A Hybrid High Performance Intelligent Computing Approach of CACNN and RNN for Skin Cancer Image Grading

Manimurugan S (✉ mmurugan@ut.edu.sa)

University of Tabuk <https://orcid.org/0000-0003-1837-6797>

Research Article

Keywords: Melanoma, Histology, Context Aware, CNN, RNN, VGG-19, Inception, ResNet50, DarkNet-53

Posted Date: October 25th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-986978/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Soft Computing on March 29th, 2022. See the published version at <https://doi.org/10.1007/s00500-022-06989-x>.

A Hybrid High Performance Intelligent Computing Approach of CACNN and RNN for Skin Cancer Image Grading

Manimurugan S¹

¹Industrial Innovation & Robotics Center, Faculty of Computers and Information Technology, University of Tabuk, Tabuk City, Kingdom of Saudi Arabia (KSA)

Abstract—Skin cancer is characterized as the uncontrollable growth of skin cells caused by unrepairable DNA damage. Melanoma is the deadliest form of skin cancers caused by melanocyte and early diagnosis supports therapists in curing it. Computational pathology offers a one-of-a-kind ability to spatially dissect certain interfaces on digitized histology images. A hybrid context-aware convolutional neural networks with recurrent neural network (CA-CNN-RNN) based on skin cancer histological images is proposed in this research. The proposed model encodes a histology image's local representation into higher-dimensional features first, then aggregated the feature by consider their spatial arrangement to enable the final predictions. In this research, H&E-stained sectioned images from the Cancer Genome Atlas are used as the dataset for assessment. From 58 images, 37 images were used for training and 21 images are used for testing. The process on histology images of melanoma skin cancer was analyzed and validated with various classifiers such as VGG-19, Inception, ResNet50, and DarkNet-53 using the hybrid CA-CNN-RNN model. The dataset is used to generate the results, which are then analyzed based on criteria such as accuracy, recall, precision, and F-score. The performance analysis shows that the proposed CA-CNN-RNN with different classifiers has performed better and among the classifiers the DarkNet-53 model has the better performance in all the parameters.

Keywords: Melanoma, Histology, Context Aware, CNN, RNN, VGG-19, Inception, ResNet50, DarkNet-53.

I. INTRODUCTION

Melanoma, the harmful type of skin cancers, has risen by around 50 percent in the last decade to over 287,0001 cases. This equates to over 60,0001 melanoma-related deaths per year. The World Health Organization estimates that over a million non-melanoma skin cancer cases are diagnosed every year [1]. Though, most of the countries are not legally reporting the cases of non-melanoma cancer, the real incidence was widely considered to turn into several millions. As a result, skin cancer is the most widespread cancer in the world. Skin cancer is classified into two sorts: non-melanoma and melanoma [2]. Melanoma is a most lethal type of skin cancers. Basal cell carcinoma and squamous cell carcinoma are the most prominent non-melanoma tumors. Merkel cell cancer is a rare, highly active, and rapidly developing cancer [3].

According to the most recent WHO data, the number of melanoma-related deaths will rise by 20% by 2025, increasing to 74% by 2040. Around 2008 and 2018, the incidence of melanoma skin cancer increased by 44 percent, while deaths increased by 32 percent. Australia, Norway, Denmark, Sweden, and Germany are the five countries with the highest rates of skin cancer [4].

Melanoma develops in pigment cells (melanocytes). Unlike other forms of skin cancers, it can quickly expand (metastasizes) to other glands. Cancer may spread across tissue, the lymph system, or the bloodstream. When spreading through skin, it can only disperse through surrounding areas; but, if it penetrated the lymph systems or blood vessel, it could spread to other body tissues [5]. Melanoma-spread tissue become a tumor proliferation, which was hard to treat. Well, since the malignant

development happens on the skin's surfaces, it may be detected by a quick visual examination and fully cured if detected early. A dermatologist's visual examination of the suspicious skin area is the initial step in identifying a malignant lesion. Because several forms of lesions are identical, a proper diagnosis is important [6].

Table I
MELANOMA STAGES WITH DESCRIPTION [9]

Stages	Description
Stage 0	Melanocytes that are abnormal can be located in the skin's outer layer (epidermis). These cells have the potential to develop into tumor. This was also known as "melanoma in situ."
Stage 1	Cancer was developed. Consider the size and ulcerations. This level was split into two stages: IA and IB. Stage 1A: The cancer was less than 1 mm diameter and has no ulcerations. The formations of a skin break. Stage 1B: Whether the cancer has ulcerations but was less than 1 mm thickness, or it has a diameter among 1 mm and 2 mm but no ulcerations.
Stage 2	Considering the size and ulcerations. This stage was split into three sections: Stages 2A, 2B, and 2C. Stage 2A: Whether the cancer has ulcerations and was more than 1 mm thickness but not more than 2 mm thickness, or it was among 2 mm and 4 mm thickness but has no ulcerations. Stage 2B: Whether the cancer has ulcerations and was more than 2 mm thickness but not more than 4 mm thickness, or the cancer was more than 4 mm thickness but has no ulcerations. Stage 2C: The cancer was thicker than 4 mm and has ulcerations.
Stage 3	Consider metastasis. Tumors of any size, without or with ulcerations might exist. It might have expanded to more than one lymph nodes, or tumor cells may be at least 2 cm far from the main cancer region and might be in a lymph vessel among the main tumor and the neighbouring lymph nodes, or there may be smaller tumors in a 2 cm area around the main tumor on/below the skin.
Stage 4	Tumor may also expand to other areas of the body, such as the lungs, brain, or livers, and may be far from the main tumor.

Fig.1. Melanoma Stages [https://bit.ly/3gAIdhT]

Unfortunately, only after the suspicious lesions (or mole) was biopsied or excised can the stage of a melanoma be established. Four specific features are known to decide the stage: tumor diameter, ulceration, and spreads to lymph node or different areas of the body. Staging is critical in designing an adequate recovery plan and assessing the prognosis. Melanoma is classified into five stages, including Stages 0-4 [7-8]. Pathologists use histology slides to examine the microanatomy of cell and tissue under the microscope. Anyway, new advances in digital imaging solution have digital histology slides, allowing pathologists to perform the same study on a computer screen [10-11]. The CA-CNN-RNN based on skin cancer histological images is proposed in this work. The proposed model encodes a histology image's local representation into higher-dimensional feature first, then aggregated the feature by considered their spatial arrangement to enable the final predictions. From the Cancer Genome Atlas 58 H&E-stained sectioned images from formalin-fixed, paraffins-embedded diagnosed block of melanoma skin cancers are used as dataset. Hematoxylin and Eosin (H&E) staining is commonly utilized in histopathology studies because it allows the researcher/pathologist to see the tissue in precise detail. This is accomplished by visibly labelling cell features such as the organelles, nucleus, cytoplasm, and extracellular components. Using a context-aware CNN-RNN architecture, this proposed model assesses melanoma skin cancer from histopathological images.

The remaining part of this research is presented in following sections as, section II discusses the related works, section III presents the proposed method, section IV presents the performance analysis, and section V presents the conclusion and future extension of the research.

II. RELATED WORKS

To develop cells classification method that relied exclusively on cells nuclei morphology, Konstantinos Z P et al. proposed a hierarchical structure that replicated how pathologists viewed tumor architecture and identified tumor heterogeneity. To present the global context, the SLIC super pixel algorithm was utilized for segmentation and identify tumor region in low resolutions H&E-stained histology images of melanoma cancers. In 58 whole-tumor images from the melanoma data set, super pixel classifications classified the images into tumor, stroma, epidermis, and lumen/white space yielded 97.70 percent accuracy in training and 95.70 percent accuracy in testing [12].

Zhiying X et al. proposed a computer-aided methodology for detecting skin cancer from dermoscopy images. To improve diagnostic accuracy, the input images were first denoised using a median filter. To distinguish the cancerous area from the context, CNN based on the SBO algorithm was used. The segmented images were then used to remove a variety of features. An

optimal approach based on the SBO algorithm was used for optimal feature selection and reducing the order of features to refine the classification process. To perform the final classification, the extracted features from the images were categorized using the SBO algorithm and an optimized SVM classifier [13].

Youssef Filali et al. proposed an effective skin-lesions classifications method based on the combination of handcrafted feature (shapes, skeleton, colours, and textures) and feature derived from the very effective deep learning architecture. Then, feature engineering was used to exclude unnecessary features and pick only those that were appropriate. This was attributed, first, to the convergence of handcraft and pre-trained feature, and second, to the utilize of the genetic algorithms for feature selections, which demonstrated the usefulness of those chosen as the better and most important one. For better results, this model could be enhanced by extracting more features from the data [14].

Ammara M and Adel AJ implemented a new method for automated segmentation of histopathological images of skin without the need for user interference. For fine segmentation of cancerous regions, the approach merged the advantages of clustering and evolutionary algorithms such as level set. The improved fuzzy C means clustering with orientations sensitivity aided in the better approximation of the tumor field. Novel sensitivity to orientation Fuzzy C means clustering was utilized to produce the original coarse segmentations and to quantify the level sets evolutions governing parameters. Hence, to complete the segmentation process, a refined quick level set based algorithm was used [15]. Ni Zhang et al. proposed an optimal CNN approach for skin cancer early detection. An improved whale optimization strategy was used to optimize the CNN. The primary goal of CNN training was to improve outcomes for layer parameters, which created a strong relationship between the layers to ensure proper recognition. The optimization algorithm was used to optimize the collection of weights and biases in the network in order to decrease the error between the network output and the desired output [16].

Marwan AA proposed the new prediction technique that classified skin lesion as benign or cancer depended on the regularizer method. Deep CNNs was used to achieve strongly differentiated and theoretically common tasks over the finely grained target classified. As a result, this was the binary classifiers that distinguished among benign and malignant lesion. Furthermore, this CNN model was reviewed on various use cases and yielded positive AUC-ROC data. In this work, this novel regularizer was not appropriate for features selection or features reduction [17]. Muhammad S et al. presented a new context-aware DNN for tumor grading that could add 64 times more contexts than traditional CNN-based patch classifier. The network was suited well for the cancer grading role that was

based on detecting tumor in glandular structure. Two stacked CNNs were used in the proposed context-aware network. The first LR-CNN was used to learn the histological image's local representations. RA-CNN then aggregated the studied local representation based on its spatial pattern. The context-aware model proposed was tested for cancer grading and classifications. For patient survival analyses, this method can conduct downstream analysis at the digital WSI level [18].

K.M. Hosny et al. used AlexNet and transfer learning to classify skin lesions using a deep neural network. To compare with the state of the art, the proposed approach was trained and evaluated using the public dataset ISIC 2018. This approach has the potential to classify seven different types of lesions [19]. M.A. Kassem et al. classified skin lesions using transfer learning and a pre-trained model using GoogleNet. The ISIC 2019 dataset was used to evaluate the proposed model's capacity to classify various types of skin lesions. To solve the problem of image imbalance between classes, this method's performance improved as the quantity of images in each class decreased [20]. K.M. Hosny et al. developed a new deep convolutional neural network-based classification technique for investigating and examining skin cancer. An alternative technique to augmentation, such as rotation and translation, was applied to the segmented images to solve the problem of class imbalance and overfitting [21].

Fig.2. Histology Images of Skin Lesions: (A & B) Benign Melanoma and (C & D) Malignant Melanoma [https://bit.ly/3gAIdhT]

III. PROPOSED METHOD

Because of the large amount of pixel data found in digital histology images, CNN-RNN can be used to analyze them. In this research a hybrid CA-CNN-RNN model was proposed which focused on images to provide a large context. The proposed model encodes a histology image's local representation into higher-dimensional feature first, therefore aggregated the feature by evaluating its spatial arrangement to enable the final predictions.

Fig.3. Proposed Model Workflow

From the Cancer Genome Atlas 58 H&E-stained image from formalin-fixed, paraffins-embedded diagnosed block of melanoma skin cancers were used as dataset. Using CA-CNN-RNN, this proposed model assesses melanoma skin cancer from histopathological images. Figure 2 depicts pathological representations of the benign pigmented nevus and the cancerous melanoma (20x magnification). (A and B) Pigmented nevus that is benign. (C & D) Cancerous melanoma. (A) and (C) have less pigmentations level, while (B) and (D) have high pigmentations level.

Figure 3 shows the processing stages for classifying skin cancer in the proposed model. Initially, at the preprocessing step, dataset samples were preprocessed in order to determine the real features of the images. The dataset was then divided into training and testing sets. The pre-trained CNN architecture combined with an RNN classifier was utilized to classify the test dataset.

a. Hybrid CA-CNN-RNN Model

In this research, a hybrid model combining CNN and RNN architecture is presented to classify skin cancer from histopathological images. The proposed model's first stage encoded the input image Z^i into the features-cube E^i . Most of the input images were patch-based interpreted by the LR-CNN. The proposed architecture is adaptable enough to allow the use of any image classifier as an LR-CNN, like VGG-19, Inception, ResNet50, and DarkNet-53. This adaptability further allows it to utilize pre-trained weight in the part of the small data set. Furthermore, the LR-CNN can be trained independently before being integrated into the proposed model, allowing it to learn concrete representations and resulting in early convergences of the model's context-aware learnings stage. The output function's spatial dimensions e_{ij}^t of a patch z_{ij}^t which differ depending on the dimensions of the input patches and the network architectures for features extraction [18]. Figure 4 represents the detailed architecture of the proposed CA-CNN-RNN hybrid model.

The preceding layer's output was given to the subsequent layers by the operator ($\rightarrow$), and preceding layer's output was represented by the operator ($\bullet$). Since, the proposed model's feedback has the relatively broad spatial dimensions, certain parts of the image may be irrelevant for image labels prediction. Attention block was implemented to reduce the importance of trivial features and vice versa. The weighted feature-cube E' was termed as (Equation 1):

$$E' = L_c^{1 \times 1}(E, \theta_c) \rightarrow L_s(\bullet) \otimes E \quad (1)$$

where $L_c^{1 \times 1}$ and θ_c represent the 1×1 convolution layers and its parameter, accordingly. L_s indicates the soft-max layer, and operator $\otimes$ represents the Hadamard product.

After the LR-CNN was utilized for encoding the essential patch-based representation of image into the features-cube, the primary context block's (CB) goal was to know the spatial contexts among the features cube. The CB discovers the relationship among image patch feature based on their spatial position. Three CB architectures were implemented, each with a different level of sophistication and capacity to collect context data. The first CB is made up of a 3×3 convolution layer, which was then accompanied by ReLU activations and batch normalizations. The next CB used a residual blocks architecture of two distinct filters

Fig.4. Architecture of the Proposed Hybrid Model

size. In contrast to the prior two contexts blocks, the third context block processed the input features-map in paralleled with various filtered size for capturing contexts from differing receptive parts. The features extracted were given as input to the RNN layer.

Using internal memory, the RNN network can handle long-term dependencies. Nodes between layers in CNN's fully connected networks are connectionless and process only one input, but nodes in RNN are connected from a directed graph and process an input in a specified sequence [18]. The cascaded sequence of three contexts blocks ($C(\bullet)$) of the similar form was used. The final prediction V' was computed from the feature of the input image Z as (Equation 2):

$$V' = C(E', \theta_c) \rightarrow L_p^g(\bullet) \rightarrow L_e(\bullet, \theta_{e'}) \rightarrow L_s(\bullet) \quad (2)$$

Using the traditional CNN structure, a CNN tries to produce coarse and fine labels. The softmax loss function was employed during the training phase to jointly optimize the coarse and fine label predictions, as described in (1) (Equation 3).

$$Loss = -\frac{1}{N} \sum_{i=1}^N (\sum_{j=1}^C 1\{x^i = j\} \log p_j + \sum_{k=1}^F 1\{y^i = k\} \log p_k) \quad (3)$$

Where $1\{\cdot\}$ is the indicator function. The characters N , C , and F represent the collection of images, coarse categories, and fine categories, respectively. The softmax probabilities of the coarse and fine categories are denoted by p_j and p_k , respectively. The trained network may be utilized to determine coarse and fine labels at the same time during the inference phase.

Function CA-CNN-RNN Model (x)

weights Define weights

biases Define biases

x Reshape (x , shape [Pixel_X, Pixel_Y, Slice_Count])

conv1 ReLU_Activation_Func (Conv3(x , weights [0]) + biases [0])

conv1 Dropout (0.5)

conv1 Max_pool (conv1)

conv2 ReLU_Activation_Func (Conv3(x , weights [1]) + biases [1])

conv1 Dropout (0.5)

conv1 Max_pool (conv2)

fully_connect Reshape (conv2, Inverse (weights [2]))

fully_connect

ReLU_Activation_Func

*(fully_connect*weights [2] + biases [2])*

fully_connect Dropout (0.8)

*output fully_connect*weights [3] + biases [3]*

Return output

End Function

A recurrent neural network is a feedforward neural network with one or more feedback loops that is designed to process sequential data. Given an input sequence ($x_1, \dots, x_t$), an RNN creates an output sequence ($y_1, \dots, y_t$). RNN is utilized anytime an input-output relationship is discovered based on time and the ability to deal with long-term dependencies. The modelling sequence method is to use an RNN to feed the input sequence to a fixed-sized vector, and then map the vector to a softmax layer. Unfortunately, RNN has a difficulty when the gradient vector increases and decreases exponentially over lengthy sequences. The vanishing gradient and exploding issue make it difficult to learn long-term associations from the RNN architecture's sequences.

However, the Long Short-Term Memory (LSTM) is capable of successfully solving such a long-distance dependencies problem. The primary distinction between LSTM and RNN is because LSTM adds a separate memory cell state to hold long-term states, which it updates or exposes as needed. The LSTM is made up of three gates: an input gate, a forget gate, and an output gate. These gates are used to regulate how much the input should be read (input gate i), if the current cell value should be forgotten (forget gate f), and whether the new cell value should be produced (output gate o).

These gates allow the input signal to prognosticate via the recurrent hidden states without influencing the output; as a result, LSTM can deal with vanishing and exploding gradients and successfully represent long-term temporal dynamics that RNN cannot learn [22].

LSTM neurons are used as recurrent neurons in this research. The gates' definition and the LSTM update at timestep k are as described in the following equations 4 to 9:

$$i_k = \sigma(Z_{yi}y_k + Z_{di}d_{k-1} + Z_{ui}u + w_i) \quad (4)$$

$$f_k = \sigma(Z_{yf}y_k + Z_{df}d_{k-1} + Z_{uf}u + w_f) \quad (5)$$

$$o_k = \sigma(Z_{yo}y_k + Z_{do}d_{k-1} + Z_{uo}u + w_o) \quad (6)$$

$$g_k = \varphi(Z_{ys}y_k + Z_{ds}d_{k-1} + Z_{us}u + w_s) \quad (7)$$

$$s_k = f_k \odot s_{k-1} + i_k \odot g_k \quad (8)$$

$$d_k = o_k \odot \varphi(s_k) \quad (9)$$

Where $\odot$ denotes the product operation, σ was the sigmoid function $\sigma(x) = (1 + e^{-x})^{-1}$, and φ was the hyperbolic tangent function $\left(\varphi(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}\right)$. Other symbols are i_k , f_k , o_k , g_k which stand for input gate, forget gate, output gate, and input modulation gate. The input vector, hidden state, image visual feature, and memory cell are represented by y , d , u , and s , respectively. When updating the LSTM, the image visual feature v was proposed to be imposed at each timestep. The weights and biases that must be learnt are denoted by Z and w . Because the CA-CNN-RNN classifies skin cancer and also trains and predicts skin cancer images, there is no need to construct separate networks for different levels of classification. As a result, the hybrid CA-CNN-RNN is robust and can classify any image category.

IV. EXPERIMENT ANALYSIS

The proposed model's results are experimented using the MATLAB/Simulink tool. Based on histology images of melanoma skin cancer, a context aware CNN-RNN hybrid model for skin cancer grading was proposed in this work. The dataset was used to generate the results, which are then analyzed based on parameters such as accuracy, recall, precision, and F-score. H&E-stained images from the Cancer Genome Atlas were used for the assessment. The melanoma cancer grading was the key focus of this study, and the model can be used to do different cancer grading for any severe disease by using different datasets in the process.

A. Description of Dataset

The Cancer Genome Atlas provided 58 H&E-stained images from formalin-fixed, paraffins-embedded diagnosed block of melanoma cancers. Using Bio-Format, each digitized histology image was scaled to 20x magnifications with the pixel size of $0.504\mu\text{m}$. The expert pathologists presented annotation on the slide for four separate areas to fix the ground truths for regional classifications: tumor region, normal stroma, normal epidermis, and lumen/white space. From 58 images, 37 images selected randomly were used for training and 21 images were used for testing [23].

B. Performance Assessment

By using the parameters like accuracy, recall, precision, and F-score, the results are evaluated based on various networks models such as VGG-19 [24], Inception [25], ResNet50 [26], and DarkNet-53 [27]. The melanoma cancer histology images are split into patch size 1800×1800 , and the labels of every patch was predicted utilizing the proposed model with the level of 224×224 . The performance analysis of these classifier models for melanoma cancer grading is tabulated in table II (Equations 10 to 13).

Table II
MEAN AND STANDARD DEVIATION PERFORMANCE OF CLASSIFIERS

Model	Fold-1	Fold-2	Fold-3	Mean	Std. Dev
VGG-19	92.18	92.54	91.01	90.26	2.06
Inception	94.84	95.47	91.15	93.87	2.65
ResNet50	95.56	92.35	90.74	92.61	3.12
DarkNet-53	94.46	92.29	92.38	93.02	0.99

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

$$Precision = \frac{TP}{TP+FP} \quad (11)$$

$$Recall = \frac{TP}{TP+FN} \quad (12)$$

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (13)$$

Whereas the efficiency of every classifier was equal, Inception outperforms them all by having the high mean accuracy. DarkNet-53 model, on the contrary, exhibits constant efficiency towards 3 folds with minimum standard deviation.

Table III
Performance Evaluation of Classifiers in Training

Model	Accuracy	Precision	Recall	F-score
VGG-19	95.52	94.68	96.81	94.70
Inception	96.76	95.36	97.58	95.05
ResNet50	96.03	95.74	97.02	94.98
DarkNet-53	98.39	97.15	98.92	97.85

Table III and IV represents the performance evaluation of the proposed model with different transfer learning models.

Table IV
Performance Evaluation of Classifiers in Testing

Model	Accuracy	Precision	Recall	F-score
VGG-19	93.72	92.30	94.35	93.16
Inception	95.86	94.01	95.98	94.65
ResNet50	94.56	93.28	95.05	93.08
DarkNet-53	97.14	96.49	98.21	96.50

Fig.5. Accuracy Comparison of Training and Testing Performance

Figure 5 represents the accuracy comparison of training and testing performance based on the classifiers. The accuracy of data from the classifiers represents that the DarkNet-53 model has the highest performance in training and testing with 98.39% and

97.14% and the VGG-19 has the least performance in both cases 95.52% and 93.72%.

Fig.6. Precision Comparison of Training and Testing Performance

Figure 6 represents the comparison of precision evaluation of classifiers from the training and testing data. The DarkNet-53 model has the highest performance in both training and testing evaluation with 97.15% in training and 96.49% in testing. The Inception and ResNet50 obtained similar results in training data with 95.36% and 95.74% and different scenario in testing data.

The comparison of recall performance evaluation on both training and testing data is represented in figure 7. The performance of DarkNet-53 model has delivered the best performance in both the cases with 98.92% and 98.21%. The Inception model has the second leading performance with 97.58% in training and 95.98% in testing.

Fig.7. Recall Comparison of Training and Testing Performance

Figure 8 represents the comparison of F-score performance analysis of training and testing data. In this comparison, the DarkNet-53 model has the highest performance 97.85% in training and 96.50% in testing. The least performance was obtained by VGG-19 in training with 94.70% and ResNet50 in testing with 93.08%.

Fig.8. F-score Comparison of Training and Testing Performance

By using the CA-CNN-RNN model, the feature extraction and classification process on histology images of melanoma skin cancer were analyzed and validated with different classifiers like VGG-19, Inception, ResNet50, and DarkNet-53. This proposed model is adaptable, allowing it to use any network architecture based on local representation learning. The performance review revealed that the proposed CA-CNN-RNN with various classifiers worked higher, with the DarkNet-53 model outperforming the others in each parameter.

V. CONCLUSION

Based on histology images of melanoma skin cancer, a context aware CNN-RNN hybrid model for skin cancer grading was proposed in this research. The proposed model is intended for the classification of larger input images. In this work, H&E histology images from The Cancer Genome Atlas were used as a dataset for assessment. From the dataset of 58 images, 37 images were selected randomly and used for training phase and 21 images were used for testing phase. The process on histology images of melanoma skin cancer was analyzed and validated with various classifiers such as VGG-19, Inception, ResNet50, and DarkNet-53 using the CA-CNN-RNN model. The dataset was used to generate the results, which are then analyzed based on parameters such as accuracy, recall, precision, and F-score. The melanoma cancer grading was the key focus of this work, and the proposed model can be used to do different cancer grading for any severe disease by using different datasets in the process. This proposed model is adaptable, allowing it to use any network architecture based on local representation learning. According to the performance review, the proposed CA-CNN-RNN with different classifiers performed better, and among the classifiers, the DarkNet-53 model performed better in every parameter, as discussed in the results section. In future, the advanced deep learning models can be included with this proposed model for enhanced performance and this proposed model can be used for various cancer grading process using suitable datasets.

ACKNOWLEDGEMENT

We would like to thank the University of Tabuk, Kingdom of Saudi Arabia for providing the support for our research work.

DECLARATIONS

- *Funding:* None
- *Conflicts of interest/Competing interests:* The authors declare no conflict of interest, financial or otherwise.
- *Availability of data and material:* The authors confirm that the data supporting the findings of this research are available within the article.
- *Code availability:* Custom code
- *Authors' Contributions:* There is an author in this article.
- *Human And Animal Rights:* No animals/humans were used for studies that are the basis of this research.
- *Ethics approval:* Not applicable.
- *Consent to participate (include appropriate statements):* Not applicable.
- *Consent for publication (include appropriate statements):* Not applicable.

REFERENCES

- [1]. Manu G, Thomas K, Shaofeng Y, and Saeed H, "Artificial intelligence-based image classifications methods for diagnosis of skin cancers: Challenges and opportunities," *Comput Biol Med*, Vol.127, 104065, 2020.
- [2]. Khushboo M, Hassan E, Afsheen A, Fabrizio F, and Antonello R, "Cancer Diagnosis Using Deep Learning: A Bibliographic Review," *Cancers*, Vol.11, 1235, 2019.
- [3]. Erdem O and Mehmet T, "A survey on automated melanoma detections," *Eng Appl Artif Intel*, Vol.73, pp.50–67, 2018.
- [4]. Peizhen X et al., "Interpretable Classifications from Skin Cancers Histology Slides Using Deep Learning: A Retrospective Multicenter Study," *Tissues and Organs; Artificial Intelligence*, arXiv:1904.06156, 2019.
- [5]. Titus J.B et al., "Deep neural networks are superior to dermatologists in melanoma images classifications," *Euro J Cancer*, Vol.119, pp.11-17, 2019.
- [6]. M. A. Kassem, K. M Hosny, R. Damaševičius, and Mohamed M. Eltoukh, "Machine Learning and Deep Learning Methods for Skin Lesion Classification and Diagnosis: A Systematic Review," *Diagnostics*, Vol.11, no. 8, pp. 1390, 2021.
- [7]. Andre E et al., "Dermatologist-level classifications of skin cancers with deep neural network," *Nature*, 542(7639):115-118, 2017.
- [8]. Sumaiya D, Maha M K, and Saiful I, "Cancer diagnosis in histopathological images: CNN based approach," *Inform Med Unlocked*, Vol.16, 100231, 2019.
- [9]. Yan Xu et al., "Large scale tissues histopathology image classifications, segmentations, and visualizations via deep convolutional activation features," *BMC Bioinformatics*, 18(281):1-17, 2017.
- [10]. Achim H et al., "Pathologist-level classifications of histopathological melanoma image with deep neural network," *Euro J Cancer*, Vol.115, pp.79-83, 2019.
- [11]. Jwan N.S and Subhi R.M.Z, "Skin Lesion Classifications Based on Deep Convolutional Neural Network Architecture," *J Appl Sci Technol Trend*, 2(1):41-51, 2021.
- [12]. Konstantinos Z-P et al., "Capturing global spatial contexts for accurate cell classifications in skin cancers histology," In: OMIA, COMPAY. *Lecture Notes in Computer Sciences*, Vol.11039, pp.52-60, 2018.
- [13]. Zhiying X, Fatima R S, Noradin G, and Navid R, "Computer-aided diagnosis of skin cancers based on soft computing techniques," *Open Med*, Vol.15, pp.860–871, 2020.
- [14]. Youssef Filali, Hasnae EL Khoukhi, My Abdelouahed Sabri and Abdellah Aarab. "Efficient fusions of handcrafted and pre-trained CNNs features to classify melanoma skin cancers," *Multimedia Tool Appl*, Vol.79, pp.31219–31238, 2020.
- [15]. Ammara M and A. Al-Jumaily, "Orientation Sensitive Fuzzy C Means Based Fast Levels Set Evolutions for Segmentations of Histopathological Images to Detect Skin Cancers," In: Madureira A., Abraham A., Gandhi N., Varela M. (eds) *Hybrid Intelligent System. HIS 2018. Advance in Intelligent System and Computing*, Vol.923, pp.501-510, 2020.
- [16]. Ni Zhang et al, "Skin cancers diagnosis based on optimized convolutional neural networks," *Artif Intell Med*, Vol.102, 101756, 2020.
- [17]. Marwan A.A, "Skin Lesions Classifications using Convolutional Neural Networks with Novel Regularizer," *IEEE Access*, Vol.7, pp.38306-38313, 2019.
- [18]. Muhammad S et al., "Context-Aware Convolutional Neural Networks for Grading of Colorectal Cancers Histology Images," *IEEE T Med Imag*, 39(7):2395-2405, 2020.
- [19]. M. A. Kassem, K. M Hosny, and M. M. Foud, "Classification of Skin Lesions into Seven Classes Using Transfer Learning with AlexNet," *Journal of Digital Imaging*, Vol. 33, No. 5, pp. 1325-1334, 2020.
- [20]. M. A. Kassem, K. M Hosny, and M. M. Foud, "Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning", *IEEE Access*, Vol. 8, pp. 114822-114832, 2020.
- [21]. M. A. Kassem, K. M Hosny, and M. M. Foud, "Skin melanoma classification using ROI and data augmentation with deep convolutional neural networks," *Multimedia Tools and Applications*, Vol. 79, No. 33, pp. 24029-24055, 2020.
- [22]. Y. Guo et al. "CNN-RNN: a large-scale hierarchical image classification framework," *Multimed. Tools Appl.*, 77:10251–10271, 2018. <https://doi.org/10.1007/s11042-017-5443-x>.
- [23]. <https://www.openmicroscopy.org/bio-formats/>
- [24]. S. A. Kamran and A. S. Sabbir, "Efficient Yet Deep Convolutional Neural Networks for Semantic Segmentation," 2018 International Symposium on Advanced Intelligent Informatics (SAII), 2018, pp. 123-130.
- [25]. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818-2826.
- [26]. K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770-778.
- [27]. Redmon, Joseph, and A. Farhadi. "Yolov3: An incremental improvement," *arXiv preprint: arXiv:1804.02767*, 2018.