Multimedia Tools and Applications (2022) 81:10279–10297
https://doi.org/10.1007/s11042-022-12229-z

Classification of malignant tumors by a non-sequential recurrent ensemble of deep neural network model

Dipanjan Moitra[1] · Rakesh Kr. Mandal[1]

Received: 24 September 2020 / Revised: 12 January 2021 / Accepted: 14 January 2022 /
Published online: 14 February 2022

Abstract
Many significant efforts have so far been made to classify malignant tumors by using various machine learning methods. Most of the studies have considered a particular tumor genre categorized according to its originating organ. This has enriched the domain-specific knowledge of malignant tumor prediction, we are devoid of an efficient model that may predict the stages of tumors irrespective of their origin. Thus, there is ample opportunity to study if a heterogeneous collection of tumor images can be classified according to their respective stages. The present research work has prepared a heterogeneous tumor dataset comprising eight different datasets from The Cancer Imaging Archives and classified them according to their respective stages, as suggested by the American Joint Committee on Cancer. The proposed model has been used for classifying 717 subjects comprising different imaging modalities and varied Tumor-Node-Metastasis stages. A new non-sequential deep hybrid model ensemble has been developed by exploiting branched and re-injected layers, followed by bidirectional recurrent layers to classify tumor images. Results have been compared with standard sequential deep learning models and notable recent studies. The training and validation accuracy along with the ROC-AUC scores have been found satisfactory over the existing models. No model or method in the literature could ever classify such a diversified mix of tumor images with such high accuracy. The proposed model may help radiologists by acting as an auxiliary decision support system and speed up the tumor diagnosis process.

Keywords Malignant · Tumor · Classification · Deep learning · Staging

✉ Dipanjan Moitra
tataijal@gmail.com

Rakesh Kr. Mandal
tataijal@yahoo.in

[1] University of North Bengal, Siliguri, India

1 Introduction

Many notable works have so far been carried out to classify malignant tumors by using different machine learning techniques [45]. Most of the studies have tried to predict whether a tumor is benign or malignant [62]. Their subjects have also been homogeneous in terms of tumor origin and scanner modalities [41]. Thus, domain-centric prediction of tumors has been much exercised than to propose an algorithm for predicting a tumor accurately irrespective of its originating organ. The real-life prognosis of a tumor is much more complicated than just to predict it as benign or malignant. The American Joint Committee on Cancer (AJCC) [18] has propounded the popular Tumor-Node-Metastasis (TNM) staging system that depicts how much a tumor has already spread in the body. Table 1 shows an example of how the TNM staging of renal cancers has been accomplished.

The TNM stage may vary from one tumor-type to another. The overall pathological staging brings TNM stages under a uniform prognostic group. In Table 2, the overall AJCC staging of renal tumors has been shown. The present study has proposed a model that can predict the overall pathological staging of a tumor irrespective of its genre.

The deep neural network (DNN) has been a prevalent technology in computer vision and biomedical image processing. DNN has a wide variety of applications ranging from the estimation of blood pressure [52, 70] to the detection of COVID-19 infections [1, 2]. Unlike the traditional machine learning algorithms, explicit image pre-processing [37], segmentation [38], and manual feature crafting are not required in deep learning. Thus, image processing and re-generation becomes easier and effective with deep learning techniques [44]. However, it has been observed that the traditional sequential models might mislay important imagery features during the down-sampling phase. The sequential models also suffer from the high variance that may affect generating a consistent accuracy. As the current problem at hand is a multiclass problem having heterogeneous imagery, the present study has adopted a non-

Table 1 Definition of TNM staging of renal tumors as per AJCC 7th Edition manual

Primary Tumor (T)	TX	Primary tumor cannot be assessed
	T0	No evidence of primary tumor
	T1	Tumor size <=7 cm
	T1a	Tumor size <=4 cm
	T1b	Tumor size >4 cm and<=7 cm
	T2	Tumor size >7 cm
	T2a	Tumor size >7 cm and<=10 cm
	T2b	Tumor size >10 cm
	T3	Tumor extends into major veins or perinephric tissues but not into the ipsilateral adrenal gland and not beyond Gerota's fascia
	T3a	Tumor grossly extends into the renal vein or its segmental branches, or tumor invades perirenal and/or renal sinus fat but not beyond Gerota's fascia
	T3b	Tumor grossly extends into the vena cava below the diaphragm
	T3c	Tumor grossly extends into the vena cava above the diaphragm or invades the wall of the vena cava
	T4	Tumor invades beyond Gerota's fascia
Regional Lymph Nodes (N)	NX	Regional Lymph Nodes cannot be assessed
	N0	No Regional Lymph Node metastases
	N1	Metastasis in regional lymph node(s)
Distant Metastasis (M)	M0	No Distant Metastasis
	M1	Distant Metastasis

Table 2 Overall AJCC staging of renal tumors as per AJCC 7th Edition manual

Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T1 or T2	N1	M0
	T3	N0 or N1	M0
Stage IV	T4	Any N	M0
	Any T	Any N	M1

sequential paradigm of deep learning. The model has been developed by combining branched, re-injected, and bidirectional recurrent layers [13]. The success of the model would pave a new revolution in cancer treatment as there would be no need to rely on different models for staging different cancer. A single model would work as a decision support system for staging tumors of different genres and would help radiologists to affirmatively decide on the treatment plan.

2 Objective

The study aims to first prepare an image collection containing eight different cancers: bladder, liver, renal, head & neck, breast, thyroid, uterus, and lungs. These cancers have been selected as they belong to the leading cause of cancer-related deaths both in developed, developing, and under-developed countries [10]. In this way, the image dataset prepared would have a varied mix of tumor images as per originating organs, imaging modalities, subject demography, and treatment strategy. The next aim is to develop a deep neural network model capable of classifying the AJCC staging of such a varied mix of tumors. The problem at hand is complex relative to other contemporary efforts where homogeneous image collections have been classified. The present study conducts experiments with both sequential and non-sequential models. The final aim is to compare the results obtained from different models and to select the best model. The rest of the study is divided into the following sections: related work, data acquisition, methodology, discussion, and conclusion.

Table 3 Recent significant studies on bladder cancer

Study	Year	Details	Outcome
Xu et al. [65]	2017	used MRI radiomic features for differentiation between Non-Muscle Invasive Bladder Cancer (NMIBC) and Muscle Invasive Bladder Cancer (MIBC)	AUC of 0.8610
Ikeda et al. [26]	2018	used a pre-trained network with 1.2 million images from the ImageNet to distinguish between bladder tumors and healthy urothelium	93.0% sensitivity
Eminaga et al. [19]	2018	used 18,681 images from 479 cystoscopy videos in an Xception-based model	99.52% accuracy
Cha et al. [11]	2018	investigated 123 bladder cancer subjects where AUC in clinical assessment was 0.74	AUC of 0.8
Shkolyar et al. [55]	2019	used CNNs to detect bladder tumors from cystoscopic videos of 100 patients	90% sensitivity
Zheng et al. [73]	2019	extracted radiomic features and used the nomogram as a bladder cancer classifier	AUC of 0.876
Lin et al. [31]	2019	used the least absolute shrinkage and selection operator (LASSO) method with radiogenomics to predict the survival of patients	AUC of 0.956

Table 4 Recent significant studies on liver cancer

Study	Year	Details	Outcome
Vivanti et al. [61]	2017	detected new liver tumors by using deep learning and CT images	86% accuracy
Sabut et al. [49]	2018	classified liver cancer subtype by using texture features and Deep Neural Network (DNN) along with 225 CT images	99% accuracy
Ben-Cohen et al. [6]	2018	proposed a CNN for the detection of liver metastasis	94% accuracy
Bharti et al. [7]	2018	proposed a model ensemble to differentiate four liver stages	96.6% accuracy
Frid-Adar et al. [21]	2018	used Generative Adversarial Networks (GANs) to reconstruct tumor images	78.6% sensitivity
Romero et al. [48]	2019	adopted a deep learning approach for classification of Colorectal liver lesions	96% accuracy
Sato et al. [51]	2019	predicted hepatocellular carcinoma using real-world data	87% accuracy

3 Related work

Recent notable studies have been included in the review of the literature to compare the limitations and to bridge the research gap.

Tables 3, 4, 5, 6, 7, 8, 9, 10 reveals that, on a considerable number of occasions, non-invasive approaches have successfully overshadowed the in-vitro diagnosis of tumors. Machine learning, especially, deep learning has emerged as a seminal technique for CAD-based tumor prognosis. It has also been found that a model ensemble performs better than a single model. Most of the Researchers have so far concentrated on the classification of a single tumor genre. This has elevated the performance of domain-specific classification of tumors. However, the initiative to automate pathological staging has not been seen very often. Existing

Table 5 Recent significant studies on renal cancer

Study	Year	Details	Outcome
Ing et al. [27]	2017	used a machine learning framework and The Cancer Genome Atlas (TCGA) data for prediction of disease-free-survival	AUC 0.79
Ali et al. [3]	2018	kidney cancer subtype classification by using Long Short Term Memory (LSTM) along with the TCGA	95% accuracy
Bektas et al. [5]	2018	Support Vector Machines (SVM) classified nuclear grades	85% accuracy
Han et al. [25]	2019	did a renal cancer subtype classification by using the deep learning method	85% accuracy
Zhou et al. [74]	2019	used a deep learning model to classify renal tumors into benign and malignant	93% accuracy
Tabibu et al. [50]	2019	used CNN to predict renal cancer histology subtypes	94% accuracy
Tian et al. [57]	2019	did the grading of renal carcinoma by using The Cancer Genome Atlas (TCGA) and the Lasso model	84.6% sensitivity
Kocak et al. [29]	2019	predicted the mutation status in renal cell carcinoma by using the Random Forest (RF)	AUC 0.98

Table 6 Recent significant studies on head & neck cancer

Study	Year	Details	Outcome
Halicek et al. [23]	2017	developed a CNN classifier to detect differences between normal and cancer affected tissue	81% accuracy
Ma et al. [33]	2017	A CNN-based model that could distinguish between normal and cancerous tissues	91% accuracy
Gupta & Malhi [22]	2018	used a deep learning framework to detect head & neck tumors	98.8% accuracy
Lo et al. [59]	2018	predicting metastasis of a malignant tumor by proposing a Support Vector Machines (SVM) based method	100% accuracy
Li et al. [30]	2018	trained an Artificial Neural Networks (ANN) based model to classify HN tumors	0.812 AUC
Halicek et al. [24]	2019	differentiated head & neck squamous cell carcinoma from normal tissues by using CNN	0.916 AUC
Diamant et al. [16]	2019	used deep learning to predict clinical outcomes of head & neck cancer by considering 300 cases from TCIA	0.92 AUC
Ma et al. [34]	2019	differentiated benign and malignant tumors by using an auto-encoder network	92.32% sensitivity

studies have mostly been engaged in distinguishing between benign or malignant tumors. Accuracy levels dropped whenever the problem at hand went beyond simple binary classification. Many of the studies have been semi-automated where manual feature extraction created significant processing overhead. The use of transfer learning has created resource-consuming architecture in many of the recent studies. Many research works relied on a single database for carrying out the learning process and ended up with less trustworthy results. Although the efforts to classify histological subtypes or grading have been identified in some cases, they are also confined to some particular tumor-type. It has also been observed that many of the studies have considered a single scanner modality. As a result, the existing studies have created different models that may detect a particular type of tumor having a certain type of scanner modality. Thus, there is a great scope for developing a model that can identify different tumors having dissimilar scanner modalities. The present study bridges the research gaps found in the related studies and proposes a new emerging scientific model for automated detection of malignant tumors of different genres.

Table 7 Recent significant studies on breast cancer

Study	Year	Details	Outcome
Dhungel et al. [15]	2017	used CNN for classifying benign and malignant tumors	90% accuracy
Xu et al. [66]	2018	used CNN and INbreast datasets to estimate breast density	93% accuracy
Shen et al. [54]	2019	applied a convolutional model on the InBreast dataset with 0.95 AUC (Single model)	0.98 AUC (model ensemble)
Reddy et al. [60]	2020	used Deep Neural Network with Support value (DNNS)	97% accuracy

Table 8 Recent significant studies on thyroid cancer

Study	Year	Details	Outcome
Torab-Miandoab et al. [58]	2017	applied image enhancement, image segmentation, and feature extraction to determine cold thyroid nodules automatically	99% accuracy
Farihah et al. [20]	2018	aimed to test the reliability of the Ultrasound Classification system in predicting thyroid malignancy	93% accuracy
Park et al. [47]	2019	developed an ultrasound-based deep learning model for the prognosis of thyroid nodules	90% sensitivity
Wang et al. [63]	2019	did the histological subtype classification of thyroid tumor with VGG-19	97% accuracy
Zhang et al. [72]	2019	used a random forest algorithm to diagnose thyroid nodules (0.834 AUC in clinical diagnosis)	0.924 AUC

4 Data acquisition

A dataset is prepared from eight different datasets from The Cancer Imaging Archive (TCIA) [14] representing different tumors. TCGA-BLCA dataset represents Urothelial Bladder Carcinoma (BLCA). The dataset comprises 111,781 images of 120 numbers of patients. Major imaging modalities are Computed Tomography (CT), Magnetic Resonance (MR), Computed Radiography (CR), Positron Emission Tomography (PET), and Digital Radiography (DX). TCGA-KIRP depicts cervical renal papillary cell carcinoma. It has 33 cases comprising 376 series and 26,667 images. Major imaging modalities are CT, MR, and PT. TCGA-LIHC is the Liver Hepatocellular Carcinoma (LIHC) image dataset. It has 97 cases with 1688 series having a total number of 125,397 images. Major imaging modalities are CT, MR, and PT. Non-Small Cell Lung Cancer (NSCLC) radiogenomics dataset has a cohort of 211 subjects. The dataset comprises Computed Tomography (CT), and Positron Emission Tomography (PET)/CT images. TCGA-THCA represents thyroid cancer, having 6 cases in the image set with 28 series and 2780 numbers of images. Major imaging modalities are CT and PET. TCGA-UCEC represents the Uterine Corpus Endometrial Carcinoma. There are 65 cases including 912 series having 75,829 images. Major imaging modalities are CT, CR, MR, and PT. Head & Neck radiomics collection contains clinical data and computed tomography (CT) from 137 head and neck squamous cell carcinoma (HNSCC) patients treated by radiotherapy. TCGA-BRCA represents Breast Invasive Carcinoma. It has 164 cases with 1877 series containing 230, 167 images. Imaging modalities are MR and mammography (MG).

Table 9 Recent significant studies on uterine cancer

Study	Year	Details	Outcome
Malek et al. [35]	2018	proposed a computer-assisted method for distinguishing uterine sarcoma from leiomyomas	92% accuracy
Sun et al. [56]	2019	developed a CADx by using CNN and attention mechanisms which predicted subtypes of the endometrial tumor	84% accuracy
Santhi et al. [53]	2019	used a deep convolutional neural network for malignancy detection in uterine cancer, and their contour extracted images	92.14% accuracy

Table 10 Recent significant studies on lung cancer

Study	Year	Details	Outcome
Mao et al. [36]	2018	used a deep auto-encoder that was used for the four-type classification task	95% accuracy
Liu et al. [32]	2018	trained CNNs for lung nodule classification with ELCAP dataset	90.3% accuracy
Bhatia et al. [8]	2019	did Lung Cancer Detection using Deep Learning Residual Approach with LIDC-IRDI dataset	84% accuracy
Moitra & Mandal [42]	2019	Did the grading of Non-Small Cell Lung Cancer (NSCLC) using the Fuzzy-Rough Nearest Neighbour (FRNN) method	0.96 AUC
Moitra& Mandal [43]	2019	Did AJCC staging of NSCLC Radiogenomics dataset using One-dimensional CNN & Gated Recurrent Unit (GRU)	97% accuracy

The final image acquisition has been carried out by retrieving images from all the aforementioned collections (Fig. 1). Each subject with pre-surgical DICOM images stored in TCIA is identified with a Patient ID that is identical to the Patient ID of the same subject in TCGA. Twenty best scans from each case having pathological data from all the eight collections have been taken to form the final image collection. In this way, 717 cases where the supportive clinical and pathological data are available to have been considered, and from each such case thirty best scans are extracted. Thus, 14,340 radiological images have been collected to form the new image dataset. This newly prepared image collection is heterogeneous to image modalities, cancer types, cancer stages/grades, and demographic characteristics of patients.

5 Methodology

Equations 1 and 2 represent two key techniques used in the study, namely, branching and re-injection [4], respectively.

$$F^{(i)}(W,X) = O_l^{(i)} = f\left(W_l^{(i)} X_l + b_l^{(i)}\right) \tag{1}$$

Here X is the input vector; W is the weight vector; b is the bias; l is the corresponding layer number; O is the output; i is the branch number and $f(\ldots)$ is the non-linear activation function.

$$O_n^{(k)} = O_l^{(i)} + O_m^{(j)} \tag{2}$$

Where i, j, k are different branches of different convolutional layers l, m, n ($n > l \geq m$, and $k \neq i \neq j$).

The concatenation of all the branches [46] is done by using Eq. 3:

$$Y = F'(W,X) = g_c\left(\left\{F^{(j)}(W,X)\right\}\right) \tag{3}$$

In Eq. 3, $\{F^{(j)}(...)\}$ is the collection of output tensors emanating from j ($j = 1, 2\ldots n$; $n > 0$) branches and g_c is the concatenation operation via the axis c of the tensor.

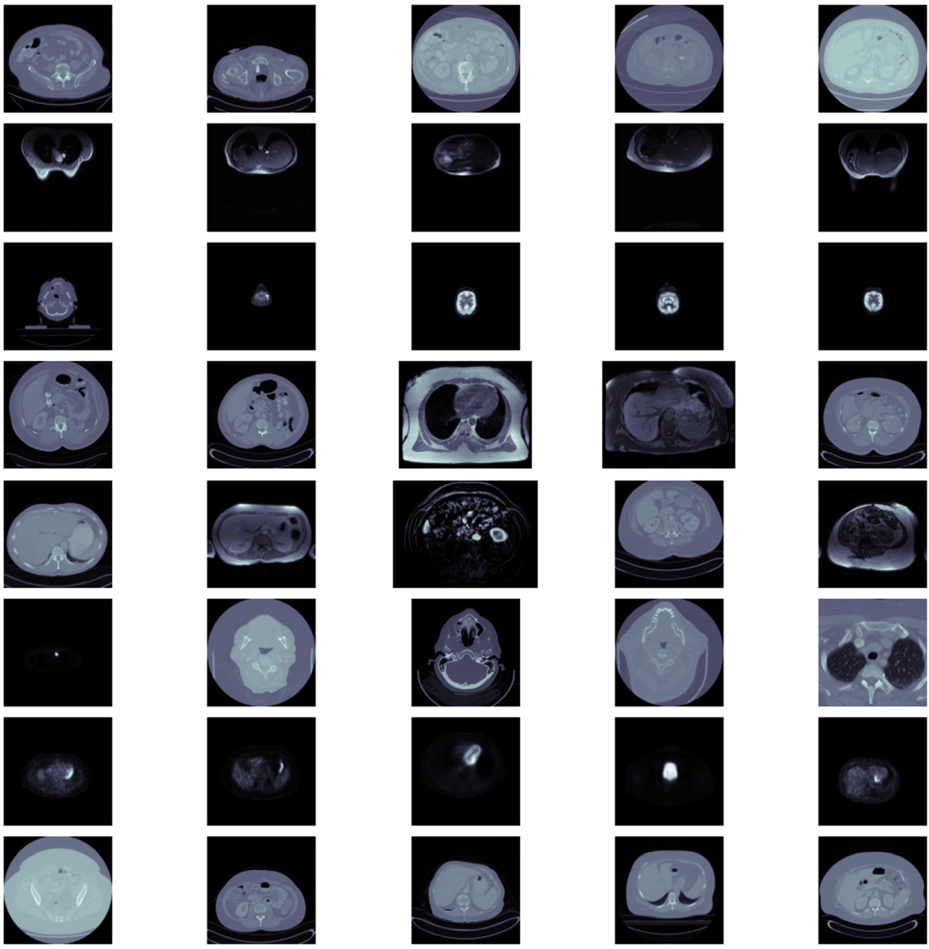

Fig. 1 Glimpse of the final image collection (1st row represents TCGA-BRCA, 2nd row represents Head & Neck Radiomics, 3rd row represents TCGA-BLCA, 4th row represents TCGA-KIRP, 5th row represents TCGA-LIHC, 6th row represents TCGA-UCEC, 7th row represents TCGA-THCA and 8th row represents NSCLC radiogenomics)

Bidirectional LSTM [69] is described in Eq. 4:

$$\begin{aligned} \overrightarrow{H}_t &= \left(X_t \;\; W_{xh}^{(f)} + \overrightarrow{H}_{t-1} \;\; W_{hh}^{(f)} + b_h^{(f)}\right) \\ \overleftarrow{H}_t &= \left(X_t \;\; W_{xh}^{(b)} + \overleftarrow{H}_{t+1} \;\; W_{hh}^{(b)} + b_h^{(b)}\right) \\ y_t &= g\left(W_y \left[\overrightarrow{H}_{t-1} \;\; , \overleftarrow{H}_{t+1} \;\; \right] + b_y\right) \end{aligned} \tag{4}$$

Where t is the timestamp; X_t is the mini-batch input; h is the number of hidden units; $\overrightarrow{H}_t$is the forward and $\overleftarrow{H}_t$is the backward hidden states; ∅ is the layer activation function.

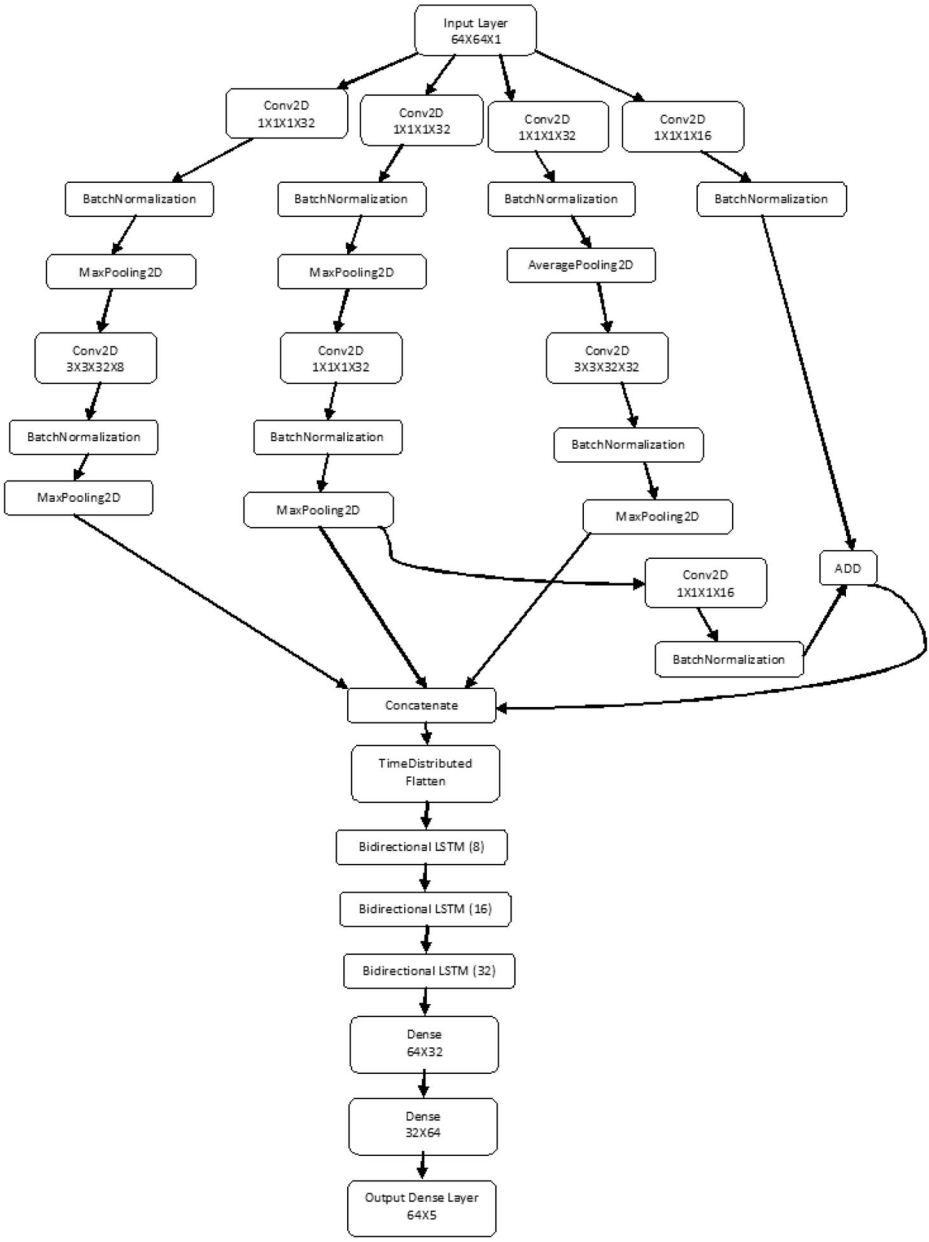

Fig. 2 Proposed non-sequential recurrent deep neural network model

Dense layers [67] are expressed in Eq. 5:

$$Y_i = Wy_t + b \tag{5}$$

bcr_patient_uuid	ajcc_tumor_pathologic_pt	ajcc_nodes_pathologic_pn	ajcc_metastasis_pathologic_pm	ajcc_pathologic_tumor_stage
a6003b1c-56a9-430a-a5e2-b70af3f81bdb	T1	NX	M0	Stage I
ebc094a6-cdc9-4c02-b8ec-a181b25a364d	T3	N0	M0	Stage III
556fcbc8-172a-4af1-8822-ae036e8d68e8	T4	N2	M0	Stage IV
0F4FB205-B13C-4A85-97B3-779429E6CCDD	T3	N0	M0	Stage III
C1AFC5EB-8C1D-4B6C-8453-48728FC023BA	T4a	N0	MX	Stage III
68bc786e-07c7-4964-bc88-5e2a8633626a	T4a	N0	MX	Stage III
22473ad4-df39-4716-b558-3c1409d5fa0e	T3b	N0	M0	Stage III
74b1e72d-fdd0-4d84-8858-3bb32d58ff7e	T2b	N0	MX	Stage II
ebedab7c-dd7c-448e-8324-78c86f13be8b	T3a	N0	MX	Stage III
c0e7fbe7-b9cb-4e65-941b-cd7d581c54ec	T3a	N0	MX	Stage III
f48937ed-e294-41f0-8872-7010391167a9	T3a	N0	M0	Stage III
419deaac-ea45-4bdd-9fa0-b5cd8429b44f	T3b	N2	M0	Stage IV
f8780a4b-7f55-41d1-bf92-8513955d150e	T3b	N1	M0	Stage IV
096b4f32-10c1-4737-a0dd-cae04c54ee33	T3b	N1	M0	Stage IV

Fig. 3 Glimpse of clinical data from TCGA-BLCA

Softmax [64] is used for detecting the class scores from the final layer outcome (Eq. 6):

$$\hat{Y}_i = \text{softmax}\ (Y_i) \tag{6}$$

In Eq. 6, $\hat{Y}_i = (\exp(o_i)/\Sigma_j \exp.(o_j))$ and o_i is the relative levels of confidence [12] for belongingness to each class I ($0 < \hat{Y}_i < 1$).

The most likely class may be found in Eq. 7

$$\hat{i}(o) = i\text{argmax}o_i = i\text{argmax}\hat{Y}_i \tag{7}$$

The cross-entropy loss [71] is expressed as Eq. 8:

$$l(Y, \hat{Y}) = -\Sigma_j\ Y_j\ \log\ \hat{Y}_j \tag{8}$$

Where Y is the actual value and $\hat{Y}$ is the predicted value. The ultimate objective is to minimize the negative log-likelihood [68] or to maximize the accuracy (Eq. 9):

$$L^*(Y, \hat{Y}) = \text{arg}\hat{Y}_i\text{min}\sum_{i=1}^{n} H\left[p(Y_i), p(\hat{Y}_i)\right] \tag{9}$$

Where H[p] is the entropy of distribution [28] p and is calculated as Eq. 10:

$$H[p] = \sum_j -p(j)\log\ p(j) \tag{10}$$

The proposed model (Fig. 2) may be described with the help of steps 1 through 9:

Step 1. Input tensor is fed in four varied parallel convolutional branches (Eq.1).
Step 2. Each convolutional layer is followed by pooling and normalization layers.
Step 3. Layer 2 is added with layer 4 (Eq.2).
Step 4. All the branches are concatenated (Eq.3).
Step 5. The concatenated output is vectorised with time-step.

Image Array	[[[[0.0000e+00]	[[2.4000e+01]	[[1.6424e+04]		[[2.9725e+04]	[[2.3581e+04]	[[2.5629e+04]
	[0.0000e+00]	[2.4000e+01]	[1.6407e+04]		[2.9725e+04]	[2.3581e+04]	[2.5629e+04]
	[0.0000e+00]	[2.4000e+01]	[1.6407e+04]		[2.9725e+04]	[2.3581e+04]	[2.5629e+04]
	...	...	...	...	...	...	...
	[4.9176e+04]	[2.3000e+01]	[2.4000e+01]		[2.2045e+04]	[2.8701e+04]	[1.4879e+04]
	[4.9176e+04]	[2.3000e+01]	[2.4000e+01]		[2.2045e+04]	[2.8701e+04]	[1.4879e+04]
	[1.6184e+04]]	[1.6183e+04]]	[1.6184e+04]]		[1.6285e+04]]	[1.6285e+04]]	[-1.6465e+04]]]]
AJCC Staging	1	1	3		1	1	2

Fig. 4 An instance of pixel arrays and their corresponding AJCC class labels extracted from clinical data

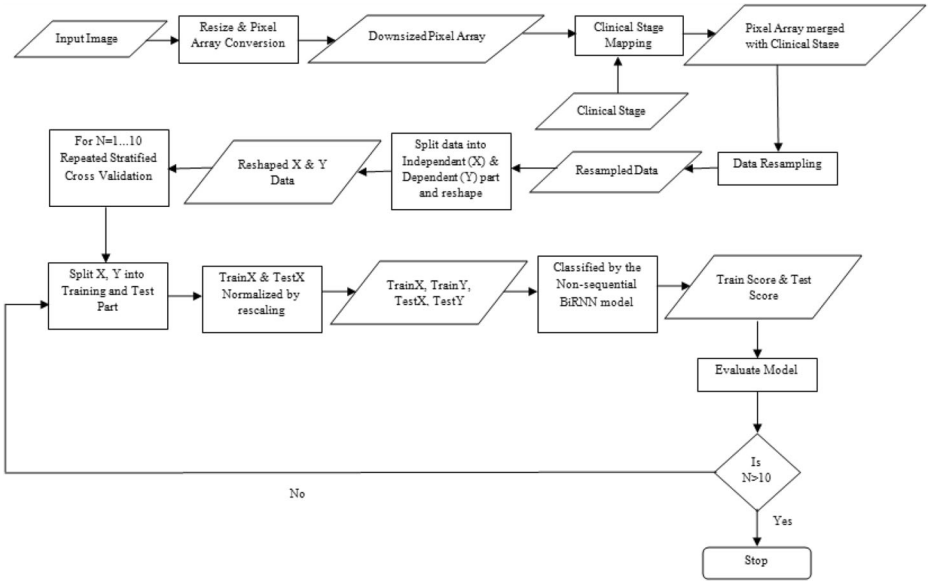

Fig. 5 Flow diagram of the overall classification method

Step 6. The flattened output is injected in bidirectional recurrent layers (Eq.4).
Step 7. The recurrent layer is followed by fully connected dense layers (Eq.5).
Step 8. Class scores and the most likely class are measured by using Eq.6and Eq.7.
Step 9. Loss is measured and minimized by Eq.8, Eq.9, and Eq.10.

Unlike a typical sequential model, the combination of branching and re-injecting layers keep important features alive in the system. In each branch, the initial point-wise Convolutional layer determines features that mix information from the channels of the input tensor. Four dissimilar branches form the heterogeneous ensemble that helps in surpassing the limitation of a typical sequential model. Reinjection makes sure that, even if the output of a layer becomes tiny after activation or down-sampling, it gets regenerated from the original layer output. These layers perform steps like segmentation, feature selection implicitly. Had it

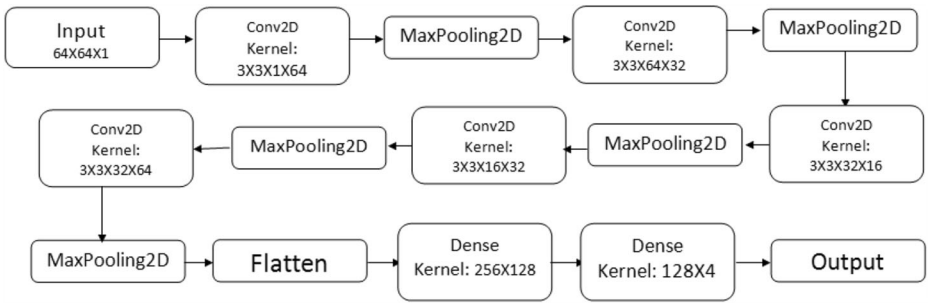

Fig. 6 Sequential CNN model used in the study

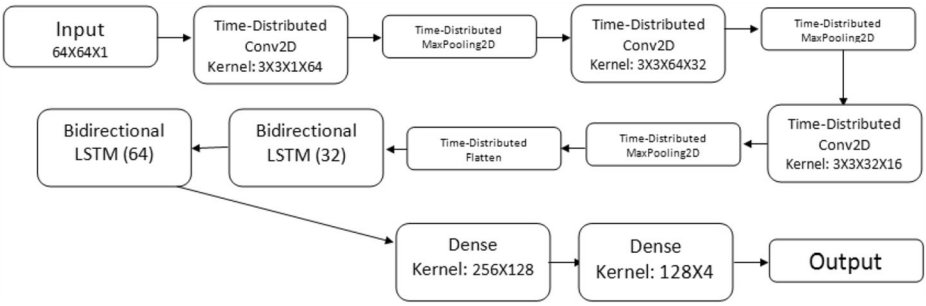

Fig. 7 Sequential CNN + BiRNN model used in the study

been a traditional machine learning model, these steps would have to be done manually. The time-distributed flattened layer vectorises the concatenated output and appends the time-step feature needed by the following bidirectional recurrent layers. The output of the flattened layer passes through two bi-directional Long Short Term Memory (LSTM) layers. The bi-directional LSTM layers can learn from their previous and successive layers. This increases the strength of the classifier as it can adjust the weights and bias from both directions. Finally, the fully connected dense layers produce the class scores (Fig. 2).

All the images are resized to 64*64 for convenience in processing. After getting converted to pixel array, the input dataset typically takes the shape of rank 4 tensor: (number of samples, image height, image width, number of color channels). From the available clinical data (Fig. 3), the AJCC label corresponding to each patient has been tied with the respective image array. The whole dataset has been compressed and loaded for the experiment (Fig. 4). The class imbalance issue has been resolved by Synthetic Minority Over-Sampling (SMOTE) [9]. The input has been fed into the proposed model and has been run for 2000 epochs with an *early stopping* callback value and 200 as *patience* value. The experiment has been done with 10-Fold Repeated *Stratified* Cross-Validation with a repeat value of 10 (Fig. 5). It repeats Stratified K-Fold 10 times with different randomization in each repetition. Here, the number of folds is 10 and the cross-validator gets repeated 10 times with a random state value of 999 for each repetition. It reduces preprocessing bias and correlation between data so that the accuracy never gets artificially inflated. Training and validation data have been rescaled for standardization. Each of the convolutional layers has been regularized by L2 or Euclidean norm and followed by *batchnormalisation* and *MaxPooling* layers or *AveragePooling* layers for normalizing and *down-sampling* the spatial features. Default *padding* and *strides* are used. The batch size used is 128 and *adam* is used as the optimizer with a learning rate of 1e-4. All the hyper-parameters have been determined by conducting a prolonged experiment. At last, the training and validation accuracies are measured and other evaluation metrics [39] like ROC-AUC score, Kappa Statistics, and F1-score are fetched from the confusion matrix. Similar experiments are carried out with other sequential models that performed well in the past in a

Table 11 Best evaluation results of different models measured by various metrics

Model	Validation Accuracy	F1-Score	Cohen's Kappa	ROC AUC
Proposed Non-sequential Model	0.93	0.9	0.86	0.98
Sequential CNN+BiRNN Model	0.84	0.81	0.75	0.9
Sequential CNN Model	0.66	0.64	0.61	0.76

Table 12 Average results (with standard deviations) of different models evaluated by various metrics

Model	Validation Accuracy	F1-Score	Cohen's Kappa	ROC AUC
Proposed Non-sequential Model	0.89 ± 0.001	0.87 ± 0.004	0.84 ± 0.005	0.96 ± 0.002
Sequential CNN+BiRNN Model	0.66 ± 0.01	0.62 ± 0.03	0.65 ± 0.01	0.75 ± 0.02
Sequential CNN Model	0.59 ± 0.03	0.56 ± 0.02	0.57 ± 0.01	0.69 ± 0.01

similar domain [40]. All the experiments have been performed with Python 3.6.8 (IPython 7.5.0) [17].

6 Discussion

The proposed Non-Sequential Recurrent Model Ensemble (NSRME) has been run on the newly formed dataset along with other models like a sequential CNN model (Fig. 6) and a sequential CNN model combined with bi-directional Recurrent Neural Network (CNN + BiRNN) (Fig. 7). The latter two models were quite successful in classifying NSCLC TNM staging and histology subtypes by using the TCIA Radiogenomics dataset, respectively. The best and average results attained by these models are compared and analysed (Tables 11 and 12).

The best training and validation accuracy of the proposed model is found in iteration 2 at epoch 123. Whereas, the best accuracy of CNN + BiRNN is found in iteration 4 at epoch 216 and the same for the CNN model is found in iteration 5 at epoch 267. From Table 11, it may be interpreted that the proposed model's performance is ahead of other sequential models. Kappa statistics nearing 1 is quite encouraging, so are the high ROC-AUC score and high F1-Score.

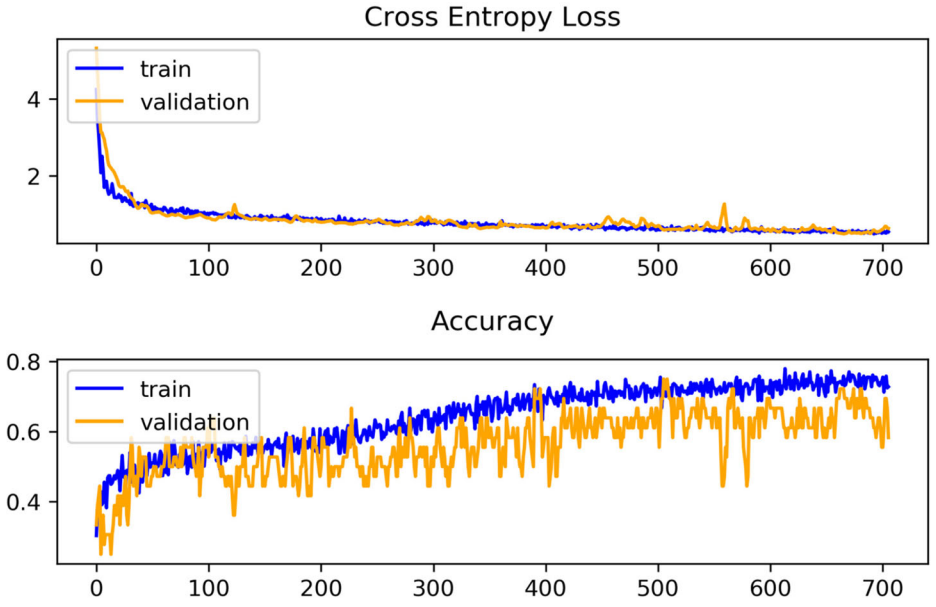

Fig. 8 Training and validation accuracy and loss of the sequential CNN model

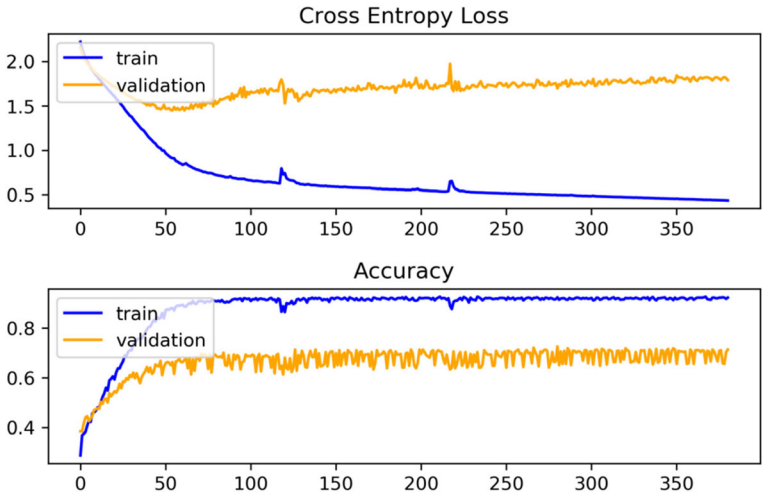

Fig. 9 Training and validation accuracy and loss of the sequential CNN + BiRNN model

Evaluation results (Tables 11 and Table 12) depict a high True Positive Rate (TPR) and less type-I and type-II errors i.e., less False Positives (FPs) and less False Negatives (FNs).

From Figs. 8, 9, and 10 it may be observed that the average validation accuracy and loss of the proposed model are better than other sequential models. From Table 12, it has been found that the average ROC-AUC score of the proposed model is higher than the average ROC-AUC scores of the CNN + BiRNN model and the sequential CNN model. Deviations are also less with the proposed non-sequential model (Table 12). These results imply that classification results have a less rate of miss and fewer false alarms. It has happened as the preprocessing layers of the proposed non-sequential model have not let the important features die out of down-sampling and the bidirectional LSTM layers memorized important features emanated from both the forward and backward path. The average memory usage during the non-sequential model execution also got decreased to 50% which was around 80% during

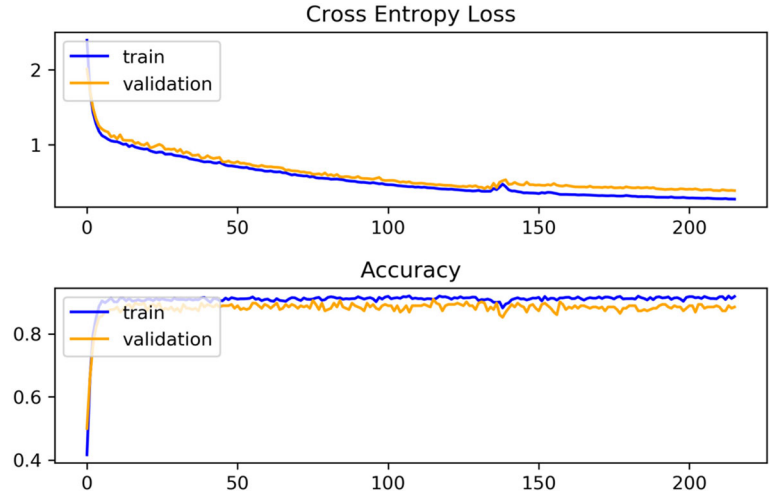

Fig. 10 Training and validation accuracy and loss of the proposed non-sequential model

sequential model execution. This has happened as the inception layers acted as cheaper filters and fewer numbers of time-distributed layers have been used than the CNN + BiRNN model. Thus, it may be concluded that the proposed model has performed steadily in classifying heterogeneous classes of tumors.

When the results of the newly proposed model have been compared with the recent notable studies (Table 13), it has been found that the non-sequential recurrent ensemble of deep neural network has performed satisfactorily.

In Table 13, the recent prominent studies have been compared with the proposed one by considering one of the most trustworthy parameters i.e. *Area under the ROC Curve (AUC)* that depicts the aggregated performance of a classifier against all possible threshold values. From Table 13, it is evident that the performance of the proposed model has indeed matched the top performers in various genres. Most of the existing studies are based on a single tumor-type and a single imaging modality. They often considered a single database and tried to detect subtypes or grades of a particular cancer type. It has been observed that many of them relied on manually crafted features. Thus many important features were ignored and whenever the number of classes increased, the performance was affected. These problems were mitigated by training the proposed model from the scratch and by using automated features. With the newly proposed model, the dataset was a mix of eight databases, imaging modalities were also diverse, and the task was more complicated than the grading of tumors as the number of target classes was more. Despite such intricacy, the non-sequential recurrent model ensemble (NSRME) has truly matched the performance of the leading recent studies. This speaks in favour of the promises made by the proposed model. The study may be considered as a momentous step towards making a revolutionary model that eliminates the need for having different models for identifying different tumor-types.

7 Conclusion

No other model in the existing literature could classify such a varied mix of malignant tumor imagery with such high accuracy. Here lies the novelty of the study. The scientific contribution of the study is also manifold. Unlike the existing models, it helps in determining the overall

Table 13 Comparison of leading studies concerning the parameter Area under the ROC Curve (AUC)

Study	Year	Cancer Type	#Class	Prediction	#Database	Modality	Method	AUC
Lin et al. [31]	2019	Bladder	Two	Survival	Single	Single	LASSO	0.96
Romero et al. [48]	2019	Liver	Two	Metastasis	Single	Single	Inception & Residual	0.97
Kocak et al. [29]	2019	Kidney	Two	Mutation	Single	Single	RF	0.98
Diamant et al. [16]	2019	Head&Neck	Five	Subtype	Four	Single	CNN	0.92
Shen et al. [54]	2019	Breast	Two	Malignancy	Single	Single	CNN Ensemble	0.98
Zhang et al. [72]	2019	Thyroid	Two	Malignancy	Single	Single	RF	0.92
Sun et al. [56]	2019	Uterine	Four	Grade	Single	Single	CNN & Attention	0.95
Moitra & Mandal [42]	2019	Lung	Four	Grade	Single	Single	FRNN	0.96
Proposed Model	**2020**	**All Above**	**Five**	**Stage**	**Eight**	**Multi**	**NSRME**	**0.98**

prognostic group of a tumor irrespective of its type and imaging modality. Once the overall pathological staging is determined, the TNM staging of the respective tumor may also be detected easily. The proposed model may determine histopathological grades and subtypes of different tumors with little customization. In this way, the present study may help medical personnel in determining the stage or grade of tumors in a more assenting way. The present study could not include many tumor genres for a lack of clinical data. In the future, many other types of tumors may be brought under the periphery of the study. In the future, the model may also diagnose blood cancer or leukemia, where no tumors are formed. Experiments may be carried out in the future with different hyper-parameters and meta-learners to improve the model further.

Author's contribution Dipanjan Moitra: Conceptualization, Methodology, Data preparation, Visualization, Investigation, Formal analysis, Writing - original draft, Writing - review & editing.
Rakesh Kr. Mandal: Supervision.

Declarations

Conflict of interest None.

Ethical approval For this type of study formal consent is not required.

Informed consent NOT APPLICABLE.

References

1. Alghamdi AS, Polat K, Alghoson A, Alshdadi AA, Abd El-Latif AA (2020) A novel blood pressure estimation method based on the classification of oscillometric waveforms using machine-learning methods. Appl Acoust 164:107279, ISSN 0003-682X. https://doi.org/10.1016/j.apacoust.2020.107279
2. Alghamdi A, Hammad M, Ugail H, Abdel-Raheem A, Muhammad K, Khalifa HS, Abd el-Latif AA (2020) Detection of myocardial infarction based on novel deep transfer learning methods for urban healthcare in smart cities. Multimed Tools Appl. https://doi.org/10.1007/s11042-020-08769-x
3. Ali AM, Zhuang H, Ibrahim A, Rehman O, Huang M, Andrew W (2018) A machine learning approach for the classification of kidney cancer subtypes using miRNA genome data. Appl Sci 8. https://doi.org/10.3390/app8122422
4. Alom MZ, Hasan M, Yakopcic C, Taha TM, Asari VK (2020) Improved inception-residual convolutional neural network for object recognition. Neural Comput & Applic 32:279–293. https://doi.org/10.1007/s00521-018-3627-6
5. Bektas C, Kocak B, Yardimci AH, Turkcanoglu M, Yucetas U, Koca S, Erdim C, Kilickesmez O (2018) Clear Cell Renal Cell Carcinoma: Machine learning-based quantitative computed tomography texture analysis for prediction of fuhrman nuclear grade. Eur Radiol. https://doi.org/10.1007/s00330-018-5698-2
6. Ben-Cohen A, Klang E, Kerpel A, Konen E, Amitai M, Greenspan H (2018) Fully convolutional network and sparsity-based dictionary learning for liver lesion detection in CT examinations. Neurocomputing: 1585–1594
7. Bharti P, Mittal D, Ananthasivan R (2018) Preliminary study of chronic liver classification on ultrasound images using an ensemble model. Ultrason Imaging 40(6):357–379
8. Bhatia S, Sinha Y, Goel L (2019) Lung cancer detection: a deep learning approach. Soft Computing for Problem Solving 817:699–705
9. Blagus R, Lusa L (2013) SMOTE for high-dimensional class-imbalanced data. BMC Bioinformatics 14: 106. https://doi.org/10.1186/1471-2105-14-106
10. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68:394–424. https://doi.org/10.3322/caac.21492

11. Cha KH, Hadjiiski LM, Cohan RH, Chan HP, Caoili EM, Davenport MS, Samala RK, Weizer AZ, Alva A, Kirova-Nedyalkova G, Shampain K, Meyer N, Barkmeier D, Woolen S, Shankar PR, Francis IR, Palmbos P (2018) Diagnostic accuracy of CT for prediction of bladder cancer treatment response with and without computerized decision support. Acad Radiol 26:1137–1145. https://doi.org/10.1016/j.acra.2018.10.010
12. Chen D, Wang Y, Wang C, Shi C, Xiao B (2020) Selective feature connection mechanism: concatenating multi-layer CNN features with a feature selector. Pattern Recogn Lett 129:108–114, ISSN 0167-8655. https://doi.org/10.1016/j.patrec.2019.11.015
13. Chollet, François (2018) Deep learning with Python. Manning Publications Co., ISBN: 9781617294433
14. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F (2013) The cancer imaging archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 26(6):1045–1057
15. Dhungel N, Carneiro G, Bradley A (2017) A deep learning approach for the analysis of masses in mammograms with minimal user intervention. Med Image Anal 37:114–128. https://doi.org/10.1016/j.media.2017.01.009
16. Diamant A, Chatterjee A, Vallières M, Shenouda G, Seuntjens J (2019) Deep learning in head & neck cancer outcome prediction. Sci Rep 9
17. Dipanjan M, Samanta RK (2015) Performance evaluation of BioPerl, biojava, BioPython, BioRuby and BioSmalltalk for executing bioinformatics tasks. Int J Comput Sci Eng 03(01):157–164
18. Edge S, Compton C (2010) The American joint committee on Cancer: the 7th edition of the AJCC Cancer staging manual and the future of TNM. Ann Surg Oncol 17(6):1471–1474
19. Eminaga O, Eminaga N, Semjonow A, Breil B (2018) Diagnostic classification of cystoscopic images using deep convolutional neural networks. JCO Clinical Cancer Informatics 2:1–8. https://doi.org/10.1200/CCI.17.00126
20. Farihah AG, Nurismah MI, Husyairi H, Shahrun Niza AS, Radhika S (2018) Reliability of the ultrasound classification system of thyroid nodules in predicting malignancy. Med J Malaysia 73:9–15
21. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H (2018) GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing:321–331
22. Gupta P, Kaur Malhi A (2018) Using deep learning to enhance head and neck cancer diagnosis and classification. In: IEEE international conference on system, computation, automation and networking (icscan), Pondicherry, pp 1–6
23. Halicek M, Lu G, Little JV, Wang X, Patel M, Griffith CC, El-Deiry MW, Chen AY, Fei B (2017) Deep convolutional neural networks for classifying head and neck cancer using hyperspectral imaging. J Biomed Opt 22:060503
24. Halicek M, Shahedi M, Little JV, Chen AY, Myers LL, Sumer BD, Fei B (2019) Head and neck cancer detection in digitized whole-slide histology using convolutional neural networks. Sci Rep 9
25. Han S, Hwang S, Lee HJ (2019) The classification of renal cancer in 3-phase CT images using a deep learning method. J Digit Imaging 32:638–643
26. Ikeda A, Hoshino Y, Nosato H, Kojima T, Kawai K, Ohishi Y, Sakanashi H, Murakawa M, Yamanouchi N, Nishiyama H (2018) Objective evaluation for the cystoscopic diagnosis of bladder cancer using artificial intelligence. Eur Urol 17:e1230–e1231. https://doi.org/10.1016/S1569-9056(18)31702-0
27. Ing N, Huang F, Conley A, You S, Ma Z, Klimov S, Ohe C, Yuan X, Amin MB, Figlin R, Gertych A, Knudsen BS (2017) A novel machine learning approach reveals latent vascular phenotypes predictive of renal cancer outcome. Nat Sci Rep 7:13190. https://doi.org/10.1038/s41598-017-13196-4
28. Johnson RW (1979) Determining probability distributions by maximum entropy and minimum cross-entropy. SIGAPL APL Quote Quad 9, 4-P1 (June 1979), 24–229. DOI:https://doi.org/10.1145/390009.804434.
29. Kocak B, Durmaz ES, Ates E, Ulusan MB (2019) Radiogenomics in clear cell renal cell carcinoma: machine learning–based high-dimensional quantitative CT texture analysis in predicting PBRM1 mutation status. Am J Roentgenol 212:W55–W63. https://doi.org/10.2214/AJR.18.20443
30. Li S, Wang K, Hou Z, Yang J, Ren W, Gao S, Meng F, Wu P, Liu B, Liu J, Yan J (2018) Use of radiomics combined with machine learning method in the recurrence patterns after intensity-modulated radiotherapy for nasopharyngeal carcinoma: a preliminary study. Front Oncol 8. https://doi.org/10.3389/fonc.2018.00648
31. Lin P, Wen DY, Chen L, Li X, Li SH, Yan HB, He RQ, Chen G, He Y, Yang H (2019) A radiogenomics signature for predicting the clinical outcome of bladder urothelial carcinoma. Eur Radiol 30:547–557. https://doi.org/10.1007/s00330-019-06371-w
32. Liu XL, Hou F, Hao A (2018) Multi-view multi-scale CNNs for lung nodule type classification from CT images. Pattern Recogn 77:262–275
33. Ma L, Lu G, Wang D, Xu W, Chen ZG, Muller S, Chen A, Fei B (2017) Deep learning based classification for head and neck cancer detection with hyperspectral imaging in an animal model. SPIE–the International

Society for Optical Engineering, Medical Imaging : Biomedical Applications in Molecular, Structural, and Functional Imaging 10137. https://doi.org/10.1117/12.2255562
34. Ma L, Guolan Lu, Dongsheng Wang, Xulei Qin, Zhuo Georgia Chen & Baowei Fei. (2019) Adaptive deep learning for head and neck cancer detection using hyperspectral imaging. Visual Computing for Industry, Biomedicine, and Art 2.
35. Malek M, Gity M, Alidoosti A, Ebrahimi SMS, Tabibian E, Oghabian MA (2018) A machine learning approach for distinguishing uterine sarcoma from leiomyomas based on perfusion weighted MRI parameters. Eur J Radiol 110:203–211. https://doi.org/10.1016/j.ejrad.2018.11.009
36. Mao KM, Tang RJ, Wang XQ, Zhang WY, Wu HX (2018) Feature representation using deep autoencoder for lung nodule image classification. Complexity. 2018:1–11
37. Moitra D (2017) Segmentation strategy of pet brain tumor image. Indian J Comput Sci Eng 0976–5166(8): 575–577
38. Moitra D (2018) Comparison of multimodal tumor image segmentation techniques. Int J Adv Comput Res 9. https://doi.org/10.26483/ijarcs.v9i3.6010
39. Moitra D (2019) Classification of malignant tumors: a practical approach, LAP LAMBERT Academic Publishing, ISBN: 978-613-9-47500-1
40. Moitra D, Kr R (2020) Mandal classification of non-small cell lung cancer using one-dimensional convolutional neural network. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2020.113564
41. Moitra D, Mandal R (2017) Review of Brain tumor detection using pattern recognition techniques. Int J Comput Sci Eng 5(2):121–123
42. Moitra D, Mandal RK (2019) Automated grading of non-small cell lung cancer by fuzzy rough nearest neighbour method. Netw Model Anal Health Inform Bioinforma 8:24. https://doi.org/10.1007/s13721-019-0204-6
43. Moitra D, Mandal RK (2019) Automated AJCC (7th edition) staging of non-small cell lung cancer (NSCLC) using deep convolutional neural network (CNN) and recurrent neural network (RNN). Health Inf Sci Syst 7:14. https://doi.org/10.1007/s13755-019-0077-1
44. Moitra D, Mandal RK (2020) Prediction of non-small cell lung cancer histology by a deep ensemble of convolutional and bidirectional recurrent neural network. J Digit Imaging 33:895–902. https://doi.org/10.1007/s10278-020-00337-x
45. Munir K, Elahi H, Ayub A, Frezza F, Rizzi A (2019) Cancer diagnosis using deep learning: a bibliographic review. Cancers 11:1235
46. Noreen N, Palaniappan S, Qayyum A et al (2020) A Deep Learning Model Based on Concatenation Approach for the Diagnosis of Brain Tumor[J]. IEEE Access 8:55135–55144
47. Park VY, Han K, Seong YK, Park MH, Kim E-K, Moon HJ, Yoon JH, Kwak JY (2019) Diagnosis of thyroid nodules: performance of a deep learning convolutional neural network model vs. radiologists. Sci Rep 9
48. Romero FP, Diler A, Bisson-Gregoire G, Turcotte S, Lapointe R, Vandenbroucke-Menu F, Tang A, Kadoury S (2019) End-to-end discriminative deep network for liver lesion classification. In: 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019), Venice, Italy, pp 1243–1246. https://doi.org/10.1109/ISBI.2019.8759257
49. Sabut S, Das A, Acharya UR, Panda S (2018) Deep learning based liver cancer detection using watershed transform and Gaussian mixture model techniques. Cogn Syst Res 54:165–175. https://doi.org/10.1016/j.cogsys.2018.12.009
50. Sairam T, Vinod PK, Jawahar CV (2019) Pan-renal cell carcinoma classification and survival prediction from histopathology images using deep learning. Sci Rep 9
51. Sato M, Kentaro Morimoto, Shigeki Kajihara, Ryosuke Tateishi, Shuichiro Shiina, Kazuhiko Koike & Yutaka Yatomi (2019) Machine-learning approach for the development of a novel predictive model for the diagnosis of hepatocellular carcinoma. Nat Sci Rep9.
52. Sedik A, Iliyasu AM, Abd El-Rahiem B, Abdel Samea ME, Abdel-Raheem A, Hammad M, Peng J, Abd El-Samie FE, Abd El-Latif AA (2020) Deploying machine and deep learning models for efficient data-augmented detection of COVID-19 infections. Viruses 12(7):769. https://doi.org/10.3390/v12070769
53. Shanthi PB, Faruqi F, Hareesha KS, Kudva R (2019) Deep convolution neural network for malignancy detection and classification in microscopic uterine cervix cell images. Asian Pac J Cancer Prev 20:3447–3456. https://doi.org/10.31557/APJCP.2019.20.11.3447
54. Shen L, Margolies LR, Rothstein JH, Fluder E, McBride R, Sieh W (2019) Deep learning to improve breast cancer detection on screening mammography. Sci Rep 9:12495. https://doi.org/10.1038/s41598-019-48995-4
55. Shkolyar E, Jiacd X, Chang TC, Trivedi D, Mach KE, Meng MQ-H, Xing L, Liao JC (2019) Augmented bladder tumor detection using deep learning. Eur Urol 76:714–718. https://doi.org/10.1016/j.eururo.2019.08.032

56. Sun H, Xianxu Zeng, Tao Xu, Gang Peng & Yutao Ma. (2019). Computer-aided diagnosis in histopathological images of the endometrium using a convolutional neural network and attention mechanisms. https://arxiv.org/ftp/arxiv/papers/1904/.
57. Tian K, Rubadue CA, Lin DI, Veta M, Pyle ME, Irshad H, Heng YJ (2019) Automated clear cell renal carcinoma grade classification with prognostic significance. PLoS ONE:14. https://doi.org/10.1371/journal.pone.0222641
58. Torab-Miandoab A, Rezaei-hachesu P, Samad T, Habibi-Chenaran S, Slemani (2017) Image processing techniques for determining cold thyroid nodules. In: International Conference on Current Research in Computer Science and Information Technology (ICCIT), pp 133–136
59. Tzu-Yun Lo, Peiyin Wei, Chiaheng Yen, Jiing Feng Lirng, Muhhwa Yang, Penyuan Chu, Shinn-Ying Ho. (2018). Prediction of Metastasis in Head and Neck Cancer from Computed Tomography Images. ICRAI 2018: Proceedings of the 2018 4th International Conference on Robotics and Artificial Intelligence. pp. 18–23. https://doi.org/10.1145/3297097.3297108.
60. Vaka AR, Badal Soni, Sudheer Reddy K (2020) Breast cancer detection by leveraging machine learning, ICT Exp 6(4):320–324, ISSN 2405–9595, https://doi.org/10.1016/j.icte.2020.04.009.
61. Vivanti R, Szeskin A, Lev-Cohain N, Sosna J, Joskowicz L (2017) Automatic detection of new tumors and tumor burden evaluation in longitudinal liver CT scan studies. Int J Comput Assist Radiol Surg 12:1945–1957
62. Wang X, Mao K, Wang L, Yang P, Lu D, He P (2019) An appraisal of lung nodules automatic classification algorithms for CT images. Sensors 19:194
63. Wang Y, Guan Q, Lao I, Wang L, Wu Y, Li D, Ji Q, Yu W, Zhu Y, Lu H, Xiang J (2019) Using deep convolutional neural networks for multi-classification of thyroid tumor by histopathology: a large-scale pilot study. Ann Transl Med 7:468. https://doi.org/10.21037/atm.2019.08.54
64. Wu Q, Wang F (2019) Concatenate convolutional neural networks for non-intrusive load monitoring across complex background. Energies 12:1572
65. Xu X, Liu Y, Zhang X, Tian Q, Wu Y, Zhang G, Meng J, Yang Z, Lu H (2017) Preoperative prediction of muscular invasiveness of bladder cancer with radiomic features on conventional MRI and its high-order derivative maps. Abdominal Urology 42:1896–1905. https://doi.org/10.1007/s00261-017-1079-6
66. Xu J, Li C, Zhou Y, Mou L, Zheng H, Wang S. (2018). Classifying mammographic breast density by residual learning. https://arxiv.org/abs/1809.10241
67. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9:611–629. https://doi.org/10.1007/s13244-018-0639-9
68. Yao H, Zhu D, Jiang B, Yu P (2020) Negative log likelihood ratio loss for deep neural network classification. In: Arai K., Bhatia R., Kapoor S. (eds) Proceedings of the Future Technologies Conference (FTC) 2019. FTC 2019. Advances in intelligent systems and computing, vol 1069. Springer, Cham https://doi.org/10.1007/978-3-030-32520-6_22
69. Zhou Yu, Vikram Ramanarayanan, David Suendermann-Oeft, Xinhao Wang, Klaus Zechner, Lei Chen, Jidong Tao, Aliaksei Ivanou and Yao Qian (2015) "Using bidirectional lstm recurrent neural networks to learn high-level abstractions of sequential features for automated scoring of non-native spontaneous speech, " 2015 IEEE workshop on automatic speech recognition and understanding (ASRU), Scottsdale, AZ, pp. 338–345, https://doi.org/10.1109/ASRU.2015.7404814.
70. Zebin T, Rezvy S (2021) COVID-19 detection and disease progression visualization: deep learning on chest X-rays for classification and coarse localization. Appl Intell 51:1010–1021. https://doi.org/10.1007/s10489-020-01867-1
71. Zhang Z, Sabuncu MR (2018) Generalized cross entropy loss for training deep neural networks with noisy labels. In: Proceedings of the 32nd international conference on neural information processing systems (NIPS'18). Curran associates Inc., red hook, NY, USA, pp 8792–8802
72. Zhang B, Tian J, Pei S, Chen Y, He X, Dong Y, Lu Z, Mo X, Huang W, Cong S, Zhang S (2019) Machine learning–assisted system for thyroid nodule diagnosis. Thyroid Radiol Nuclear Med 29. https://doi.org/10.1089/thy.2018.0380
73. Zheng J, Kong J, Wu S, Li Y, Cai J, Yu H, Xie W, Qin H, Wu Z, Huang J, Lin T (2019) Development of a noninvasive tool to preoperatively evaluate the muscular invasiveness of bladder cancer using a radiomics approach. Cancer. 125:4388–4398. https://doi.org/10.1002/cncr.32490
74. Zhou L, Zhang Z, Chen Y-C, Zhao Z-Y, Yin X-D, Jiang H-B (2019) A deep learning-based radiomics model for differentiating benign and malignant renal tumors. Transl Oncol 12:292–300. https://doi.org/10.1016/j.tranon.2018.10.012