combined using different fusion techniques. A better clinical accuracy, reduced data biasness and increased robustness was observed. Overall, this method could help in better clinical decision making and support automation in healthcare sector. However, to create such extensive model a lot of data is required in the whole process and multiple training needs to be performed which would result in high computational cost.

In a related study, (Imrie et al., 2025) explored the use of automation and multimodal ensemble learning in healthcare sector. The study aims to structure and combine clinical data and medical images to get a better diagnosis. Then it discusses the use of automated machine learning to optimize these models by hyperparameter training to improve the overall performance and output. The automated machine learning is carried out by using a framework called as AutoPrognosis-M. The overall process till fusion and optimization of models is same, however in this study the use of Automated ML is demonstrated where the combination of best model is detected by AutoML. The overall benefit of this technique is that it removes human bias and manual work by automating the process of selecting the best combination of models. This model is also a generalized approach and hence can be applied on many diseases. But, with great automation comes great computation. This is one of the major limitations of these techniques. As if the computation required for training the models and applying fusion and multi modal, the AutoML take things to next level. Also, since the technique is generalized, hence it required a huge dataset to train and perform better.

VI. Methodology

- **Research Design:**

  This study follows an experimental, quantitative approach. As ML techniques are used in this research to predict the stage of Cancer in the Lung Cancer patients and a comparative study is being conducted to see if the techniques used are useful. The research is put under experimental category because it discusses a new methodology. The paper unfolds the effectiveness of this approach as it goes forward by comparing with currently deployed or well researched methods on the basis of accuracy, precision, F1 score and recall score, and the design uses image data set from secondary authentic sources to get more quantity of data to treat high variance and bias.

- **Participants/Subjects:**

  The model is trained with the data of patients of Hospital of Harbin Medical University in Harbin, Heilongjiang Province, China. These patients are anonymous to us and only their Lung CT scans, PET scans, their Smoking history, T-Stages, N-Stages, M-Stages and Histopathological grades along with other general data. All the subjects have cancer but of different stages. The dataset was taken from CANCER Imaging Archive and is available for public use. All the terms and conditions of the Dataset provider are being upheld by our team for the usage of the data of the subjects and all the rights are reserved with the publisher of dataset.

- **Data Collection Instruments and Procedures:**

  Data collected in this research is from secondary sources, namely, Cancer Imaging Archive, which is made publicly available. Due to large size of Dataset no download link was provided instead the files were downloaded from the official NBIA Data Retriever; the official download portal provided by the Cancer Imaging Archive. After the data was downloaded, it was analyzed by the team using Microsoft Excel and Microsoft Edge to read the annotation files. Pandas' library was used to read its columns in the model and it is a strong library to work with csv files and xlxs files. Pandas were also used to link the annotations with the splices of the PET and CT scans.

- **Data Analysis:**

  For the initial data analysis, the paper discusses the implementation of descriptive data management approach. A code snippet first iterates through all the splice files (DICOM format) and reads the metadata present in the splice files. It makes a clean metadata file of csv format which links the splice to its Series UID. And these Series UID are present in the filename of the Annotation files. This metadata file tends to link the Annotation with the splice image.

These Annotation files contain a mask which get on the splice image it is linked with stating at what position the tumor is present. This whole process further compliments the structure of the supervision machine learning models. So instead of marking whole image as Cancerous, we only have to mark a specific part and in this order more of the features recorded are useful which in turn addresses High Bais.

Then all the Patients have 64 images each under one folder of a single CT scan. To understand this, first we need to understand that CT and PET Scans are 3D images and as these 3D images are difficult to present in digital format for the model to learn, these 3D images are divided into splices which are some millimeters apart from each other mentioned in metadata. And stacking these splices on top of each other we get the whole of the 3D image of the Lung. And for a single patient we collect all the splices and train on these with markings from annotation files stating which of the splice files have tumor in them. As not all the splices of the Lung CT/PET won't have tumors in them.

- **Ethical Considerations:**

  The data is collected from secondary source so the patients are anonymous to us thus making us exempted from considering any Ethical considerations. For the sake of publishers, we have made it necessary to give all the credit of the dataset to publishers.

VII. Results

The whole classification of the Cancer Stage is done currently with the help of mainly image classification techniques, namely, ResNet18 which is trained on general data and DenseNet121 which has layers trained largely on medical data. Then the Ensemble of these two are also used. And we will compare the performance of these with the multimodal to test the performance of multimodal technique.

The dataset we have used for the training purpose of the various models is the same and the data has not been manipulated in any way. However, due to the limitations of the dataset the results are affected which is being discussed in later portion.

1. ResNet 18:

For ResNet we have trained it with the following parameters:

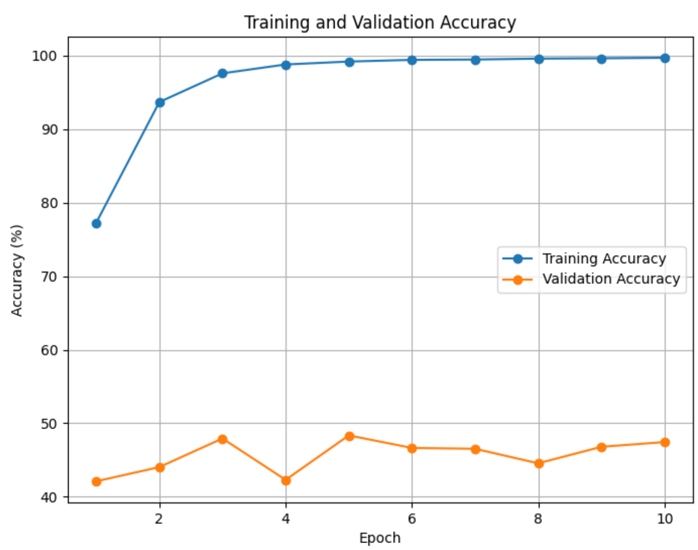
Epochs:10

Batch size: 32:

Num Workers: 8

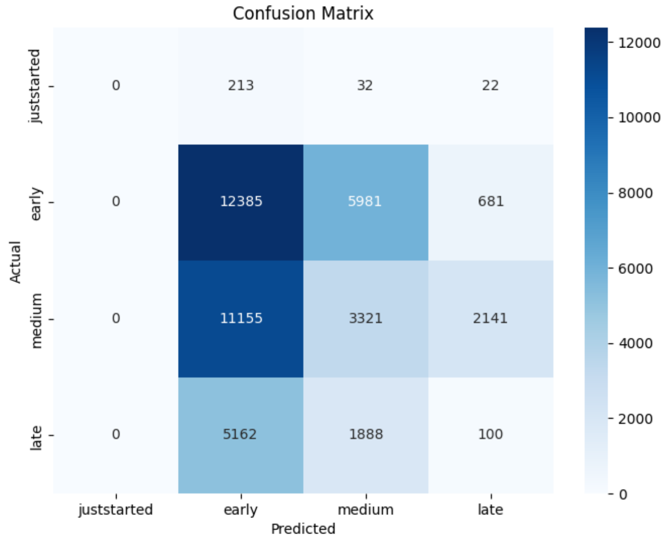
Learning rate= 0.001

Accuracy per epoch (training vs validation)



The validation accuracy is way lower that than of training accuracy. Indicating high variance.

Confusion matrix:



2. DenseNet 121:

For DenseNet we have trained it with the following parameters:

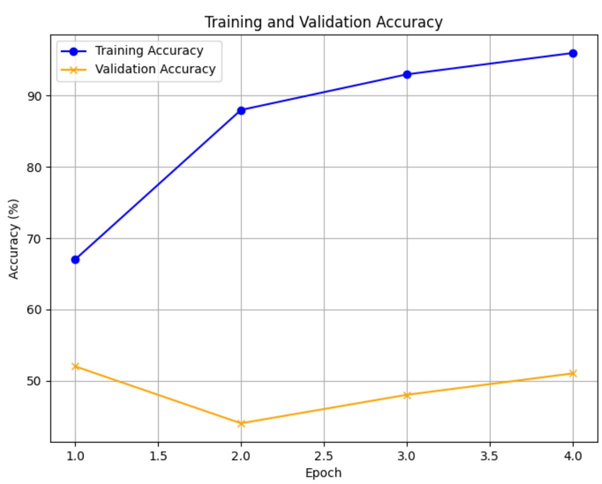
Epochs:4

Batch size: 32:

Num Workers: 8

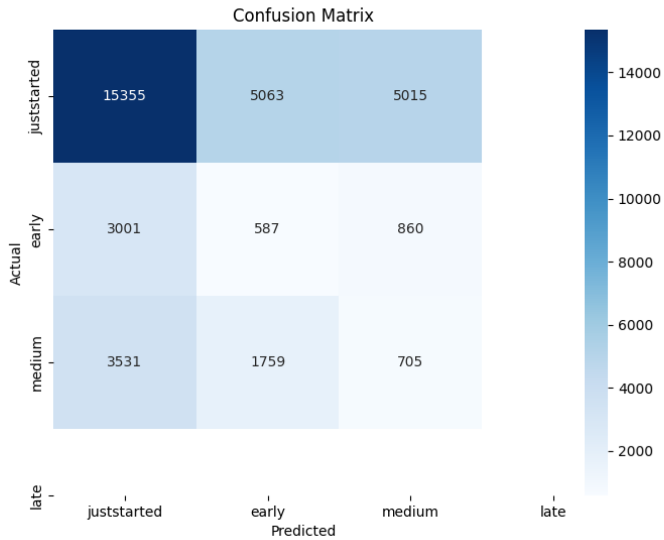
Learning rate= 0.001

Accuracy per epoch (training vs validation)



The validation accuracy is still largely below training accuracy still indicating high variance.

Confusion Matrix:



3. Ensemble (Stacking) of ResNet18 and DenseNet121:

For this ensemble we have trained it with the following parameters:

Epochs:5

Batch size: 32:

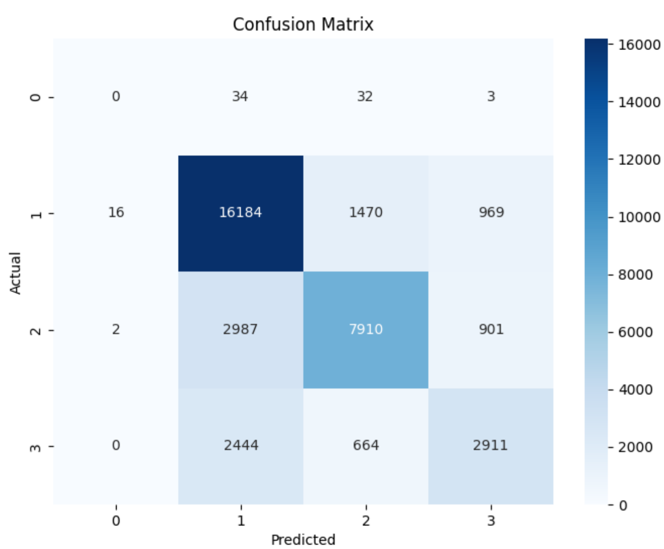
Num Workers: 8

Learning rate= 0.001

Accuracy per epoch (training vs validation)

The validation is now close to accuracy here and it indicates that stacking can solve the overfitting problem to some extent.

Confusion Matrix:



4. DenseNet 121 in MultiModal:

For this Multimodal we have trained it with the following parameters:

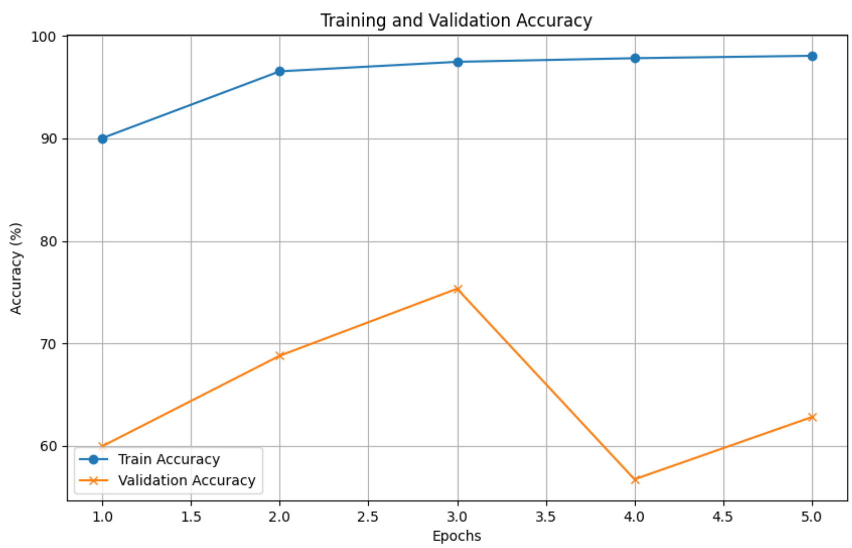
Epochs:5

Batch size: 32:

Num Workers: 8

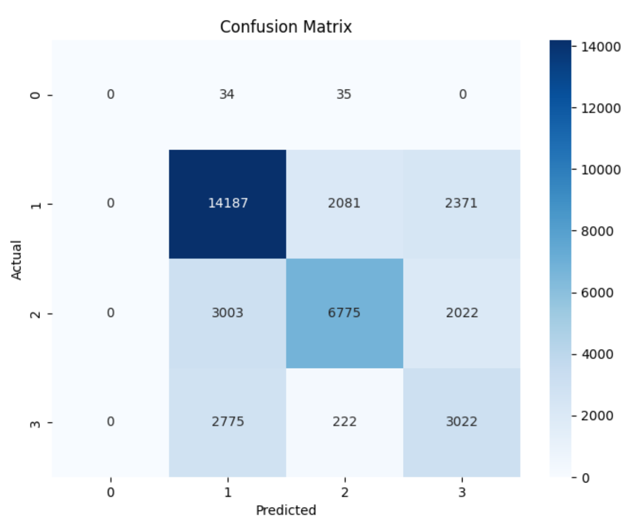
Learning rate= 0.001

Accuracy per epoch (training vs validation)



The validation is below accuracy curve indicating high variance.

Confusion matrix:



5. Ensemble in MultiModal:

For this ensemble we have trained it with the following parameters:

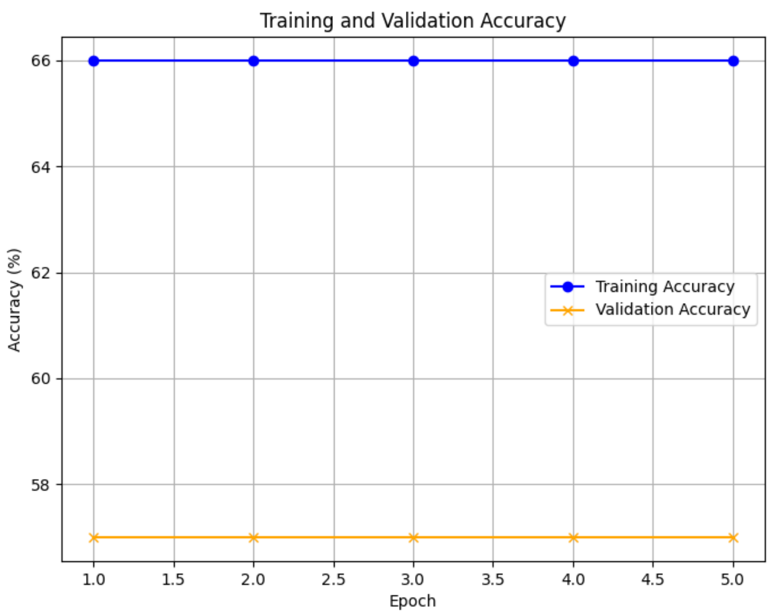
Epochs:4

Batch size: 32:

Num Workers: 8

Learning rate= 0.001

Accuracy per epoch (training vs validation):



VIII. Discussion

- **Interpretation of Findings:**

To understand finding, the research first highlights the current dataset and the characteristics of this dataset and how it affected the finding in the research. The current dataset partially incomplete so the results are affected.

As seen from the above result, for classifying just images DenseNet121 performs better than Resnet18 simply because, the original DenseNet121 model is trained on medical images only, whereas the Resnet18 model is trained on different objects like humans, cat, dog etc.

In the Ensemble Stacking Accuracy v/s Epoch plot we can see the cross validation is significantly improving as compared to just ResNet18 or DenseNet121 because it uses both of these as a weak learner for training itself and then applied logistic regression top of the given result to classify using the logit function and cross entropy loss function. Now the dataset doesn't contain the whole 3-D images since it's practically impossible to have thousands of 3-D images and then process them. So as mentioned above the images are spliced into multiple planes for each patient so that 1 CT scan results into numerous splices like 64. The tumor part in the image also gets spliced. The ResNet18 and DenseNet121 does not work properly on such type of spliced images without annotation because it considers the whole set of splices as cancerous instead of marking only those splices that contains tumor. Whereas in stacking, with the help of generalization of ResNet18 and specialization of DenseNet121 the model becomes more sensitive towards the tumor part.

- **Implications of the Findings:**

The approach of Hierarchical Ensembling in Multi Modal with Late Fusion model and the comparative study discussed in this research shows the practical use of our approach. The comparative study performed helps visualizing the accuracies and limitations of multiple models. As shown in results the pre trained DenseNet121 model on medical image also showed high variance meaning for each dataset the model has to be fine-tuned properly to get the desired results. Our study also shows how multi modal improves cross validation accuracy as compared to the normal DenseNet121 on medical image classification. However, it is still not enough to convert high variance into a best fit. The Hierarchical Ensembling in Multi Modal with Late Fusion model shows us how it can drastically reduce the high variance that is the overfitting done by the model. But due to limitations of dataset it reflects high bias that is the model is too simple for the dataset.

This research can overall help medical professionals in making efficient and improved decision making. If tweaked right the model could be used for prediction of other diseases as well.

- **Limitations of the Study:**

  1. The dataset used was downloaded from the Cancer Image Archive and is partially wrong. And all the other datasets on the Cancer Imaging archive are incomplete or partially wrong. The main dataset we needed for the validation of the findings is that we have a dataset that have image data, clinical lab data and reports, all linked for the same patient. What were downloaded was a dataset which said to have medical image data, annotations and some clinical notes. However, the annotations were not correctly mapped with the given medical images.
  2. Due to lack of annotations, the medical images were not properly identified for cancerous part. As a result, all the splices were considered to have tumor while only some had in a set for a particular patient. This resulted in weak supervision model.
  3. The dataset used was of 128 Gigabytes. Even after huge size it still had data of only 356 patients which is very less to predict in real world.
  4. To get improved and highly efficient results we need to run the whole dataset for multiple epochs until perfect weights are identified in the neural network. However, due to the size of dataset and the complexity of convolution neural networks, multi modal, hierarchical Ensembling and fusion model, we had a constraint of hardware and as a result we were not able to run any model for more than 10 epochs. Even for those 10 epochs the training, and testing time was around 5 hours for each model.
  5. If the desired hardware is acquired then the accuracy and efficiency of multi modal can be improved by running it with more layers and for more epochs.

- **Suggestions for Future Research:**

  1. If dataset with proper annotations for medical images and clinical lab reports is provided, then the working of model and efficiency can be further improved mainly by addressing the underfitting which occurs due to model being too simple.

2. Once the Hierarchical Ensembling in Multi Modal with Late Fusion model is completely tuned, we plan to develop a user-friendly web-application, Ingenious Health Management System (IHMS), which will help users to upload and manage their medical information.

IX. Conclusion

The research shows how normal stacking improves the overall accuracy, but Hierarchical Ensembling in Multi Modal with Late Fusion model addresses the issue of overfitting. But the high variance situation turns into a high bias situation due to lack of proper dataset. The study will significantly help the medical professional in making an improved and efficient decisions. The multi modal is still under research and in past few years a lot of research has been done on it. But it is yet to be applied in real world as mentioned theoretically. Our study also helps to understand the comparative difference between multiple models which helps gain understanding of the use of models according to the dataset. The IHMS will also help patients around the world and specifically the patients residing in rural areas to do early detection of cancer which cold help in preventing loss of life.

To summarize the whole study, the Hierarchical Ensembling in Multi Modal with Late Fusion model addresses the occurrence of overfitting, but it is only recommended to use the technique with a proper labeled and annotated dataset with both medical images and clinical lab report and a complex model which will be able to detect the underlying patters in the data. To handle this situation more sophisticated multi model is required for which more layers are to be added in the neural networks. It demands more resources and we are lacking resources and due to these constraints, this study has to wind-up with this conclusion.

X. References

Munir, K., Elahi, H., Ayub, A., Frezza, F., & Rizzi, A. (2019). Cancer diagnosis using deep learning: A bibliographic review. Cancers, 11(9), 1235. https://doi.org/10.3390/CANCERS11091235

Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., ... & Denniston, A. K. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. Lancet Digital Health, 1(6), e271–e297. https://doi.org/10.1016/S2589-7500(19)30123-2

Zeng, K., Pan, Z., Xu, Y., & Qu, Y. (2020). An ensemble learning strategy for eligibility criteria text classification for clinical trial recruitment: Algorithm development and validation. JMIR Medical Informatics, 8(7), e17832. https://doi.org/10.2196/17832

Rajaraman, S., Siegelman, J., Alderson, P. O., Folio, L. S., Folio, L. R., & Antani, S. K. (2020). Iteratively pruned deep learning ensembles for COVID-19 detection in chest X-rays. IEEE Access, 8, 115041–115050. https://doi.org/10.1109/ACCESS.2020.3003810

Muller, D., Soto-Rey, I., & Kramer, F. (2022). An analysis on ensemble learning optimized medical image classification with deep convolutional neural networks. IEEE Access, 10, 66467–66480. https://doi.org/10.1109/ACCESS.2022.3182399

Kline, A., Walcott, B. P., Mowery, D. L., Zhu, V., Xue, J., & He, J. (2022). Multimodal machine learning in precision health: A scoping review. npj Digital Medicine, 5(1), 1–14. https://doi.org/10.1038/s41746-022-00712-8

Mahajan, P., Uddin, S., Hajati, F., & Moni, M. A. (2023). Ensemble learning for disease prediction: A review. Healthcare, 11(12), 1808. https://doi.org/10.3390/HEALTHCARE11121808

Chakravarty, D., Thomas, A. M., & Vivek, V. (2023). A survey on decentralization and virtualization of medical trials: An approach through ensemble learning models and convolutional neural networks. In 2023 International Conference on Computer Science and Emerging Technologies (CSET). https://doi.org/10.1109/CSET58993.2023.10346978

Reddy, S. (2024). Generative AI in healthcare: An implementation science informed translational path on application, integration and governance. Implementation Science, 19(1), 1–15. https://doi.org/10.1186/S13012-024-01357-9

Krones, F., Marikkar, U., Parsons, G., Szmul, A., & Mahdi, A. (2025). Review of multimodal machine learning approaches in healthcare. Information Fusion, 114, 102690. https://doi.org/10.1016/J.INFFUS.2024.102690

Imrie, F., Denner, S., Brunschwig, L. S., Maier-Hein, K., & van der Schaar, M. (2025). Automated ensemble multimodal machine learning for healthcare. IEEE Journal of Biomedical and Health Informatics, 1–14. https://doi.org/10.1109/JBHI.2025.3530156

https://www.iarc.who.int/ (WHO statistics)

yuvraj vishwas

R&D report -Yuvraj & Vishwas(PDF).pdf

NIIT University, Neemrana

Document Details

Submission ID

trn:oid:::7692:93303686

Submission Date

Apr 28, 2025, 4:09 PM GMT+5:30

Download Date

Apr 28, 2025, 4:10 PM GMT+5:30

File Name

R&D report -Yuvraj & Vishwas(PDF).pdf

File Size

513.0 KB

25 Pages

6,048 Words

31,757 Characters

7% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Cited Text

Match Groups

37 Not Cited or Quoted 7%
Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

4% Internet sources

4% Publications

4% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

yuvraj vishwas

R&D report -Yuvraj & Vishwas(PDF).pdf

NIIT University, Neemrana

Document Details

Submission ID

trn:oid:::7692:93303686

Submission Date

Apr 28, 2025, 4:09 PM GMT+5:30

Download Date

Apr 28, 2025, 4:10 PM GMT+5:30

File Name

R&D report -Yuvraj & Vishwas(PDF).pdf

File Size

513.0 KB

25 Pages

6,048 Words

31,757 Characters

0% detected as AI

The percentage indicates the combined amount of likely AI-generated text as well as likely AI-generated text that was also likely AI-paraphrased.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Detection Groups

0 AI-generated only 0%
Likely AI-generated text from a large-language model.

0 AI-generated text that was AI-paraphrased 0%
Likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.

Disclaimer
Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (it may misidentify writing that is likely AI generated as AI generated and AI paraphrased or likely AI generated and AI paraphrased writing as only AI generated) so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?
The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?
Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.