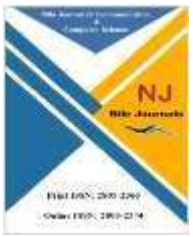
Nile Journal of Communication & Computer Science

Volume 9, June 2025

Journal Webpage: https://njccs.journals.ekb.eg



Attended CNN-LSTM for Prediction Bladder Cancer Recurrence and Response to Treatment

Taghreed S. Ibrahim*, M. S. Saraya, Ahmed I. Saleh, and Asmaa H. Rabie
Computers and Control Dept. faculty of engineering Mansoura University, Mansoura, Egypt

Abstract
One of the most prevalent cancers is bladder cancer, and non-muscle-invasive bladder cancer (NMIBC) has a high recurrence rate, therefore early detection is essential for efficient patient care. This work combines longitudinal clinical data and histological pictures to provide a deep learning-based method for bladder cancer recurrence prediction. Convolutional neural networks (CNNs), which is fine-tuned VGG16 and long short-term memory networks (LSTMs), which is stacked bidirectional GRU-LSTM were combined in a hybrid model that was improved by an attention mechanism to collect temporal and spatial data. Large-scale datasets were used for training and validation, and the model performed better than conventional techniques, achieving 90% accuracy, 88% precision, 85% recall, and 86% F1-measure. In accordance with clinical findings, the model identified vital factors such as tumor size, recurrence intervals, and treatment protocols. Attention maps, which highlighted important visual areas and temporal points, substantially improved interpretability. By facilitating individualized treatment planning, this method helps physicians to optimize therapeutic treatments and stratify patients according to recurrence risk.

Keywords: Bladder cancer, Deep Learning, CNN, LSTM, Attention Mechanism, Feature Fusion.

1. Introduction

Bladder cancer (BL) costs the European Union's healthcare system more than 50 billion euros a year and is the sixth most frequent disease worldwide. Surface transitional cell carcinomas account for more than 70% of initial bladder tumors, and their recurrence rates range from 50% to 75%. Although cystoscopy can aid in the diagnosis of BL, a biopsy, anesthesia, and pathologist categorization are necessary for confirmation. One of the most common cancers in the urinary tract, urinary bladder cancer, is more common in men [1]. It starts with the unchecked proliferation of mucosal cells, which can extend to other organs. BL comes in several forms, such as sarcoma, adenocarcinoma, small cell carcinoma, urothelial carcinoma, and squamous cell carcinoma. Originating from urothelial cells at the inner surface of the bladder, urothelial carcinoma, sometimes referred to as transitional cell carcinoma (TCC), is the most prevalent kind. The illness is marked by a high risk of recurrence, particularly in high-grade bladder tumors and carcinomas in situ (CIS) (61% in the first year, 78% in five years). It is vital to identify the right type of bladder cancer during diagnostic procedures since CIS forms frequently present as erythematosus flat lesions of the urinary bladder mucosa [2].

In its early stages, bladder cancer does exhibit symptoms similar to those of other disorders, including trauma, benign prostatic hyperplasia, kidney stones, interstitial cystitis, prostate cancer, and kidney cancer [3]. During the early stages of tumor development prior to clinical diagnosis, bladder cancer may be mistakenly identified as

infection or cystitis [4]. Similar symptoms to those of chronic cystitis and in situ bladder cancer include hematuria, frequent urination, lower abdominal pain, and, in some cases, urine incontinence. Cancer cells will seize the chance to spread if a patient has had antibiotic treatment for chronic cystitis. Interstitial cystitis and other bladder inflammatory illnesses might be difficult at times, although they are frequently misdiagnosed [5].

In the medical domain, artificial intelligence (AI) has demonstrated exceptional precision in forecasting the recurrence and development of diseases, especially in computer-aided diagnostic (CAD) and computer-aided predictive (CAP) systems. Machine learning (ML) is a significant area of artificial intelligence (AI) that has emerged to tackle medical issues [6, 7]. Training, validating, and testing are the three stages involved in training machine learning models or algorithms. The machine learning model adjusts to the input data during training and validation, producing outputs for regression or classification [8, 9]. The majority of medical machine learning applications, such as those used for non-muscle-invasive bladder cancer (NMIBC), are made for categorization through supervised learning. ML methods, including support vector machines (SVM), random forests (RF), artificial neural networks (ANN), and deep learning (DL), have been used in the development of several CAD and CAP systems.

These techniques help diagnose NMIBC, segment the bladder, and detect BC malignancies [10, 11]. AI has also been applied to individualized management, tumor staging and grading, and the prediction of recurrence rates, survival rates, and responsiveness to certain chemotherapies. Recurrence and risk categorization in machine learning algorithms have been enhanced by combining radiomic, clinical, pathological, imaging, and genetic indicators [12, 13].

DL methods have transformed the area of survival analysis throughout the last ten years. DL models that demonstrate how neural networks can handle censored data and nonlinear relationships, which increases the precision and effectiveness of survival analysis. Additionally, novel deep learning frameworks for survival analysis of recurring events with numerous competing hazards have been presented by researchers. Long Short-Term Memory (LSTM) networks are used in the frameworks to enhance the prediction of event timings and their causes. Convolutional neural network (CNN) architecture that has also been applied for image-based cancer detection. Several architectures are used, including VGG16, on blue-light cystoscopy images for the detection of BL [14, 15]. CNNs are deep learning systems that possess the ability to learn invariant characteristics. For various object identification tasks, including detection, segmentation, prediction, and classification, CNNs' filter banks, feature pooling layers, batch normalization layers, dropout layers, and dense layers cooperate to produce patterns. Throughout the training process, the input distribution varies in CNNs' multilayer hierarchies. In order to achieve improved performance across jobs, preprocessed inputs—such as those derived from the whitening process, etc.—are desired. Among the various variations of CNNs are some that provide shorter connections, which has benefits for feature circulation and offers significant hyperparameter reduction to create effective topologies [16].

Missing value scenarios have been effectively handled by recurrent neural networks (RNNs), including LSTM and gated recurrent units (GRUs). Thus, these machine learning models are potential tools for managing very large real-world datasets with irregular data, such as time-series data, where the interval between numerous tests may change, and single time-point values, where not all individuals have all values accessible. Because it has internal gating mechanisms to prevent the vanishing and exploding gradient calculation, LSTM in particular can be modified to take advantage of missing value patterns, time intervals, and complex temporal dependencies in irregular univariate and multivariate time series data. Because of these advantages, LSTM is a suitable approach for handling both missing data and lack of regularity in intervals [17, 18].

In this study, the proposed hybrid model (CNN-LSTM) is introduced. This novel strategy integrates CNN for feature extraction from histopathological images with LSTM for temporal pattern analysis using TCGA clinical data to predict BL outcomes like recurrence, progression, and response to the treatments. In this framework, TCGA dataset's histopathology images will be investigated with the pre-trained VGG16 model. The attention mechanism enhances the capability of pre-trained CNN models by giving attention to the most informative region

of the image and suppressing the influence of irrelevant features. Similar to the human visual system, it gives precedence to a portion of the critical area of the input data; this enhances the performance of CNNs in tasks such as classification, segmentation, and object detection, adding to their interpretability. On the other side, temporal data from TCGA is investigated by a stacked bidirectional GRU-LSTM model, which is great for analysis and prediction based on sequential or time-dependent data. In the context of BL prediction, LSTM networks are important to process temporal data: patient histories, biomarker trends, and treatment responses over time. When making predictions. In this work, LSTM networks are improved by the attention mechanism, which enables them to selectively concentrate on the most pertinent segments of the input sequence. Lately, the hybrid approach included a feature fusion layer, which serves as a cornerstone in modern predictive models, enabling them to make more informed and reliable predictions by leveraging diverse and complementary data types. The results of a hybrid model offer the clinician armamentarium for early prediction, precise risk profiling, and personalized treatment strategies. This model, powered by new opportunities offered by artificial intelligence and integrated data, will provide a perspective toward better patient outcomes and enhanced efficiency in healthcare delivery. The following are the research contributions of this paper:

- This model overcomes the limitations of individual modalities by taking the strengths from CNNs to analyze images and from LSTMs for temporal data.
- Enhanced attention mechanisms along with feature fusion further enhance the model's power for the task of predicting recurrence in BL cancer.
- Integrating multimodal data, it offers a comprehensive framework for precision medicine by helping clinicians make informed treatment decisions.
- It has a feature fusion mechanism that selects the most relevant features from each modality for decision-making by balancing the contributions of histopathological and temporal data.
- The model is designed as an end-to-end learning pipeline, where all the components are trained simultaneously, reducing the risk of information loss in feature extraction and integration.
- Combining spatial and channel attention by CBAM with temporal attention mechanisms provides insights about which features and time steps provide the most significant contribution to the predictions.
- Variability in histopathological image staining due to different protocols of tissue preparation and scanning has been handled for consistent and reliable feature extraction.

The rest of the current paper is organized as follows: Section 2 shows previous efforts about the prominence of breast cancer detection. Section 3 concentrates on the proposed CNN-LSTM model with attention mechanism techniques to predict BL cancer recurrence and response to therapies. Section 5 depicts the experimental result. Section 6 concludes the paper and illustrates the future work of the paper.

2. Related Work

This section will review previous endeavors on BL cancer recurrence prediction. BL cancer recurrence prediction and response to treatments are of utmost importance in improving patient outcomes and optimizing therapeutic strategies. BL cancer, especially NMIBC, has a high recurrence rate, which calls for the need to develop robust predictive models. Large, sophisticated deep learning models, mainly convolutional neural networks, have promised a great deal in analyzing such images for the estimation of recurrence risks. Combined with clinical data such as tumor stage, grade, patient demographics, and treatment history, these models result in better accuracy and reliability. These techniques also aid in the identification of biomarkers linked to treatment outcomes, such as the reaction to chemotherapy or intravesical Bacillus Calmette-Guérin (BCG) therapy.

In [19], a pathomics model (RDLPI), which is the prediction of non-muscle invasive bladder cancer recurrence using deep learning of pathology images. This model developed a deep learning model for predicting the recurrence of non-muscle-invasive bladder cancer using pathology images. Their study utilized convolutional neural networks (CNNs) to extract features from histopathological images and achieved high predictive accuracy. The integration of clinical data with imaging further enhanced model performance, indicating the importance of multimodal approaches. In [20], a deep learning pipeline called Bladder4Net is developed to categorize whole-slide histopathology images of BL cancer into two groups: low-risk (combined PUNLMP and low-grade tumors) and high-risk (combined invasive and high-grade tumors). This pipeline addresses the challenges of recognizing intrusive and PUNLMP classes by using four convolutional neural network (CNN)-based classifiers. In [21], this study used images of bladder tissue and artificial intelligence (PSBAI) to create an interpretable grading system. Two separate risk categories with varying outcomes in cases of high-grade bladder cancer were found. There were several molecular characteristics and gene alterations linked to the score system. By determining which patients require more molecular testing, this method helps save costs and facilitate clinical decision-making. In [22], this study used full-slide digitized histological pictures from two cohorts to create poorly supervised deep learning models for bladder cancer diagnosis and overall survival prediction in patients with muscle-invasive bladder cancer. Based on encouraging results, our models can help physicians diagnose bladder cancer accurately. They can also help patients with muscle-invasive bladder cancer make more customized treatment decisions by facilitating differential risk assessment. To further boost the quantity of information collected from pathological pictures, the region most pertinent to diagnosis or prognosis can be further examined.

In [23], using high-resolution 512 × 512-pixel samples, the suggested DCEAC model uses a completely unsupervised two-step learning process to distinguish between mild, infiltrative, and non-tumor patterns. By including a convolutional attention module, our methodology surpasses earlier clustering-based techniques by allowing the latent space's features to be refined before the classification phase. In [24], machine learning of quantitative morphological traits is used to try and predict early recurrence of NMIBC. To identify cancer atypia, structural, cellular, and nuclear atypia are often assessed. However, in this work, nuclear atypia is used and examined using feature extraction and classification using support vector machine and random forest machine learning methods because it is challenging to precisely quantify structural atypia from TUR specimens. In [25], the study analyzed CT detection outcomes in 60 bladder cancer patients using Western Blot technology and deep learning algorithms. Results showed algorithm-based CT detection was more accurate and reliable for clinical therapy and preoperative staging. The study also found that glucose ceramide synthase (GCS) alterations are linked to bladder cancer development and prognosis. Further investigation is needed to determine GCS efficacy in bladder cancer treatment.

In [26], in order to identify bladder tumors in clinical settings, this study employed convolutional neural networks. In the classification test of cystoscopic pictures, these three networks (LeNet, AlexNet, and GoogleNet) performed well despite having a very simple network design. The recognition efficiency of the deep learning system was on par with that of seasoned medical professionals. Based on the cystoscope, this study demonstrated the efficacy of the convolutional neural network for bladder tumor diagnosis. In [27], improved bladder cancer staging accuracy by CECT is demonstrated using a hybrid machine/deep learning model (HDML). This model will help with the proper clinical care of patients with bladder cancer, ultimately leading to better patient outcomes. Radiologists may be able to identify bladder cancer more precisely if they use radiomics to help interpret CT scans. In the end, this can improve patients' clinical results by enabling the prompt use of medical resources and consultation with urologists and oncologists. In [28], the study analyzed cystoscopic pictures of bladder cancers using a convolutional neural network based on mask regions. The sensitivity, specificity, and accuracy of the AI's diagnostic performance were assessed. With a DSC of 95.0%, the study discovered great sensitivity, specificity, and accuracy. With a 90% accuracy rate for carcinoma in situ, the AI demonstrated efficacy in the diagnosis of benign tumors and chronic non-specific inflammation.

In [29], comparing the deep learning-based model (DLRP-NMIBC) with clinical data or image data alone, the

combination of digital histopathology slides and clinical data improves the recurrence prediction (within 5 years). Recurrence prediction in bladder cancer patients is improved by utilizing deep learning to integrate histopathological pictures and patient record data. 359 and 281 individuals with recurrence rates of 27% and 63%, respectively, were included in the research. For 1- and 5-year recurrence predictions, the model that included clinical data with digital histopathology slide data exhibited better AUCs.

In [30], the deep neural learning cancer prediction model (DNLC) trains a deep neural network using genomic or clinical data samples, selects the best features from datasets using a deep network, and assesses the model's ability to detect cancer in its early stages. The model classifies five cancer datasets: leukemia, breast, squamous cell carcinoma, lung adenocarcinoma, and colon. There are training and testing sets inside the dataset. In [31], to improve performance in assessing recurrent occurrences with dynamic temporal information, this work combines the Cox proportional hazards model with the benefits of deep learning models for processing long-sequence data. This LSTM-Cox model also evaluates how well various models can extract and use characteristics from time-dependent clinical recurrence data. In [32], this research aims to characterize the current state of artificial intelligence and its future potential, with a particular with a focus on how AI might revolutionize the detection and treatment of bladder cancer, this study aims to identify the current state of AI and its promise for the future. Artificial intelligence (AI) has the potential to revolutionize healthcare by enhancing diagnosis, prognosis, and therapy of urological illnesses, including bladder cancer. Its subclasses, deep learning, machine learning, and artificial neural networks, can draw conclusions from large datasets and refine prediction models, reshaping traditional forecasting methods like nomograms. This review critically examines recent research on AI's potential in healthcare settings.

The proposed hybrid approach (CNN-LSTM) not only identifies high-risk patients for early intervention but also informs treatment personalization, reducing recurrence rates and optimizing therapeutic efficacy. Such advancements highlight the transformative potential of combining histological and clinical data for bladder cancer management.

3. The proposed Attended CNN-LSTM model

This section illustrates the novel approach, which integrates a pre-trained CNN model with a GRU-LSTM model. The hybrid approach is based on the attention mechanism technique, as seen in Fig. 1. Attention mechanisms strengthen CNNs and RNNs by enhancing interpretability, dynamic focusing, and dealing with input patterns that are complicated. It emphasizes spatial relevance in CNNs and solves the long-term dependency issues with temporal focus in RNNs. The attention mechanism has proven to be transformative across domains such as NLP, computer vision, and time-series analysis, enhancing the robustness, efficiency, and interpretability of models.

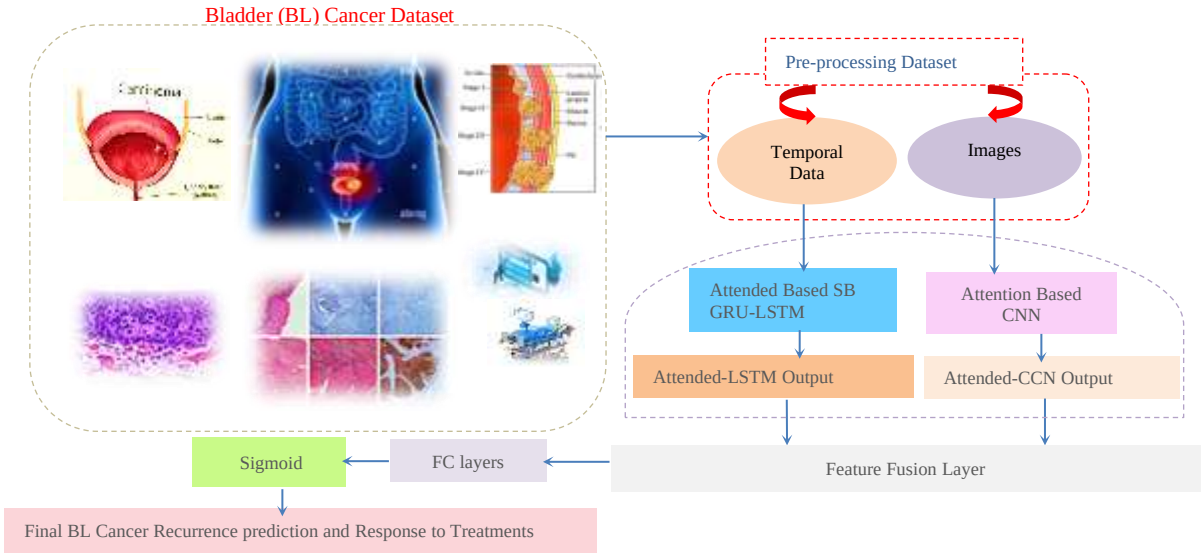


Fig.1.A Hybrid Model for BL Cancer Recurrence Prediction and Response to Treatments

3.1.Attention based Pre-trained CNN (VGG16) model

In this section, the pre-trained model VGG16 will be introduced as the first model in the novel hybrid approach to analyze histopathological images of the TCGA dataset. The structure of this model consists of preprocessing the dataset and then feeding it to the attended fine-tuned VGG16. Finally, the attended pre-trained VGG16 output is fed to the next layer of the proposed hybrid model, which is the feature fusion layer, as shown in Fig. 2.

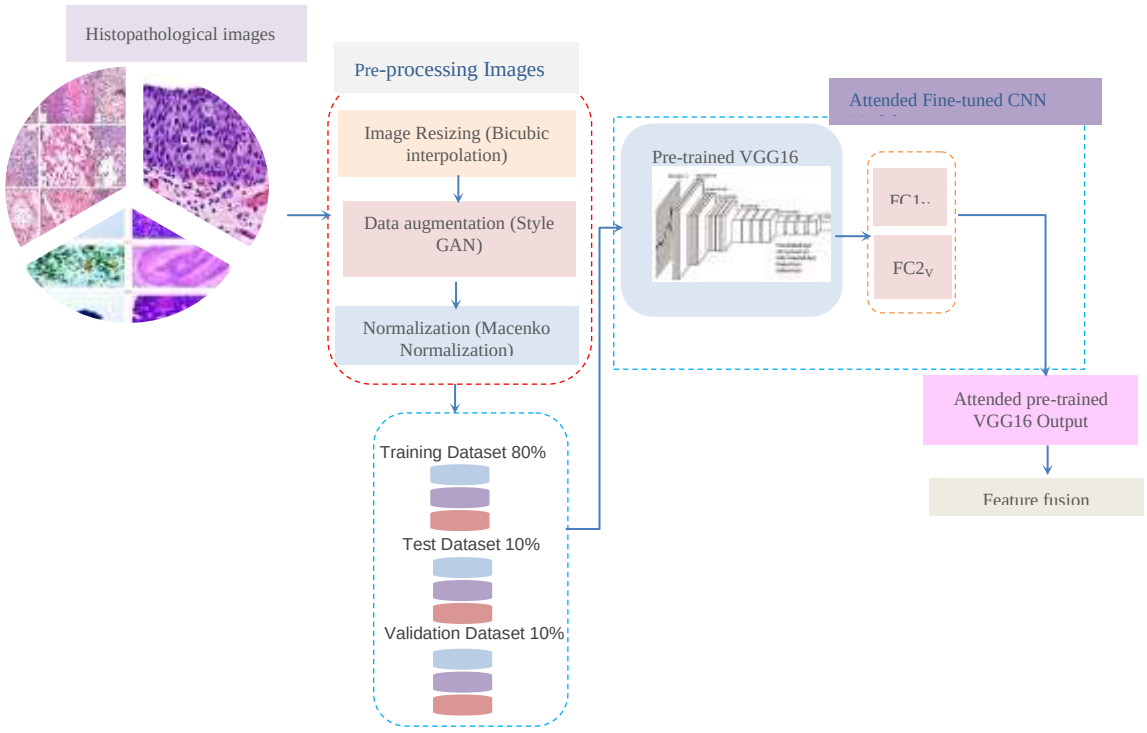


Fig.2. The proposed Pre-trained CNN model

3.1.1. Pre-processing Histopathological Images

The first step is a preprocessing and preparing dataset which is obtained from TCIA before feeding it to the VGG16.Pre-processing data is a crucial procedure before training the fine-tuned CNN model. In order to supply neural network tiers in the deep learning framework reliable characteristics data, preprocessing step is an essential approach. It consists of several steps, which are arranged as follows:(i) Resizing the images, which is a straightforward approach, consistently resizes the image to the necessary dimensions. Every image was shrunk to a size of 224 × 224 pixels. It can improve model generalization by reducing the image resolution, introducing the essential features to the model, and reducing the risk of overfitting to specific details or noise in the data. Implementing image resizing method by applying bicubic interpolation algorithm, which relies on the distance to each pixel, the nearest 16 pixels are used in a bicubic interpolation to calculate the new pixel value. This algorithm produces pictures with more smooth edges. It offers an appropriate ratio between preservation of detail and smoothness. (ii)Data augmentation by applying Style GAN, which can indeed be utilized to assist a fine-tuned CNN model in recognizing the intricate patterns found in histopathological images [33]. It can enhance image accuracy by expanding the type and amount of training data, diminishing overfitting, and enhancing generalization to unknown data. The extra images that Style GAN produces can strengthen your model's resistance to variations in real-world data. (iii) Stain Normalization (Macenko or Reinhard) Because staining procedures, preparation of tissue techniques, and scanning tools differ, histopathological images frequently have different colors [34]. In order to ensure that differences in staining do not impair model performance, stain normalization approaches particularly tackle this by normalizing the color distribution across all photos. Macenko is the best strategy for standardizing color variations in histopathology images in the TCGA dataset while maintaining significant tissue characteristics.

3.1.2. The Attended Fine-tuned VGG16

The pre-trained model depends on the transfer learning (TL) approach, so there is no starting from scratch. The use of a pre-trained model offers several advantages. Firstly, expensive computational power is the primary necessity of training huge models on large datasets. Secondly, it may take many weeks to train huge models. Finally, a pre-trained model can speed up convergence and improve in-network generalization. The structure of pre-trained VGG16 is introduced [35], as seen in Fig. 3. The VGG16 structure is presented with an added CBAM after convolution block 5, followed by a flatten layer and two custom dense blocks.

Convolutional layers' blocks 1-3 are frozen for the purpose of extracting generic features, avoiding overfitting, and lowering training computational overhead. These layers pull out low-level characteristics that are applicable to different applications and datasets, such as edges, textures, and patterns. Despite not being trainable, the batch normalization (BN) and convolutional layers in Blocks 1-3 are used in forward propagation. Fine-tuning Convolutional Blocks 4-5 and custom layers is crucial for high-level feature adaptation, attention enhancement, and custom dense layers. Blocks 4-5 extract task-specific semantic features, while adding Custom Blocks 5 allows the network to focus on spatial and channel-wise features. CBAM integrated both channel and spatial attention processes. It employs spatial attention to focus on crucial areas of the image after implementing channel attention to highlight significant feature channels. This dual strategy enabled the system to concurrently concentrate on the most pertinent channels and spatial region. For histopathological images, the CBAM method works extremely well since distinct features like texture or color, as well as their precise location, are significant. The forecasting performance as a whole may get better. High-level semantic characteristics that reflect intricate patterns like objects or forms are produced by Block 5. These features are refined by CBAM, which concentrates on pertinent portions before sending them to feature fusion or downstream dense layers. When merged, these characteristics can enhance LSTM-derived features, which are essential for classification tasks. The Flatten layer is a crucial part of the updated VGG16 architecture, preparing the Convolutional Bayes Algorithm (CBAM) output for input into dense layers. It converts multi-dimensional input into a 1D vector, bridging the gap between dense and convolutional layers. The Flatten layer retains feature values retrieved and refined by convolutional

layers and CBAM, protecting feature information. The output vector from the Flatten layer is used in the final job, providing necessary information for dense layers [36]. It is crucial after CBAM for further calculations.

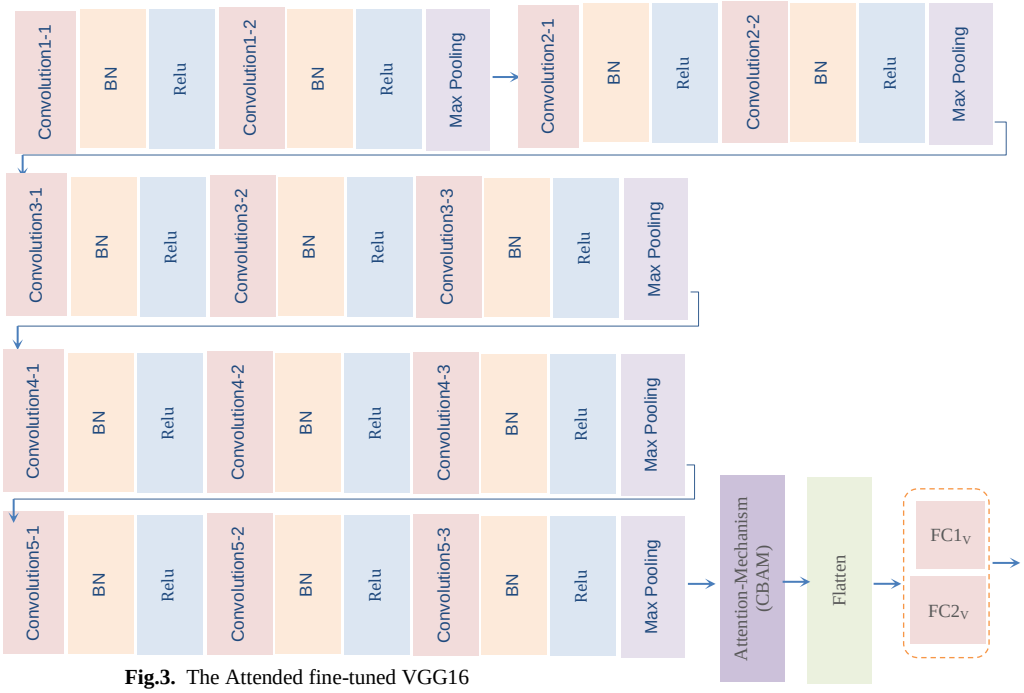


Fig.3. The Attended fine-tuned VGG16

Using dense serial blocks of data (FC_{ix}), the refined VGG16 algorithm is utilized. The size of the dataset and the dependability of the features determine whether FC layers are required to transform the basic characteristics into a more task-driven representation or if the current features are sufficient. Since extra layers and factors may make the design harder to understand, its interpretability and intricacy will be assessed. Consequently, the refined VGG16 framework has two adjacent blocks (FC_2) that are fully linked, where i = 2. Figure 4 illustrates that each of them is made up of dense, $Relu$ activation, batch-normalization, and dropout layers, respectively. The activation process, normalization of batches, dropping out layers, and extremely dense layers constitute the primary and secondary dense blocks, respectively, as seen in Fig. 4. There are (512) nodes in the first layer's dense (N) layer. The dense (N) layer in the second block has 128 nodes. The model known as VGG16 that has already been trained is subjected to the Adam optimizer. In the proposed hybrid model, the attended pre-trained VGG16 output will be sent into the feature fusion layer, where it will be mixed with the LSTM outputs.

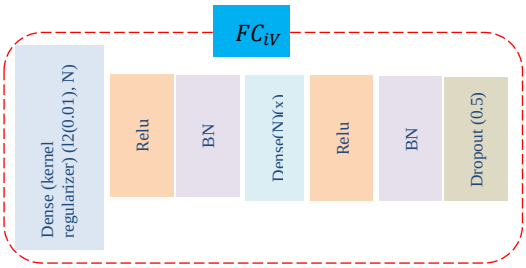


Fig.4. The Modified Dense layers of the EnhancedVGG16

3.2. A Stacked Bidirectional GRU-LSTM (SB GRU-LSTM) with Attention Mechanism

Once data are collected from TCGA, Pre-processing approaches are employed in order to acquire a balanced dataset. Pre-processing techniques contain (i) data collection, which aims to specify significant temporal data that varies over time for prediction tasks from the TCGA dataset, such as tumor size measurements, treatment histories, or biomarker values. For each individual, the data should be arranged chronologically, emphasizing the order of observations or occurrences, including baseline, follow-up appointments, and therapies. (ii) Handling Missing Data that involves recognizing missing data, particularly in the patient's time series statistics. There are three methods for interpreting numerical data: forward/backward filling, interpolation, and statistical imputation, which compensate for missing data in time steps and between known locations. (iii) Aligning time step, which is establishing standard time intervals for patient data, such as every three months, and resampling the data if necessary to align with these intervals. Sequence padding also ensures uniform duration across all sequences, even with varying time steps, by appending zero or unique values to the start or finish of the series. (v) feature scaling, which standardizes or normalizes characteristics. Normalization Adjust and reduce the characteristics to a range of 0 to 1 by applying min-max scaling. Standardization Assigns a mean of 0 and a standard deviation of 1 to the characteristics. (iv) Converting data into sequences that sort the preprocessed data into sequences, with each one representing the feature's temporal evolution for a particular patient. Finally, the input data format is (number of samples, time steps, number of features). Dividing the data into training: 80%, validation: 10%, and test: 10% to make sure that the model generalization is better to the new data, as shown in Fig. 5.

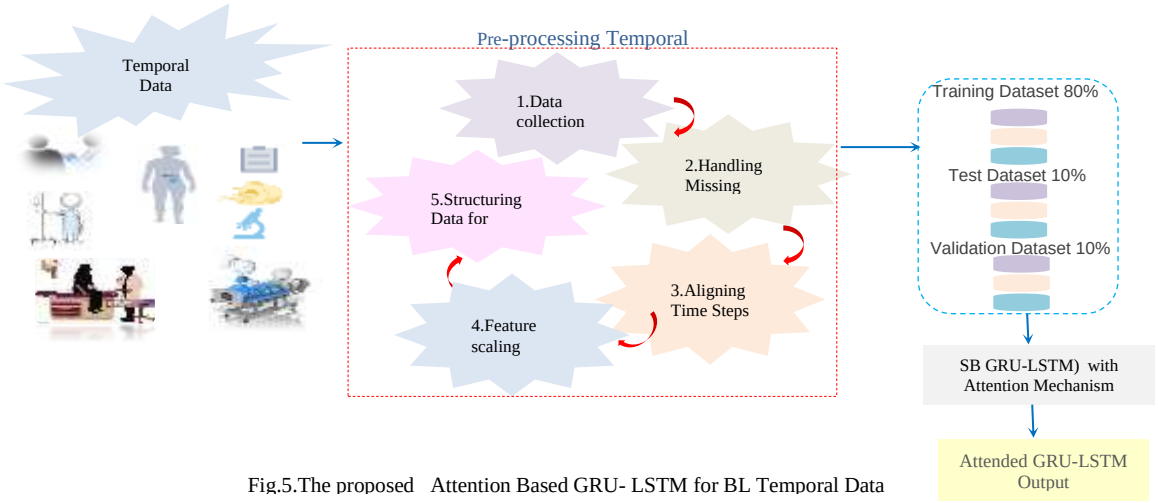


Fig.5.The proposed Attention Based GRU- LSTM for BL Temporal Data

In this work, the proposed stacked bidirectional GRU-LSTM model is an interleaving sequence of GRU and LSTM layers [37]. Every recurrent layer is implemented bidirectionally. As a result, it provides the possibility of catching dependencies in both forward and backward directions within sequence data. There are four bidirectional GRU and LSTM layers in total, with 128, 64, 32, and 16 units for each of them, respectively. These are followed by a standard dropout layer with a rate of 20%. It helps prevent overfitting, as the dropout is random in nature, by dropping units during training. There is also the incorporation of recurrent dropout within the GRU and LSTM layers to regularize the connections across time steps. After these stacked recurrent layers, the model applies a temporal-attention mechanism to the output from this last LSTM layer, whereby it prioritizes the time steps that are important to the sequence, as shown in Fig. 6.

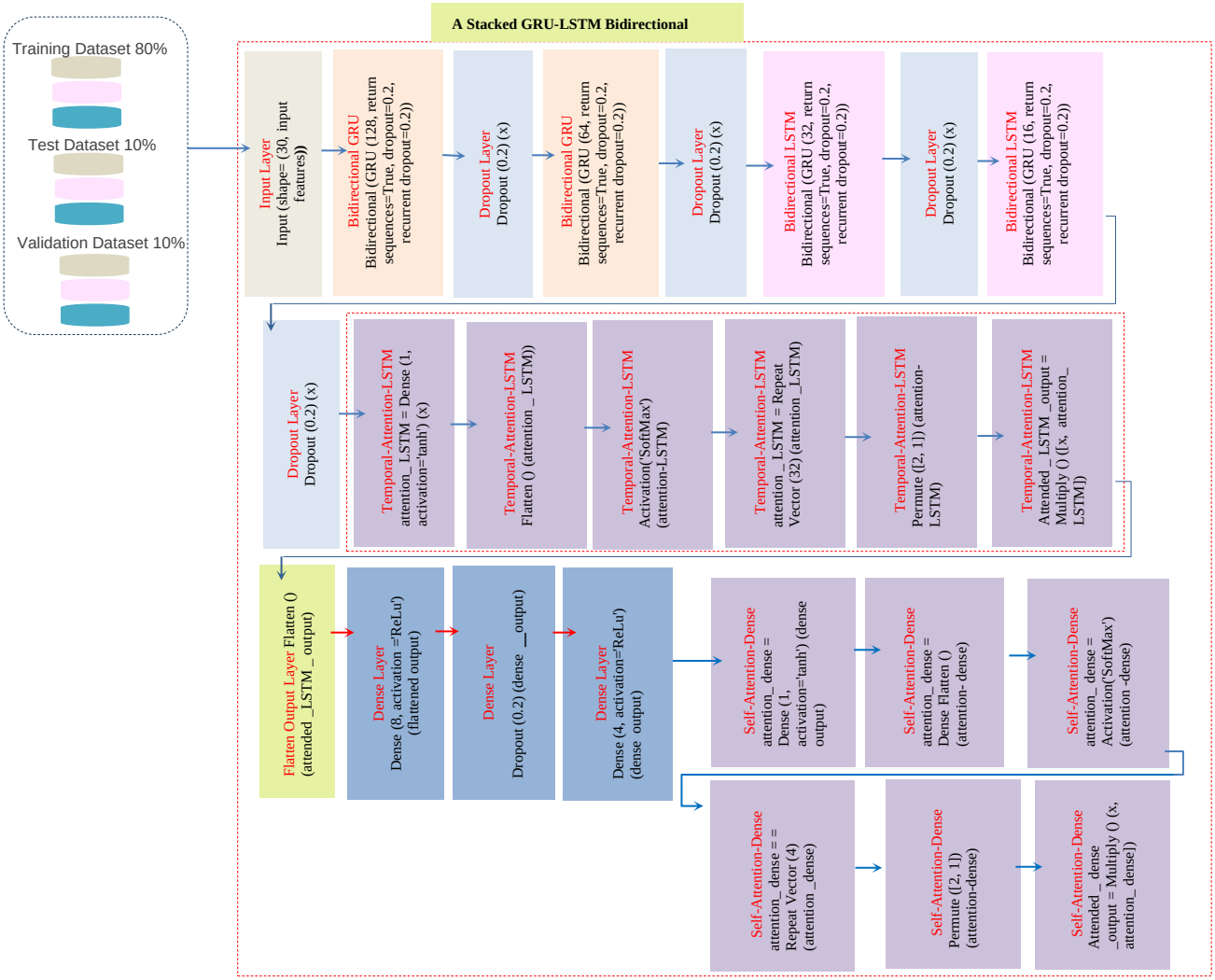


Fig.6. The structure of A Stacked Bidirectional (GRU-LSTM) with Attention Mechanism

After the application of this attention mechanism, the output is then flattened and goes through two dense layers of 8 and 4 units, respectively, each of which is followed by a respective activation function. After these dense layers, a second attention mechanism (self-attention mechanism) is applied, which allows the model to refine its focus more consciously on the significant features generated by these dense layers. In the end, this last attention mechanism enhances the model's attention to detail and complex feature interactions before finally coming up with the output. All these justify how the introduction of multiple attention mechanisms with a stacked bidirectional GRU-LSTM architecture and dropout layers adds to the model's strength, flexibility, and effectiveness in dealing with sequential data, making it very apt for complex prediction tasks.

In order to make sure that the model takes into account all elements of the input data (images, genetic data, temporal data, and clinical data) when generating predictions, the feature fusion layer fuses the characteristics from several models. By progressively decreasing the dimensionality and concentrating on the most pertinent features, FC layers enable the model to learn intricate patterns. FC contains three dense blocks, as shown in Fig. 7. Three dense blocks are the optimal option, based on this challenge, since they minimize the danger of overfitting while enabling the model to acquire a suitably advanced representation of the fused features. The last

layer of output represents the binary classification task. The sigmoid activation function is a great option as it offers a probability score for the presence or absence of BL cancer recurrence.

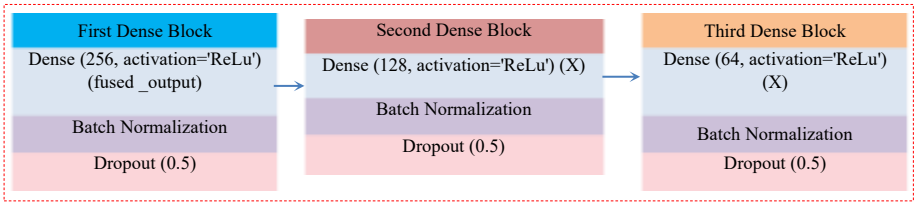


Fig. 7. Three Dense Blocks of FC

4. The Study Results

The implementation of the suggested attended CNN-LSTM, involving the use of an innovative design of pre-trained VGG16 to which CBAM is introduced after convolution block 5, and stacked bidirectional GRU-LSTM, which is based on the attention mechanism, will be addressed in this section. Recurrence prediction for BL cancer can be greatly aided by a hybrid model that uses both temporal data from the clinical dataset and histological pictures. This new hybrid approach's ability to combine patient clinical and histopathological data can merge deep learning with clinical decision-making, providing a vital tool for BL cancer research and therapy progress. The suggested plan of action will be implemented in a critical situation. In this case, the proposed strategy is used to the TCGA-bladder urothelial carcinoma (BLCA) dataset and contrasted with other advanced methods. The effectiveness of the Attended CNN-LSTM will be evaluated using the accuracy, precision, recall, and F1-score metrics.

4.1. The Description of BL Cancer Dataset

This study contains all BL cancer cases from The Cancer Genome Atlas (TCGA). Histopathological images and clinical details are included in the extensive dataset for BLCA provided by the Cancer Genome Atlas (TCGA). Digital histopathology slides of bladder cancer tissues are included in the collection, which offers comprehensive cellular morphology and tissue architecture for pathological evaluation and tumor study. The clinical data includes demographics of the patients, clinical outcomes, treatment details, and genomic and molecular profiles, such as age, gender, disease status, survival rates, treatment details, and gene mutations [36, 37].

4.2. Performance Metrics for Attended CNN-LSTM

This portion will evaluate the suggested approach, called Attended CNN-LSTM, for BL cancer recurrence and response to treatment using a range of measures, all of which will be presented. Measures like as F1-score, accuracy, recall, and precision are frequently used to assess systems for deep learning. The confusion matrix is used to determine the outcomes for the variables that are called false positive (FP), false negatives (FNs), true positives (TPs), and true negatives (TNs). The final results that are essentially positive and that the framework predicts will be positive are known as TPs. When assessing a predictive design's effectiveness in BL cancer recurrence prediction, the evaluation criteria of precision, recall, accuracy, and F1-score are crucial. These measurements demonstrate different aspects of the model's ability to predict cancer instances (e.g., detecting persons with BL cancer from a dataset). The percentage of all accurate predictions—both true positives and true negatives—out of all the predictions produced using (1) may be used to calculate accuracy. The percentage of

real BL cancer cases that the hybrid model correctly detected using (2) may be used to compute sensitivity, also known as recall. Additionally, precision may be calculated using (3) by dividing the number of genuine positive predictions (i.e., accurately predicted BL cancer cases) by the total number of positive predictions the model produced. The F1-score may be calculated as the sum of recall and accuracy using equation (4).

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

$$Recall = \frac{TP}{(TP+FN)} \quad (2)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (3)$$

$$F1 - score = 2 * \left(\frac{Precision*Recall}{Precision+Recall}\right) \quad (4)$$

To evaluate the performance of the attended CNN-LSTM approach for BL cancer recurrence prediction based on attended fine-tuned VGG16 and attended stacked bidirectional GRU-LSTM. Digitized slides from The Cancer Genome Atlas (TCGA) database (https://www.cancer.gov/tcga) were used as a basic benchmark. The performance of this method was evaluated using four metrics: accuracy (Acc), precision (P), recall (R), and the F1-score.

4.3. Testing A hybrid Model (Attended CNN-LSTM) against other Strategies

The effectiveness of the BL cancer recurrence prediction approach, known as Attended CNN-LSTM, may be shown in this paragraph by testing and contrasting it with other modern prediction techniques. RDLPI [19], HDML [27], DNLC [30], LSTM-Cox [31], PSBAI [21], and DLRP-NMIBC [29] are some other prediction techniques. The accuracy, precision, recall, and F1-score of the BL cancer prediction techniques are shown in Figs. (8–11) and Table 1. The recommended attended CNN-LSTM actually outperforms other BL cancer recurrence strategies, and it may produce the greatest performance outcomes. At an average number of practice cases = 100, this hybrid approach provides the highest accuracy, precision, recall, and F1-score values, which translate to 94.45%, 92.28%, 90.5%, and 91.38%, respectively. Attended CNN-LSTM outperforms RDLPI, HDML, DNLC, LSTM-Cox, PSBAI, and DLRP-NMIBC, as shown in Figs. (8–11) and Table 1. On the other hand, RDLPI, HDML, DNLC, LSTM-Cox, PSBAI, and DLRP-NMIBC have accuracy values of 100 for many training examples, which are 90%, 88.72%, 86.83%, 85.61%, and 78.72%, respectively, and precision values of 88%, 86.23%, 85%, 83.92%, 81.63%, and 80.02%. The F1 rating scores are 86.47%, 85.16%, 83.90%, 81.98%, 80.5%, and 78.48%, respectively, however the percentages for different outcome measures, such as recall at a total number of prior training instances = 100, are 85%, 84%, 82%, 80.13%, 79.43%, and 77%. As a result, the DLRP-NMIBC method produces the lowest degrees of accuracy, precision, recall, and F1-score ratios. Attended CNN-LSTM is the technique's representative, while DLRP-NMIBC is the worst. In certain locations, the attended CNN-LSTM approach produces the best results. In addition, HDML is the third-best solution with regard to of precision and recall values, while RDLPI ranks in second. Consequently, DNLC ranked fourth, LSTM-Cox fifth, and PSBAI sixth. Attended CNN-LSTM represents the best approach, whilst DLRP-NMIBC represents the worst.

Table .1: Evaluation Indicators for BL Cancer Recurrence prediction

Bladder Cancer (BL) Recurrence prediction	Accuracy	Precision	Recall	F1-score
DLRP-NMIBC	78.72	80.02	77	78.48
PSBAI	80.5	81.63	79.43	80.59
LSTM-COX	85.61	8883.92	80.13	81.98
DNLC	86.83	85	82.83	83.90
HDML	88.72	86.23	84	85.16
RDLPI	90	88	85	86.47
Attended CNN-LSTM	94.45	92.28	90.5	91.38

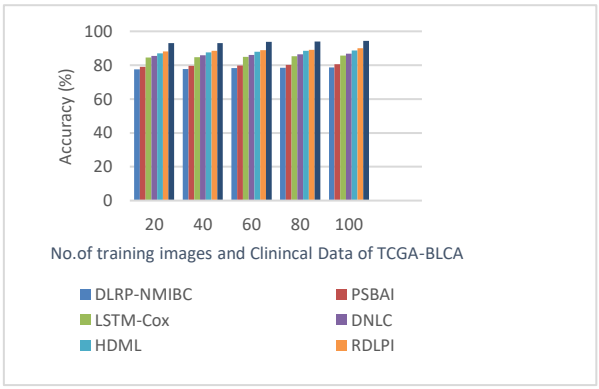


Fig.8. Accuracy of Attended CNN-LSTM

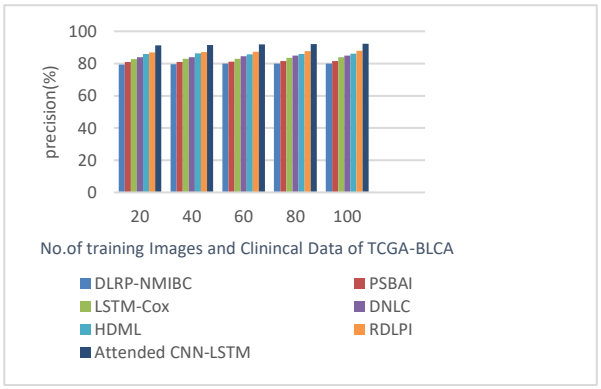


Fig.9. Precision of Attended CNN-LSTM

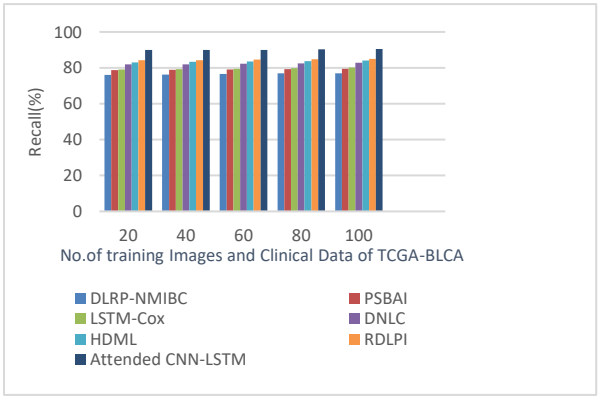


Fig.10. Recall of Attended CNN-LSTM

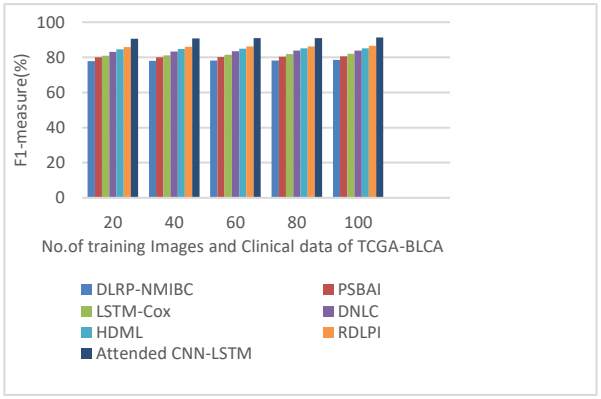


Fig.11. F1-measure of Attended CNN-LSTM

A hybrid approach (attended CNN-LSTM) shows the highest overall performance with the best accuracy, recall, and F1-score. The proposed approach can also surpass other models for many reasons. The comparison highlights that the integration of VGG16, stacked bidirectional GRU-LSTM, and attention mechanisms contributes to improved spatial and temporal feature extraction, resulting in higher predictive accuracy and balanced precision-recall performance. Pretrained VGG16: Enhanced spatial feature extraction from histopathological images compared to custom CNNs used in other methods. Bidirectional GRU-LSTM with Attention: Superior temporal data handling, capturing recurrence patterns more effectively than traditional RNNs or unidirectional LSTMs. Attention Mechanism: Highlights critical spatial regions and temporal intervals, improving interpretability and clinical relevance. The attended CNN-LSTM approach addresses critical clinical challenges in BL cancer recurrence prediction, as follows: **(i) Identification of Vital Factors: (**a) Tumor **Size**: Tumor size is a key indicator in predicting recurrence and progression in bladder cancer. The model leverages histopathological images processed through the VGG16 backbone to extract spatial features that capture fine-grained details such as tumor morphology and cellular architecture. By focusing on these features, the model identifies tumor size-related patterns that correlate with recurrence risk, making predictions clinically meaningful. (b) **Recurrence Intervals:** The use of stacked bidirectional GRU-LSTM networks enables the model to process temporal clinical data (e.g., follow-up records and time intervals between recurrences). This allows the model to recognize trends and intervals in recurrence patterns, which are often indicative of disease progression. For instance, shorter recurrence intervals may suggest aggressive tumor behavior, and the model's ability to detect these patterns helps guide timely interventions.

(c) **Treatment Protocols**: The inclusion of clinical data (e.g., details of chemotherapy, immunotherapy, or surgical interventions) helps the model evaluate how different treatment regimens impact recurrence risks. This provides a personalized prediction that incorporates both historical treatment data and patient-specific responses. (ii) **Personalized Treatment Planning**: (a) **Patient Stratification**: By accurately identifying high-risk and low-risk patients based on recurrence predictions, the model enables clinicians to stratify patients into distinct groups. High-risk patients can be prioritized for intensive monitoring, additional diagnostic procedures, or aggressive therapies, while low-risk patients may require less frequent follow-ups, reducing unnecessary interventions. (b) **Interpretable Attention Maps**: The attention mechanism incorporated in the model improves interpretability by highlighting critical spatial regions in histopathological images (e.g., areas with dense cellular atypia) and temporal points in clinical data (e.g., periods of high recurrence likelihood). This helps clinicians understand the "why" behind the model's predictions, fostering trust and adoption in clinical settings. (c) **Optimization of Therapeutic Interventions**: With predictions that consider factors such as tumor size, recurrence intervals, and prior treatments, the model supports the design of tailored therapeutic strategies. For instance, patients with frequent recurrences might benefit from immunotherapy or systemic chemotherapy. Patients with isolated large tumors might be candidates for localized surgical resection or radiotherapy. Consequently, the proposed model achieves improving Accuracy: The model's ability to consider both spatial and temporal data ensures more accurate recurrence predictions and timely decision-making: By identifying high-risk intervals and features, the model enables timely interventions, potentially improving patient outcomes and enhancing efficiency: Stratifying patients and reducing unnecessary interventions helps optimize healthcare resources.

RDLP ranked second as it incorporates a two-phase approach (patch-level and WSI-level prediction), transfer learning, and model interpretability. Generalization might be impacted by possible data variability in the TCGA images of this model. Dataset size of this experiment can limit model adaptability, especially for testing outside. In the integration of multimodal data, like attended pre-trained VGG16 and attended GRU-LSTM in the proposed model, robust techniques for feature fusion are needed to prevent overfitting or redundancy. The third model is HDML, which has high accuracy through hybrid feature extraction. Tailored treatment planning through accurate staging; efficient and cost-effective diagnosis leverages pre-trained DNNs for quality features. Effective for limited data with statistical ML. Multiple staging tasks are comprehensively addressed. It suffers from the following: computation-intensive and time-consuming; deep learning steps not interpretable; and validation and generalization challenges. Potential bias and ethical concerns.

Thus, this hybrid framework, DNLC, will provide enhanced bladder cancer diagnosis, improved classification performance oriented by the tasks, personalized clinical decisions, and will be more versatile. Having employed the pre-trained model of extracting high-quality features from CTs by the proposed model reduces unnecessary time consumption and resources by not demanding extensive training on large-sized medical images. Classified using robust and interpretable statistical machine learning classifiers, namely support vector machine or random forest. This framework reduces dependency on manual staging and is cost-effective, thus performing faster and reliable diagnostics. Its disadvantages include the following: high-quality dependence on CT scans, variability in imaging, preprocessing, challenges in classification between PTC versus MIBC, computation complexity, limited interpretability, clinical adoption, generalization of validation, and ethics. Overcoming these drawbacks is crucial for this hybrid approach to realize its fullest potential in bladder cancer management.

The advantages of this hybrid framework (LSTM-COX) in diagnosing BL cancer are, in general, the increase of classification accuracy, the best task-oriented performance, the clinically tailored decision, and flexibility. Feature extraction involves the usage of pre-trained models to reduce time consumption during the training process. Statistical machine learning techniques, like support vector machine or random forest, give robustness and interpretability to the classification process. The proposed framework is independent of manual staging, which makes it cost-effective, too. This will accelerate the diagnosis effectively. It faces a number of challenges in relation to dependence on good-quality CT scans, variation in imaging, the need for preprocessing, and

difficulty in the distinguishing of post-treatment changes from Muscle Invasive Bladder Cancer (MIBC) classification. In addition, the high computational complexity can be a serious concern as it involves the usage of resource-intensive pre-trained DNNs for feature extraction. Another challenge is that this deep learning feature extraction step might reduce interpretability, and clinicians may hesitate to use the model without clear insights.

The AI-based scoring of MIBC in PSBAI model has a number of advantages, including appropriate risk stratification, making personalized treatment decisions, cost savings by identifying those most likely to derive a benefit, and encouraging shared clinical decision-making with improved collaboration between the physician and the patient. Interpretable AI techniques are used to generate a scoring system linked to specific histopathological patterns and molecular features. It also integrates multi-omics data for comprehensive insight into tumor behavior and potential biomarkers. The data-driven strategy used therein further enhances the prognostic accuracy. On the other hand, this model needs high-quality H&E-stained whole-slide images for its performance, requiring a wide range of validation across populations and clinical workflows. Besides, it is with high computational complexity and resource intensity, long training time, and poor generalization to unknown features. It relies on molecular features. This can easily lead to overfitting to molecular subtypes and confounding factors that reduce generalization to broader clinical scenarios. Interpretability: Poor human readability, possibly misaligned with clinical judgment in complex cases. Ethical and regulatory—the considerations also have to be taken care of. DLRP-NMIBC combines histopathology and clinical data. It improves long-term predictions. The method can perform patient monitoring and treatment. It provides interpretable segmentation results. Limitations: This method contains a three-step process, which is computationally intensive; it requires extensive external validation. Risk of class imbalance affecting recall

The attention mechanism enhances model interpretability by providing visual explanations for spatial and temporal features critical to prediction. (i) Spatial Interpretability: (a) In the histopathological images processed by the VGG16 component, the attention mechanism generates heatmaps to focus on regions with significant cellular and morphological changes, such as areas with high tumor cell density, nuclear atypia, or structural irregularities. (b)These heatmaps provide visual explanations that enable pathologists and clinicians to validate the model's focus on clinically meaningful regions, fostering confidence in the model's predictions. (ii) Temporal Interpretability: (a) For clinical data, the attention mechanism in the bidirectional GRU-LSTM highlights specific time intervals that are most predictive of recurrence risk. For example, it may emphasize a short recurrence interval or a particular treatment cycle, helping clinicians understand why the model predicts a high or low risk of recurrence. (b) By identifying critical temporal points, the model allows clinicians to contextualize predictions within the patient's treatment history and disease progression. (iii) Clinical Trust and Adoption: (a) Interpretability features such as heatmaps (for spatial data) and attention weight visualizations (for temporal data) bridge the gap between complex machine learning outputs and the decision-making needs of clinicians. These features provide transparency, enabling clinicians to verify the model's focus on biologically and clinically relevant data, which builds trust and increases the likelihood of clinical adoption.

To facilitate practical deployment, the hybrid model can be implemented as a cloud-based API integrated with existing healthcare systems, providing clinicians with seamless access to recurrence predictions. Deploying the hybrid model as a cloud-hosted RESTful API that allows seamless integration with existing hospital information systems (HIS) or electronic medical record (EMR) platforms. The API will take histopathological images and clinical data as input, process them through the hybrid model, and return recurrence risk scores and attention visualizations (heatmaps and timelines) to clinicians. This approach ensures that the model's predictions can be accessed directly from clinicians' existing workflows without requiring significant changes to the current infrastructure. Hosting the model on cloud platforms such as AWS, Google Cloud, or Microsoft Azure offers scalable compute resources, enabling the system to handle large datasets and concurrent requests. Cloud-based deployment also facilitates centralized updates, allowing for easy maintenance and improvement of the model without requiring local installations in each hospital. To address privacy concerns and comply with regulations

like HIPAA and GDPR, we have also considered deploying the model on local edge devices (e.g., hospital servers or secure workstations). This ensures that patient data remains within the hospital's network while still benefiting from the model's predictions.

A deployment pipeline will be established to pre-process input data, execute the model's predictions, and post-process results for easy interpretation by clinicians. Tools such as Docker and Kubernetes will be used to containerize and orchestrate the deployment, ensuring flexibility and ease of scaling. There are some challenges, like: (i) Interoperability: Integration with diverse healthcare IT systems can be challenging due to differences in data formats. To address this, we will implement standard data exchange formats such as HL7 FHIR for compatibility. (i) Training and Adoption: Clinicians may require training to understand and utilize the model's outputs effectively. Providing user-friendly interfaces and detailed documentation will help address this. (iii) Latency Concerns: For real-time predictions, we will optimize model inference through quantization and caching techniques to ensure rapid response times.

4.3.1. External Validation and Robustness

In addition to the TCGA dataset, the model was tested on an external dataset obtained from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, available through the Cancer Data Access System (CDAS) [39]. This comprehensive dataset includes histopathological images and clinical data from a large cohort of participants, providing a diverse sample for external validation. PLCO Bladder datasets are available for delivery on CDAS, which includes a diverse patient population and different imaging techniques. The model achieved 92.86% accuracy, 90.56% precision, 88.3% recall, and 89.5% F1-score on this dataset. This resource can further support the evaluation of your model across different datasets. It can also highlight its ability to generalize across different clinical and institutional settings. This validation underscores the potential for broader applicability of the proposed approach."

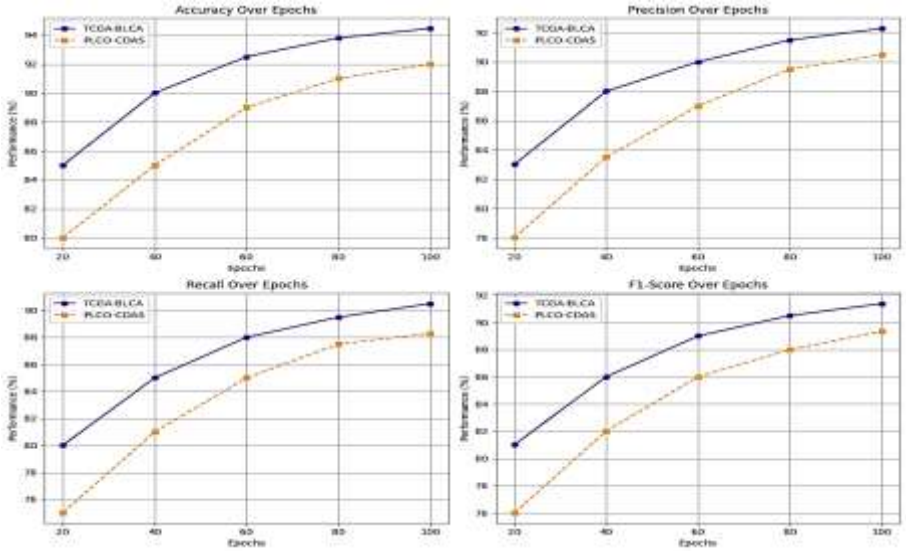


Fig.12. Attended CNN-LSTM Validation with New Dataset

Scalability remains a critical consideration for the hybrid model, particularly when applied to larger datasets. To address this issue, key techniques to optimize the hybrid model combining a pretrained VGG16 with attention mechanisms and stacked bidirectional GRU-LSTMs are presented as follows: First, **model pruning** is applied to remove less significant weights from VGG16's fully connected layers and GRU's input-hidden weights,

reducing model size and computational overhead while retaining critical functionality. Second, **quantization** is used to convert the model's parameters to a lower precision format (e.g., INT8), significantly reducing memory usage and improving inference speed without sacrificing accuracy. Third, **distributed training** leverages multiple GPUs using PyTorch's Distributed Data Parallel (DDP), allowing parallel processing of data to accelerate training for larger datasets. Finally**, data prefetching** is implemented Data Loader to load batches of data into memory ahead of time, minimizing data transfer bottlenecks and maximizing GPU utilization. Together, these techniques ensure efficient training and scalability of the hybrid model for larger datasets like TCGA and PLCO while maintaining robust performance and interpretability.

5. Conclusion and Future Work

The innovative hybrid model demonstrates strong potential in BL cancer recurrence and treatment response prediction, utilizing an attention-guided stacked bidirectional GRU-LSTM for temporal feature learning and an attention-enhanced pre-trained VGG16 for spatial feature extraction. The model achieves good accuracy, precision, recall, and F1-measure by combining spatial patterns from histological pictures with longitudinal clinical data. This results in a complete and comprehensible answer for clinical decision-making. By concentrating on crucial areas in photos and significant moments in clinical sequences, the attention mechanism improves model performance even more by bringing the model's predictions into line with clinically significant variables. This hybrid concept has the potential to greatly improve bladder cancer management results by facilitating individualized treatment planning and improving recurrence risk assessment.

In the future, this strategy may leverage a cloud-based API or service for model deployment, allowing integration into hospital systems. Model interoperability may be improved by applying explainable AI (XAI), which improves clinician trust and adoption by providing clear explanations of predictions. Hybrid transformer-LSTM may be applied to robust temporal modeling with efficient feature extraction.

References

[1] K. Takahashi, G. Kimura, Y. Endo, J. Akatsuka, T. Hayashi, Y. Toyama, et al., Urothelial carcinoma of the bladder, lipid cell variant: A case report and literature review, Journal of Nippon Medical School, vol. 86, 2019.

[2] A. Dotson, A. May, F. Davaro, S. J. Raza, S. Siddiqui, Z. Hamilton, Squamous cell carcinoma of the bladder: poor response to neoadjuvant chemotherapy, International Journal of Clinical Oncology, vol. 24, no. 6, 2019, pp. 706-711.

[3] V.J. Sharp, K.T. Barnes, B.A. Erickson, Assessment of asymptomatic microscopic hematuria in adults, American Family Physician, vol. 88, 2013, pp. 747-754.

[4] A.F. Kantor, P. Hartge, R.N. Hoover, A.S. Narayana, J. Sullivan, J.R.J. Fraumeni, Urinary tract infection and risk of bladder cancer, American Journal of Epidemiology, vol. 119, 1984, pp. 510-515.

[5] K. Thorsteinsson, J.B. Jensen, Misinterpretation resulting in a diagnosis of bladder cancer–A case emphasizing the value of diagnostic reconsideration, Urology Case Reports, vol. 40, 2022, article 101928.

[6] I. Kononenko, Machine learning for medical diagnosis: History, state of the art, and perspective, Artif. Intell. Med., vol. 23, 2001, pp. 89–109. *https://doi: 10.1016/S0933-3657(01)00077-X.*

[7] I. El Naqa, M. J. Murphy, what are machine and deep learning, in: M. J. Murphy (Ed.), Machine and Deep Learning in Oncology, Medical Physics and Radiology, Springer, Cham, Switzerland, 2022, pp. 3–15.

[8] S. Borhani, R. Borhani, A. Kajdacsy-Balla, Artificial intelligence: A promising frontier in bladder cancer diagnosis and outcome prediction, Crit. Rev. Oncol., vol. 171, 2022, p. 10360.

[9] C. Gandi, L. Vaccarella, R. Bientinesi, M. Racioppi, F. Pierconti, E. Sacco, Bladder cancer in the time of machine learning: Intelligent tools for diagnosis and management, Urol. J., vol. 88, 2021, pp. 94–102.

[10] S. Chen, L. Jiang, X. Zheng, J. Shao, T. Wang, E. Zhang, F. Gao, X. Wang, J. Zheng, Clinical use of machine learning-based pathomics signature for diagnosis and survival prediction of bladder cancer, Cancer Sci., vol. 112, 2021, pp. 2905–2914.
[11] I.-J. Tsai, W.-C. Shen, C.-L. Lee, H.-D. Wang, C.-Y. Lin, Machine learning in prediction of bladder cancer on clinical laboratory data, Diagnostics, vol. 12, 2022, p. 203.
[12] P.-N. Yin, K. C. KC, S. Wei, Q. Yu, R. Li, A. R. Haake, H. Miyamoto, F. Cui, Histopathological distinction of non-invasive and invasive bladder cancers using machine learning approaches, BMC Med. Inform. Decis. Mak., vol. 20, 2020, p. 162.
[13] Y. Yang, X. Zou, Y. Wang, X. Ma, Application of deep learning as a noninvasive tool to differentiate muscle-invasive bladder cancer and non–muscle-invasive bladder cancer with CT, Eur. J. Radiol., vol. 139, 2021, p. 109666.
[14] A. Gu, T. Dao, Mamba: Linear-time sequence modeling with selective state spaces, IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 1, 2023, pp. 110-121. arXiv preprint arXiv:2312.00752.
[15] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, W. Zhang, Informer: Beyond efficient transformer for long sequence time-series forecasting, Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, no. 1, 2021, pp. 11106–11115.
[16] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 30, no. 1, 2017, pp. 4700-4708.
[17] Z. Che, S. Purushotham, K. Cho, D. Sontag, Y. Liu, Recurrent neural networks for multivariate time series with missing values, Scientific Reports, vol. 8, no. 1, 2018, pp. 6085-6096.
[18] R. DiPietro, G.D. Hager, Deep learning: RNNs and LSTM, Handbook of Medical Image Computing and Computer-Assisted Intervention, Elsevier, 2019, pp. 503-519. *https://doi: 10.1016/B978-0-12-816176-0.00026-0.*
[19] G.Y. Wang, J.F. Zhu, Q.C. Wang, Prediction of non-muscle invasive bladder cancer recurrence using deep learning of pathology images, in: Sci Rep, vol. 14, Springer, Cham, Switzerland, 2024, pp. 18931. *https://doi.org/10.1038/s41598-024-66870-9.*
[20] W. Barrios, B. Abdollahi, M. Goyal, Q. Song, M. Suriawinata, R. Richards, B. Ren, A. Schned, J. Seigne, M. Karagas, S. Hassanpour, Bladder cancer prognosis using deep neural networks and histopathology images, in: J. Pathol. Inform., vol. 13, Elsevier, 2022, pp. 100135. *https://doi.org/10.1016/j.jpi.2022.100135.*
[21] O. Eminaga, S.-R. Leyh-Bannurah, S.F. Shariat, L.-M. Krabbe, H. Lau, L. Xing, M. Abbas, Artificial Intelligence reveals distinct prognostic subgroups of muscle-invasive bladder cancer on histology images, in: Cancers, vol. 15, MDPI, Basel, Switzerland, 2023, pp. 4998. *https://doi.org/10.3390/cancers15204998.*
[22] Q. Zheng, R. Yang, X. Ni, S. Yang, L. Xiong, D. Yan, L. Xia, J. Yuan, J. Wang, P. Jiao, Accurate diagnosis and survival prediction of bladder cancer using deep learning on histological slides, in: Cancers, vol. 14, MDPI, Switzerland, 2022, pp. 5807. *https://doi.org/10.3390/cancers14235807.*
[23] G. García, A. Esteve, A. Colomer, D. Ramos, V. Naranjo, A novel self-learning framework for bladder cancer grading using histopathological images, in: Comput Biol Med, vol. 138, Elsevier, 2021, p. 104932. *https://doi.org/10.1016/j.compbiomed.2021.104932.*
[24] N. Tokuyama, A. Saito, R. Muraoka, S. Matsubara, T. Hashimoto, N. Satake, J. Matsubayashi, T. Nagao, A.H. Mirza, H.P. Graf, E. Cosatto, C.L. Wu, M. Kuroda, Y. Ohno, Prediction of non-muscle invasive bladder cancer recurrence using machine learning of quantitative nuclear features, in: Int J Urol, vol. 35, no. 4, Springer, Tokyo, Japan, 2022, pp. 533-538.
[25] Y. Xu, J. Lou, Z. Gao, M. Zhan, Computed tomography image features under deep learning algorithm applied in staging diagnosis of bladder cancer and detection on ceramide glycosylation, in: Comput. Intell. Neurosci., Hindawi, 2022, Article ID 7979523. *https://doi.org/10.1155/2022/7979523.*

[26] R. Yang, Y. Du, X.D. Weng, Z.Y. Chen, S.S. Wang, X.H. Liu, Automatic recognition of bladder tumours using deep learning technology and its clinical application, in: Vol. 17, Issue 2, 2021, e2194. *https://doi.org/10.1002/rcs.2194.*

[27] S. Sarkar, K. Min, W. Ikram, R.W. Tatton, I.B. Riaz, A.C. Silva, T. Wu, performing automatic identification and staging of urothelial carcinoma in bladder cancer patients using a hybrid deep-machine learning approach, in: Cancers, vol. 15, MDPI, Basel, Switzerland, 2023, pp. 1673. *https://doi.org/10.3390/cancers15061673.*

[28] J.W. Yoo, K.C. Koo, B.H. Chung, S.Y. Baek, S.J. Lee, K.H. Park, K.S. Lee, Deep learning diagnostics for bladder tumor identification and grade prediction using RGB method, in: Sci Rep, vol. 12, Springer, Cham, Switzerland, 2022, pp. 17699. *https://doi.org/10.1038/s41598-022-22797-7.*

[29] M. Lucas, I. Jansen, T.G. van Leeuwen, J.R. Oddens, D.M. de Bruin, H.A. Marquering, Deep learning-based recurrence prediction in patients with non-muscle-invasive bladder cancer, in: Eur Urol Focus, vol. 8, Elsevier, 2022, pp. 165-172. *https://doi.org/10.1016/j.euf.2020.12.008.*

[30] H. Elwahsh, M.A. Tawfeek, A.A. Abd El-Aziz, M.A. Mahmood, M. Alsabaan, E. El-shafeiy, A new approach for cancer prediction based on deep neural learning, in: J. King Saud Univ. - Comput. Inf. Sci., vol. 35, Elsevier, Amsterdam, Netherlands, 2023, pp. 101565. *https://doi.org/10.1016/j.jksuci.2023.101565.*

[31] R. Zhang, J. Jiang, X. Shi, Modeling Long Sequences in Bladder Cancer Recurrence: A Comparative Evaluation of LSTM, Transformer, and Mamba, in: arXiv preprint, vol. 2405.18518v2 [cs. LG], 2024, pp. 1–20. *https://arxiv.org/abs/2405.18518v2*.

[32] A. Bashkami, A. Nasayreh, S.N. Makhadmeh, A review of Artificial Intelligence methods in bladder cancer: segmentation, classification, and detection, in: Artif Intell Rev, vol. 57, Springer, Cham, Switzerland, 2024, pp. 339. *https://doi.org/10.1007/s10462-024-10953-6.*

[33] H. Shi, J. Lu, and Q. Zhou, "A novel data augmentation method using style-based GAN for robust pulmonary nodule segmentation," 2020 Chinese Control and Decision Conference (CCDC), August 2020, pp. 2486-2491.

[34] D. Ivanov, C. A. Barbano, M. Grangetto. Multi-target stain normalization for histology slides. *arXiv preprint*, 2024, arXiv:2406.02077.

[35] A. Faghihi, M. Fathollahi, R. Rajabi, "Diagnosis of skin cancer using VGG16 and VGG19 based transfer learning models," *Multimedia Tools and Applications*, vol. 83, no. 19, pp. 57495-57510, 2024.

[35] Y. Tashtoush, R. Obeidat, A. Al-Shorman, O. Darwish, M. Al-Ramahi, and D. Darweesh, "Enhanced convolutional neural network for non-small cell lung cancer classification," *International Journal of Electrical and Computer Engineering*, vol. 13,2023, pp. 1024-1038. *https://doi: 10.11591/ijece.v13i1.*

[36] H. S. Abdullah, N. H. Ali, N. A. Abdullah, "Evaluating the Performance and Behavior of CNN, LSTM, and GRU for Classification and Prediction Tasks," *Iraqi Journal of Science*, vol. 65, no. 3, 2024.

[37] *https://www.cancerimagingarchive.net/collection/tcga-blca*.

[38] *https://datacatalog.mskcc.org/dataset/10403*.

[39] *https://cdas.cancer.gov/datasets/plco/18/.*