[1]Salini S Nair

M. Subaji[2*]

Grading Breast Ductal Carcinomas using Spatial Activation Map Coupled Vision Transformers

Abstract: - Invasive Ductal Carcinoma (IDC) is the most common form of breast cancer. Accurate grading of IDC is crucial for determining patient prognosis and treatment options. However, subtle differences between grades present a challenge for traditional classification methods. This paper proposes a novel deep-learning architecture for robust IDC grading. The proposed approach tackles the challenge of low inter-class variance by incorporating a pre-processing step followed by a feature extraction pipeline. The pipeline utilizes a sparse autoencoder to capture low-level features and a Vision Transformer (ViT) for high-level feature extraction. The extracted features are then fed into a dense neural network for classification into three distinct grades. The paper compares the performance of our proposed architecture with Convolutional Neural Networks (CNNs), transfer learning, and ViT alone. The proposed method achieves a superior accuracy of 99% on unseen test data, demonstrating its effectiveness in overcoming the limitations of traditional classification methods for grading fine-grained variations in IDC.

Keywords: demonstrating, incorporating, architecture, IDC, traditional

I. INTRODUCTION

Cancer, a multifaceted disease characterized by uncontrolled cell growth, manifests in various forms across the human body. Among these, breast cancer, particularly Breast IDC [2], as described in Figure 1, stands as one of the most prevalent and menacing types, affecting millions of women globally. Its impact is profound, not only due to its prevalence but also because of its potential to metastasize, spreading beyond the breast tissue. Within the realm of breast cancer diagnosis and treatment, grading plays a pivotal role. Grading determines the aggressiveness and severity of the cancer cells, guiding clinicians in formulating tailored treatment plans. This process involves evaluating the microscopic features of cancer cells and tissues to categorize them into different grades based on their abnormality and differentiation. One of the primary objectives of cancer grading is to stratify tumors [1] into distinct categories that reflect their potential for growth, spread, and response to therapy. Typically, cancer grading systems assign tumors a numerical or descriptive grade, with higher grades indicating more aggressive behavior and poorer prognosis. For instance, in the widely used Gleason grading system [23] for prostate cancer, tumors are assigned a score ranging from 1 to 5 based on the patterns of tumor growth observed under the microscope.

Grading also serves as a critical tool for standardizing communication among healthcare professionals and researchers, facilitating the comparison of treatment outcomes and the development of new therapeutic approaches. Cancer grading presents several challenges despite its significance, including interobserver variability [21] and subjective interpretation of histological features. Efforts to address these challenges have led to the development of standardized grading systems and the incorporation of molecular markers [9] and advanced imaging techniques into the grading process. Figure 2 shows different grades [28] of Breast IDCs.

[1] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore 632014, India

Email: salinis.nair2015@vit.ac.in

[2*]VIT Business School and Institute for Industry and International Programme, Vellore Institute of Technology, Vellore 632014, India

***Corresponding Author:**

*Email: msubaji@vit.ac.in

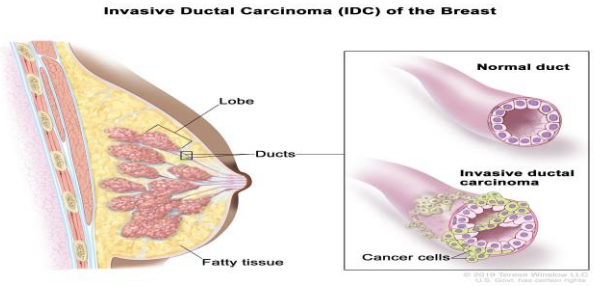

Fig. 1. Breast invasive ductal carcinomas

In recent years, there has been growing interest in leveraging artificial intelligence and machine learning algorithms to enhance the accuracy and reproducibility of cancer grading [26]. These technologies hold the potential to automate the grading process, reduce variability, and uncover novel prognostic factors, ultimately improving patient stratification and personalized treatment approaches. Artificial Intelligence (AI) and Machine Learning (ML) techniques have revolutionized

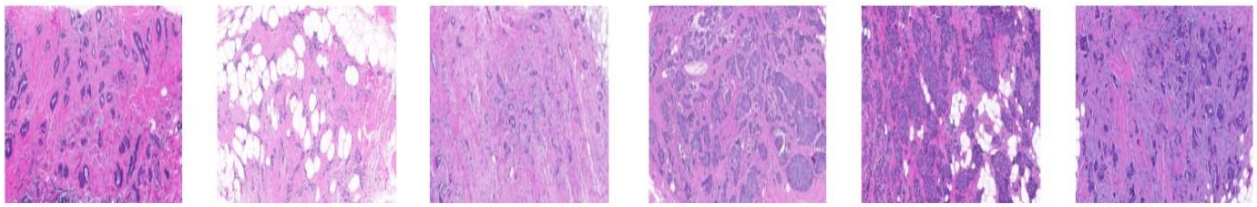

Fig. 2. Invasive ductal carcinomas pathological images of different grades

cancer grading [10] by providing powerful tools for automated, standardized, and accurate assessment of tumors. CNNs are particularly adept at extracting features from medical images, such as histopathological slides [14] and MRI scans, enabling precise grading based on visual characteristics. Support Vector Machines (SVM) and Random Forests, both supervised learning methods, excel in classifying tumor grades by analyzing diverse data sources, including imaging data and molecular profiles. Deep learning architectures like Recurrent Neural Networks (RNNs) and Generative Adversarial Networks (GANs) offer additional capabilities, capturing temporal dependencies and generating synthetic samples [22] for comprehensive analyses. Feature selection and dimensionality reduction techniques, such as Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE) [18], help identify informative features and improve model efficiency. Transfer learning enables the adaptation of pre-trained models to specific grading tasks, while clustering algorithms like k-means clustering aid in identifying distinct tumor subtypes. Together, these techniques enhance the efficiency, accuracy, and standardization of cancer grading, facilitating more informed clinical decision-making and ultimately improving patient outcomes.

In this paper, we delve into the intricacies of cancer grading, particularly Breast IDC, exploring the principles underlying current grading systems, their challenges, and the emerging role of AI and ML in revolutionizing cancer grading practices. Through a comprehensive review of the literature and case studies, we aim to shed light on the importance of accurate and reproducible cancer grading in optimizing patient care and advancing cancer research.

Major Contributions

1) Addresses the grading of very particular breast cancer variants, IDC
2) Modeled the grading problem as a classification using novel architectures.
3) Utilized the potential of Sparse AutoEncoder and Vision Transformer for feature extraction.

The rest of the paper describes the literature review that critically analyses the state-of-the-art techniques in breast cancer grading, followed by the preliminary that describes the fundamental theories involved in the architecture.

The methodology section elaborates on the architecture proposed, and the results and discussion section show the effectiveness of the proposed architecture through critical analysis and evaluation.

II. LITERATURE REVIEW

Paul et al. [20], complicated manual mitotic counting is required to grade histopathological malignancies. We provide an automatic detection technique based on random forest categorization and area morphological scale space. On more than 450 photos, our method improves the F1 score by 12%. We build relative entropy maximized scale space for precise segmentation using edge-preserving filters. An experimental study validates the superiority of our method. Among our future ambitions are submitting entries to medical imaging challenges and using our edge-preserving filter for Synthetic-Aperture Radar (SAR) image processing. Initial SAR experiments are encouraging. One exciting avenue for future investigation is generalized entropy-scale space.

Dalle et al. [3], breast cancer grading, a crucial clinical practice, burdens pathologists with 100 cases daily. This paper proposes an automatic grading method combining Nottingham criteria. Tubule formations are detected in low-resolution images, while individual cells are classified in high-resolution frames. Test results align closely with pathologists' grades. Encouraged by these findings, an automatic grading system can assist pathologists efficiently.

Tabar et al. [27], using data from 1,973 breast tumors in women aged 40–69, investigated histological type's impact on prognosis and sojourn time. Ductal carcinoma in situ, invasive ductal carcinoma (grade 1), mucinous carcinoma, and tubular carcinoma correlated with favorable survival. Grade 3 ductal carcinoma was linked to poor survival, while grade 2 ductal carcinoma, lobular, and medullary carcinoma showed intermediate survival. Sojourn time was shorter for lobular and medullary carcinoma (50–69 age group) and longer for grade 1 ductal carcinoma and tubular carcinoma. Evidence supported potential dedifferentiation, particularly in ductal tumors (40–54: 91%, 55–69: 38%).

Silverstein et al. [25], modern techniques are compiled in the literature review on image-detected breast cancer detection and therapy, which experts in the area wrote. Prominent medical experts make contributions to improvements in the treatment of breast cancer. The synopsis of the abstract emphasizes image-detected breast cancer: state-of-the-art diagnosis and therapy utilizing innovative methods. The technique combines cutting-edge treatment strategies with an image-detected breast cancer diagnosis.

Naik et al. [16], an automated approach for identifying and classifying nuclear and glandular structures in histopathology images is presented in this research. It combines data from several scales and estimates likelihood pixel-by-pixel using a Bayesian classifier for object boundary identification, template matching, and a level-set method. Structural limitations ensure the accurate identification of structures of interest. The method shows effectiveness in classifying specimens from breast histology that are benign or malignant, as well as in grading prostate and breast cancer. Comparisons between the classification accuracies for different cancer grades, histological distinctions, and manual segmentation findings indicate strong performance. The method's evaluation of a more significant image cohort will be the main focus of future work.

Henson et al. [8] evaluate the relationship between histologic grade, stage of disease, and outcome in a large cohort of breast cancer patients. The study demonstrates the importance of considering both histologic grade and stage of disease in predicting patient outcomes. It suggests the development of a prognostic index based on these factors to improve outcome prediction in breast cancer patients.

Ma et al. [13] introduce a five-gene Molecular Grade Index (MGI) and evaluate its prognostic utility in early-stage breast cancer. The study demonstrates that MGI and HOXB13:IL17BR provide improved prognostic information compared to traditional histologic grades. The authors propose the combined use of MGI and HOXB13:IL17BR for enhanced risk assessment and treatment planning in early-stage breast cancer patients.

Pan et al. [19], this paper investigates the survival outcomes of breast cancer patients in Ottawa, Canada, considering factors such as age, stage, histology, grade, and treatment. The study highlights the significant influence of age, stage, treatment, and grade on patient outcomes. It emphasizes the importance of histologic grade, particularly in conjunction with the stage of disease, for predicting patient prognosis and guiding treatment decisions.

Elston et al. [5], this paper evaluates the prognostic value of histological grade in breast cancer using a modified grading technique. The study demonstrates a strong correlation between histological grade and patient prognosis, suggesting that consistent and objective grading protocols can provide important prognostic information. The authors advocate for the inclusion of histological grade in multifactorial prognostic indices to stratify breast cancer patients for appropriate therapy.

Westenend et al. [29], this paper investigates the prognostic significance of tumor size, vascular invasion, and histologic grade in predicting distant metastasis in stage I breast cancer. The study highlights the importance of considering histologic grade, tumor size, and vascular invasion for predicting metastatic risk. It suggests the potential use of histologic grade to distinguish between early and late metastasis in breast cancer patients.

Ogston et al. [17] evaluate a new histological grading system to assess the response of breast cancers to primary chemotherapy and its prognostic value. The study demonstrates a significant correlation between histological response and patient survival, suggesting the potential use of this grading system for predicting outcomes in breast cancer patients undergoing primary chemotherapy.

Gandhi et al. [7], this paper investigates the correlation between Robinson's cytological grading and Elston and Ellis' Nottingham modification of Bloom Richardson score for breast carcinoma. The study demonstrates a correlation between cytological and histological grading, highlighting the prognostic value of both grading systems in breast cancer.

Liu et al. [12], this paper proposes a classifier-combined method based on Dempster-Shafer theory for grading breast cancer. The study demonstrates the effectiveness of combining classifiers using Dempster-Shafer's theory to improve grading accuracy. It suggests the potential use of this method for enhancing breast cancer prognostication and treatment planning.

Koteles et al. [11], this paper compares¨ a deep-learning model with general pathologists' assessment for grading breast cancer. The study demonstrates the high precision and accuracy of the deep-learning model in grading breast cancer, suggesting its potential for enhancing grading reproducibility and accuracy.

Mehak et al. [15], this paper proposes an automated grading method using a multilayered autoencoder for breast cancer histopathology images. The study demonstrates the effectiveness of the proposed method in accurately grading breast cancer, highlighting its potential for improving grading efficiency and accuracy.

Xi et al. [30], this paper discusses the automated classification of breast cancer histologic grade using multiphoton microscopy and generative adversarial networks. The study demonstrates the feasibility of using deep-learning algorithms for automated breast cancer histologic grade classification, suggesting their potential for enhancing diagnostic accuracy and efficiency.

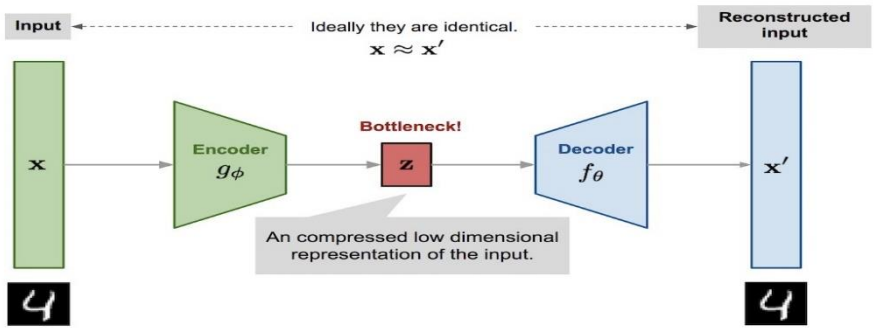

Fig. 3. Sparse autoencoder

Elsharawy et al. [4], this paper explores AI grading based on nucleolar prominence for refining breast cancer prognostic classification. The study demonstrates the significance of AI grading in predicting patient outcomes and guiding treatment decisions, suggesting its potential for enhancing prognostic precision and management strategies.

Dooijeweert et al. [24], this paper reviews the Nottingham Histologic Grading System for grading invasive breast carcinoma. It emphasizes the importance of accurate grading in breast cancer prognosis and discusses strategies for improving grading reproducibility and reliability.

III. PRELIMINARIES

A. Deep Sparse AutoEncoder

A deep sparse autoencoder is an artificial neural network used for unsupervised learning and dimensionality reduction, as shown in Figure 3. It belongs to the family of autoencoders, which are neural networks trained to copy their input to their output. The "sparse" aspect refers to imposing sparsity constraints on the hidden layer activations, meaning that only a few neurons are allowed to be active during the encoding process. Components:

1) Encoder: This part of the network maps the input data to a hidden representation. In a deep sparse autoencoder, there are typically multiple layers of encoding, each progressively extracting higher-level features from the input data. The sparsity constraint is often enforced at this stage, encouraging the activation of only a few neurons in each layer.
2) Decoder: The decoder takes the encoded representation and attempts to reconstruct the original input data. Like the encoder, the decoder typically consists of multiple layers, with each layer gradually transforming the encoded representation back into the original input space.
3) Sparsity Constraint: In a deep sparse autoencoder, a sparsity constraint is imposed on the activations of the hidden layers during training. This constraint encourages the network to learn a compact and efficient representation of the input data by promoting the activation of only a small subset of neurons in each layer.
4) Objective Function: The training objective of a deep sparse autoencoder typically involves minimizing a reconstruction error term, which measures the difference between the input data and the reconstructed output, along with a regularization term that enforces sparsity in the hidden layer activations as shown in equation 1.

Deep sparse autoencoders have applications in various domains, including feature learning, dimensionality reduction, data denoising, and anomaly detection. They are instrumental when dealing with high-dimensional data or when a compact representation is desired. The loss function of the autoencoder is composed of two different parts. The first part is the loss function (e.g., mean squared error loss), calculating the difference between input and output data. In contrast, the second term would act as a regularization term, which prevents the autoencoder from overfitting.

$$Obj = L(x,x\hat{}) + regularizer \tag{1}$$

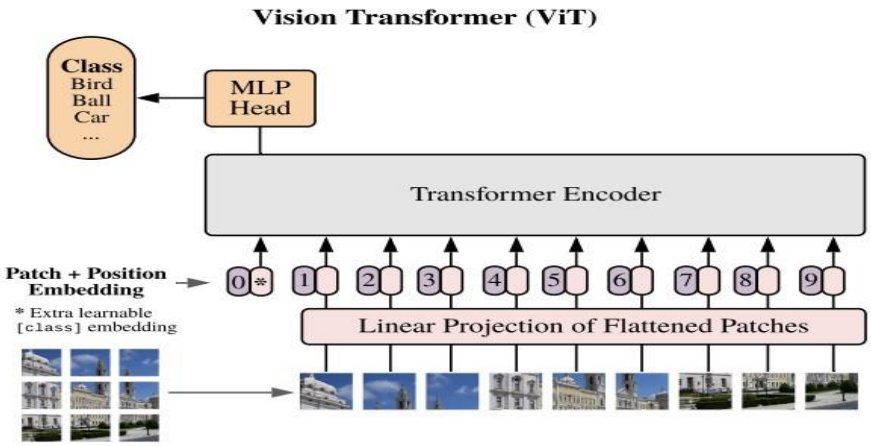

Fig. 4. Vision transformer

L1 and L2 regularization are widely used in machine learning and deep learning. L1 regularization adds the "absolute value of magnitude" of coefficients as a penalty term, while L2 regularization adds the "squared

magnitude" of coefficients as a penalty term. Consider that we have two loss functions, L1 and L2 as depicted in equation 2, which represent L1 regularization and L2 regularization, respectively.

$$L_1 = \|w\|,\ L_2 = w^2 \quad (2)$$

B. Vision Transformer

The ViT is a deep learning architecture that applies the transformer model, initially developed for Natural Language Processing (NLP), to image recognition tasks. Its overall architecture is described in Figure 4. Traditional CNNs have long been the dominant architecture for computer vision tasks. However, ViT represents a departure from this trend by demonstrating strong performance without convolutions.

1) Patch Embeddings: Instead of processing entire images directly, ViT divides the input image into fixed-size patches and flattens each patch into a one-dimensional vector. These patch embeddings serve as the input to the transformer model.
2) Positional Encodings: Like in NLP transformers, ViT requires positional information to be incorporated into the input embeddings since transformers do not inherently understand spatial relationships. Positional encodings are added to the patch embeddings to provide this spatial information.
3) Transformer Encoder: The patch embeddings and positional encodings are fed into a transformer encoder. The transformer encoder consists of multiple layers of self-attention mechanisms and feedforward neural networks, allowing the model to capture long-range dependencies in the input data.
4) Classification Head: A classification head is attached to make predictions at the end of the transformer encoder. Image classification tasks typically involve a simple classifier, such as a fully connected layer, followed by a softmax activation function.

Position embeddings are added to the patch embeddings to retain positional information. In Computer Vision, these embeddings can represent either the position of a feature in a 1-dimensional flattened sequence or a 2-dimensional position of a feature (1-dimensional: a sequence of patches works better, 2-dimensional: X-embedding and Y-embedding, Relative: define the relative distance of all possible pairs).

$$PE(pos,2i) = \sin\left(\frac{pos}{10000^{2i/dmodel}}\right) \quad (3)$$

$$PE(pos,2i+1) = \cos\left(\frac{pos}{10000^{2i/dmodel}}\right) \quad (4)$$

Position Embedding formula as per attention mechanism in equation (3) and (4), where i is the index of the position vector, dmodel represents input dimension and pos denotes relative position.

IV. METHODOLOGY

This section describes the methodology of the proposed work and is organized as the description of the proposed architecture and the specific implementation.

A. Proposed Architecture

The architecture of this work is shown in Figure 5. The architecture has two modules: one is the training, and the other is the inference. The training module explores a dataset of three grades of IDC to build a classification model. As the data has less variance among the classes, typical image classification models fail to achieve good performance. To handle this issue, the work proposed a novel architecture that utilizes Sparse Autoencoder and Vision transformers for feature extraction and then uses a dense network for the classification. A set of pre-processing techniques were employed to improve the model's performance. After the training, the model is used to infer and grade the pathological images of IDC.

B. Algorithm

Algorithm 1: Breast Cancer Image Grade Classification
Input:
I: Set of breast cancer images.

L: Corresponding labels for the images.

Output:

y^: Predicted class labels for the test images.

Steps:

Load and Pre-process Images:

1.1. Load images I from the dataset.

1.2. Rescale images to a fixed size of 224×224 pixels:

$I_{rescaled}$=resize(I,224,224)

1.3. Convert images to RGB model:

I_{RGB}=convert_to_RGB($I_{rescaled}$)

1.4. Apply histogram equalization:

$I_{equalized}$=histogram_equalization(I_{RGB})

1.5. Denoise images:

$I_{denoised}$=denoise(Iequalized)

1.6. Normalize pixel values to [0,1] range:

$I_{normalized}$=$I_{denoised}$/255.0

Split Dataset:

2.1. Split dataset into training and testing sets with a 70-30 ratio:

(I_{train},I_{test})=train_test_split($I_{normalized}$,0.7)

Encode Labels:

3.1. Encode class labels to integers:

$L_{encoded}$=label_encoding(L)

3.2. One-hot encode labels:

$L_{one-hot}$=one_hot_encoding($L_{encoded}$)

Sparse Autoencoder for Feature Extraction:

4.1. Define the encoder with multiple convolutional layers and SELU activation:

E(x)=SELU(Conv(MaxPool(x)))

4.2. Define the decoder with upsampling layers to reconstruct input:

D(z)=SELU(UpSample(Conv(z)))

4.3. Train the sparse autoencoder:

$F_{Autoencoder}$ =Autoencoder(x)=D(E(x))

Train on I_{train} for 80 epochs

Vision Transformer (ViT) for Feature Extraction:

5.1. Load pre-trained ViT model on ImageNet.

5.2. Extract features using ViT:

F_{ViT}=ViT(I_{train})

Combine Features:

6.1. Combine features from the sparse autoencoder and ViT:

$F_{combined}$=[$F_{Autoencoder}$,F_{ViT}]

Fully Connected Neural Network for Classification:

7.1. Define the fully connected neural network:

f(x)=ReLU(Dropout(BatchNorm(Dense(x))))

7.2. Compile the model with Adam optimizer and categorical cross-entropy loss:

model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])

7.3. Train the model:

model.fit($F_{combined}$,L_{onehot},epochs=150,batch_size=32)

Testing and Prediction:

8.1. Pre-process the input test image:

$I_{test\_processed}$=preprocess(I_{test})

8.2. Extract features using the trained models:

F_{test_conv}=Conv($I_{test\_processed}$)

F_{test_pool}=Pool($I_{test\_processed}$)

8.3. Combine features:

$F_{test\_combined}$=[F_{test_conv},F_{test_pool}]

8.4. Classify the image using the trained neural network:

y^=model.predict($F_{test\_combined}$)

C. Implementation

The methodology proposed starts by loading and pre-processing breast cancer images, organized into subdirectories representing different grades (Grade 1, Grade 2, and Grade 3) (https://www.kaggle.com/datasets/bolhasani/breast-cancer-grading) and BreakHis (Breast Cancer Histopathological Database (BreakHis) – Laboratório Visão Robótica e Imagem (ufpr.br)) [31]. Every image is then rescaled to a fixed size of 224 × 224 pixels and normalized to the RGB model in both the validation and training sets. Furthermore, certain image pre-processing steps such as histogram equalization, denoising, and resizing are used to perform image enhancement, remove noise, and any form of blurring to improve the overall clarity of the image or vision that may be significant to the application. Following this raw data pre-processing, the images are imported from their intended folder with appropriate resizing and conversion to maintain proper formatting for further processing.

This pre-processing pipeline prepares the images for the subsequent feature extraction and classification processes. The structured dataset enables efficient learning of multiple aspects of cancer grade. After completing the pre-processing steps, the dataset is divided into training and testing subsets using a 70-30 ratio. This division enables the model to be trained with a wide range of image input samples while having new data that the model has not seen before to test its ability to generalize on other instances. In addition, to ensure model updates are consistent and accurate, the pixel values of the images are divided by 255.0. This normalization process helps minimize the effect of pixel intensity variation, which may mislead the model during training and later during the testing phase, thus significantly enhancing the model's effectiveness in learning from the data. Before applying the class labels to the training of the neural network, an encoding approach is utilized to transform the string formats of the class labels into integers. This conversion helps portray categorical data in a numerical format that the model can comprehend during training.

Further, the label mapping method is used to achieve a correspondence between the string labels and their corresponding integer values. Such a mapping proves helpful in encoding the information and in making sure that there is a coherent transition from one label format to another. After that, the label mapping converts the string labels to integers, which are then transformed into one-hot encoded vectors using TensorFlow's utility functions. This encoding scheme enables the classification of breast cancer images based on their grades, leveraging the neural network's interpretation of categorical class labels. During the feature extraction, a sparse autoencoder model, consisting of an encoder and a decoder part, is carefully designed. The encoder section is proposed to contain multiple convolutional layers, where the filter size gradually increases, and the activation function is SELU. These convolutional layers are crucial in extracting detailed features inside the breast cancer images. In addition, max-pooling layers are incorporated at appropriate intervals to enforce the systematic downsampling of the input data to extract the essential features without compromising efficiency.

On the other hand, the decoder segment has a similar structure to the encoder segment but includes upsampling layers to reconstruct the input images from the encoded features. This maintained symmetry guarantees that the original image information can be reconstructed from the compressed representation obtained by the encoder. After that, the sparse autoencoder is trained carefully using the normalized training dataset for 80 epochs. At this training phase, a validation split is applied to carefully observe and fine-tune the model to become more efficient at feature extraction and reconstruction. Another model employed for feature extraction is a pre-trained ViT, which is used for extracting features from images. ViT model is trained on the ImageNet dataset, which consists of extracting features from the image at a high level. The features obtained from the sparse autoencoder and the ViT are then joined horizontally to get the final features of the training and testing data. This integration is built on the performances of both convolutional and transformer-based architectures.

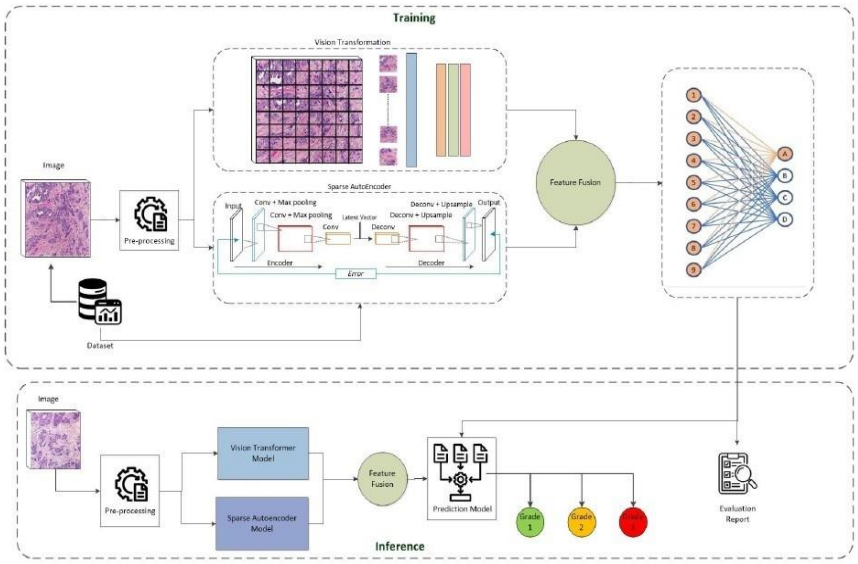

Fig. 5. Proposed architecture

TABLE I

Model summary of dense network

Layer (type)	Output Shape	Param
(Dense)	(None, 128)	198784
(Dropout)	(None, 128)	0
(BatchNormalization)	(None, 128)	512
(Dense)	(None, 64)	8256
(Dropout)	(None, 64)	0
(BatchNormalization)	(None, 64)	256
(Dense)	(None, 32)	2080
(Dropout)	(None, 32)	0
(BatchNormalization)	(None, 32)	128
(Dense)	(None, 3)	99
Total params: 210115 (820.76 KB)		
Trainable params: 209667 (819.01 KB)		
Non-trainable params: 448 (1.75 KB)		

A fully connected neural network is then devised to classify the three classes, and its model summary is shown in Table I. It consists of multiple layers with ReLU activation functions, dropout layers to avoid overfitting, and batch normalization layers to provide a stable solution. Finally, the model is fitted with an 'Adam' optimizer and 'categorical cross-entropy loss' for 150 epochs with a batch size 32. Training loss and training accuracy are used to show the model performance measures at various epochs. During the testing phase, when an image is inputted into the breast cancer image classification model, it undergoes pre-processing to align with the model's requirements. The pre-processed image is fed into the trained deep-learning model for feature extraction. Within this stage, convolutional and pooling layers analyze the image to extract relevant features reflecting various cancer grades. Subsequently, these extracted features are passed through the prediction model to classify the images.

V. RESULTS AND DISCUSSION

Table II demonstrates the high effectiveness of the classification model across three grades. The model achieves perfect precision for grade 2 and 3 and nearly perfect precision for grades 1, indicating minimal false positives. The recall is perfect for grades 1 and 2, ensuring no false negatives, while grade 3 has a slightly lower recall of 0.98, suggesting some instances are missed. The f1-scores are exceptionally high, with 0.99 for grade 1, 1.00 for

grade 2, and 0.99 for grade 3, reflecting a strong balance between precision and recall. The support values show that the dataset is reasonably balanced across the grades, enhancing the reliability of these metrics. Overall, the model exhibits robust performance, effectively distinguishing between the different grades with high accuracy.

TABLE II

Performance metrics of the classification model

Grade	precision	recall	f1-score	support
1	0.980000	1.000000	0.990000	41
2	1.000000	1.000000	1.000000	56
3	1.000000	0.980000	0.990000	53

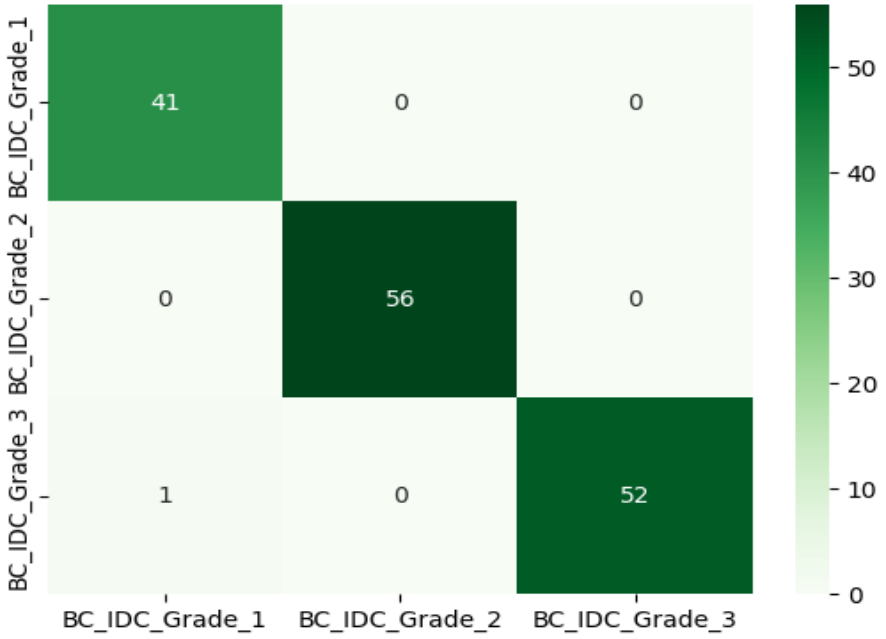

Fig. 6. Confusion matrix

Figure 6 illustrates the classification model's high accuracy across three grades. The model perfectly identifies all instances of Grade 1 and Grade 2 with 41 and 56 true positives, respectively, and no false positives or false negatives, reflecting its robustness. For Grade 3, the model correctly classifies 52 instances but misclassifies one instances (1 as Grade 1), resulting in minor inaccuracies. These results align with the performance metrics, showing excellent precision and recall for Grade 2 and slightly lower recall for Grade 1 and Grade 3. The confusion matrix confirms the model's effectiveness in distinguishing between grades with minimal errors.

The ROC curve in Figure 7 shows the performance of the classification model across three grades. Each grade's ROC curve reaches the top-left corner of the plot, indicating near-perfect classification capabilities. The Area Under the Curve (AUC) for each grade is 1.00, demonstrating that the model has excellent discriminatory power with no trade-off between the true and false positive rates. Additionally, the macro-average AUC of 1.00 confirms that the model performs uniformly well across all classes. This comprehensive performance indicates an exceptionally accurate model with strong predictive capability for all three grades, aligning with the earlier analysis of precision, recall, and the confusion matrix.

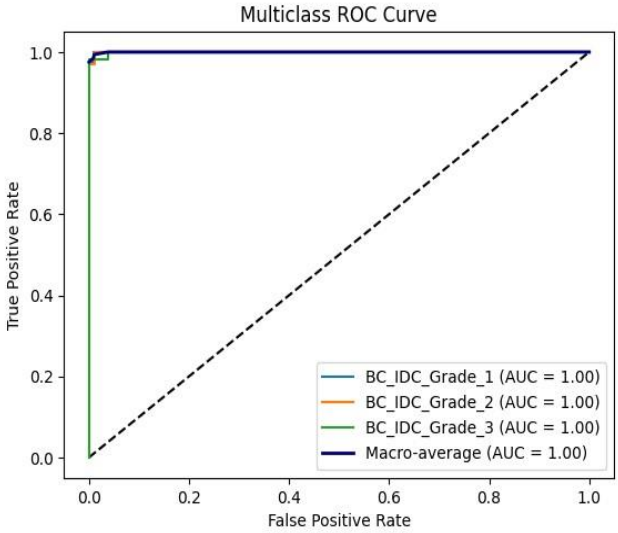

Fig. 7. ROC-AUC curve

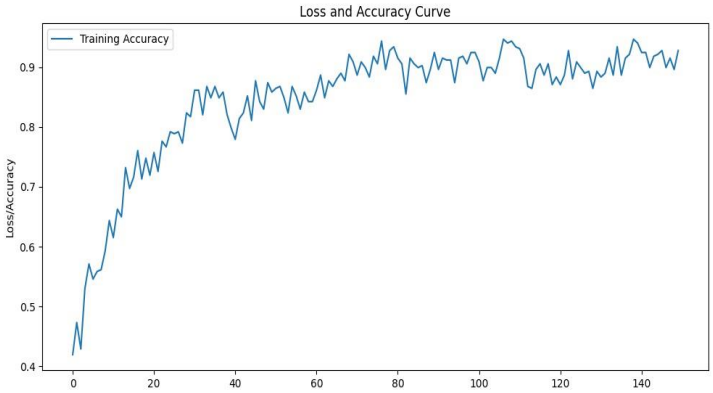

Fig. 8. DNN loss/accuracy curve

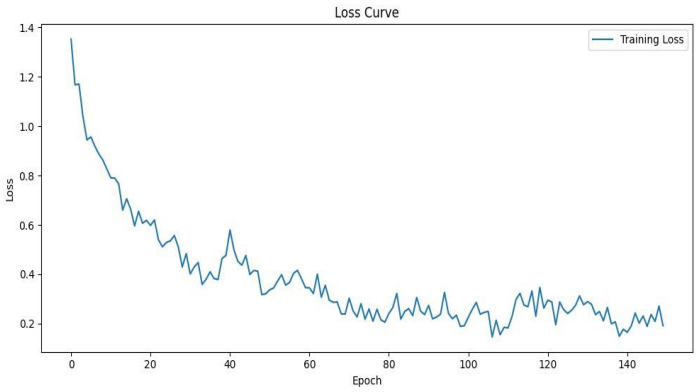

Fig. 9. DNN loss curve

Figure 8 illustrates the training accuracy of a classification model over 150 epochs, presenting the loss and accuracy curve. The curve shows a steady increase in accuracy from the initial epochs, starting around 0.4 and quickly rising, indicating the model's rapid learning phase. As the training progresses, the accuracy improves, albeit with some fluctuations, eventually stabilizing around 0.9. The minor oscillations observed in the later stages suggest slight variability but overall high performance. This trend indicates that the model effectively learns the underlying patterns in the training

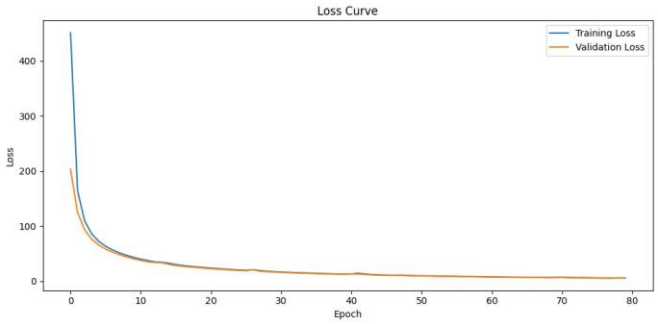

Fig. 10. Sparse autoencoder loss curves

TABLE III

Cross-validation results

Fold	Accuracy
1	99.2%
2	98.8%
3	100%
4	98.1%
5	99.9%
6	99.5%
7	99.0%
8	98.3%
9	99.1%
10	99.7%

TABLE IV

Performance comparison

	Proposed Model	CNN	Transfer Learning	Vision Transformer
Accuracy	**99.00%**	76.00%	82.00%	84.00%

data, achieving high accuracy and indicating robust training progression. The consistent improvement and final stabilization reflect the model's convergence and readiness for deployment or further validation with test data.

The loss refers to the difference between the model's predictions and the actual values. Figure 9 and Figure 10 illustrate the corresponding loss curves for the DNN and sparse autoencoder models, respectively. In the DNN curve, the training loss decreases over time, but it is not smooth, indicating some variability in the learning process. In contrast, the sparse autoencoder curve shows a smooth and consistent reduction in loss, suggesting that the sparse autoencoder model is steadily improving its performance as it is trained on more data.

The Table III displays the accuracy outcomes derived from a ten-fold cross-validation procedure, illustrating how the model performs across various dataset segments. Accuracy rates span from a minimum of 98.1% (observed in fold 4) to a maximum of 100% (observed in fold 3), with most folds achieving accuracy levels surpassing 99%. The average performance remains consistently strong, suggesting the model can generalize effectively across diverse data partitions. Minor fluctuations in accuracy among folds imply inherent variability in performance, as expected due to disparities within each subset.

Table IV compares the accuracy of different models. The models evaluated are the proposed model, CNN, transfer learning, and ViT. The proposed model achieves the highest accuracy of 99.00%, followed by the ViT with an

accuracy of 84.00%. The transfer learning model performs slightly better than the CNN, with accuracies of 82.00% and 76.00%, respectively.

This analysis shows that the proposed model outperforms the other models in terms of accuracy. The ViT also demonstrates competitive performance, while transfer learning and CNN lag in accuracy. However, further analysis beyond accuracy, such as computational efficiency or generalizability, would provide a more comprehensive understanding of the model's overall performance.

VI. CONCLUSION

This research has demonstrated the power of deep learning architectures in overcoming the inherent challenges associated with grading IDC. By leveraging a combination of sparse autoencoders, ViTs, and dense neural networks, we achieved a remarkable accuracy of 99% on unseen data. This signifies a significant leap forward in the quest for ever-more precise cancer diagnosis. However, pursuing perfect accuracy in cancer diagnosis is not merely a technical endeavor. It represents a profound aspiration to navigate the inherent uncertainty accompanying the disease. IDC, like many cancers, presents a spectrum of presentations, blurring the lines between grades. Our pursuit of finer distinctions reflects a deep-seated human desire for control, for knowledge that empowers us to predict and potentially conquer a formidable foe. However, the nature of biological processes, with their inherent variability, reminds us of the limitations of absolute certainty. This research's success lies in achieving high accuracy and acknowledging the nuanced complexities of cancer biology. By embracing ever-more sophisticated tools, we can move closer to a future where uncertainty is not a barrier but an impetus for continued exploration and discovery. This journey towards precision medicine ultimately serves a greater philosophical purpose: to empower patients and clinicians in making informed decisions, ultimately fostering hope and a path toward improved patient outcomes.

REFERENCES

[1] Beton- und stahlbetonbau aktuell, "*Beton- und Stahlbetonbau aktuell 12/2022*," pp. 117:851–859, 2022.

[2] Arps DP, Healy P, Zhao L, Kleer CG, and Pang JC, "Invasive ductal carcinoma with lobular features: a comparison study to invasive ductal and invasive lobular carcinomas of the breast," *Breast Cancer Research and Treatment*, 138(3), pp. 719–726, 2013.

[3] J. R. Dalle, W. K. Leow, D. Racoceanu, A. E. Tutac and T. C. Putti, "Automatic breast cancer grading of histopathological images," 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, pp. 3052-3055,2008.

[4] Elsharawy KA, Gerds TA, Rakha EA, and Dalton LW, "Artificial intelligence grading of breast cancer: a promising method to refine prognostic classification for management precision," Histopathology, 79(2), pp. 187-199, 2021.

[5] Elston CW, and Ellis IO, "Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up," Histopathology, 19(5), pp. 403-410, 1991.

[6] Elston GN, "Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function," Cereb Cortex,13(11), pp. 1124-1138, 2003.

[7] Gandhi H, Maru A, Shah N, Mansuriya RK, Rathod G, and Parmar P., "Correlation of Robinson's Cytological Grading with Elston and Ellis' Nottingham Modification of Bloom Richardson Score of Histopathology for Breast Carcinoma," Maedica (Bucur), 18(1), pp.55-60, 2023.

[8] Henson DE, Ries L, Freedman LS, and Carriaga M., "Relationship among outcome, stage of disease, and histologic grade for 22,616 cases of breast cancer. The basis for a prognostic index." Cancer, 68(10), pp. 2142-2149, 1991.

[9] M. Kashani-Sabet, "Molecular markers in melanoma," *British Journal of Dermatology*, 170(1), pp. 31–35, 2014.

[10] Kryvenko ON, and Epstein JI., "Changes in prostate cancer grading: including a new patient-centric grading system," *The Prostate*, 76(5), pp. 427– 433, 2016.

[11] Köteles MM, Vigdorovits A, Kumar D, Mihai IM, Jurescu A, Gheju A, Bucur A, Harich OO, and Olteanu GE, "Comparative evaluation of breast ductal carcinoma grading: A deep-learning model and general pathologists' assessment approach", *Diagnostics*, 13(14), pp. 2326, 2023.

[12] Liu Z, Lin F, Huang J, Wu X, Wen J, Wang M, Ren Y, Wei X, Song X, Qin J, Lee EY, Shao D, Wang Y, Cheng X, Hu Z, Luo D, and Zhang N., "A classifier-combined method for grading breast cancer based on Dempster-Shafer evidence theory," *Quantitative Imaging in Medicine and Surgery*, 13(5), pp. 3288-3297, 2023.

[13] Ma XJ, Salunga R, Dahiya S, Wang W, Carney E, Durbecq V, Harris A, Goss P, Sotiriou C, Erlander M, and Sgroi D, "A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer," *Clinical cancer research*, 14(9):2601–2608, 2008.

[14] Babawale M, Gunavardhan A, Walker J, Corfield T, Huey P, Savage A, Bansal A, Atkinson M, Abdelsalam H, Raweily E, Christian A, Evangelou I, Thomas D, Shannon J, Youd E, Brumwell P, Harrison J, Thompson I, Rashid M, Leopold G, Finall A, Roberts S, Housa D, Nedeva P, Davies A, Fletcher D, and Aslam M, "Verification and Validation of Digital Pathology (Whole Slide Imaging) for Primary Histopathological Diagnosis: All Wales Experience," Journal of Pathology Informatics, 12(1), 2021.

[15] Mehak S. Ashraf M. U., Zafar R., Alghamdi A., Alfakeeh A, Alassery F., Hamam H., and Shafiq M, "Automated Grading of Breast Cancer Histopathology Images Using Multilayered Autoencoder. Computers, Materials and Continua," 71, pp. 3407-3423, 2021

[16] S. Naik, S. Doyle, S. Agner, A. Madabhushi, M. Feldman and J. Tomaszewski, "Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology," 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Paris, France, pp. 284-287, 2008.

[17] Ogston, Keith N, Iain D Miller, Simon Payne, Andrew W Hutcheon, Tarun K Sarkar, Ian Smith, A Schofield, and Steven D Heys. "A New Histological Grading System to Assess Response of Breast Cancers to Primary Chemotherapy: Prognostic Significance and Survival," The Breast, 12(5), pp. 320–327, 2003.

[18] Coghill, G. M., "Large margin distribution machine recursive feature elimination," 2017 4th International Conference on Systems and Informatics (ICSAI),2017.

[19] Pan SY, Johnson KC, Ugnat AM, Wen SW, and Mao Y; "Canadian Cancer Registries Epidemiology Research Group. Association of obesity and cancer risk in Canada," *American journal of epidemiology*, 159(3):259–268, 2004.

[20] Paul A, and Mukherjee DP, "Mitosis Detection for Invasive Breast Cancer Grading in Histopathological Images, *IEEE Transactions on Image Processing*," 24, pp. 4041–4054, 2015.

[21] Polderman KH, Jorna EM, and Girbes AR, "Inter-observer variability in APACHE II scoring: effect of strict guidelines and training," *Intensive Care Medicine*, 27(8), pp. 1365–1369, 2001.

[22] J. Priesnitz, C. Rathgeb, N. Buchmann, and C. Busch, "SynCoLFinGer: Synthetic contactless fingerprint generator," Pattern Recognition Letters, 157, pp. 127-134, 2022.

[23] Pudasaini S and Subedi N, "Understanding the gleason grading system and its changes," Journal of Pathology of Nepal, 9(2), pp. 1580–1585, 2019.

[24] van Dooijeweert C, van Diest PJ, and Ellis IO, "Grading of invasive breast carcinoma: the way forward", Virchows Arch., 480(1) pp. 33-43, 2022.

[25] Silverstein MJ, Lagios MD, Recht A, Allred DC, Harms SE, Holland R, Holmes DR, Hughes LL, Jackman RJ, Julian TB, Kuerer HM, Mabry HC, McCready DR, McMasters KM, Page DL, Parker SH, Pass HA, Pegram M, Rubin E, Stavros AT, Tripathy D, Vicini F, and Whitworth PW, "Image-detected breast cancer: state of the art diagnosis and treatment," *Journal of the American College of Surgeons*, 201(4), pp. 586–597, 2005.

[26] Stenkvist B, Westman-Naeser S, Vegelius J, Holmquist J, Nordin B, Bengtsson E, Eriksson O, "Analysis of reproducibility of subjective grading systems for breast carcinoma", *Journal of Clinical Pathology*, 32, pp. 979–985, 1979.

[27] Tabar L, Fagerberg G, Chen HH, Duffy SW, and Gad A, "Tumour development, histology and grade of breast cancers: prognosis and progression," *International Journal of Cancer*, 66(4), pp. 413–419, 1996.

[28] Weigelt B, Geyer FC, Horlings HM, Kreike B, Halfwerk H, and Reis-Filho JS. "Mucinous and neuroendocrine breast carcinomas are transcriptionally distinct from invasive ductal carcinomas of no special type", *Modern Pathology*, 22()11, pp. 1401–1414, 2009.

[29] Westenend PJ, Meurs CJ, and Damhuis RA, "Tumour size and vascular invasion predict distant metastasis in stage I breast cancer. Grade distinguishes early and late metastasis," Journal of Clinical Pathology, 58(2), pp. 196-201, 2005.

[30] G. Xi, Q. Wang, H. Zhan, D. Kang, Y. Liu, T. Luo, M. Xu, Q. Kong, L. Zheng, G. Chen, J. Chen, and S. Zhuo, "Automated classification of breast cancer histologic grade using multiphoton microscopy and generative adversarial networks," *Journal of Physics D: Applied Physics*, 56(1), 2022.

[31] Spanhol, F., Oliveira, L. S., Petitjean, C., Heutte, L., "A Dataset for Breast Cancer Histopathological Image Classification," IEEE Transactions on Biomedical Engineering (TBME), 63(7):1455-1462, 2016.