nCD and Clipped RBM Based Multimode DBN for Optimal Classification of Heterogeneous Images in Big Data

[1]Neha Ahlawat and [2]Franklin Vinod D
[1, 2]Department of Computer Science and Engineering, Faculty of Engineering and Technology,
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India.
[1]nehablow@gmail.com, [2]datafranklin@gmail.com

Correspondence should be addressed to Franklin Vinod D : datafranklin@gmail.com

Article Info
Journal of Machine and Computing (https://anapub.co.ke/journals/jmc/jmc.html)
Doi : https://doi.org/10.53759/7669/jmc202505054
Received 25 April 2024; Revised from 18 September 2024; Accepted 20 January 2025.
Available online 05 April 2025.

Abstract – The scientific community has shown a keen interest in the application of big data analytics in the healthcare industry. The management of healthcare records is extremely challenging not just because of the sheer volume of these records but also due to the multifaceted nature of the data sets and the high dimensions of these records. Recently, it has been demonstrated that deep learning models are extremely powerful generative models that are able to progressively separate features and provide great predictive execution. When it comes to medical image processing, traditional algorithms have been established for a particular modality and a specific condition. Because of the high memory and processing needs of each neural network, it is challenging to build a system that utilises a large number of neural networks and a wide range of specialised image-processing algorithms. This study proposed a C-RBM and nCD-based multimode DBN method. In the first stage, we use nCD (neutral Contrastive Divergence) to train unimodal CRBM pathways, and in the second stage, we build a multimode DBN architecture using only the shared representations of the two pathways. This method fundamentally comprises two-stage learning techniques. A computerised method for the classification of breast and brain cancer was represented by the multimode technique that was mentioned above. When it comes to accuracy, the recommended methodological setups perform better than the alternatives that are considered to be state-of-the-art.

Keywords – RBM (Restricted Boltzmann Machine, Deep Learning, Big Data, Deep Belief Network (DBN)), Computed Tomography (CT), Transfer Learning.

I. INTRODUCTION

Big data is becoming a vital component of all sectors due to the development of contemporary technology. Currently, it is a tedious endeavour to extract information and useful data from a variety of data sources. An increase in data storage capacity, a surge in computing power, and more accessibility of data have collectively encouraged the growth of big data. Volume, Volatility, Variety, Veracity, Validity, and Velocity are the six core problems that are being addressed by most of the solutions that are now being employed to solve Big Data concerns. Volume and Variety are the two Vs of Big Data features that Deep Learning (DL) primarily works with. This indicates that (DL) are well suited for analyzing and deriving information from massive volumes of data as well as data gathered from various sources [1]. Unstructured data having a variety of unknown tempos falls under the topic of heterogeneous data. The massive amounts of data created make it impossible to manage, store, handle, understand, and analyze using conventional methods. Current data analysis tools are having some limitations to handle enormous data. DL (Deep Learning) is best suitable for big data because it can extract crucial information from complex and diverse data. DL can help us to discover previously unattainable and practical patterns in Big Data. The sheer volume of medical data has risen considerably, and the detection of numerous ailments has become more difficult. In context with this, deep learning has demonstrated tremendous success in diagnosing or predicting several serious or terminal conditions. In this study, we mainly focused on deep learning methods that are vital for the early detection of breast cancer and brain cancer, two of the most prevalent cancers in the world [2].

Cancer occurs when cells in the body proliferate out of control and invade nearby tissues. One in every six fatalities is caused by cancer. In a healthy organism, the production of new cells serves to replace ageing tissues. Sometimes, the process can go wrong. While new cells grow even when we might not need them, old cells would not perish when they

should. These excessive cells can clump together to create a tumour. The most often examined malignant tumour in women is probably breast cancer (BC). One out of every three women who are afflicted will die from this disease, which is a higher mortality rate than other cancers [3]. By identifying high-risk individuals early on and treating them properly, the death rate from breast cancer can be progressively lowered. It's a disorder where cells around the breast grow uncontrollably, typically in the milk-producing ducts called lobules. Common symptoms include breast tenderness, edoema, tissue thickening, bloody nipple discharge, and others. At the early stages, there are no outward signs of breast cancer. As soon as symptoms appear, it means the patient's breast cancer has progressed to an advanced stage. Because of this, it is critical to closely monitor symptoms and report them right once to prevent any surprises. Recent years have seen a decline in breast cancer-related deaths due to the successful use of AI advancements, which can predict the presence of cancer cells in a patient's body long before they show any symptoms [4]. Throughout their lives, one in nine women will be affected by breast cancer. Breast cancer is the second most common type of cancer, according to study, however exact statistics on the disease's incidence are unavailable. The fear of dying and the trauma of having a mastectomy make breast cancer one of the illnesses that appears to have a significant psychological impact on sufferers [5]. It also contains lymphatic tissue, which is part of the immune system and is responsible for draining waste products and cellular fluids [6]. Many distinct types of breast cancer can develop in various areas of the breast. For instance, fibrocystic change is a non-cancerous illness that affects women and causes them to develop cysts, fibrosis, lumpiness, and areas of stiffness, soreness, or breast pain. It's also characterised by the production of scar-like connective tissue [7].

Histopathology images are regarded as the highest quality among existing methods for enhancing the examination's precision in patients who have undergone sufficient mammography and other examinations. In addition, histopathological evaluation can provide more precise and comprehensive data for assessing the effects of cancer on neighbouring tissues and conducting an analysis of the disease itself. To overcome this obstacle, an increasing number of studies are examining histopathological images with DL techniques to increase the accuracy of malignancy detection. However, pathologists find it exceedingly difficult to separate massive patterns of harmless areas from harmful (malignant) regions. The display of a model in directed (supervised) learning depends on the amount of marked (labelled) data [8].

After breast cancer, brain tumours are the second most deadly form of the disease. Brain tumour patients often have a wide range of symptoms, including but not limited to: disorientation, memory loss, behavioural abnormalities, seizures, and more. The intricate anatomy of the brain makes accurate tumour identification challenging. Tumours of the brain and other malignant diseases are characterised by the unanticipated proliferation of cells throughout the body. Problematic cell growth in the human brain has the potential to spread to other parts of the body and disrupt the normal functioning of important organs. The two main categories of brain tumours are primary tumours, which are considered benign or noncancerous, and secondary tumours, which are considered metastatic. Glioma is a common kind of cancer that, when left untreated, can cause a rapid decline in quality of life. Depending on their sort, location, and form, the symptoms could be different. The rapid infiltration of various brain tissues by this sickness exacerbates a patient's condition [9].

To accurately classify brain images, it is necessary to segment them into cancer and non-tumor areas. This makes research into brain image segmentation and classification an important area of study. Segmentation is essential for the extraction of significant properties, which is the first step in correct categorization [10]. The information-rich and spatially superior characteristics of magnetic resonance imaging (MRI) make it an important tool in medical imaging for the detection of brain tumours. An important part in making accurate diagnoses of brain disorders is the automated imaging technologies. The data acquired by magnetic resonance imaging (MRI) is very real and exact as it does not involve radiation, which is a major advantage. Unfortunately, the tumor's size and shape render identification useless. The magnetic resonance imaging (MRI) photographs used to detect brain tumours include four slices: T1, T2, T1 contrast, and FLAIR images. Consequently, MRI generates a plethora of information on the brain structures that can be used to detect a brain tumour [11].

Recent years have seen a flurry of activity in the field of medical image processing, which has spurred ground-breaking developments in related technologies and tools. Medical professionals use a wide variety of diagnostic tools to learn about an organ's structure, function, and disease. These tools include X-rays, ultrasound, mammography, MRI, positron emission tomography (PET) etc. Each of the mentioned modalities presents data on human organs from a unique perspective. Medical professionals use the term "multimodal medical image" to describe a composite picture that incorporates data from several different types of imaging scans (e.g., PET, SPECT, CT, MRI, etc.) to reveal more detailed information about the patient's anatomy and spectral properties than would be possible with any one scan alone. The major goal of this method is to improve the image's quality while keeping the best and most relevant parts of each element. In turn, this makes the medical pictures more useful for diagnosis in the clinic. Under the broader concept of "multimodal imaging," several images are combined using generic information fusion techniques to address medical issues posed by modern medical imaging technology. 'Multimodal medical picture' and other new medical imaging technologies are changing the way we record medical history, make diagnoses, and conduct investigations. Computer advancements and state-of-the-art imaging techniques have made it possible to objectively assess medical imaging images. As a result, clinicians are better able to assess patients, reach valid conclusions, and decide on the most appropriate treatment faster [12].

This paper is set up as follows: sections 1 and 2 impart the theoretical background or introduction to the paper. Section 3 focuses on the inspiration for the theoretical features of the existing tumour detection investigation. Section 4 discuss

about the modifications made to the existing methodology and detailed the proposed system. Section 5 discusses the evaluation of the model. Section 6 is devoted to results and discourse, and section 7 is the conclusion.

II. RELATED RESEARCH

To familiarize ourselves with the fundamental concepts and ideas, we first review previous research on multiple feature selection, integration, and classification. This will set the stage for our heterogeneous feature selection system.

Recent years have seen the development of several new learning frameworks [13] that aim to maximize the use of multi-modal data. For example, several techniques based on kernel learning have been suggested [14,15]. By combining several physiological data, multimodal fusion can produce results for emotion recognition [16]. Multimodal fusion emotion recognition research has recently emerged, with some researchers focusing on the unique ways in which each modality expresses emotions. Accurate identification results and excellent robustness are guaranteed by data-driven multiplication of several combination modes. The multi-modal emotion detection technique, which leverages deep autoencoders to capture brain electrical interaction expressions, was introduced in reference [17] as an advanced solution aimed at addressing the performance shortcomings of several existing methods in emotion recognition. First, we use the decision tree to identify which traits are most crucial. The test sample is subsequently classified into one of the face expression categories using the solution vector coefficients, which are derived from the sparse representation of facial expression characteristics. Next, the bimodal deep autoencoder is utilized to combine the data from the electroencephalogram and the facial expressions. Accurate emotion identification is made possible by this. Processing big sample data will incur increased recognition costs because of feature fusion in deep learning models [18].

A technique based on neighborhood rough sets was suggested by Hu et al. [19] to choose heterogeneous feature subsets. All these approaches work on the premise that a common feature space, where each feature is represented by a feature vector, may be used to find associations between the original, separate features. As a result, developing a more solid framework for picking heterogeneous traits is crucial.

Bolei Xu et al. put up a novel breed strategy for BC categorization in [20]. Using a hard visual consideration (attention) method, it efficiently selects a series of raw locations from the raw image; subsequently, it may use a different mechanism (soft-attention) to investigate the uncommon aspects of each location. Afterwards, a recurrent network is trained to decide how to characterise the image segment and also to predict which area (locale) of the segment would be examined in the subsequent time step. With its non-differentiability, the area choice connection uses reinforcement to smooth out the network, making it easier to handle an optimum method for area classification.

A study by Liang et al. [21] employed an attention mechanism to categorise breast cancer pictures using a CNN technique. Photos taken by Pcam seemed to have been artificially stained with hematoxylin (H) and eoxin (E). However, shadowing is meaningless, and different images had different colours since the incisions are hand-stained by administrators. Hence, we normalised each picture separately by dividing by its standard deviation and deleting the average of all pixels. To address the issue of tumours having too few pixels, rearrange the pixels such that two are aligned in the middle of each pixel and resize the pictures from 96x96px to 288x288px.

To attain sparsity inside and between groups, the group lasso model proposed by Friedman et al. [22] use an L2-norm regularisation. In their study, Syahmi et al. proposed comparing breast illness identification using two pre-existing DL network models [23]. Preprocessing images, classifying them, and then evaluating their performance are the standard steps in the technique. In order to determine if a DL model network is present, two models were utilised: ResNet50 and VGG16. These models were trained to distinguish between typical tumours and those that were considered odd using the IRMA dataset. Since the Adam enhancer was utilised and there was no completely linked layer, no regularisation technique was employed. Out of all the recommended models except Inception-ResNet-V2 and ResNeXt-101(32 × 4d), information augmentation techniques such as rotation, flat, and vertical flipping might affect the results. When comparing the two models' accuracy levels, we find that VGG16 reaches 94% and ResNet50 only manages 91.7%.

To identify the most significant predictors in an integrated genomes investigation, Peng et al. [24] utilised a similar approach. A dual-purpose was proposed by Shahidi et al. [25]. The primary objective is to investigate various learning methods for breast cancer histopathology image characterization. The best accurate models for two-, four-, and eight-order breast cancer histopathology image databases are recognised in this review. We also drew ideas for our synthesis from contemporary models that have received scant or no attention in earlier research. Newly updated ImageNet informative index findings have been associated with models like NASNet, Dual-Path Net, and ResNeXt SENet. The two-overlap and eight setups were tested on BreakHis using these models.When using sparse group lasso, characteristics are removed from both the deleted and remaining groups.For the computer vision challenge of gesture identification in depth map sequences, Azad.R et al.[26] suggested a motion energy-based multilevel temporal sampling (MTS) method. Results on three publicly accessible datasets show that the suggested strategy is quite accurate, and testing has been conducted on a large scale.

H. Sultan *et al.* [27] developed a model that uses two publicly available datasets from General Hospital, TMU, and Nan Fang, correspondingly, to diagnose distinct brain lesions using a convolutional neural network. The first information rested on tumour growths of various varieties. Another separates the three glioma grades into various Grading divisions. For such two data sets, their projected engineering utilized a maximum perfection of 96.13% and 98.7%. This simulation model is formed by multiple layers, commencing with the information layer that contains the previously treated images and progressing via several convolution layers.

In [28], Kamanasish Bhattacharjee *et al.* investigated the Classification method of sub-atomic brain tumours, which is completed by designing a multi-facet perceptron, was evaluated (MLP). Particle Swarm Optimization and Hybrid Genetic Algorithm (HGA) are used to prepare the MLP (multi-facet perceptron) (PSO). The six benchmark datasets are then effectively grouped using the constructed MLP. The Sigmoid capacity dataset underwent grouping by GA with 100% accuracy, much as the XOR. Nevertheless, when datasets get more complex, such as the iris and breast cancer datasets, the accuracy decreases to 90% when using PPSOGA2 and 90% when using the iris dataset.

In this paper [29], Mohamed arbane *et al.* devised a convolutional neural network (CNN) based transfer perceptual approach for predicting the order of cortical tumours from MRI images. The deployed research analyzes alternative CNN architectures, emphasizing on Res.Net, Xception, and Mobil-Net-V2. To develop, support, and analyse the intended DL simulation, the dataset employed in this study is categorized into three parts. Eighty per cent of the entire sample is contained in the essential subgroup, which has been altered to match the simulation. The excess is allocated accordingly for certifying and putting the structure through its trials. This framework yielded the highest results, scoring 98.25% for accuracy and 98.43% for F1-score, correspondingly.

Hasan Ucuzal et al. expounded upon the notion of electronic programming capable of diagnosing brain tumours (glioma, pituitary, meningioma) through the utilisation of convolutional-neural organisation derived from deep learning calculations and high-precision T-1 difference enticing reverberation images. All established metrics for classifying the diverse array of cerebrum malignancies on the designated preparation dataset exceed 98%, as indicated by the performance outcomes. In total, 3061 MR image filters are available to detect growths such as pituitary cerebrum, meningioma, and glioma. 466 of these MR images are utilised during the inspection phase, while 2590 are utilised during the preparation phase [30].

III. BACKGROUND

Introduction of Basic DBN

An essential unassisted DL model is a Deep Belief Network (DBN), which consists of equipped restricted Boltzmann machines (RBMs). It is a vital autonomous DL model, and its visualization is essential to the creation of DBN. RBM is an instance of recurrent probabilistic Neural Network (NN) that can learn a probability distribution from its inputs.

The Binary[2] RBM has two-fold parameters as apparent (visible) and concealed (hidden) [14], which is the most prominent type of model. It consists of p apparent parameter $v = (v_{.1}, \dots\dots, v_{.p})$ and q concealed parameter $h = (h_{.1}, \dots\dots, h_{.q})$. The two sporadic variables (v, h) assume the values $(v, h) \in \{0,1\}^{p+q}$ in the current RBM for display the two-fold information of layers. Equation (1) gives the Energy function of EBM as follows:

$$E(\mathrm{v}, \mathrm{h}) = -\sum_{i.=1}^{p} b_{i.} v_i - \sum_{j=1}^{q} c_j . h_j - \sum_{i=1}^{p} \sum_{j=1}^{q} v_{i.} W_{ij} h_j \quad (1)$$

where c_j stands for the j_{th}bias term associated with the j_{th} concealed parameter h_j and b_i stands for the i_{th} bias term associated with the i_{th} observable parameter v_i.W_{ij} also denotes the weight associated with the j_{th} concealed parameter h_j and the i_{th} visible parameter v_i.

Nodes that comprise the same layer have no links in an RBM. As a result, a node within a visible/hidden layer is not linked to another node of the same layer. From the perspective of possibility or likelihood, this constraint is beneficial since the concealed variables are independent of the state of the visible units and vice versa, which are illustrated as in equation (2) as well as equation (3).

$$P\left(h_{j.} = \frac{1}{v}\right) = sd.\left(-\ln\left(\frac{1}{(b_{j.} + \sum_{i \in v} v_{i.} W_{ij})} - 1\right)\right) \quad (2)$$

$$P\left(v_{i.} = \frac{1}{h}\right) = sd.\left(-\ln\left(\frac{1}{(c_{i.} + \sum_{j \in h} h_{j.} W_{ij})} - 1\right)\right) \quad (3)$$

where, $(sd.)$= sigmoid activation function (σ) = $\frac{1}{1+\exp.(-x)}$, W_{ij}: the weight matrix connecting the visible units to the hidden units.

Deep Belief Network (DBN) is a potent computational model which utilizes an advanced structure made up of various stacks of Restricted Boltzmann machines (RBM). The overall DBN addition is equivalent to the primary (first) layer RBM's input and its output is comparable to the last layer RBM's output. The output of the earlier RBM gives the contribution of each RBM after the first layer. As stated in **Fig 1**, DBN is a deep design that relies on the fine-tuning or adjusting phase, which optimizes the weights by reducing the entropy erroneous.

The goal of the RBM's learning process is to maximize the log-likelihood such that the dispersion distribution it learns is as like the distribution of the input data. A vanishing gradient issue was met by the subordinate for the sigmoid stimulation function, which caused erratic behavior throughout RBM training. Deep neural networks (DNN) commonly encounter issues related to vanishing and exploding gradients, which can significantly hinder the training process and affect

model performance. The proposed methodology mentioned in section 4 covers the solution to overcome the gradient problem in detail.

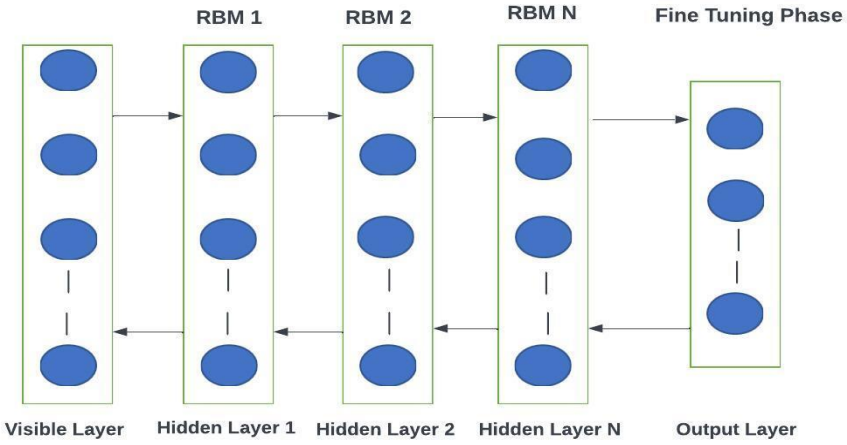


Fig 1. DBN Architecture.

Contrastive Divergence Algorithm (CD)

Contrastive Divergence (CD) is an efficient learning algorithm for probabilistic graphical models that uses Gibbs sampling to approximate the log-likelihood gradient, minimizing the difference between observed data and the model's distribution to update parameters effectively.CD uses Markov-Chain Monte Carlo (MCMC.) to optimize the difference between observed data and samples from the current model distribution. In the context of training Restricted Boltzmann Machines (RBMs), the key focus lies in understanding the data type and its statistical properties, and the tuning of RBM hyperparameters. All typical RBM training strategies use statistics to approximate the log-probability gradient and then apply gradient descent to those approximations. The recent research has revealed that RBMs may be successfully educated via CD (Contrastive Divergence). However, certain areas, such as weight regularisation, require further refinement provides a solution to the provided problem.

Spark Framework

Correct data analysis will lower expenses and raise healthcare quality. With conventional hardware and software platforms, analysis is rendered impossible. It is crucial to pick the best platform for managing this sort of data. Apache Spark is one of the most potent frameworks for distributed computing and is employed in the big data environment. Spark provides a unified framework for processing diverse datasets—such as graph, image/video, and text data—across various sources, including batch and real-time streaming, efficiently addressing the complexities of big data management. Distributing DL's processes is a good idea because it requires a lot of computing, and Apache Spark is one of the simplest ways to do it. There are several approaches to implement DL in Apache Spark, as shown by the following examples: Distributed DL using Keras and PySpark is used by Elephas, Tensor Flow on Spark by Yahoo Inc., Distributed Keras by CERN, and Tutorial Keras with Spark by Qubole. A further open-source tool called DL Pipelines provide high-level APIs for seamlessly integrating deep learning workflows within Apache Spark. In our implementation tensorflow, Pyspark, Keras and Elephas python libraries are used to construct end to end deep learning pipeline.

IV. PROPOSED METHODOLOGY

This section first addresses the clipped approach for firing the neurons and modified contrastive divergence approach. Both approaches are employed to train hidden layers and and optimize the learning process across bimodal pathways specifically for brain and breast cancer. After that, provide an illustration of the construction of Multi-mode DBN and discuss about Spark Framework in detail.

Clipped Restricted Boltzmann Machine (C-RBM)

The clipping RBM concept, in conjunction with ReLU activation, is used to activate the neurons of hidden units to address the gradient problem. When masses gradually disperse, gradients become less noticeable; conversely, when masses explode, gradients become instantly noticeable. The primary idea behind this notion is to establish rules to prevent gradients from exploding.

A clipped ReLU layer thresholds activations by setting values below zero to zero and limiting values exceeding a specified threshold to the clipping. Employing the clipping strategy for train RBM hidden layers can assist to mitigate the problem of exploding gradients. ReLU following the basic rule outlined in Equation (4). Its ability to introduce non-

linearity without requiring complex mathematical operations makes it a preferred choice for many deep learning models, balancing performance and computational efficiency.

$$\textit{Rectified Linear Unit} = f.(y) = max.\ (0, y) \quad \begin{matrix} y;\ for\ y > 0 \\ 0;\ for\ y <= 0 \end{matrix} \tag{4}$$

Gradient clipping, as shown in Equation (5), limits gradients to a specified threshold z(>0). If the gradient exceeds this value, it is scaled to stay within the range:

$$Clipped.\ Unit.\ (y,z) = min.(max.((0, y),z) \tag{5}$$

Neutral Contrastive Divergence (nCD)

An established method for training energy-based inert variable models is the CD methodology, which has been widely used in a variety of deep learning models, such as deep belief networks and limited Boltzmann machines. Standard algorithms are adversely affected by certain challenges. We employ two methods to train RBMs: choosing relevant data types and adjusting hyperparameters. The log-probability gradient is approximated using statistics in all common RBM training procedures, and gradient descent is then applied to those approximations. Important elements of the suggested neutral CD algorithm are the concept of unbiased MCMC, the clipped approach (outlined in section 4) and weight decay regularization to address bias and the vanishing gradient issue.

$$\text{Weight Decay } (Wd1) = \sum_{i.} \theta_i \tag{6}$$

$$\text{Weight Decay } (Wd2.) = \sum_{i.} \theta_i{}^2 \tag{7}$$

We utilised the nCD method with weight decay regularisation, as shown by equations 6 and 7, to balance the sensitivity and complexity of the model. Equation 8 demonstrates how this neutralising behaviour can also be used to enhance the loss regularisation (L_r).

$$L_r = L + \frac{\alpha}{2}.\sum_{i=1}^{p_0} \sum_{j=1}^{p_1} (W_{ij}.m)^2 \tag{8}$$

Multi-Mode DBN (Deep Belief Network)

The DBN employs numerous stacked Restricted Boltzmann Machines in hierarchical framework and proceeds with a layer-by-layer consecutive learning process as already discussed in Background section. When neurons are being trained, traditional DBN encounters problems including the gradient problem, hyper-tuning of parameters, weight regularization etc. The difficulties described above are resolved, and the learning rate is also improved, by employing both CRBM and nCD approach as discussed in section 3. The proposed multimode DBN (mDBN) model is presented here for classifying the diverse cancer images. In this model, the pre-processing stage receives the input depicts initially, including images of breast and brain cancer. The preliminary and most critical stage is employing unique modality with latent variable to get low-level representations of each data modality in a distinct manner. To achieve this, create first-rate generative models for each unique modality using massive amounts of unlabeled data. Unannotated data is readily obtainable across an extensive range of domains, encompassing machine vision, text retrieval, speech perception, and medical image classification. This network uses two bimodal pathways of RBMs, one for brain and one for breast cancer images, to process the data. Beginning with multimodal data inputted into lower-level RBM layers, RBMs with hidden variables at their top-level indicate similar characteristics across modalities in multi-platform data.

Let V_l & V_k denote the visible distributions corresponding to brain and breast-specific images, as specified in Equations (9) and (10):

$$P\left(\frac{V_k}{h(0)}\right) = \frac{P(V_k)}{\sum_{h(0),h(1).} P(h_1,h_0)} \tag{9}$$

$$P\left(\frac{V_l}{h(0)}\right) = \frac{P(V_l)}{\sum_{h(0),h(1).} P(h_1,h_0)} \tag{10}$$

The proposed method employs a two-stage learning process: initially, unsupervised learning generates separate representations for each unimodal RBM pathway **Figs 2 and 3**, followed by fine-tuning to integrate shared representations into a multimodal DBN architecture. The brain-specific **Fig 2** and breast-specific **Fig 3** RBMs utilize clipped and neutral divergence to model real-valued image data distributions.

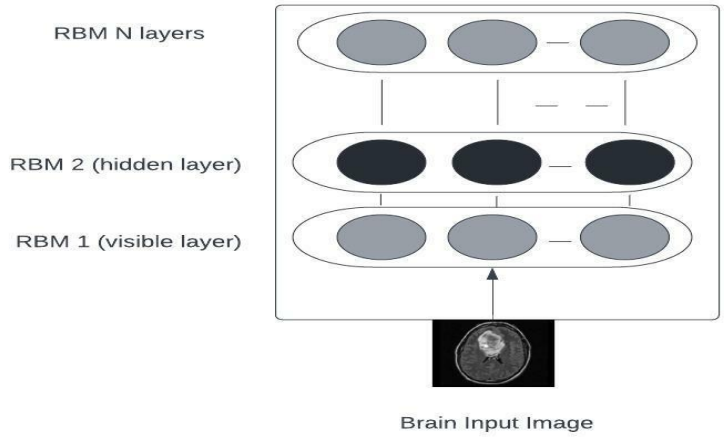


Fig 2. RBM- Brain Pathway.

After the unsupervised phase, the Multimode DBN fine-tuning proceeds as follows: (a) Training the RBM layer with original and scaled data. (b) Using representative data from the CRBM's first layer as input for the second RBM layer. (c) Training subsequent CRBM layers similarly to ensure stable initialization. (d) Fine-tuning the network parameters with a supervised softmax classifier.

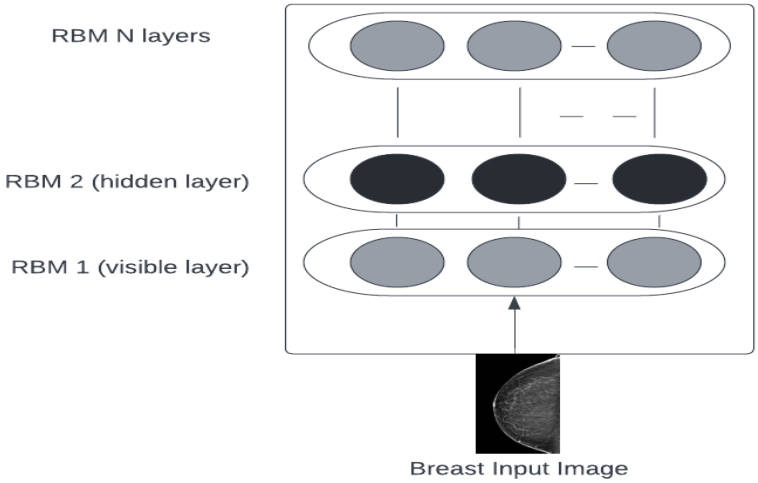


Fig 3. RBM- Breast Pathway.

Fig 4 displays the architecture of multimode-DBN deep learning model for image classification. Training is done using a stack of CRBMs with non-overlapping blocks in the input layer. This process consists of two parts: training and classification. For the brain pathway and the breast pathway, separate RBM structures are established.

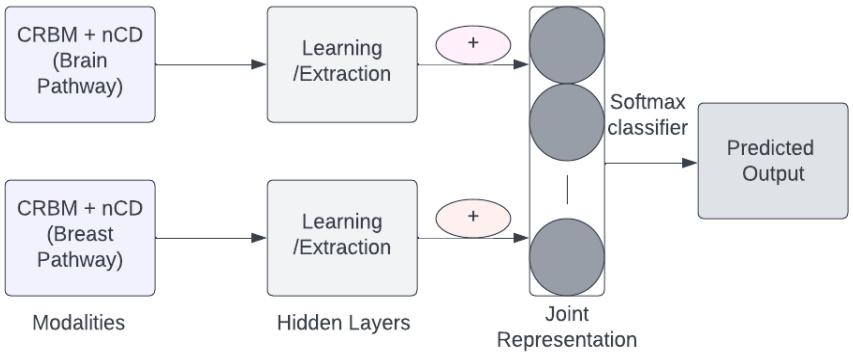


Fig 4. Multimode- DBN.

Each modality (Brain and Breast) requires its own two-layer Deep Belief Network (DBN), which must be constructed separately before the multimode DBN can be assembled. For the Brain and Breast specific, we employ real-valued image inputs v_k and v_l, whose probabilities are calculated by equations (11)

$$P(v_k, v_l, \theta) = W.\sum_{h_{k^2}, h_{l^1}, h_3} P(h_k, h_{k^1}, h_{k^2}) \sum_{h_{k^1}} P(\ v_k, h_{k^0}, h_{k^1}))(\sum_{h_{l^1}} P(\ v_l, h_{l^0}, h_{l^1})) \quad (11)$$

The approach leverages a binary layer of hidden units distributed across multiple undirected graphs, multimodal DBNs capable to learn joint representations from multiple modalities. As undirected models, they allow for mutual influence between the low-level representations of different modalities during collaborative training, enabling the model to capture complex relationships and dependencies across diverse data sources.

Spark Based Multimode DBN

The proposed model for brain and breast cancer prediction is trained using the BRCA and BRATS dataset. The datasets should be checked before pre-processing and discretizing it for cleaning. To eliminate redundant and missing data, as well as to correct inconsistencies, data must be cleaned. The following phase involves choosing the columns to work with, cleaning up missing data, and splitting the data. Our Spark-based architecture is made up of two basic parts: a Spark master and one or more Spark workers. The Spark driver, initialized by the master node, manages the execution of multiple partial models. It coordinates task distribution to Spark workers, which process the tasks in parallel, enabling efficient computation and model training across the cluster. Each worker node trains a partial deep model on an extremely small segment of the data and transmits it to the deep learning algorithm during each iteration **Fig 5**.

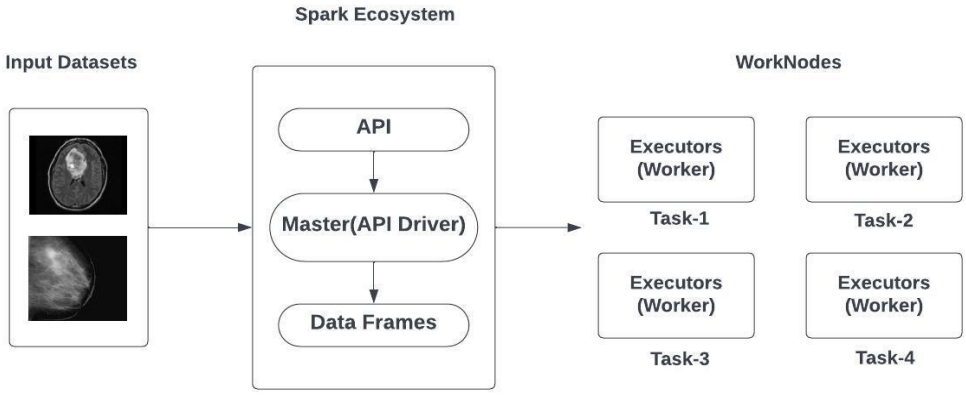


Fig 5. Spark Based Multimode DBN Framework.

The master node initializes the network's parameters, such as the number of neurons in the visible and hidden layers, biases, learning rate, and epoch frequency, for the shared architecture. It then randomly splits the dataset into subsets, broadcasting them to all worker nodes along with the preset parameters. The number of splits is automatically determined based on the training setup. During each epoch, each worker uses Gibbs sampling to compute gradient approximations for their respective data subset. The master obtains a copy of the trained network when the training is over and initials it for a subsequent layer of RBM. The input dataset is altered appropriately, but the training of every other level of RBMs is essentially identical to that of the bottom-level RBM. The given Algorithm describes how to build a spark based multimode DBN.

Algorithm for Spark Based Multimode DBN

Input: Train set Y= {(y_1), (y_2), ..., (y_m)}, e (no. of epoch), N (no. of layers), hyper-parameters
Output: The trained multimode DBN model
1: while N is not equal to 0 do
2: Set the network hyper-parameters at the master node based on the model configuration.
3: Divide the training set's subgroups into multiple sets.
4: Distribute a copy of the master's configuration settings to each worker.
5: Distribute a subset of the dataset to every worker (Executor-er).
6: Check for condition for both i and j.
7.for i= 1 to er do
8: for j= 1 to e do
9: To get the approximate gradients of W, p, and q, use Gibbs sampling.
10: Average the parameters on a regular basis and deliver the findings to the master.

11: end for
12: The trained network should be copied and sent to the master.
13: end for
14: Configure another layer of RBM.
15: Broadcast to all workers the global parameters, which were set as the average values derived from each worker.
16: Return to step-5.
17: End while.
18: The master determines the final learned network by aggregating the global parameters.
The distribution of the training work is carried out using a data parallel technique based on the above-mentioned algorithm. On each worker computer, we keep a copy of the complete model and process various subsets of the training data set.
c over its portion of the split.

V. EXPERIMENTATION AND MODEL EVALUATION

Experimental Setup

The model has been validated by several experimental evaluations. The tests were executed on an Ubuntu-based personal computer that had a 3.40 GHz Core-i7 CPU and 4 GB of RAM. The Python 2.7 platform was used, along with a few programmes for basic data analysis such as TensorFlow, Keras, and Scikit-learn, as well as numpy, scilpy, and matplotlib. Our experimental approach is validated by comparing the suggested image classifier.

Dataset and Image Processing

Data from the TCGA-BRCA and BRATS datasets are used to offer experimental proof that the suggested model is successful. The clinical image data set includes 85 multi-contrast MR imaging scans from glioma patients; 28 of these scans came from patients with mild gliomas (defined as astrocytomas or oligoastrocytomas according to histology) and 41 from patients with more advanced gliomas (defined as anaplastic astrocytomas and glioblastoma multiforme tumours). The Cancer Genome Atlas Breast Cancer (TCGA-BRCA) dataset contains 1,063 breast cancers along with additional histologic type annotations.

Deep Learning Image Processing Pipeline

Step 1. Preprocess images to get one-dimensional vectors.
Step 2. To build the mDBN model, first load the network with typical imaging data, and then train each CRBM independently.
Step 3. The Energy statistic is computed using common image data, and kernel density estimation is then utilised to determine the Energy control limit.
Step 4. Applying bilinear interpolation for depict scaling and Otsu thresholding for grayscale image thresholding.
Step 5. To extract features from test images, gather them and input them into a multimode-DBN.

Parameters Setup

The suggested model relies on well-chosen parameters. Both the Brain RBM and the Breast RBM routes utilized the values shown in **Table 1** for their respective parameters.

Table 1. List The Values for Various Parameters

Parameters:	Brain RBM	Breast RBM
Visible Units:	1200	1200
Hidden Units:	850	850
Maximum no. of Epochs:	120	120
Rate of learning:	0.1	0.1
Model:	Generative	Generative

VI. RESULT

Several experiments were conducted to see how well Multimode-DBN based classifiers work. All the studies were conducted using the Python programming environment. After developing the nCD-CRBM-based multimode-DBN model, we can test its data reproducibility using the nCD-CRBM function. The enhanced performance of the model on the brain and breast cancer datasets is demonstrated in **Table 2**, which illustrates the precision of the dataset across various types of hidden layers using the multimodal-DBN function. In this test, we use several types of hidden layers consisting of 55,105,205,305,405,505,605,705 nodes in two and four layers. During each epoch, the input layer computes a loss function to evaluate how well the hidden layer is performing. Based on this assessment, it signals the necessary adjustments to each node, guiding the optimization process to improve the model's accuracy and reduce the error.

Table 2. Classification Precision by Using the Ncd-CRBM Tool With Several Sorts of Hidden Layers

Number	Iterations	No. of nodes in hidden layers	Layers	Precision	Layers	Precision
1	120	55	2	0.8561	4	0.8650
2	120	105	2	0.8623	4	0.8956
3	120	205	2	0.9008	4	0.9217
4	120	305	2	0.9258	4	0.9179
5	120	405	2	0.9285	4	0.9231
6	120	505	2	0.9516	4	0.9560
7	120	605	2	0.9508	4	0.9791
8	120	705	2	0.9681	4	0.96880

Table 2 shows how the suggested model compares to current traditional models in terms of performance. The four criteria used to measure performance are accuracy, specificity, sensitivity, and F-score. These performance requirements were established by the equations 12, 13, 14, and 15 listed below.

$$Accuracy = (T.P + T.N) / (T.P + T.N + F.P + F.N) \tag{12}$$

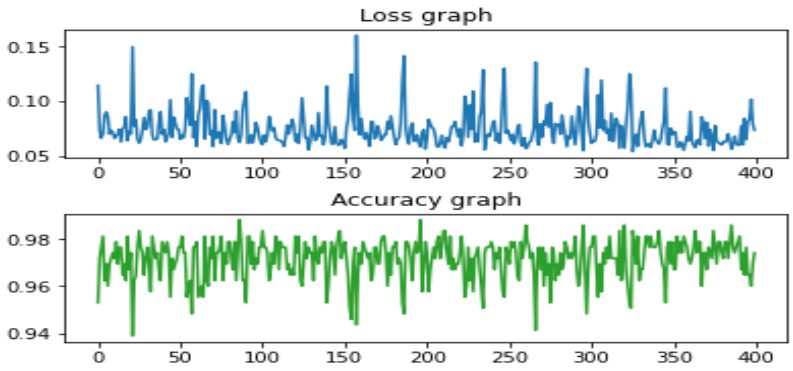
$$Sensitivity = T.P / (F.N + T.P) \tag{13}$$

$$Specificity = T.N / (F.P+T.N) \tag{14}$$

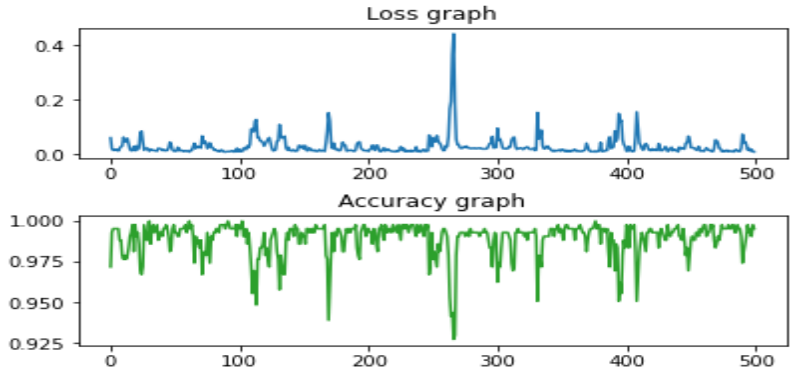
$$F1\text{-}score = 2T.P/(2T.P+F.P+F.N) \tag{15}$$

In equations (12), (13), (14), and (15), T.P, T.N, F.P, and F.N represent True Positives, True Negatives, False Positives, and False Negatives, respectively. These metrics offer comprehensive view of each method's strengths and limitations, helping to evaluate their overall effectiveness and performance in classification tasks.

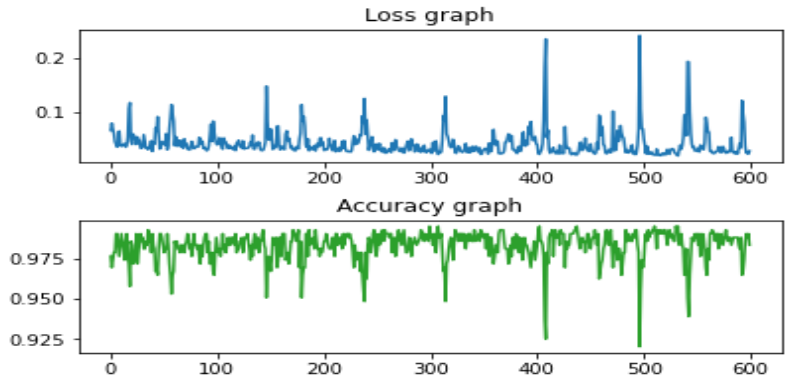
The representation of the average loss along with accuracy curves are shown in **Fig 6** for the multimode classification. **Fig 6 (a, b, and c)** depicts the accuracy and the loss curve of the proposed model over the 400,500 and 600 epochs. It is apparent that the accuracy curve has dramatically risen, and the model's performance significantly increases with increasing epoch denotation. The loss value of the model reduced with each epoch, which indicates that the performance of the model improved. This contrasts with the accuracy curve, which grew continuously.



(a)



(b)



(c)

Fig 6. Accuracy and Loss Graph for Multimode Classification.

VII. CONCLUSION

The primary goal of this research is to create a model that uses histology and MRI images to identify brain and breast tumours. The CRBM and multimode-Deep Belief Network (DBN) method for brain and breast tumour identification has been effectively implemented and applied to pictures. The model offers a joint density model that is specified over the entire space of possible combinations of modalities for missing data. It also performs effectively on activities involving discrimination. Nevertheless, it surpasses unimodal models that are trained on only one modality when there is only one modality available for testing. The primary objective of parameter initialization is to prevent the occurrence of exploding or vanishing gradients in the activation outputs of layers during forward propagation. The network will take a long time to converge if one of these issues occurs, because the loss gradients (inclination) will be excessively big or insufficient. The design we provided effectively yielded improved outcomes and efficiently addressed the difficulties. The multimodal DBN framework is inherently scalable and versatile, making it exceptionally suited for large-scale, complex problems.

CRediT Author Statement

The authors confirm contribution to the paper as follows:
Conceptualization: Neha Ahlawat and Franklin Vinod D; **Methodology:** Franklin Vinod D; **Software:** Neha Ahlawat and Franklin Vinod D; **Data Curation:** Neha Ahlawat and Franklin Vinod D; **Writing- Original Draft Preparation:** Franklin Vinod D; **Visualization:** Neha Ahlawat and Franklin Vinod D; **Investigation:** Franklin Vinod D; **Supervision:** Neha Ahlawat and Franklin Vinod D; **Validation:** Franklin Vinod D; **Writing- Reviewing and Editing:** Neha Ahlawat and Franklin Vinod D; All authors reviewed the results and approved the final version of the manuscript.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding agency is associated with this research.

Competing Interests

There are no competing interests

References

[1]. H. Sun, Z. Liu, G. Wang, W. Lian, and J. Ma, “Intelligent Analysis of Medical Big Data Based on Deep Learning,” IEEE Access, vol. 7, pp. 142022–142037, 2019, doi: 10.1109/access.2019.2942937.

[2]. Z. Chen, F. Zhong, X. Yuan, and Y. Hu, “Framework of integrated big data: A review,” 2016 IEEE International Conference on Big Data Analysis (ICBDA), pp. 1–5, Mar. 2016, doi: 10.1109/icbda.2016.7509815.

[3]. I. Hirra et al., “Breast Cancer Classification From Histopathological Images Using Patch-Based Deep Learning Modeling,” IEEE Access, vol. 9, pp. 24273–24287, 2021, doi: 10.1109/access.2021.3056516.

[4]. Y. Liang, J. Yang, X. Quan, and H. Zhang, “Metastatic Breast Cancer Recognition in Histopathology Images Using Convolutional Neural Network with Attention Mechanism,” 2019 Chinese Automation Congress (CAC), pp. 2922–2926, Nov. 2019, doi: 10.1109/cac48633.2019.8997460.

[5]. S. S. Aboutalib, A. A. Mohamed, W. A. Berg, M. L. Zuley, J. H. Sumkin, and S. Wu, “Deep Learning to Distinguish Recalled but Benign Mammography Images in Breast Cancer Screening,” Clinical Cancer Research, vol. 24, no. 23, pp. 5902–5909, Dec. 2018, doi: 10.1158/1078-0432.ccr-18-1115.

[6]. A. A. Alhussan, N. M. A. Samee, V. F. Ghoneim, and Y. M. Kadah, "Evaluating Deep and Statistical Machine Learning Models in the Classification of Breast Cancer from Digital Mammograms," International Journal of Advanced Computer Science and Applications, vol. 12, no. 10, 2021, doi: 10.14569/ijacsa.2021.0121033.

[7]. J. Peng et al., "Regularized multivariate regression for identifying master predictors with application to integrative genomics study of breast cancer," The Annals of Applied Statistics, vol. 4, no. 1, Mar. 2010, doi: 10.1214/09-aoas271.

[8]. J. Tong, Y. Zhao, P. Zhang, L. Chen, and L. Jiang, "MRI brain tumor segmentation based on texture features and kernel sparse coding," Biomedical Signal Processing and Control, vol. 47, pp. 387–392, Jan. 2019, doi: 10.1016/j.bspc.2018.06.001.

[9]. S. Yazdani, R. Yusof, A. Karimian, M. Pashna, and A. Hematian, "Image Segmentation Methods and Applications in MRI Brain Images," IETE Technical Review, vol. 32, no. 6, pp. 413–427, Jul. 2015, doi: 10.1080/02564602.2015.1027307.

[10]. M. Chen, Q. Yan, and M. Qin, "A segmentation of brain MRI images utilizing intensity and contextual information by Markov random field," Computer Assisted Surgery, vol. 22, no. sup1, pp. 200–211, Oct. 2017, doi: 10.1080/24699322.2017.1389398.

[11]. G. Litjens et al., "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.

[12]. Y. Yang, "Medical Multimedia Big Data Analysis Modeling Based on DBN Algorithm," IEEE Access, vol. 8, pp. 16350–16361, 2020, doi: 10.1109/access.2020.2967075.

[13]. Y. Liu, X. Chen, J. Cheng, and H. Peng, "A medical image fusion method based on convolutional neural networks," 2017 20th International Conference on Information Fusion (Fusion), pp. 1–7, Jul. 2017, doi: 10.23919/icif.2017.8009769.

[14]. W. Xue-jun and M. Ying, "A Medical Image Fusion Algorithm Based on Lifting Wavelet Transform," 2010 International Conference on Artificial Intelligence and Computational Intelligence, pp. 474–476, Oct. 2010, doi: 10.1109/aici.2010.337.

[15]. T. Mittal, U. Bhattacharya, R. Chandra, A. Bera, and D. Manocha, "M3ER: Multiplicative Multimodal Emotion Recognition using Facial, Textual, and Speech Cues," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 02, pp. 1359–1367, Apr. 2020, doi: 10.1609/aaai.v34i02.5492.

[16]. M. Egger, M. Ley, and S. Hanke, "Emotion Recognition from Physiological Signal Analysis: A Review," Electronic Notes in Theoretical Computer Science, vol. 343, pp. 35–55, May 2019, doi: 10.1016/j.entcs.2019.04.009.

[17]. M. Ashwin Shenoy and N. Thillaiarasu, "Enhancing temple surveillance through human activity recognition: A novel dataset and YOLOv4-ConvLSTM approach," Journal of Intelligent & Fuzzy Systems, vol. 45, no. 6, pp. 11217–11232, Dec. 2023, doi: 10.3233/jifs-233919.

[18]. X. Tai and W. Song, "An Improved Approach Based on FCM Using Feature Fusion for Medical Image Retrieval," Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007), pp. 336–342, 2007, doi: 10.1109/fskd.2007.160.

[19]. B. Xu et al., "Attention by Selection: A Deep Selective Attention Approach to Breast Cancer Classification," IEEE Transactions on Medical Imaging, vol. 39, no. 6, pp. 1930–1941, Jun. 2020, doi: 10.1109/tmi.2019.2962013.

[20]. F. Shahidi, S. Mohd Daud, H. Abas, N. A. Ahmad, and N. Maarop, "Breast Cancer Classification Using Deep Learning Approaches and Histopathology Image: A Comparison Study," IEEE Access, vol. 8, pp. 187531–187552, 2020, doi: 10.1109/access.2020.3029881.

[21]. R. Azad, M. Asadi-Aghbolaghi, S. Kasaei, and S. Escalera, "Dynamic 3D Hand Gesture Recognition by Learning Weighted Depth Motion Maps," IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 6, pp. 1729–1740, Jun. 2019, doi: 10.1109/tcsvt.2018.2855416.

[22]. H. H. Sultan, N. M. Salem, and W. Al-Atabany, "Multi-Classification of Brain Tumor Images Using Deep Neural Network," IEEE Access, vol. 7, pp. 69215–69225, 2019, doi: 10.1109/access.2019.2919122.

[23]. K. Bhattacharjee and M. Pant, "Hybrid particle swarm optimization-genetic algorithm trained multi-layer perceptron for classification of human glioma from molecular brain neoplasia data," Cognitive Systems Research, vol. 58, pp. 173–194, Dec. 2019, doi: 10.1016/j.cogsys.2019.06.003.

[24]. M. Arbane, R. Benlamri, Y. Brik, and M. Djerioui, "Transfer Learning for Automatic Brain Tumor Classification Using MRI Images," 2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH), pp. 210–214, Feb. 2021, doi: 10.1109/ihsh51661.2021.9378739.

[25]. H. Ucuzal, S. YASAR, and C. Colak, "Classification of brain tumor types by deep learning with convolutional neural network on magnetic resonance images using a developed web-based interface," 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pp. 1–5, Oct. 2019, doi: 10.1109/ismsit.2019.8932761.

[26]. G. E. Hinton, S. Osindero, and Y.-W. Teh, "A Fast-Learning Algorithm for Deep Belief Nets," Neural Computation, vol. 18, no. 7, pp. 1527–1554, Jul. 2006, doi: 10.1162/neco.2006.18.7.1527.

[27]. W. Zhang, L. Ren, and L. Wang, "A Method of Deep Belief Network Image Classification Based on Probability Measure Rough Set Theory," International Journal of Pattern Recognition and Artificial Intelligence, vol. 32, no. 11, p. 1850040, Jul. 2018, doi: 10.1142/s0218001418500404.

[28]. H. Jang, H. Choi, Y. Yi, and J. Shin, "Adiabatic Persistent Contrastive Divergence learning," 2017 IEEE International Symposium on Information Theory (ISIT), pp. 3005–3009, Jun. 2017, doi: 10.1109/isit.2017.8007081.

[29]. T. Tlusty, G. Amit, and R. Ben-Ari, "Unsupervised clustering of mammograms for outlier detection and breast density estimation," 2018 24th International Conference on Pattern Recognition (ICPR), pp. 3808–3813, Aug. 2018, doi: 10.1109/icpr.2018.8545588.

[30]. S. A. Ali et al., "An Optimally Configured and Improved Deep Belief Network (OCI-DBN) Approach Disease Prediction Based on Ruzzo–Tompa and Stacked Genetic Algorithm," IEEE Access, vol. 8, pp. 65947–65958, 2020, doi: 10.1109/access.2020.2985646for Heart.