International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Deep Learning for Holistic Mental Illness Diagnosis with Multimodal Imaging

Vijay Kumar Ravi[1], Kavita Ravat[2]

[1,2]Department of Information Technology, UIT Barkatullah University, Bhopal, Madhya Pradesh, India

Abstract

Detecting and diagnosing mental illnesses is a complex task due to the heterogeneous nature of these conditions, which encompass various types and subtypes. Timely identification and diagnosis are essential for advancing mental health research, improving early intervention, and developing effective therapies. However, existing methods have primarily focused on categorizing and diagnosing single mental illness types, often prioritizing patient survival over comprehensive treatment. This research introduces a deep learning-based approach to establish a universal framework for the detection, diagnosis, symptom-based prediction, and screening of diverse mental illnesses, their subtypes, and associated factors using a diverse dataset comprising images and clinical scans. The proposed architecture leverages a VGG-19-based 3D-convolutional neural network for robust feature extraction and employs a random forest algorithm for regression tasks. To create a comprehensive dataset, we incorporate results from the DAIC-WOZ laboratory dataset, imaging studies, and biopsy reports, moving beyond sole reliance on clinical images. The initial step involves distinguishing between healthy and unhealthy depression, sleep disorder, Alzheimer's, schizophrenia, and epilepsy conditions and subsequently classifying depression, sleep disorder, Alzheimer's, schizophrenia, and epilepsy cases into their specific subtypes while assessing their severity and growth patterns. Our model is designed to predict risk levels at various time intervals, utilizing the available pertinent factors consistently across different diagnostic tools. Our results demonstrate promising accuracy, with a 98.80% success rate in classifying them into their respective subtypes, a sensitivity of 98.24%, and a specificity of 98.49%. This research represents a significant step toward a more holistic and precise approach to mental illness detection, diagnosis, and management, performing 4.74% better than other state-of-the-art AI models and offering potential benefits for individuals and mental health professionals alike.

Keywords: Computer-aided Diagnosis; Electroencephalogram; Mental Illness; Recurrent Neural Network

1. Introduction

Artificial intelligence (AI), for many years, has caught society’s imagination and created a passion for its potential to better our lives. AI has played an important role in our day-to-day lives and our interplay with transportation, media, communications, and health. There is an increment in interest in the use of AI in healthcare for the betterment of mental illness diagnosis, management, and the evolution of successful therapies [1]. Hence, with the significant amounts of data generated, there is a particular interest in the use of AI for better oncologic care. Machine learning has diverse possibilities in diagnosis,

detection, and symptoms-based prediction, screening, identification, and detection of mental illness using medical images [2].

Image processing, clinical photographs, radiographic imaging, and digital pathology have proved to be among the most successful methods by which AI has influenced everyone. Given the huge amount of digital imaging data present inside medicine, there is an increase in excitement about the use of the same techniques for imaging in oncology. This uprising in image analysis was sped up by the development of a specific deep learning architecture, the CNN [3]. CNN examines pixel-level data from the images. The additional advantage of CNNs as compared to other DL configurations is their capacity to account for the adaptation of the pixels concerning one another. This successfully permits the CNN to value lines, curves, and, at the end, objects inside images. CNN-based models have recently been shown to be equal to humans in picture classification and object detection [4].

We examine the situation of the art in mental illness systems and analyze current practices, problems, and chances of image acquisition, segmentation, selection, extracting features, categorization of dermoscopic visuals, and pre-processing. The paper explains the results of the most important execution recorded to date. We collate the workings of some classifiers mainly made for lesion diagnosis and debate the corresponding findings.

Lately, depression malignancy is one of the most perilous kinds of disease. On the off chance that the location of depression disease in the beginning phase is done, it very well may be useful to fix it. In clinical picture finding, computer vision can assume a significant role and is demonstrated by many existing frameworks. For recognizing depression malignancy, the scientists have found a PC-supported technique. A picture-handling instrument is utilized. The information will be the mental injury picture. A novel picture preparation procedure will be applied to that picture, and it will analyze it to determine the presence of mental depression, sleep disorder, Alzheimer's, schizophrenia, and epilepsy growth. The various limits are analyzed, like jaggedness, dimensions, diameter, color, etc., by the sensitive image assessment contraptions. The element boundary that is taken out is utilized for the characterization of the picture as sound brain and melanoma, depression, sleep disorder, Alzheimer's, schizophrenia, and epilepsy growth [5].

Malignancy is a broadened illness containing diverse subtypes. The beginning stage finding and forecast of a disease type is significant in depression, sleep disorder, Alzheimer's, schizophrenia, and epilepsy growth research, as it can facilitate the clinical administration of patients. The significance of the classification of disease victims into high- or slim-likelihood bunches has driven a few examination groups, from the biomedical and bio-informatics fields, to notice the utilization or use of AI strategies. Thus, the accompanying procedure has been utilized as an objective for the therapy of dangerous conditions. Besides, the limit of ML instruments to discover key highlights from complex datasets uncovers their significance. A different strategy that incorporates Bayesian networks, support vector machines, artificial neural networks, and decision trees has been generally remembered in disease research for creating prescient models that bring about exact and fruitful dynamics. By clearly utilizing ML strategies to improve our comprehension of depression, sleep disorders, Alzheimer's, schizophrenia,

and epilepsy growth movement, what's required is an appropriate degree of approval so these techniques can be remembered in everyday clinical practice [6].

The researchers had presented a review of progressing ML-based methodologies used in exhibiting the development of tumors. The prescient architectures considered here rely upon different oversaw ML techniques, as well as various data features and data tests. Given the design of the usage of ML systems in danger research, presented here are the latest circulations that use these techniques as a goal to display hazardous or patient results. A pathway-driven view acknowledges the heterogeneity among genomic profiles from various tumor victims while assuming that the changed qualities are generally possible to have a place with a similar pathway and cause similar sickness aggregates. Genuinely, network-driven methods have ended up being helpful for identifying genotypic reasons for infections, arranging illness sub-types, and sorting out drug points. In the examination, they talked about how organizations might be utilized to help know patient-to-patient varieties and how an individual can hold this changeability to clarify connections between malignancy drivers [7].

Mechanical advances in processing, imaging, and genomics have opened up new possibilities for examining connections between histology, subatomic events, and clinical outcomes by utilizing quantitative techniques. Slide-examining gadgets can quickly deliver gigantic advanced picture records that hold histological subtleties in high regard. Comparable advances in figuring and picture investigation calculations approve mining of records to take out depictions of histology, shifting from essential human explanations to programmed and correctly quantitative morphometric portrayals of millions of cells. These imaging capacities address a new measurement in tissue-based investigations, and when converged with genomic and clinical endpoints, they might be utilized to examine biologic highlights of the tumor miniature climate and to discover new morphologic biomarkers of hereditary changes and patient outcomes. In the paper, we break down advancements in quantitative imaging innovation and show how picture highlights can be joined with clinical and genomic information to investigate the central trouble in malignancy. Utilizing inciting models from the investigation of glioblastomas (GBMs), we show how open information from the mental illness Genome Atlas (TCGA) may fill in as an open stage to direct silico tissue-based examinations that join present information assets. We show how these methodologies might be utilized to investigate the connection of the tumor miniature climate to genomic changes and quality articulation designs and to characterize atomic morphometric attributes that are predictive of hereditary alterations and clinical outcomes. Issues, limitations, and arising chances in the fields of quantitative imaging and integrative examinations are additionally being talked about [8].

AI strategies are generally utilized in bio-informatics, computational science, and frameworks science. In this, we examine the advancement of AI strategies for protein structure forecast, quite possibly the most principal issue in underlying science and bio-informatics. Protein structure assumption is such an irksome issue that it is, generally, spoiled and experienced at four distinct levels: the 1-dimensional conjecture of essential characteristics alongside the vital course of action of amino acids; the 2-dimensional figure of the spatial connection between amino acids; the 3-dimensional assumption for the third development of a protein; and the 4-dimensional estimate of the fourth plan of a multi-protein complex. An alternate game plan of both managed and independent AI methodologies has been set up

over various years to deal with these issues and has astoundingly driven the forefront of protein structure assumption. In the paper, we inspect the new development and vocations of hidden Markov models, neural associations, support vector machines, Bayesian methods, and gathering techniques in 1-Dimensional, 2-Dimensional, 3-Dimensional, and 4-Dimensional protein structure figures [9].

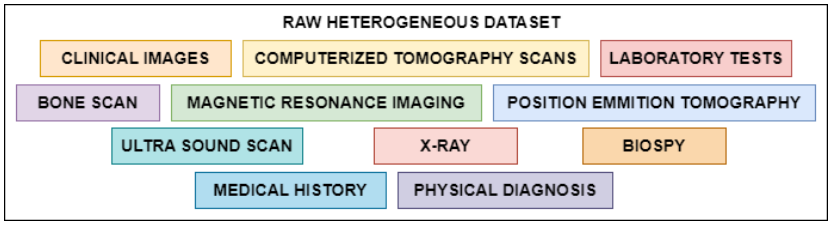
Protein structure prediction is a main point in structural bio-informatics. From the 1960s, measurable techniques, trailed by continuously troublesome ML and late DL-based strategies, are being utilized in foreseeing protein primary data at various degrees of traits. In this examination, we right away start the issue of protein structure expectation and significant components of deep learning (for example, convolutional neural networks, recurrent neural networks, and fundamental feed-forward neural networks they are established on), after which we banter the movement of prescient techniques for one-dimensional and two-dimensional protein structure annotations, from the simple factual strategies for the underlying days to the computationally thorough and exceptionally progressed deep learning calculations of late time. In the methodology, we break down the advancement of the data sets these calculations depend on and what this has meant for our ability to hold information about improvement and co-development to get improved forecasts. We finish this assessment by illustrating the most recent job of deep learning methods inside the greater pipelines to foresee protein designs and endeavoring to anticipate what incitements and chances may come straightaway [10].

2. Design Methodology

A. Data Collection

The International Collaboration on Mental Illness Research (DAIC-WOZ) dataset has been created to provide a reliable, evidence-based methodology for the outlining of mental illness and its subtypes. This dataset ensures that it contains various subtypes of mental illness lesions and EEG signals with consistency in style and content and contains every parameter needed for the management, diagnosis, and prediction of mental illness lesions and brain disorders. The dataset contains various classes containing mental illness images and scans. We have classified the data into two major categories: carcinoma and neoplastic. Now the carcinoma part generally deals with malignancy or mental illness-proliferated parts, whereas the neoplastic images contain a little outgrowth, which does not tend to mental illness at this moment. Each radiograph and clinical image is either in the format of RGB or grayscale, with a pixel resolution of 145 pixels per inch and dimensions of 2048 x 2048 pixels. We have confirmed the radiographic classification by using CT scans. Our classification of nodules and lesions as healthy or depression, sleep disorder, Alzheimer's, schizophrenia, and epilepsy was not only based on their resolution and appearance and CT scans but also focused on tested tissue samples and monitoring changes in the nodules, lesions, and EEG signals happening with time. If the nodules, lesions, or EEG signals got shrunken or disappeared due to antibodies or medical doses or didn't go severe over time, they were considered healthy.

Figure 1: Components of a Medical Diagnostic Data Set



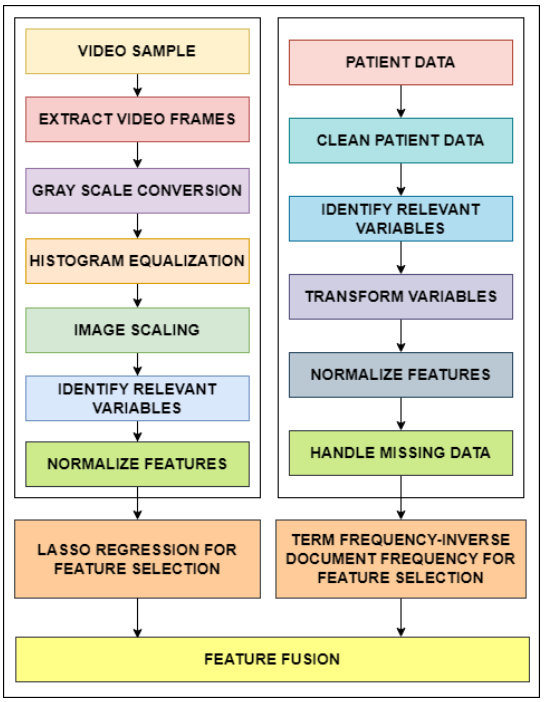
The DAIC-WOZ dataset also contains information about the size and coordinates from the center of the clinical region of importance. These values are used for constructing a square-bounding box around the affected region. In this research, we have reserved 10% of the total data as a test set and 20% of the data as a validation set. We performed the cross-validation by splitting the remaining 80% of the data into four folds. We tuned the hyper-parameters for the first three folds by training and the fourth fold by testing, repeating these steps for each fold. Once we chose the hyper-parameters, we re-trained the network using images and scans from each fold. We also used the data augmentation technique as the dataset was small. We increased the data artificially by increasing the total number of scans and images for each class distribution. As we know that the nodules can appear at various locations and may vary in size with respect to the time period within the radiograph scans, we augmented the dataset by cropping the locations, using translation, rotation, horizontal mirroring, and scaling. To use the 3D-convolutional neural network-based VGG-19 architecture on the augmented dataset, we need to modify the dataset as per the architectural requirement. The VGG-19 architecture takes an input image of size 224 x 224. We scaled and performed random dimensionality reduction on the images and scanned them between 256 x 256 and 384 x 384 dimensions, later cropping them to 224 x 224 pixels. To deal with the grayscale radiographs, the pixels were copied to the three shading channels, and the mean picture was deducted.

B. Data Preprocessing

The image dataset created is pre-processed, undergoing rotation, scaling, and translation sequentially. The images are normalized and resized according to system compatibility. We perform geometrical operations, which are scaling, rotation, and translation, on our dataset. When scaling is performed, we resize the graphical images, performing graphical transformation using raster and vector graphics without losing the image quality. Using rotation, we rotate the images to a specified degree about their center for a clarified and understandable view of the clinical regions. We either perform a forward or reverse transformation on the annotated images. Also, we perform translation by mapping the position coordinates of the clinical region and transforming them to new coordinates while keeping their dimensionality constant. The noise is removed from the image or scans, which enhances their quality and scans to make the image clear and for subsequent mental illness analysis.

Initially, we apply the salt and pepper noise removal technique and block matching with 3D filtration for noise removal and filtration of the sample image by setting the denoising strength and patch size. We perform image segmentation, in which we partition digital images into multiple pixel sets, also called image objects, and convert them into multiple segments. We apply the image segmentation to relocate the image boundaries and objects. Later, we process the images based on shapes by applying morphology to a varied set of clinical images. Based on the morphological operations, we adjust the pixel images based on other pixels in their neighbors. We also apply dimensionality reduction to convert data from high-dimensional space to low-dimensional space to reduce memory space. The low-dimensional portrayal holds some significant properties of the first information, preferably near its inherent measurement. We used this pre-processing technique only for clinical images.

Figure 2: Data Preprocessing Workflow



The dataset containing scans is created and pre-processed, undergoing rotation, scaling, and translation sequentially. The scans are normalized and resized according to system compatibility. The dataset contains bone scans, clinical images, positron emission tomography scans, computerized tomography scans, magnetic resonance imaging, ultrasonic scans, and x-ray scans. We do data securing procedures by changing over the states of being into advanced structure for additional capacity and analysis. Regularly, signals from sensors are examined, changed over to computers, and stored by an independent device. We perform geometrical operations, which are scaling, rotation, and translation, on our dataset. When scaling is performed, we resize the graphical scans, performing graphical transformation using raster and vector graphics without losing the image quality. Using rotation, we rotate the scans to a specified degree about their center for a clarified and understandable view of the clinical regions.

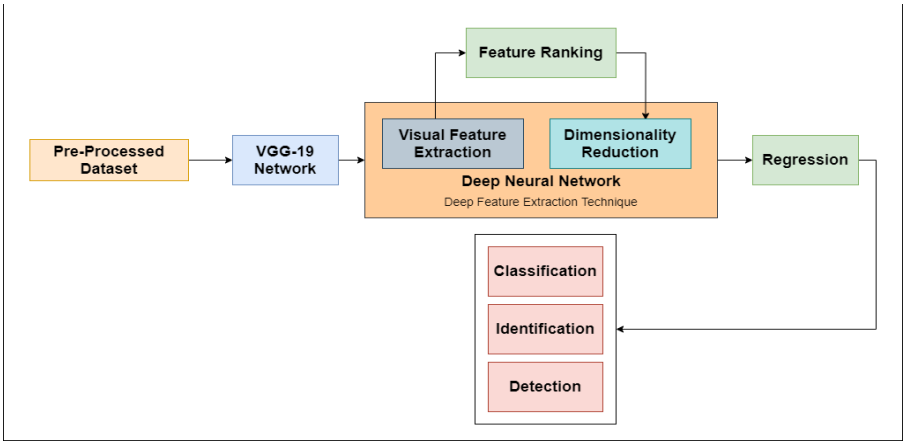
We either perform forward or reverse transformation on the annotated scans. Also, we perform translation by mapping the position coordinates of the clinical region, transforming them to new coordinates by keeping their dimensionality constant. We perform radiometric correction for calibrating and correcting pixel values, loss values and error terms in the given data. This process is used for improving the quality and degree of interpretability of the remote sensing data. We use radiometric correction and calibration to compare multiple datasets over a period of time. We also use spectral resolution for detecting and displaying the size of the smallest feature in the image. It is usually represented as a single value of the length of one side of a square. Spectral Resolution is the estimation of explicit frequencies of the electromagnetic range. The better the spectral resolution, the smaller the frequency range for a specific channel or band. We also use temporal resolution that refers to the periodic time in between the images. We performed multi-temporal analysis over clinical scans providing a detailed information of clinical changes happened over a period of time over the region of focus. We also preformed error assessment to eliminate or the probability of errors over the assessment of mental illness and affected areas.

After pre-processing the dataset, we identify the relations and their dependencies. We plot a statistical summary of all the relational dependencies identifying all the errors, missing values and null values. After error assessment, we remove all the errors, replace missing and null values with zeros and dynamic values in the frame. Hence we find error co-relation, balance the dataset and select the features having more importance. We scale the images and scans from varying pixels values to 224 X 224 as a standard RGB input.

C. Network Architecture and Feature Extraction

In this architecture, we have used VGG-19 architecture for feature extraction from clinical images and scans. We used the pre-processed dataset for extracting the features and training the model. We have partitioned the dataset into three segments as train dataset, test dataset and validation dataset in a ratio of 70:20:10. We have used random samples from the dataset and detected whether the sample is healthy or depression, sleep disorder, alzheimer's, schizophrenia, and epilepsy. Later we classify mental illnessous images into various other classes in which the input given to the system is suspected to be melanocytic lesion. We enhance the image quality by performing pre-processing on these images. We use the automated technique for calculating the edge and threshold which is used for segmentation and processing of images. Later, we perform analysis of the segmented images considering the clinical region of importance and affected region with their ABCD and geometric factors extracted by the feature extraction block which are considered as the most prominent features in detecting a mental illnessous lesion, nodule or a tumor. After extracting features we apply the regression technique of the extracted features to classify these features according to their mental illness type. Here we used a random forest algorithm for the regression. We also detect the severity and stage of the lesions, nodules and EEG signals using the feature classification block. We also compare the pre-defined thresholds with feature parameters. We perform feature extraction with dimensionality reduction using VGG-19 based 3D-Convolutional Neural Network which results in reduction of complexity in the dataset, grouping it into manageable groups for classification. The major characteristic background of this dataset is it contains a large number of parametric variables that require a lot of computation and training resources. The feature extraction process effectively reduces the amount of data to be processed, still describing the original dataset accurately. We reduce the number of resources needed for processing without loosing the important extracted parameters.

Figure 3: Feature Extraction using VGG-19 Network for Mental Illness Classification

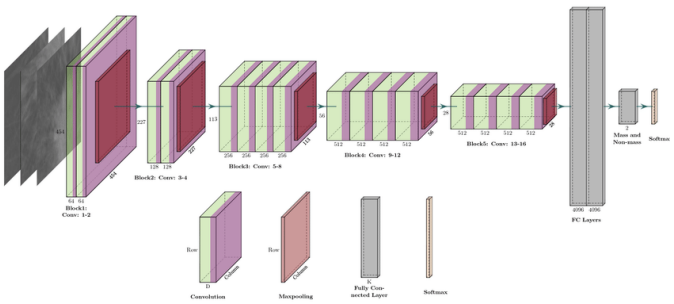


We also perform feature ranking by ordering the features according to their importance and value of score functions by calculating the feature relevance. We classify the lesions according to their severity and its developed immunity using random forest algorithm. The random forest algorithm further used for classification, identification of mental illness type and detection using regression. We apply regression on the dynamic extracted features to classify the type of mental illness according to its severity. The calculation does both line testing and segment reviewing with the choice tree as a base. Model m1, m2, m3, m4 are more not equivalent to by doing simply stowing considering fragment testing as we increase the amount of base injuries (n), the change will decrease. Right when you decrease n, distinction increases. However, inclination remains consistent for the whole cycle. n can be found using cross-endorsement. We are not worried about significance, we let them create because at end change decreases in combination. For model m1, (d-d') dataset not used in showing are out of pack dataset. they used cross-endorsement for the m1 model. Expect there are K insights and L features in the planning instructive assortment. A subset of L features is picked heedlessly and whichever feature gives the best split is used to part the center point iteratively. The tree is created to be the greatest. The advances are reiterated and assumption is given relying upon the all out of figures from various trees. As the quantity of base models expands, preparing run time increments so consistently utilize cross-approval to discover ideal hyper-boundaries.

A 3-D CNN is a powerful learning model representing volumetric data, takes an equivalent input of 3D volumetric sequences or a 2D frame data. Our model is a 3D-CNN based VGG-19 network that takes an input of RGB image as 224 X 224 pixels. The pre-processing layer takes an RGB image input with a pixel value ranging from 0-255 and subtracts the mean image values within the train dataset. We pre-process the clinical images and pass these images through weighted layers. We passed these clinical images from the train dataset through slacked convolutional neural blocks. This architecture consists of all over 19 layers in which sixteen are convolutional blocks and the remaining three are fully connected layers. This network consists of five maximum pooling layers with 4096 channels following 100 convolutional layer labels.

The last most layer consists of fully connected softmax layers used for classification purposes. The first and second layer consists of convolutional blocks with dimensions of 3 x 3 filters using 64 filters. The consequence of this is it uses dimensions of 224 x 224 x 64 convolutional volumes. With stride of 1, we use the filters of dimension 3 x 3. We have used a pooling layer with a maximum pooling size of 2 x 2 with a stride of 2. This process reduces the dimensions, height and width by a volume from 224 x 224 x 64 to 112 x 112 x 64. Following them, another two convolutional layers containing 128 filters makes a new dimension of 112 x 112x 128. After the use of the pooling layer, the volumetric dimension gets reduced to 28 x 28 x 256. The convolutional layer consisting of two more stacks separates a maximum pooling layer. Proceeding with the final pooling layer, the volumetric dimension is reduced to 7 x7 x 512 into the fully connected convolutional layer consisting of 4096 channels and 1000 classes for softmax output.

Figure 4: VGG-19 Network Architecture



We use this architecture for image based feature extraction using clinical images or imaging scans used for mental illness identification, screening and detection. The architecture takes visible input of 224 X 224 pixel RGB images. The receptive field of the convolutional layer from VGG-19 uses a very small size of 3 x 3, a much possible smaller size that captures directions, up, down, left, right.The convolutional filters of 1 x 1 dimensions followed by a ReLU unit acts as a linear transformation for the input. We fixed the convolution stride to one pixel for preserving the spatial resolutions after convolutions. All the hidden layers in VGG-19 use ReLU linear units. As the local response normalization consumes more memory with more training time in the training phase with no change in accuracy values, VGG-19 does not use LRN.

3. Result and Discussion

We have performed extensive experimentation to train the feature extraction and regression-based architecture and for the validation of the performance by these designed architectures. We perform rotation estimation up to tenfold for detection of mental illness-affected regions and the region of clinical importance on the DAIC-WOZ dataset using annotated data. We have evaluated and tested our model on the test dataset containing patient-level nodule diagnosis. We have figured our model on NVIDIA Tesla K80 GPU with a graphical memory of 12 GB in PyTorch. We have prepared our characterization model with DAIC-WOZ dataset which contains 4122 patient-named information, partitioned into preparing (3645 pictures) and approval (477 pictures) dataset with expanded information. We noticed the issue of arrangement lopsidedness in the dataset. Henceforth we performed different information argumentation methods over the dataset like interpretation, revolution, flipping, and trimming we inferred to build the preparation information and eliminate the irregularity of class. This dataset contained very much named information, which is significant for preparing the relapse-based order and highlight extraction model.

We have trained our neural network model using stochastic gradient descent values with an initial batch size of 400 epochs. We have completed the training process using three different rates of learning, initiated as 0.01 for the first hundred epochs, 0.001 for the next hundred, 0.001 for the next hundred, 0.0001 for more next hundreds, and finally 0.00001 for remaining epochs. We have considered a weight decay of 0.0001. In this architecture, we have used the VGG-19 3D-Convolutional Neural Network for improving the training and testing accuracies which contained a fully activated pre-trained network. This process lowered the training and testing errors with appropriately initialized weights. Our main purpose of training the model was to reduce the difference between the grown truth labeled data and the network's produced output. There are two main reasons for stopping the training process. Initially, the

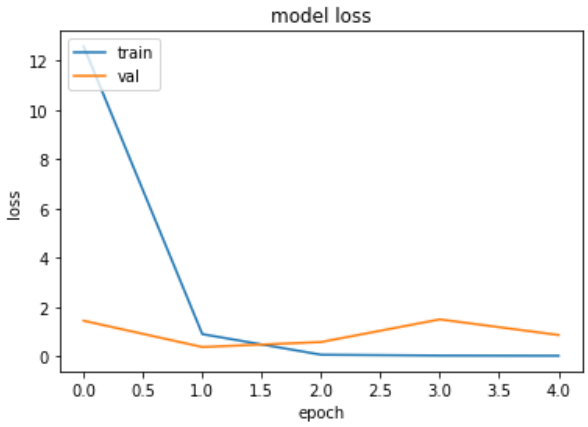
smallest error rate was reached by the network architecture, and finally, there was no change in the loss function from the validation dataset. We have trained and developed the lesion detection module on the DAIC-WOZ dataset with ten-fold patient-level split and cross-validation.

Table 1. Training Model Summary : Total Params : 7,635,264, Total Trainable Params : 7,635,264 , Total Non-Trainable Prams: 0

Convolutional Block	Output Type	Params
3D-Convolutional_1	(None, 14, 14, 512)	36928
3D-Convolutional_2	(None, 28, 28, 512)	1792
3D_Max_Pooling_1	(None, 112, 112, 64)	0
3D-Convolutional_3	(None, 112, 112, 128)	147584
3D-Convolutional_4	(None, 56, 56, 256)	73856
3D_Max_Pooling_2	(None, 56, 56, 128)	0
3D-Convolutional_5	(None, 56, 56, 256)	590080
3D-Convolutional_6	(None, 56, 56, 256)	590080
3D-Convolutional_7	(None, 112, 112, 128)	590080
3D_Max_Pooling_3	(None, 28, 28, 256)	0
3D-Convolutional_8	(None, 28, 28, 512)	2359808
3D-Convolutional_9	(None, 224, 224, 64)	1180160
3D-Convolutional_10	(None, 28, 28, 512)	2359808
3D_Max_Pooling_4	(None, 224, 224, 64)	0

We have used the free-response receiver operating characteristic curve (FROC) which is an unofficial parameter for matrix evaluation on the DAIC-WOZ dataset for validation of proposed models. The FROC curve shows the plot for sensitivity versus false positives in one scan for the proposed model in Figure 5.

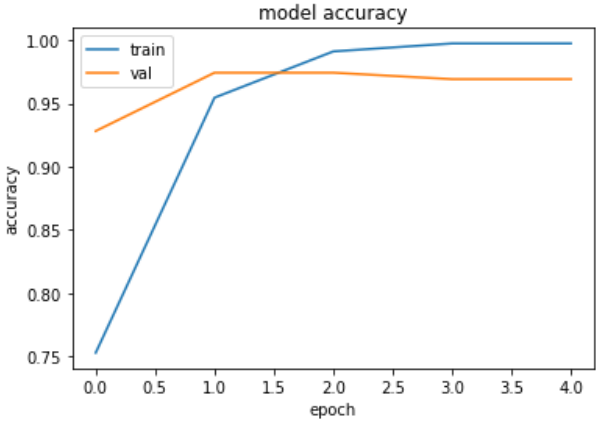
Figure 5. Model Loss



We have performed the testing phase concerning the detection of the probability of thresholds with prioritizing the sigmoid function. We also performed the intersection over unions with non-maximum suspension for threshold operations. In Figure 6, we have represented breast mental illness with the red line achieving an FROC score of 94.00% which is the highest, the lung mental illness is represented by the blue line achieving an FROC score of 93.87%, the brain mental illness is represented by magenta line achieving an FROC score of 92.34% and the bladder mental illness is represented by green line

achieving an FROC score of 92.00%. Our proposed model has achieved the following results using a lesser number of parameters.

Figure 6. Model Accuracy



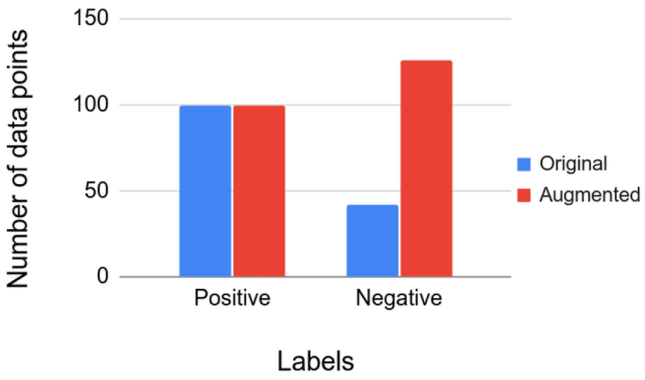
We can compute the ability of the model for the classification of images using the classification accuracy on the testing data set. We ascertain the accuracy of the model by computing the quantity of images effectively grouped or the true positives separated by the complete check of the images. The Random guessing of accuracy for two classes under clinical image classification as healthy and depression, sleep disorder, alzheimer's, schizophrenia, and epilepsy would be around 50% and around 20% under mental illness type classification. These accuracies are considered as baseline accuracies using which we are going to perform comparative analysis with over model yield accuracy. For understanding the accurate classification of images by the network architecture, precision and recall can be used. Precision values determined for each class shows the irrelevant part of clinical images that were expected to be of that class that was most likely in the class. Precision values determined for each class shows the little segment of clinical images from each class that was precisely expected to be in the class. We used the features from the maximum pooling layer and calculated precision, accuracy, F1-score, and recall values for the classes as shown in Table 1 and Table 2. The confusion matrix as shown in Figure 7 can visualize the small part of radiographs/pictures from each class in the test set that were assigned to each of the anticipated marks.

Table 2. Classification Evaluation Metrics for Mental Illness Detection

Folds	Performance Metrics			
	Specificity	Sensitivity	Accuracy	Precision
Fold-I	98.54	97.69	98.32	98.71
Fold-II	97.56	97.74	97.30	97.45
Fold-III	97.31	97.56	97.34	97.64
Fold-IV	97.07	97.31	97.54	97.21
Fold-IV	97.24	97.52	97.59	97.21
Healthy	96.54	96.65	96.25	96.52
Mental Distress	95.82	95.09	95.71	95.21
Overlapped	NULL	NULL	NULL	NULL
Average	97.64	97.50	97.57	97.45

The probability of the model to get overfit and difficulty in localization of lesions exists due to the small size of nodules and the limited amount of clinical images. It is a genuine issue to solicit to provide clinical images and four related numbers to a model and anticipate that it should confirm that these numbers restrict a segment of the picture that is around 10 x 10 pixels and we use argumentation technique for about 150 training nodule images. With not many clinical images, it is simple for the model to fit commotion to diminish the distances between bounding boxes. We also increased the regularization which reduced the overfitting in the model, but the model was unable to detect the generalizing aspects of clinical images for decreasing error in the overlapped bounding boxes from the nodules. The bright spots in the saliency maps and bounding boxes on the surface show that the nodule data was generalized in the hidden data from the network, but for better performance, we require more amount of training images.

Figure 7. Data Argumentation



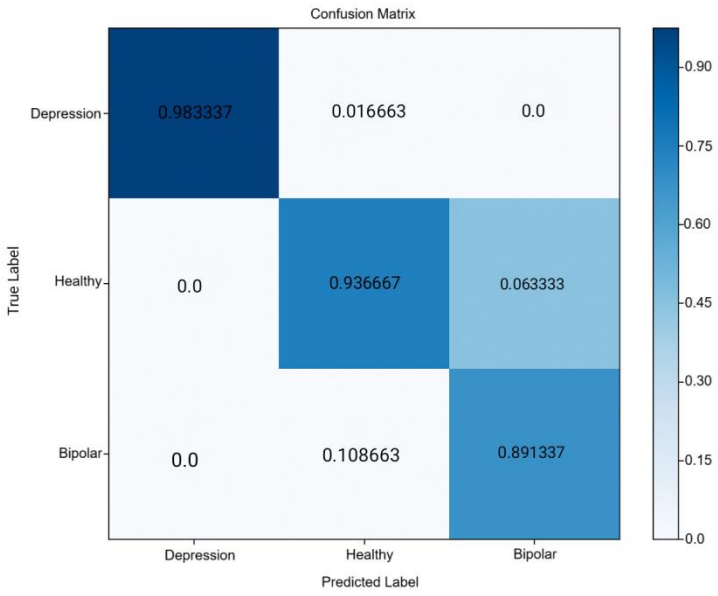
This isn't extremely astonishing since the nodules in these clinical images might be just about three pixels, so recognizing contrasts in the surface and structure of the nodules will be a tedious task for the model. It is much difficult for a radiologist to distinguish a radiograph or image as depression, sleep disorder, alzheimer's, schizophrenia, and epilepsy or healthy, and if an individual is associated to have any type of nodule in their radiograph, they are for the most part sent for additional testing utilizing a CT check or potentially assessment of tissue tests. It is not even clear from the radiographs or images whether the nodules are healthy or depression, sleep disorder, alzheimer's, schizophrenia, and epilepsy since the radiographs or images were marked as true labels based on further examination and hence we failed to know whether the radiograph contains vital information or not to make the distinction. On the off chance that we center around basically distinguishing nodules, we can consolidate depression, sleep disorder, alzheimer's, schizophrenia, and epilepsy and healthy pictures into an individual class. Doing as such, we get accuracy and precision above 95% for clinical images. For clinical analytic, it is often considered to have higher sensitivity and recall for the positive or mental illness class to the potential detriment of specificity, which estimates the review of the negative class, so as not to miss giving further testing and treatment to a patient who needs it. Here, we have a model sensitivity of 95% and a rate of false-positive as 5%. Furthermore, we can also study saliency maps to study the contribution of each part from the radiograph for better precision.

Table 3. Performance Metrics for Detection the Type of mental illness

Mental Illness	Accuracy	Precision	Recall
Depression	94.00	94.54	94.36
Sleep Disorder	93.64	92.46	93.77
Schizophrenia	93.21	93.96	92.12
Epilepsy	92.90	93.49	92.48
Alzheimer	92.05	92.16	93.83

According to our observations, there is no uniquely common visual attribute for non-nodule images or rather defined as the absence of visual attributes. Hence, the brightness is spread throughout the saliency maps for non-nodule radiographs. The saliency map contains more bright pixel spots at the center of the images for nodule images. In the radiograph, the model mostly focuses on the inner closer edge pixels of the spine ignoring the outer edge. However, the saliency maps are not able to focus on the exact nodule points and are not completely interpretable. During localization, we observe the model's strength to learn and detect nodule locations when the explicit true locations are provided while training the model. In the regression head, we fine-tuned hyper-parameters for the final fully connected convolutional block and we found that the model over-fits much easier on the training dataset.

Figure 8. Performance Matrix for Classification of mental illness Lesions



The Euclidean test data and Euclidean train data dropped significantly at first, later the testing losses started increasing and training losses dropped to zero. At the end of the training stage, the image size was reduced by 4 pixels, which was approximately 6mm or thirty percent of the size concerning average nodules present in the dataset. While in the testing set, the distance got differ from sixty pixels. Though this is a long way from covering boxes, it is likewise far superior to random sampling. Selecting two box communities consistently at arbitrary from the 224 x 224 picture, we would anticipate a mean distance

of 117 pixels, or almost twice as extraordinary as that found in the test set. This again demonstrates that the model can decide the overall locale of the nodule however can't determine its accurate area.

4. Conclusion

[1] This research suggests we used a machine learning-based methodology to develop a universal approach for diagnosis, detection, symptoms-based prediction, and screening of histopathological mental illness, their types, and subtypes using a heterogeneous dataset based on images and scans. In this architecture, we use VGG-19 based 3D-Convolutional Neural Network for deep feature extraction and later perform regression using a random forest algorithm. We create a heterogeneous dataset consisting of results from laboratory tests, imaging tests, and biopsy reports, not only relying on clinical images. We use different pre-processing techniques for images and scans to make the dataset according to the system compatibility. Initially, we categorize EEG signals and lesions as healthy or depression, sleep disorder, alzheimer's, schizophrenia, and epilepsy and classify the depression, sleep disorder, alzheimer's, schizophrenia, and epilepsy lesions into their subtypes, detecting their severity and growth rate. We designed an architecture for predicting risks at various time-points, leveraging probable factors if they are available and consistently predict across mammography machines. We found the classification accuracy for categorizing EEG signals as mental illness to be 95% whereas the accuracy for classification of depression, sleep disorder, alzheimer's, schizophrenia, and epilepsy lesions into their subtypes to be 94%. The classification results for breast mental illness are 94%, brain mental illness is 93.64%, lung mental illness is 93.21%, mental illness is 92.90%, kidney mental illness is 92.05%. We built an algorithm using a heterogeneous dataset that can classify lesions as mental illness or healthy with their growing percentage as accurately as a board-certified dermatologist, predicting if and when it will develop mental illness equitably. We can measure the development of EEG signals undergoing treatment, detect new metastases that might be overlooked and predict the tumor development according to their prognosis.

References

1. Ahmed Fares, Sheng-hua Zhong, Jianmin Jiang “EEG-based image classification via a region-level stacked bi-directional deep learning framework” IEEE International Conference on Bioinformatics and Biomedicine (2018) Madrid, Spain. 3-6 December 2018.
2. Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, Tonio Ball “Deep learning with convolutional neural networks for EEG decoding and visualization” The Authors Human Brain Mapping Published by Wiley Periodicals, Inc. 07 August 2017.
3. Yannick Roy, Hubert Banville, Isabela Albuquerque, Alexandre Gramfort, Tiago H Falk and Jocelyn Faubert “Deep learning-based electroencephalography analysis: a systematic review” Journal of Neural Engineering, Volume 16, Number 5 Published 14 August 2019
4. Muhammad Adeel Asgha, Muhammad Jamil Khan, Muhammad Rizwan, Raja Majid Mehmood and Sun-Hee Kim “An Innovative Multi-Model Neural Network Approach for Feature Selection in Emotion Recognition Using Deep Feature Clustering” Sensors (Basel). 2020 Jul; 20(13): 3765. Published online 2020 Jul 5. doi: 10.3390/s2013376

5. Jordan J. Bird, Diego R. Faria, Luis J. Manso, Anikó Ekárt, and Christopher D. Buckingham. "A Deep Evolutionary Approach to Bioinspired Classifier Optimisation for Brain-Machine Interaction" Research Article | Open Access.Volume 2019 |Article ID 4316548
6. Jason Brownlee "Ensemble Learning Methods for Deep Learning Neural Networks" Machine Learning Mastery December 19, 2018 in Deep Learning Performance
7. Raheel Zafar ,Sarat C. Dass ,Aamir Saeed Malik. "Electroencephalogram-based decoding cognitive states using convolutional neural network and likelihood ratio based score fusion" Published: May 30, 2017 https://doi.org/10.1371/journal.pone.0178410
8. Amirali Vahid, Moritz Mückschel, Sebastian Stober, Ann-Kathrin Stock & Christian Beste "Applying deep learning to single-trial EEG data provides evidence for complementary theories on action control" Communications Biology volume 3, Article number: 112 (2020) Published: 09 March 2020
9. Nir Friedman, Tomer Fekete, Kobi Gal and Oren Shriki "EEG-Based Prediction of Cognitive Load in Intelligence Tests" Front. Hum. Neurosci., 11 June 2019 | https://doi.org/10.3389/fnhum.2019.00191
10. Pranav Rajpurkar, Jingbo Yang, Nathan Dass. "Evaluation of a Machine Learning Model Based on Pretreatment Symptoms and Electroencephalographic Features to Predict Outcomes of Antidepressant Treatment in Adults With DepressionA Prespecified Secondary Analysis of a Randomized Clinical Trial" JAMA Netw Open. 2020;3(6):e206653. doi:10.1001/jamanetworkopen.2020.6653 Original Investigation |Psychiatry|June 22, 2020.
11. Zhang, B., Wei, D., Yan, G. et al. Spatial–Temporal EEG Fusion Based on Neural Network for Major Depressive Disorder Detection. Interdiscip Sci Comput Life Sci (2023). https://doi.org/10.1007/s12539-023-00567-x
12. M. Li, Y. Liu, Y. Liu, C. Pu, R. Yin, Z. Zeng, L. Deng, and X. Wang, "Resting-state EEG-based convolutional neural network for the diagnosis of depression and its severity," Frontiers in Physiology, vol. 13, 2022, Article 956254.
13. Wang B, Kang Y, Huo D, Feng G, Zhang J, Li J. EEG diagnosis of depression based on multi-channel data fusion and clipping augmentation and convolutional neural network. Front Physiol. 2022 Oct 20; 13:1029298. doi: 10.3389/fphys.2022.1029298. PMID: 36338469; PMCID: PMC9632488.
14. Zhang, B., Wei, D., Yan, G. et al. Spatial–Temporal EEG Fusion Based on Neural Network for Major Depressive Disorder Detection. Interdiscip Sci Comput Life Sci (2023). https://doi.org/10.1007/s12539-023-00567-x
15. Góngora Alonso S, Marques G, Agarwal D, De la Torre Díez I, Franco-Martín M. Comparison of Machine Learning Algorithms in the Prediction of Hospitalized Patients with Schizophrenia. Sensors (Basel). 2022 Mar 25;22(7):2517. doi: 10.3390/s22072517. PMID: 35408133; PMCID: PMC9003328.
16. Asghar MA, Khan MJ, Rizwan M, Mehmood RM, Kim SH. An Innovative Multi-Model Neural Network Approach for Feature Selection in Emotion Recognition Using Deep Feature Clustering. Sensors (Basel). 2020 Jul 5;20(13):3765. doi: 10.3390/s20133765. PMID: 32635609; PMCID: PMC7374326.

17. Bird, Jordan & Faria, Diego & Manso, Luis & Ekart, Aniko & Buckingham, Christopher. (2019). A Deep Evolutionary Approach to Bioinspired Classifier Optimisation for Brain-Machine Interaction. Complexity. 2019. 10.1155/2019/4316548.
18. Zafar R, Dass SC, Malik AS. Electroencephalogram-based decoding cognitive states using convolutional neural network and likelihood ratio based score fusion. PLoS One. 2017 May 30;12(5):e0178410. doi: 10.1371/journal.pone.0178410. PMID: 28558002; PMCID: PMC5448783.
19. Vahid, Amirali & Mückschel, Moritz & Stober, Sebastian & Stock, Ann-Kathrin & Beste, Christian. (2020). Applying deep learning to single-trial EEG data provides evidence for complementary theories on action control. Communications Biology. 10.1038/s42003-020-0846-z.
20. Friedman, Nir & Fekete, Tomer & Gal, Kobi & Shriki, Oren. (2019). EEG-Based Prediction of Cognitive Load in Intelligence Tests. Frontiers in Human Neuroscience. 13. 10.3389/fnhum.2019.00191.