Journal of Theoretical and Applied Information Technology
31st December 2023. Vol.101. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

OLGV3 NET: OPTIMIZED LIGHTGBM WITH INCEPTIONV3 FOR ACCURATE MULTI-CLASS BREAST CANCER IMAGE CLASSIFICATION

D. VETRITHANGAM[1], B. ARUNADEVI[2] , ANGATI KALYAN KUMAR[3] , S. NALINI[4], NEHA [5] , NIRMAL MAHESH[6]

[1]Professor ,Department of Computer Science & Engineering, Chandigarh University, Mohali, India.
[2] Professor ,Department of ECE , Dr. N.G.P Institute of Technology, Coimbatore, India.
[3]Assistant Professor , Department of Computer Science and Engineering(Data Science), Madanapalle Institute of Technology & Science, Madanapalle, Andhra Pradesh, India.
[4]Assistant Professor, Department of Computing Technologies, SRM Institute of Science and Technology, Kattankulathur, Chennai.
[5] Assistant Professor, Department of Computer Science & Engineering, Chandigarh University, Punjab, India.
[6]Assistant Professor, Department of Computer Engineering, Pravara Rural Engineering College, Loni, Maharashtra-413 736.

E-mail: [1]vetrigold@gmail.com , [2]arunadevi@drngpit.ac.in, [3]kalyanangatikumar@gmail.com, [4]nalini.tamilmani@gmail.com , [5]neha.arya35@gmail.com, [6]vedic1234@gmail.com

ABSTRACT

In today's world, computer models, especially those using deep learning, are helpful in diagnosing breast cancer by analyzing special images called histopathological images. Understanding and classifying these images for breast cancer diagnosis is crucial in the field of medical information technology. The existing deep learning models for breast cancer image classification include a lack of diversity in the training dataset, leading to reduced model robustness and an inability to accommodate variations in different imaging conditions. Furthermore, there exists a deficiency in the model's sensitivity and generalization capabilities, accompanied by suboptimal hyperparameter configurations. This inadequacy has the potential to hinder the model's efficiency in breast cancer classification. Additionally, the absence of regularization options heightens the susceptibility to overfitting. These identified gaps directly impact the effectiveness of current technologies in addressing crucial issues encountered in clinical practice and biomedical research concerning breast cancer diagnosis and prognosis. This research aims to overcome these challenges by focusing on important factors like making the model work well with different types of images, avoiding unnecessary information, ensuring efficient performance, and handling difficulties when there are only a few cancer cells present. The proposed solution is a new model called OLGV3 Net Classifier, which combines enhanced Inception V3 for understanding images and LightGBM for making accurate classifications. By using Sequential Model-Based Optimization (SMBO) to fine-tune the model's settings, this research achieved a remarkable accuracy of 99.80%, surpassing other models and making a significant improvement in breast cancer image classification.

Keywords: *Multi-classification, Breast cancer, Inception V3, LightGBM, Optimization*

1. INTRODUCTION

The human body consists of trillions of cells, and the term "cancer" is used to describe when a cell undergoes abnormal and uncontrolled division, potentially affecting various parts of the body [1][2][3]. The classification of this disease is based on its location within the body [4][5].As it advances and spreads to other areas, it can become life-threatening. Among the various types of cancer, breast cancer is the most frequently diagnosed form among women worldwide [6][7][8][9]. Early-stage diagnosis and treatment can substantially lower the mortality rate [10][11][12]. The gold standard for identifying breast cancer relies on a histopathological diagnosis using light microscopy [13][14].Non-invasive imaging methods might not effectively identify cancerous regions[15][16][17]. Pathologists proceed with microscopic analysis of

these slides to confirm the diagnosis of breast cancer [18][19][20]. Nevertheless, manually analyzing complex histopathological images is a time-consuming and laborious task, which can be susceptible to errors [21][22][23]. The role of computer-assisted diagnosis is crucial in assisting pathologists with the analysis of histopathology images [24][25]. These methodologies primarily depend on feature extraction techniques such as scale-invariant feature transform, speed-robust features, and local binary patterns. These methods are all based on supervised information and may be susceptible to biased outcomes when classifying breast cancer histopathology images [26][27]. Hence, the demand for effective diagnosis has driven the development of advanced computational models built upon multiple layers of nonlinear processing units, known as deep learning [28][29][30]. In recent times, deep learning models [31][32][33] have achieved significant advancements in computer vision, particularly in the realm of biomedical image processing. Their capacity to autonomously acquire complex and advanced features from images has motivated numerous researchers to explore the application of these models in breast cancer histopathology classification [34]. In particular, convolutional neural networks (CNNs) are extensively employed in image-related tasks because of their effectiveness in parameter sharing across different layers within deep learning models. Over the last few years, numerous CNN-based architectures have been introduced [35][36][37][38]. The VGG16 network with linear SVM was used for Multi-Classification of Breast Cancer Histopathology Images [39]. The Graph-Based Adaptive Regularized Learning Deep Network (GARL-Net) is employed for breast cancer classification. To train the backbone network, DenseNet121, transfer learning is applied. The model employs cross-entropy loss as a strategy to address and mitigate misclassification issues [40]. A model based on optimal feature selection has been developed for the efficient prediction of breast cancer, utilizing a modified logistic regression approach [54]. An artificial neural network model in conjunction with a metaheuristic algorithm worked well on the types of datasets for Coronary Artery Disease Prediction [55]. Multi-disease prediction was done by the use of deep reinforcement Boltzmann machines [56]. The modified ResNet152v2 model demonstrated effective performance in predicting pneumonia from chest X-rays [57]. The accuracy of brain tumor prediction is improved by enhancing the convolutional neural network layers, transfer learning fully connected layers, and weights of the layers in the VGG-19 model [58][59]. The research problem revolves around addressing critical gaps in the existing breast cancer image classification models. These gaps include insufficient diversity in the training dataset, resulting in reduced model robustness and the inability to handle variations in different imaging conditions. Additionally, the absence of an existing breast cancer classification model with enhanced sensitivity to anomalies and robustness to variations poses a challenge, potentially leading to lower diagnostic accuracy. Suboptimal hyperparameter configurations, which are crucial for model training and generalization, may further hinder the overall performance in breast cancer classification. Moreover, the inadequacy of handling categorical features and the absence of regularization options contribute to potential difficulties in capturing essential categorical characteristics in breast cancer images, consequently raising the risk of overfitting. Addressing these research gaps is crucial for advancing the effectiveness and reliability of breast cancer image classification models in real-world medical scenarios. Therefore, the objective is to create a model that not only ensures precise image classification but also showcases resilience across diverse datasets, optimally utilizes features, and performs effectively under varying magnification levels and instances where cancerous cells are limited. The model's generalizability across diverse datasets is achieved through data augmentation techniques such as rotation, shifting, flipping, shear, and zoom, which augment the training dataset and improve the model's performance on unseen data. To tackle redundancy in feature extraction and utilization, a dual-model approach is adopted, combining InceptionV3 for feature extraction and LightGBM for classification. This strategic amalgamation captures and utilizes high-level image features efficiently, enhancing the overall effectiveness of the model. Furthermore, to optimize overall performance, the research employs sequential model-based optimization (SMBO) for hyperparameter tuning, mitigating efficiency concerns by systematically and efficiently selecting the most favorable hyperparameter configurations. The proposed OLGV3 Net classifier seamlessly integrates deep learning, SMBO, and LightGBM, presenting an advanced architecture for the accurate multi-classification of breast cancer histopathological images. The improved Inception V3 plays a pivotal role in hierarchically extracting features from these images, capturing information from low-level features to high-level patterns. This comprehensive approach aims to improve model

generalization, classification accuracy, and overall efficiency in the critical task of breast cancer image classification.

2. LITERATURE REVIEW

This review section explains the various breast cancer image classification techniques. Alirezazadeh and Dornaika[41] introduced Boosted Additive Angular Margin Loss (BAM) as a means to acquire exceptionally discriminative features. One notable benefit of this approach is its consideration of the angles between deep features and non-target class weights. When applied, this model achieved accuracy rates of 99.79% at 40X magnification and 97.65% at 400X magnification. It's worth noting that results may vary across different magnification levels. Chhipa et al.[42] introduced the Magnification Prior Contrastive Similarity (MPCS) method. This approach leverages only 20% of the labels during fine-tuning, making use of magnification factors, inductive transfer, and the reduction of human priors. Remarkably, this method demonstrated substantial enhancements compared to other techniques, achieving an impressive 98.18% in fine-tuning and 96.36% in linear evaluation. However, one limitation of this method is its susceptibility to redundancy issues. Xu, C et al.[43] introduced the Multi-Dimensional Feature Fusion Network (MDFF-Net), which integrates one-dimensional and two-dimensional features within a dedicated feature fusion network. MDFF-Net demonstrates remarkable performance with an accuracy of 98.86% on the BreakHis dataset and 86.25% on the BACH dataset. However, a limitation of this model lies in the need to effectively identify and address redundancy in the fused features. Doing so is crucial to avoid unnecessary computational overhead and to enhance the interpretability of the fused features. Ahmed et al.[44] introduced a Transfer Learning approach for Breast Cancer Diagnosis, enhancing results through upsampling and image augmentation techniques. Their method achieved an impressive F1-score of 96.2%. Nevertheless, this work faces a limitation related to tissue preparation and staining duration, which can introduce variability in image appearance and quality. This variability can pose challenges in establishing consistent patterns or features for diagnosis. In their study, Vesal et al.[45] introduced a transfer learning approach leveraging the Inception-V3 network, which yielded an impressive average test accuracy of 97.08% across four distinct classes. To address color variations introduced during slide preparation, normalization techniques were employed. Feature extraction was performed using both the Inception-V3 and ResNet50 architectures. However, a noteworthy limitation of this study is the reliance on majority voting. This approach may be susceptible to instances where cancerous cells are only present in a limited portion of the image, with the remainder depicting healthy or benign tissue. Vo et al.[46] introduced an innovative boosting strategy, employing deep learning models with convolutional layers to extract highly informative visual features for breast cancer classification. This approach yielded impressive test accuracies of 95.1% at 40X, 96.3% at 100X, 96.9% at 200X, and 93.8% at 400X magnification levels. However, a limitation of this approach is the potential occurrence of feature redundancy due to feature fusion. Mehra[47] introduced a pre-trained VGG16 model coupled with a logistic regression classifier, which yielded exceptional results with an accuracy rate of 92.60%. It's worth noting that the network's capacity plays a pivotal role in influencing its performance, potentially leading to overfitting if it's excessively large or underfitting if it's too small. Nazeri et al. [48]introduced a patch-based approach wherein they employed an auto-encoder to capture the key features of image patches. Simultaneously, a second "image-wise" network was employed for the overall image classification task. The reported accuracy for this model was an impressive 95%. However, in our experimental trials, we observed that this model demands an excessively large amount of memory, rendering it unfeasible for practical applications. Koné and Boulmane[49] introduced a hierarchical system of convolutional neural networks (CNNs) achieving an impressive accuracy rate of 99%. However, it's worth noting that this model utilized a relatively small dataset, which can potentially make it susceptible to overfitting if not appropriately regularized. Xie et al.[50] introduced SHISRCNet, a novel super-resolution and Classification Network designed specifically for enhancing low-resolution breast cancer histopathology images. This comprehensive model incorporates Super-Resolution (SR) and Classification (CF) modules, achieving an impressive accuracy rate of 92.78%. However, it's important to note that the computational demands of super-resolution and classification networks may pose challenges for real-time or near-real-time processing, potentially limiting their application in scenarios where rapid diagnosis is of utmost importance. Sun et al.[51] introduced FabNet, a novel convolutional neural network (CNN) known as FabNet A, designed for the classification of multiscale breast cancer

histopathology images using feature agglomeration. This model demonstrated remarkable performance, achieving impressive 98.2% test accuracy when applied to a colon cancer dataset, all without the need for any data augmentation techniques. Furthermore, the tightly integrated architecture effectively addressed the data imbalance issue within the dataset, leading to only marginal impacts on the model's overall performance. In their study, Kode & Pranjit Das et al.[52][60] introduced a hybrid approach involving Deep Learning (DL) and Expert Knowledge-Based Feature Extraction for the evaluation of performance in classifying Breast Histopathology Images. This model demonstrated impressive results, achieving accuracy rates of up to 98% when using Neural Networks. Although expert-based systems offer greater interpretability, it's worth noting that the features selected by experts may not consistently align with the most pertinent or informative aspects of a particular classification task. Guleria et al.[53] introduced an approach that utilized a Variational Autoencoder (VAE). This VAE-based model achieved an accuracy of 73%. However, a limitation of this model is its tendency to exhibit a high loss value, which in turn leads to degradation in its overall performance. The considerations encompass a range of limitations and challenges in the context of developing and applying machine learning models for breast cancer diagnosis from histopathological images. These limitations include generalizability concerns, which pertain to the model's ability to perform well on unseen data and across different magnification levels. Susceptibility to redundancy implies the risk of duplicating or inefficiently using certain features, impacting the model's efficiency. Efficiency implications highlight the need for optimizing hyperparameters effectively. Redundancy in feature fusion emphasizes the importance of avoiding duplicative information when combining features. Variability due to tissue preparation and staining duration can introduce challenges to maintaining consistent image quality and appearance. Consistency in pattern recognition is vital for reliable diagnosis. The limited presence of cancerous cells may require specialized handling. Accuracy implications stress the need for high precision in classification. Feature redundancy due to fusion concerns duplicative information when combining different features. Misalignment with pertinent aspects underscores the importance of feature extraction aligning with the most critical aspects of the classification task. Knowledge-based feature extraction impact pertains to the reliance on expert knowledge, which may not consistently align with the most informative aspects. High loss values highlight potential performance degradation. Performance implications emphasize the need for efficient and effective model performance in breast cancer diagnosis. These challenges highlight the complexity of the task and the importance of addressing them to develop accurate and reliable diagnostic models.

3. PROPOSED METHODOLOGY

In order to address technical gaps associated with generalizability, redundancy, performance, efficiency of the deep learning models, the proposed architecture of the OLGV3 Net classifier harmoniously integrates a deep learning feature extraction model, sequential model-based optimization (SMBO) for hyperparameter optimization, and the LightGBM classifier for the classification of breast cancer histopathological images. Our OLGV3 Net classifier's design has several advantages. It improves the understanding of intricate details in breast cancer images by using the Inception V3 model for feature extraction. This makes the model better at capturing subtle and complex patterns, leading to more accurate classification. Sequential Model-Based Optimization (SMBO) is then applied to fine-tune the model's settings, improving efficiency. For the final classification, the LightGBM classifier is used, known for making quick and accurate decisions. The combination of these elements creates an advanced model specifically tailored for multi-class breast cancer classification, offering valuable insights for diagnosis.

3.1 Dataset

A comprehensive and annotated dataset was meticulously curated and made accessible to the research community for the BACH challenge. This dataset encompasses both microscopy images, geared towards a classification task, and whole-slide images, tailored for segmentation tasks. The response from the scientific community was overwhelmingly positive, with a total of 64 submissions entering the competition out of the 677 registrations received. The dataset classifies microscopy images into four categories: normal, benign, in situ carcinoma, and invasive carcinoma, according to the primary cancer type apparent in each image. For the classification task, the dataset includes two primary folders labeled 'train' and 'test.' Inside the 'image' directory, the four previously mentioned classes are discovered: benign, in situ, invasive, and normal. Furthermore, a ground truth CSV file is included to facilitate the labeling process

for these images, all of which are in the.tiff format. The dataset consists of a total of 400 microscopy images, with an equal distribution among four categories: a) Normal: 100 images, b) Benign: 100 images, c) In situ carcinoma: 100 images d) Invasive carcinoma: 100 images. These microscopy images follow specific specifications, including a color model in Red, Green, Blue (RGB), dimensions of 2048 x 1536 pixels, a pixel scale of 0.42 µm x 0.42 µm, memory space consumption ranging from approximately 10 to 20 megabytes, and they are labeled using an image-wise classification approach.

3.2 Preprocessing

Data augmentation is a process used to increase the diversity and quantity of training data for machine learning models, particularly in image classification tasks. In this context, data augmentation techniques are applied to the training data using a tool called the ImageDataGenerator. The purpose of these techniques is to introduce variations and perturbations to the original images while preserving their inherent characteristics. By doing so, the augmented dataset becomes more robust and helps improve the model's generalization, which means it can perform better on unseen or real-world data. Rotation entails the adjustment of images by different degrees, such as 90 degrees or 180 degrees, to replicate diverse orientations. Shifting introduces horizontal and vertical translations to the images. Flipping results in the creation of mirror images, either horizontally or vertically. Shearing involves distorting the image along a designated axis, producing a slanted or tilted effect. Zooming, on the other hand, includes magnifying or reducing the image, thereby simulating variations in scale. By applying these augmentation techniques, the training dataset is artificially expanded in size, creating a more extensive and diverse set of training examples. This diversity helps the deep learning model learn different patterns, textures, and variations present in the data, making it better equipped to handle various real-world scenarios and improving its overall performance in classifying breast cancer histology images.

3.3 Proposed System Model

Figure 1 illustrates the proposed architecture of the OLGV3 Net classifier. This architecture harmoniously integrates a deep learning feature extraction model, sequential model-based optimization (SMBO) for hyperparameter optimization, and the LightGBM classifier for the classification of breast cancer histopathological images.

3.3.1 Improved Inception V3

The Inception V3 model comprises several layers, each with distinct functions. It begins with a 3x3 convolution layer featuring 32 filters and a stride of 2, which reduces the feature map's spatial dimensions while capturing basic image features. This is followed by another 3x3 convolution layer with 32 filters, without any stride, focusing on more detailed patterns. A subsequent 3x3 convolution layer employs 64 filters, delving into complex and abstract feature extraction. The architecture includes a Max-Pooling layer with a 3x3 pool size and a stride of 2, which conducts down sampling to maintain essential information while reducing computational complexity.

The improved Inception V3 model incorporates a modified stem block that introduces several enhancements to elevate feature extraction. First, it increases the number of filters to 64 in each convolutional layer, widening the range for feature and pattern recognition. Batch normalization is applied after each convolutional layer, promoting training stability, accelerating convergence, and mitigating issues like vanishing gradients. The Rectified Linear Unit (ReLU) activation function follows each convolutional layer, infusing non-linearity and enabling the model to learn intricate mappings. The max-pooling layer is adjusted to employ a 2x2 pool size with a stride of 2, facilitating more aggressive spatial dimension reduction while retaining crucial information. Notably, in the initial convolutional layer, an activation function is omitted, permitting the network to capture both positive and negative features in its early stages. Overall, these modifications aim to make the stem block more effective in capturing and representing features from input images. By increasing the number of filters, adding batch normalization and ReLU activations, and fine-tuning pooling parameters, the network can extract more informative features, potentially improving its performance in subsequent layers and blocks of the architecture.

The improved Inception V3 model allows for more effective extraction of high-level features from complex structures in breast cancer images. The model's improved architecture enhances its sensitivity to anomalies or irregularities in breast tissue. This heightened sensitivity is essential for the early detection of potential cancerous regions and the classification of images, contributing to

improved diagnostic accuracy. It handles variations in image scales, orientations, and textures. This robustness is crucial when dealing with diverse datasets of breast cancer images, ensuring reliable feature extraction across different imaging conditions. The Improved Inception V3 enhances feature extraction and the capabilities of a deep learning model in discerning critical details from breast cancer images. Its sophisticated architecture, sensitivity to abnormalities, and robustness to variability contribute to improved diagnostic accuracy and performance in real-world medical scenarios.

3.3.2 Sequential Model-Based Optimization (SMBO)

The architecture of Sequential Model-Based Optimization (SMBO) comprises several integral components. SMBO, an adept technique for optimizing intricate, high-dimensional, and noisy objective functions, finds common application in the hyperparameter optimization of machine learning models. A fundamental element of SMBO is the surrogate model, typically adopting a probabilistic guise such as Gaussian Process (GP) or Random Forest. This surrogate model endeavors to approximate the objective function and discern the connection between hyperparameters and the objective function's performance. Guiding the quest for optimal hyperparameters is the acquisition function, tasked with weighing the exploration of uncharted territories against the exploitation of promising regions. Noteworthy acquisition functions encompass Probability of Improvement (PI), Expected Improvement (EI), and Upper Confidence Bound (UCB). The optimization journey commences with an initial set of hyperparameter configurations, often randomly chosen or curated via approaches like Latin Hypercube Sampling. The optimization loop, the crux of SMBO, systematically enhances the surrogate model and selects fresh hyperparameter configurations for assessment. These iterations involve updating the surrogate model post-objective function evaluation, maximizing the acquisition function to pinpoint the next configuration, assessing the selected configuration using the actual objective function (e.g., validation accuracy), and adhering to a predefined convergence criterion. Ultimately, the configuration exhibiting the highest estimated performance, as per the surrogate model, emerges as the best configuration when SMBO concludes its iterative process.

SMBO explores hyperparameter combinations that are likely to perform well with the complex feature representations extracted by the Inception V3 architecture. Applying sequential model-based optimization to hyperparameter tuning in breast cancer classification tasks following feature extraction using the Inception V3 model offers an intelligent and efficient approach. It helps navigate the complex hyperparameter space, reducing the number of model evaluations needed and ultimately improving the overall performance and accuracy of the deep learning model in classifying breast cancer images.

3.3.3 LightGBM classifier

The LightGBM classifier's architecture, used for classifying breast cancer into four distinct categories (Benign, InSitu, Invasive, and Normal), presents several significant features. First and foremost, LightGBM is rooted in the concept of gradient boosting, constituting an ensemble learning technique that amalgamates multiple weak learners, typically in the form of decision trees, to assemble a robust predictive model. Notably, LightGBM stands out due to its utilization of histogram-based splitting during the tree construction process, which involves binning continuous features into discrete bins and constructing histograms. This methodology expedites and streamlines the process of selecting optimal split points, particularly advantageous for high-dimensional datasets. Further setting LightGBM apart is its leaf-wise tree growth approach, in which the algorithm prioritizes expanding the leaves of the tree that result in the most significant reduction in the loss function. This often results in shorter trees with fewer levels, mitigating overfitting and optimizing computational efficiency. To bolster its robustness, LightGBM incorporates regularization techniques such as L1 (Lasso) and L2 (Ridge) regularization, penalizing substantial feature importance values and fostering model simplicity. This regularization is especially valuable in preventing overfitting. LightGBM also boasts native support for categorical features, obviating the need for one-hot encoding and facilitating efficient handling of non-numeric data. Its parallel and distributed training capabilities make it well-suited for large-scale datasets and parallel processing environments, potentially accelerating training. The inclusion of an early stopping mechanism further refines training efficiency by monitoring validation performance and halting training if the model's performance ceases to improve. Ultimately, LightGBM's efficiency, speed, and versatility have made it a favored choice for machine learning tasks, including classification and regression, with its architectural elements enhancing

the overall performance in the context of breast cancer classification. LightGBM model exhibits proficiency in handling substantial datasets and high-dimensional feature spaces. In the realm of breast cancer image classification, where datasets often possess extensive dimensions, LightGBM model's capacity to efficiently scale to large data sizes is particularly advantageous. Moreover, it is aptly designed for datasets containing categorical features, rendering it well-suited for scenarios where specific image characteristics assume a categorical nature. This suitability proves relevant in the context of breast cancer classification, where certain visual features, such as distinct morphological patterns, may exhibit categorical attributes. Furthermore, LightGBM offers regularization options, a critical feature in mitigating overfitting concerns in medical image classification tasks. Mitigating overfitting is essential to facilitate effective generalization of the model to unseen data.

3.4 Working Flow of the OLGV3 Net Classifier

The figure 2 represents the Workflow of the OLGV3 Net Classifier in detail. The necessary libraries, encompassing TensorFlow, Keras, ImageDataGenerator, Hyperopt, and others, have been imported for the purpose of constructing and optimizing the breast cancer image classification model. Upon loading the image dataset, the image dimensions (img_width and img_height) and batch size (batch_size) are defined, with these parameters determining the input image size and batch size for training. Data generators are then created for the train, validation, and test datasets, all facilitated by the ImageDataGenerator. The pixel values are normalized to the [0, 1] range via rescaling, serving to standardize pixel values and enhance training stability. Additionally, data augmentation techniques, encompassing rotation, shifting, flipping, shear, and zoom, are applied to the training data using the ImageDataGenerator. This augmentation process artificially amplifies the training dataset's size and contributes to the refinement of model generalization. Regarding data loading, there are no alterations required at this stage. The process will persist by loading the training and validation data via the flow_from_directory method, where the target size, batch size, and class mode (categorical, for multi-class classification) will be specified. Moving on to the model architecture, the train_model(params) function will be expanded to encompass both a deep learning model (InceptionV3) and a gradient boosting model (LightGBM). The procedure unfolds as follows: Following the neural networks (InceptionV3) training on the training data, the features (predictions) of the validation set will be extracted from the penultimate layer, which is positioned just before the final dense layer. This particular step effectively captures high-level image features that have been acquired through the neural network's learning process. Subsequently, a LightGBM classifier will be trained using these extracted features. LightGBM, being a gradient boosting framework, proves to be particularly adept at handling tabular data, rendering it well-suited for the task of managing the extracted image features. Expanding upon the hyperparameter search space, it encompasses LightGBM-specific parameters such as num_leaves, lgb_learning_rate, and lgb_num_boost_round. Within this extended space, Hyperopt will explore diverse combinations of these parameters to discern the most optimal configurations for the model. The optimization setup retains the Trials object for ongoing tracking of the optimization process. This Trials object maintains a record of the outcomes associated with each tested parameter combination during the hyperparameter optimization endeavor. Subsequently, employing the f_{min} function from Hyperopt, Sequential Model-Based Optimization (SMBO) is executed. Hyperopt diligently explores the specified hyperparameter search space, endeavoring to ascertain the most favorable hyperparameters for both the neural network and LightGBM. With the best hyperparameters secured, the final model training entails creating a fresh instance of InceptionV3. Custom classification layers are appended to the InceptionV3 base model, akin to prior steps. The final model is then configured, incorporating the best learning rate and other optimized hyperparameters. This configuration is pivotal as it tailors the model to perform optimally. Consequently, the model is trained utilizing the complete training dataset, a step that ensures the neural network is exposed to the most advantageous settings. Lastly, the evaluation of the ensemble model's performance is conducted on the test dataset, which currently encompasses images classified into four distinct categories: Benign, InSitu, Invasive, and Normal. This evaluation leverages appropriate metrics to gauge the model's proficiency in classifying these images.

Journal of Theoretical and Applied Information Technology
31st December 2023. Vol.101. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

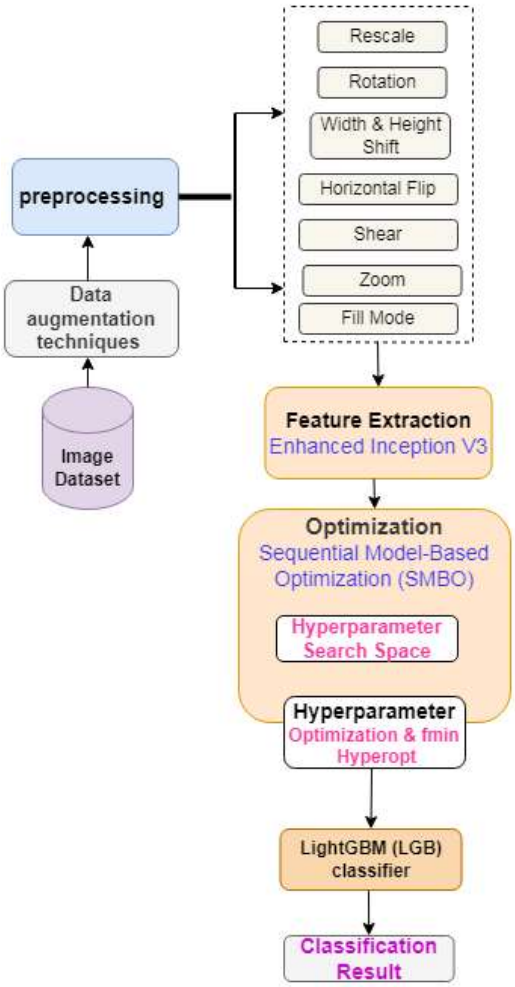

Figure 2: Flow Diagram of the OLGV3 Net Classifier

4. RESULTS AND ANALYSIS

4.1 Performance Metrics

Accuracy is determined by the correlation of samples that were precisely recognized as positive to every sample that was identified as positive. The model's accuracy measures how accurately it classifies a sample as positive as shown in equation (1).

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \quad (1)$$

"TP," "TN," "FP," and "FN" stand for "true positive," "true negative," "false positive," and "false negative," respectively. The denominator increases and the precision become low when the model only seldom classifies positive data correctly or frequently classifies positive data incorrectly as shown in equation (2).

$$\text{precision} = \frac{(\text{TP})}{(\text{TP} + \text{FP})} \quad (2)$$

Anyhow, when the conditions listed below are satisfied, the precision is high: 1.The model generates a lot of precisely positive classifications (maximize True Positive). 2. There are lesser in accurate positive classifications formed by the model (which minimizes false positives). Calculating recall involves dividing the proportion of properly labeled positive samples by the total count of positive samples. How successfully the model can detect positive samples is assessed by recall as shown in equation (3).

$$\text{Recall} = \frac{(\text{TP})}{(\text{TP} + \text{FN})} \quad (3)$$

The Recall (Re) grows in proportion to the number of positive samples found. The recall relies heavily on how the positive samples are categorized. The way the negative samples are categorized, such as for precision, has nothing to do with this. If the model correctly identifies the entire positive samples as positive. The sensitivity of a machine learning model determines how well it can locate successful examples. It is termed the "true positive rate" (TPR) or recalls. It is used to evaluate model performance because recall tells us how many samples the model was able to properly identify. A high sensitivity model might be missing some of the positive cases, according to a few false negatives. In other words, sensitivity evaluates how well a model can distinguish between good data.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (4)$$

In multiclass classification with four categories (Benign, InSitu, Invasive, and Normal), the confusion matrix is typically organized as a 4x4 matrix. Each row of the matrix represents the instances in a predicted class, and each column represents the instances in an actual (true) class. The elements of the matrix are as follows:

- **True Positives (TP)**: These are the instances where both the predicted and actual classes are the same. In other words, the model correctly predicted that an instance belongs to a particular class.
- **True Negatives (TN)**: In a multiclass problem, this term isn't applicable because we are not considering binary negative and positive outcomes for each class. Instead, we focus on the correct classification of each individual class.

- **False Positives (FP)**: These are instances where the model predicted a class, but the actual class is different. In other words, the model incorrectly predicted that an instance belongs to a particular class.

- **False Negatives (FN)**: Again, in a multiclass problem, this term isn't commonly used because we don't usually focus on binary negative and positive outcomes for each class. Instead, we focus on the correct classification of each individual class.

4.2 Results of the preprocessing Stage

The figure 3 illustrates the practice of image data augmentation, a widely adopted technique within the computer vision and deep learning domains. Its primary objective is to enhance the diversity of the training dataset, thereby contributing to the refinement of machine learning models' performance. To facilitate this, the code imports essential libraries, including numpy for numerical operations, matplotlib.pyplot for image display, ImageDataGenerator from Keras for image data augmentation, and various Keras functions like load_img, img_to_array, and array_to_img for efficient image manipulation. The code commences by initializing an ImageDataGenerator object devoid of specific augmentation parameters, which subsequently serves as the tool for applying various data augmentation techniques. These techniques include image rotation at distinct angles (0 and 40 degrees), width shifting (horizontally) employing variable shift values (0 and 0.2), height shifting in both horizontal and vertical directions with different magnitudes (0 and 0.2), shear transformations using diverse shear values (0 and 0.2), and zooming with varying factors (1 and 1.2) in both the x and y directions. Furthermore, it introduces the concept of horizontal flipping, generating mirrored versions of the original image.

4.3 Classification Results and Analysis

During the training and evaluation of a deep learning model, we take into account key metrics such as 'train accuracy,' 'validation accuracy,' 'train loss,' and 'validation loss.' The 'train accuracy' metric assesses the model's capability to classify instances within the training dataset accurately. In contrast, the 'validation accuracy' metric evaluates the model's performance on a separate validation dataset, providing insights into its accuracy when dealing with previously unseen data.

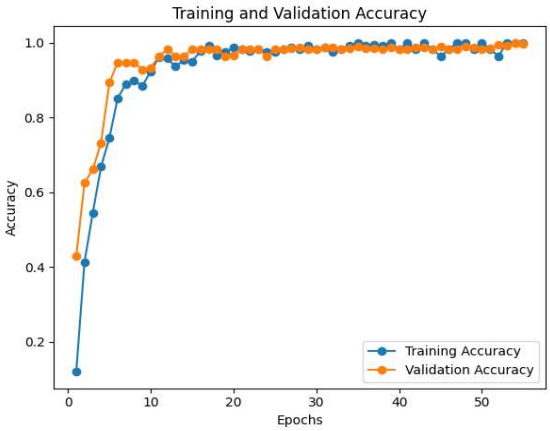

Figure 4: The OLGV3 Net Classifier's Train and Test Accuracy on the BACH Dataset for Epochs 55

As illustrated in the figure 4, the OLGV3 Net classifier achieved peak training accuracy, reaching 100%, and validation accuracy of 99.80% during Epoch 55. The training accuracy curve began at 11.91% and progressively rose to the maximum of 100%, while the validation accuracy curve began at 42.86% and reached its peak at 99.80%. In the validation process, the model's performance is assessed using the validation dataset, which contains unobserved data. The 'train loss,' also referred to as 'training loss' or 'training error,' quantifies the disparity between the expected model outputs and the actual target values within the training dataset. Similarly, 'validation loss' is computed using the validation dataset, determining the divergence between the model's predictions and the genuine target values for the validation data.

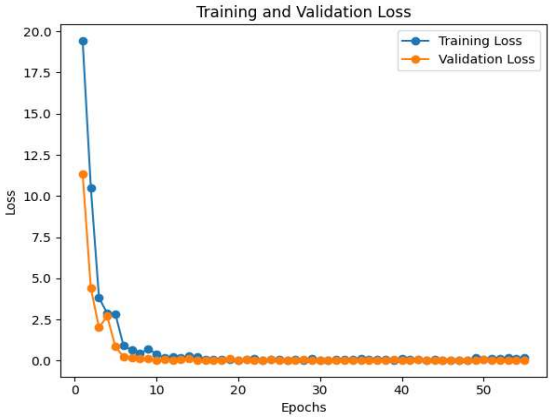

Figure 5: The OLGV3 Net Classifier's Train and Test Loss on the BACH Dataset for Epochs 55

As illustrated in figure 5 , the training loss curve initiates at 0.0033 and reaches its peak loss value of 19.4451 during the training process. Conversely, the

Journal of Theoretical and Applied Information Technology
31st December 2023. Vol.101. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

validation loss curve starts at 0.0045 and attains its highest validation loss value of 11.3472.

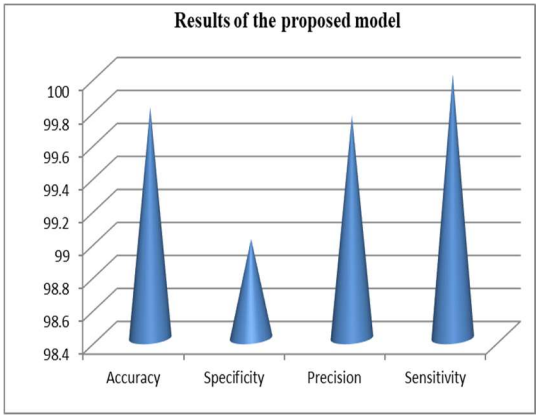

Figure 6: The Results of the OLGV3 Net Classifier on the BACH Dataset for Epochs 55

Figure 6 illustrates results of our multi-classification experimentation on the breast cancer dataset, where we strategically selected 30% of the dataset for analysis. Employing the OLGV3 Net classifier, our findings revealed outstanding performance across multiple key metrics. Most notably, our model demonstrated a remarkable sensitivity of 100%, underscoring its ability to accurately identify positive cases. In addition, it exhibited a strong specificity of 99%, highlighting its proficiency in correctly identifying negative cases. Impressively, the model achieved a precision rate of 99.75%, indicating its precision in classifying positive cases. In summary, the classifier achieved an exceptional overall accuracy of 99.8%, confirming its effectiveness in accurately classifying breast cancer instances.

In the figure 7, the cells along the diagonal represent the accurate classifications by our model for different categories. In the first cell of the diagonal, we observe that all 30 benign cases were correctly classified. Similarly, in the second cell, 29 out of 30 InSitu cases were correctly classified. The third cell in the diagonal corresponds to 30 out of 30 correctly classified invasive cases. Lastly, the fourth cell represents the accurate classification of all 30 normal cases by our model. As shown in figure 8, the construction of the normalized confusion matrix involved utilizing test set samples. The matrix's rows correspond to the actual class of each sample, while the columns represent the class assigned by the classifier. The diagonal of the matrix displays the percentage of correctly predicted movement types. In the context of our analysis, we conducted experiments on a 30% subset of the dataset, comprising 120 diverse breast images classified into four distinct categories: 'benign,' 'InSitu,' 'invasive,' and 'normal.' Our evaluation focused on comparing true values and predicted values for each category, yielding the following results: True values were 1 for all categories, including 'benign,' 'in-situ,' 'invasive,' and 'normal.' The predicted values were 1, 0.96, 1, and 1 for the corresponding categories. These findings served as the basis for creating a visual representation of the multi-classification results, offering valuable insights into the model's performance.

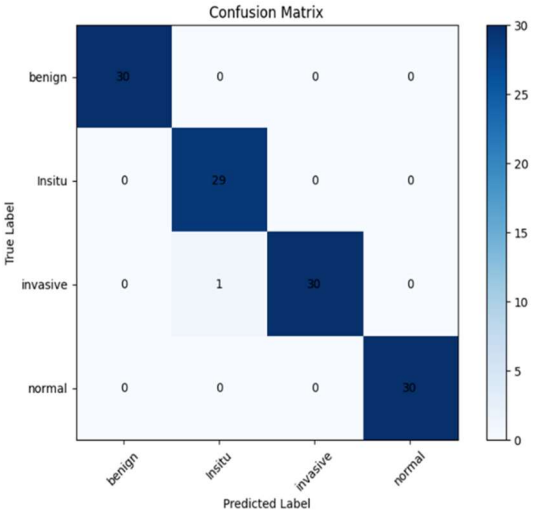

Figure 7: The Confusion Matrix for Multi-Classification by the OLGV3 Net Classifier

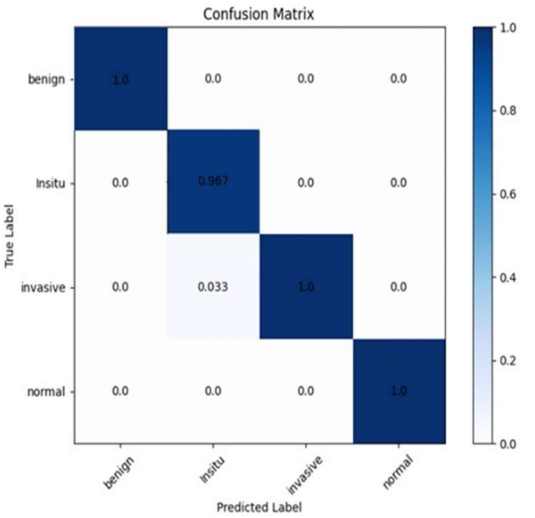

Figure 8: The Normalized Confusion Matrix for Multi-Classification by the OLGV3 Net Classifier

As illustrated in the figure 9, the proposed OLGV3 Net classifier demonstrates significant superiority over the existing models in the dataset in

terms of accuracy. The difference in percentage between the proposed OLGV3 Net and each of the other models is as follows: a mere 0.01% difference with BAM[41], a substantial 0.94% difference with MDFF-Net[43], a notable 2.72% difference with Transfer Learning[45], a substantial 7.20% difference with Logistic Regression Classifier with VGG16[47], a significant 4.80% difference with Auto-Encoder[48], a modest 0.80% difference with CNN, a considerable 7.02% difference with Super-Resolution and Classification Network, a notable 1.60% difference with CNN[51] , and a noteworthy 1.80% difference with the Hybrid Approach Involving Deep Learning (DL)[52]. The most considerable deviation in accuracy is observed with the Variational Autoencoder (VAE)[53], with a substantial 26.80% difference. These results underscore the outstanding performance of the proposed OLGV3 Net classifier, which consistently outperforms its counterparts across a range of models, often by a significant margin in terms of accuracy.

5. CONCLUSION AND FUTURE SCOPE

This research work addresses the critical challenges in multi-class breast cancer image classification using deep learning models. A lack of diversity in the training dataset, sensitivity, generalization deficiencies, feature redundancy, suboptimal hyperparameter configurations, the absence of regularization and model overfitting are addressed by the proposed OLGV3 Net Classifier, which integrates a deep learning feature extraction model, sequential model-based optimization (SMBO) for hyperparameter optimization, and the LightGBM classifier for the classification of breast cancer histopathological images. The proposed OLGV3 Net Classifier produced an overall sensitivity of 100%, a specificity of 99%, a precision rate of 99.75%, an overall accuracy of 99.8%, and outperformed all the existing models. In future research work, the proposed OLGV3 Net Classifier will undergo experimentation with diverse disease image datasets. It will be coupled with a range of deep learning models and fuzzy techniques. Additionally, exploration will extend to the integration of supplementary clinical data, including patient history, genetic information, and other pertinent factors. This is aimed at refining the model's predictive capabilities and achieving a more comprehensive understanding of breast cancer cases.

REFERENCES:

[1] Almeida CA, Barry SA. Cancer: basic science and clinical aspects. John Wiley & Sons; 2011 Aug 26. [2] Bisoyi P. A brief tour guide to cancer disease. InUnderstanding cancer 2022 Jan 1 (pp. 1-20). Academic Press.

[3] Sriharikrishnaa S, Suresh PS, Prasada K S. An Introduction to Fundamentals of Cancer Biology. InOptical Polarimetric Modalities for Biomedical Research 2023 Jul 23 (pp. 307-330). Cham: Springer International Publishing.

[4] Logeshwaran J, Adhikari N, Joshi SS, Saxena P, Sharma A. The deep DNA machine learning model to classify the tumor genome of patients with tumor sequencing. International Journal of Health Sciences. 2022 Jul;6(S5):9364-75.

[5] Carbone A. Cancer classification at the crossroads. Cancers. 2020 Apr 15;12(4):980.

[6] Sharma GN, Dave R, Sanadya J, Sharma P, Sharma K. Various types and management of breast cancer: an overview. Journal of advanced pharmaceutical technology & research. 2010 Apr;1(2):109.

[7] Akram M, Iqbal M, Daniyal M, Khan AU. Awareness and current knowledge of breast cancer. Biological research. 2017 Dec;50:1-23.

[8] Mathur G, Nain S, Sharma PK. Cancer: an overview. Acad. J. Cancer Res. 2015;8(1).

[9] Harwansh RK, Deshmukh R. Breast cancer: An insight into its inflammatory, molecular, pathological and targeted facets with update on investigational drugs. Critical Reviews in Oncology/Hematology. 2020 Oct 1;154:103070.

[10] Chu KC, Lamar CA, Freeman HP. Racial disparities in breast carcinoma survival rates: separating factors that affect diagnosis from factors that affect treatment. Cancer: Interdisciplinary International Journal of the American Cancer Society. 2003 Jun 1;97(11):2853-60.

[11] Lee K, Kruper L, Dieli-Conwright CM, Mortimer JE. The impact of obesity on breast cancer diagnosis and treatment. Current oncology reports. 2019 May;21:1-6.

[12] Kirkham AA, Beaudry RI, Paterson DI, Mackey JR, Haykowsky MJ. Curing breast cancer and killing the heart: a novel model to explain elevated cardiovascular disease and mortality risk among women with early stage

breast cancer. Progress in cardiovascular diseases. 2019 Mar 1;62(2):116-26.

[13] Lashen A, Ibrahim A, Katayama A, Ball G, Mihai R, Toss M, Rakha E. Visual assessment of mitotic figures in breast cancer: a comparative study between light microscopy and whole slide images. Histopathology. 2021 Dec;79(6):913-25.

[14] Wang P, Hu X, Li Y, Liu Q, Zhu X. Automatic cell nuclei segmentation and classification of breast cancer histopathology images. Signal Processing. 2016 May 1;122:1-3.

[15] Kim JB, Urban K, Cochran E, Lee S, Ang A, Rice B, Bata A, Campbell K, Coffee R, Gorodinsky A, Lu Z. Non-invasive detection of a small number of bioluminescent cancer cells in vivo. PloS one. 2010 Feb 23;5(2):e9364.

[16] Smith L, MacNeil S. State of the art in non-invasive imaging of cutaneous melanoma. Skin Research and Technology. 2011 Aug;17(3):257-69.

[17] Wassef C, Rao BK. Uses of non-invasive imaging in the diagnosis of skin cancer: an overview of the currently available modalities. International journal of dermatology. 2013 Dec;52(12):1481-9.

[18] Liu Y, Gadepalli K, Norouzi M, Dahl GE, Kohlberger T, Boyko A, Venugopalan S, Timofeev A, Nelson PQ, Corrado GS, Hipp JD. Detecting cancer metastases on gigapixel pathology images. arXiv preprint arXiv:1703.02442. 2017 Mar 3.

[19] Leong AS, Zhuang Z. The changing role of pathology in breast cancer diagnosis and treatment. Pathobiology. 2011 Jun 14;78(2):99-114.

[20]Steiner DF, MacDonald R, Liu Y, Truszkowski P, Hipp JD, Gammage C, Thng F, Peng L, Stumpe MC. Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. The American journal of surgical pathology. 2018 Dec;42(12):1636.

[21] Li Y, Yang Z, Wang Y, Cao X, Xu X. A neural network approach to analyze cross-sections of muscle fibers in pathological images. Computers in biology and medicine. 2019 Jan 1;104:97-104.

[22] Rączkowski Ł, Możejko M, Zambonelli J, Szczurek E. ARA: accurate, reliable and active histopathological image classification framework with Bayesian deep learning. Scientific reports. 2019 Oct 4;9(1):14347.

[23] Ji J, Wan T, Chen D, Wang H, Zheng M, Qin Z. A deep learning method for automatic evaluation of diagnostic information from multi-stained histopathological images. Knowledge-Based Systems. 2022 Nov 28;256:109820.

[24] Kaushal C, Bhat S, Koundal D, Singla A. Recent trends in computer assisted diagnosis (CAD) system for breast cancer diagnosis using histopathological images. Irbm. 2019 Aug 1;40(4):211-27.

[25] Belsare AD, Mushrif MM. Histopathological image analysis using image processing techniques: An overview. Signal & Image Processing. 2012 Aug 1;3(4):23.

[26] Koshy SS, Anbarasi LJ, Jawahar M, Ravi V. Breast cancer image analysis using deep learning techniques–a survey. Health and Technology. 2022 Nov;12(6):1133-55.

[27] Hameed Z, Garcia-Zapirain B, Aguirre JJ, Isaza-Ruget MA. Multiclass classification of breast cancer histopathology images using multilevel features of deep convolutional neural network. Scientific Reports. 2022 Sep 16;12(1):15600.

[28] Aljuaid H, Alturki N, Alsubaie N, Cavallaro L, Liotta A. Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning. Computer Methods and Programs in Biomedicine. 2022 Aug 1;223:106951.

[29] Jasti VD, Zamani AS, Arumugam K, Naved M, Pallathadka H, Sammy F, Raghuvanshi A, Kaliyaperumal K. Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis. Security and communication networks. 2022 Mar 9;2022:1-7.

[30] Hou Y. Breast cancer pathological image classification based on deep learning. Journal of X-ray science and technology. 2020 Jan 1;28(4):727-38.

[31] Kumar A, Singh SK, Saxena S, Lakshmanan K, Sangaiah AK, Chauhan H, Shrivastava S, Singh RK. Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer. Information Sciences. 2020 Jan 1;508:405-21.

[32] Han Z, Wei B, Zheng Y, Yin Y, Li K, Li S. Breast cancer multi-classification from histopathological images with structured deep learning model. Scientific reports. 2017 Jun 23;7(1):4172.

[33] Nawaz M, Sewissy AA, Soliman TH. Multi-class breast cancer classification using deep learning convolutional neural network. Int. J. Adv. Comput. Sci. Appl. 2018 Jun 1;9(6):316-32.

[34] Wang Y, Acs B, Robertson S, Liu B, Solorzano L, Wählby C, Hartman J, Rantalainen M. Improved breast cancer histological grading using deep learning. Annals of Oncology. 2022 Jan 1;33(1):89-98.

[35] Senousy Z, Abdelsamea MM, Gaber MM, Abdar M, Acharya UR, Khosravi A, Nahavandi S. MCUa: Multi-level context and uncertainty aware dynamic deep ensemble for breast cancer histology image classification. IEEE Transactions on Biomedical Engineering. 2021 Aug 30;69(2):818-29.

[36] Zhang YD, Satapathy SC, Guttery DS, Górriz JM, Wang SH. Improved breast cancer classification through combining graph convolutional network and convolutional neural network. Information Processing & Management. 2021 Mar 1;58(2):102439.

[37] Ting FF, Tan YJ, Sim KS. Convolutional neural network improvement for breast cancer classification. Expert Systems with Applications. 2019 Apr 15;120:103-15.

[38] Liu K, Kang G, Zhang N, Hou B. Breast cancer classification based on fully-connected layer first convolutional neural networks. IEEE Access. 2018 Mar 20;6:23722-32.

[39] Sharma S, Mehra R. Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images—a comparative insight. Journal of digital imaging. 2020 Jun;33:632-54.

[40] Patel V, Chaurasia V, Mahadeva R, Patole SP. GARL-Net: Graph Based Adaptive Regularized Learning Deep Network for Breast Cancer Classification. IEEE Access. 2023 Jan 25;11:9095-112.

[41] Alirezazadeh P, Dornaika F. Boosted Additive Angular Margin Loss for breast cancer diagnosis from histopathological images. Computers in Biology and Medicine. 2023 Nov 1;166:107528.

[42] Chhipa PC, Upadhyay R, Pihlgren GG, Saini R, Uchida S, Liwicki M. Magnification prior: a self-supervised method for learning representations on breast cancer histopathological images. InProceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision 2023 (pp. 2717-2727).

[43] Xu C, Yi K, Jiang N, Li X, Zhong M, Zhang Y. MDFF-Net: A multi-dimensional feature fusion network for breast histopathology image classification. Computers in Biology and Medicine. 2023 Oct 1;165:107385.

[44] Ahmed F, Abdel-Salam R, Hamnett L, Adewunmi M, Ayano T. Improved Breast Cancer Diagnosis through Transfer Learning on Hematoxylin and Eosin Stained Histology Images. arXiv preprint arXiv:2309.08745. 2023 Sep 15.

[45] Vesal S, Ravikumar N, Davari A, Ellmann S, Maier A. Classification of breast cancer histology images using transfer learning. InImage Analysis and Recognition: 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings 15 2018 (pp. 812-819). Springer International Publishing.

[46] Vo DM, Nguyen NQ, Lee SW. Classification of breast cancer histology images using incremental boosting convolution networks. Information Sciences. 2019 May 1;482:123-38.

[47] Mehra R. Breast cancer histology images classification: Training from scratch or transfer learning?. ICT Express. 2018 Dec 1;4(4):247-54.

[48] Nazeri K, Aminpour A, Ebrahimi M. Two-stage convolutional neural network for breast cancer histology image classification. InImage Analysis and Recognition: 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings 15 2018 (pp. 717-726). Springer International Publishing.

[49] Koné I, Boulmane L. Hierarchical ResNeXt models for breast cancer histology image classification. InImage Analysis and Recognition: 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings 15 2018 (pp. 796-803). Springer International Publishing.

[50] Xie L, Li C, Wang Z, Zhang X, Chen B, Shen Q, Wu Z. SHISRCNet: Super-resolution And Classification Network For Low-resolution Breast Cancer Histopathology Image. arXiv preprint arXiv:2306.14119. 2023 Jun 25.

[51] Sun K, Chen Y, Bai B, Gao Y, Xiao J, Yu G. Automatic Classification of Histopathology Images across Multiple Cancers Based on Heterogeneous Transfer

Learning. Diagnostics. 2023 Mar 28;13(7):1277.

[52] Kode H, Barkana BD. Deep Learning-and Expert Knowledge-Based Feature Extraction and Performance Evaluation in Breast Histopathology Images. Cancers. 2023 Jun 6;15(12):3075.

[53] Guleria HV, Luqmani AM, Kothari HD, Phukan P, Patil S, Pareek P, Kotecha K, Abraham A, Gabralla LA. Enhancing the Breast Histopathology Image Analysis for Cancer Detection Using Variational Autoencoder. International Journal of Environmental Research and Public Health. 2023 Feb 27;20(5):4244.

[54] Vetrithangam D, Shruti P, Arunadevi B, Himabindu R, Kumar PN, Kumar AR, Hafiar H, Kadiyono Al, Kholidil A, Prastowo AA, Sofyan D. optimum feature selection based breast cancer prediction using modified logistic regression model. journal of theoretical and applied information technology. 2023 apr 30;101(8).

[55] Vetrithangam D, Senthilkumar V, Neha A, Naresh P, Kumar Ms. coronary artery disease prediction based on optimal feature selection using improved artificial neural network with meta-heuristic algorithm. Journal of Theoretical and Applied Information Technology. 2022 Dec 31;100(24).

[56] Vetrithangam V, Devi A, Aggarwal S. Multi disease prediction based on combined deep reinforcement Boltzmann machines. InAIP Conference Proceedings 2022 Oct 21 (Vol. 2555, No. 1). AIP Publishing.

[57] Vetrithangam D, Satve PP, Kumar JR, Anitha P, Vidhya S, Saini AK. prediction of pneumonia disease from x-ray images using a modified resnet152v2 deep learning model. Journal of Theoretical and Applied Information Technology. 2023 Sep 15;101(17).

[58] Vetrithangam D, Ebin PM, Nareshkumar P, Arunadevi B, Fathima A, Kumar AR. Enhanced Vgg-19 Model for Accurate Brain Tumor Prediction. In2023 International Conference on Network, Multimedia and Information Technology (NMITCON) 2023 Sep 1 (pp. 1-6). IEEE.

[59] Chakraborty A, Vetrithangam D. ExRAN: Deep Ensemble Majority Voting using Transfer Learning for Brain tumor Identification from Magnetic Resonance Imaging. In2023 International Conference on Inventive Computation Technologies (ICICT) 2023 Apr 26 (pp. 130-134). IEEE.

[60] Das P, Vetrithangam D, Krishnan GH, Kumar A, Reddy RV. Deep Learning Techniques Revolutionizing Biomedical Applications: Arrhythmia Detection, Cardiac Sensed Signals, and Cell-Free Synthetic Biology. InApplications of Synthetic Biology in Health, Energy, and Environment 2023 (pp. 68-91). IGI Global.

Journal of Theoretical and Applied Information Technology
31st December 2023. Vol.101. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

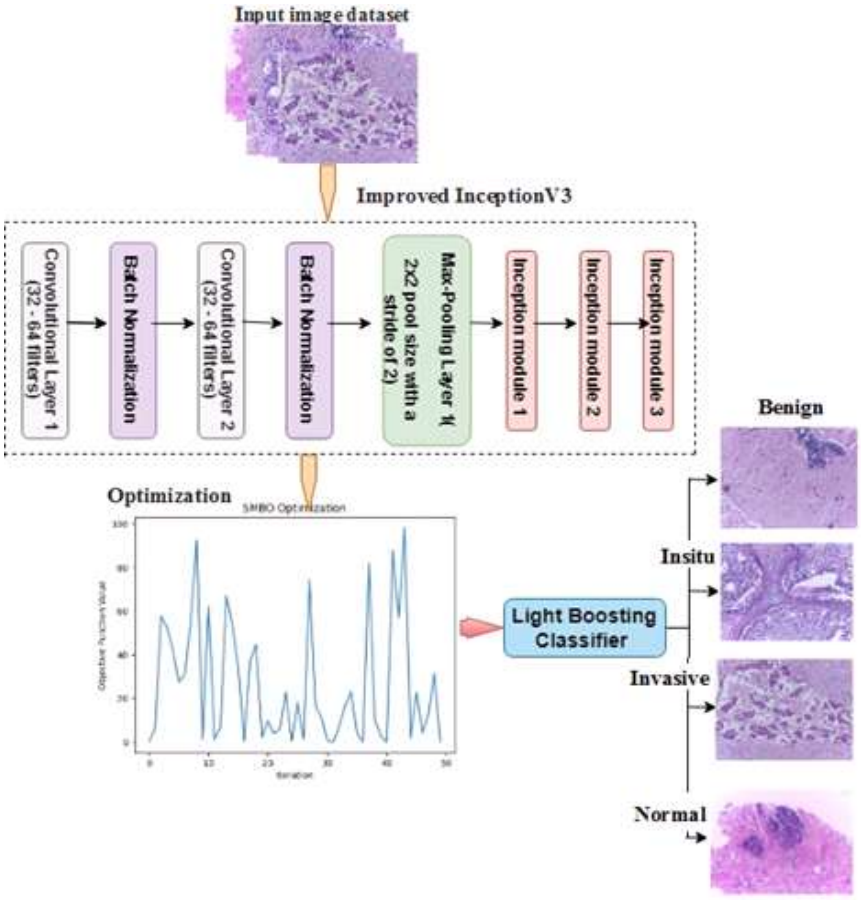

Figure 1: Proposed Architecture of the OLGV3 Net Classifier

Journal of Theoretical and Applied Information Technology
31st December 2023. Vol.101. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

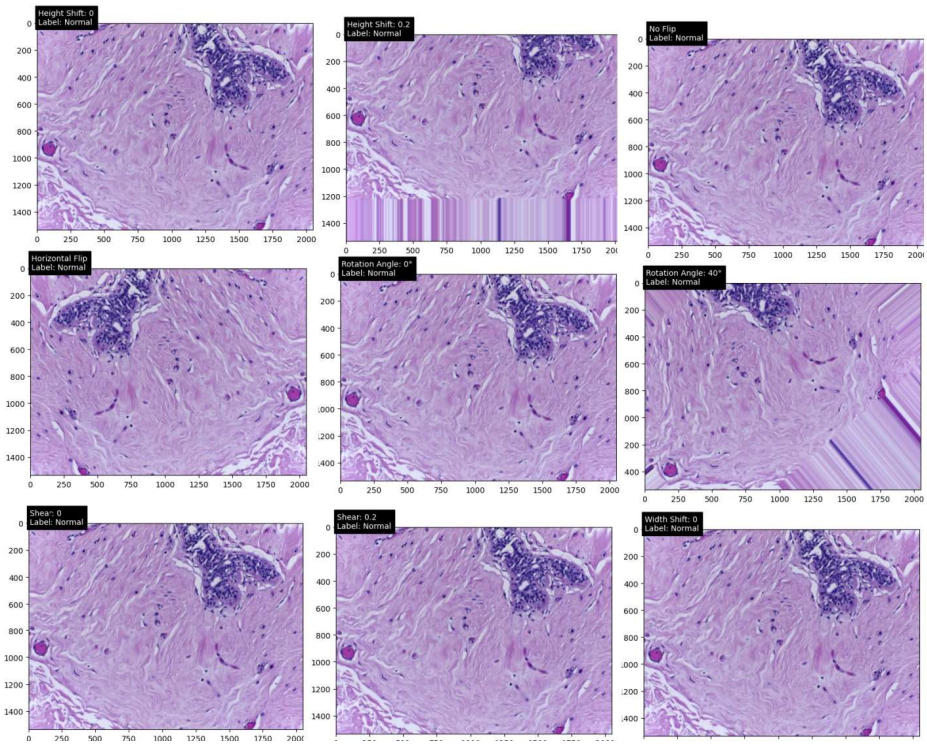

Figure 3: Sample Preprocessed BACH Dataset Images

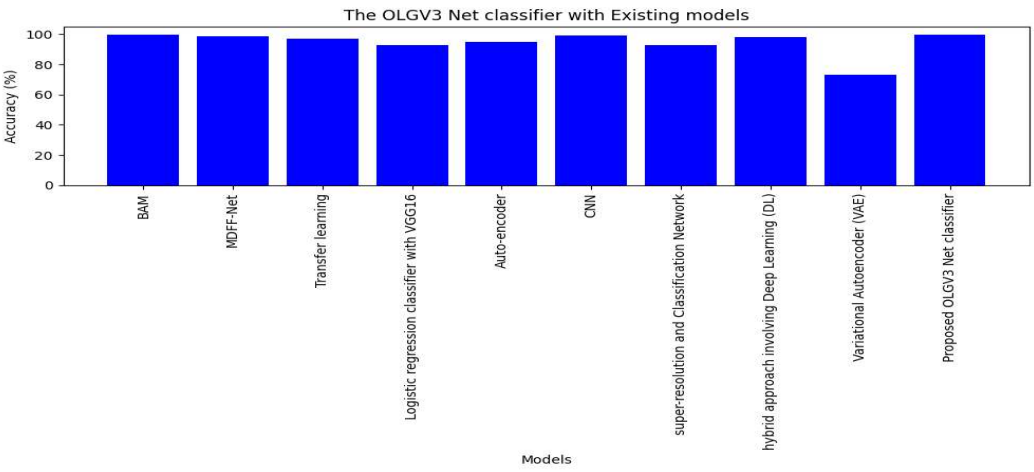

Figure 9: Comparison of the OLGV3 Net Classifier with existing models based on accuracy