Received: November 4, 2023. Revised: December 1, 2023.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Hidden Markov Random Field Model Based VGG-16 for Segmentation and Classification of Head and Neck Cancer

Ujwala Gaikwad[1]* **Kamal Shah[2]**

[1]*Departmentof Computer Engineering, Terna Engineering College, Mumbai, India*
[2]*Department of Information Technology, Thakur College of Engineering and Technology, Mumbai, India*
* Corresponding author's Email: ujwalagaikwad@ternaengg.ac.in

Abstract: The head and neck squamous cell carcinoma (HNSCC) is a group of malignant tumors that typically originates in squamous cells lining mucous membranes of head and neck regions. This paper proposed a hidden markov random field model (HMRFM) and VGG-16 for effective segmentation and classification of head and neck cancer. The preprocessing is done by using the min-max normalization method which improves the model performance and feeds into the segmentation process. In segmentation, the HMRFM is utilized to effectively identify and isolate segments in sequential data which makes it easier to analyze the accurate data from input sequence. The VGG-16 model is fine-tuned by certain convolutional layers which is utilized due to less image dataset and overcome the overfitting issues. This method is evaluated on the HNSCC-3DCT-RT dataset and attains better results with regards to AUC, accuracy, sensitivity, specificity, and F1-score values about 95.2%, 98.7%, 91.4%, 97.5% and 91.5% correspondingly. The obtained result shows that the HMRFM with VGG-16 is better compared to other techniques like Deep CNN with data augmentation, 3D-unet convolutional neural network (CNN), and voted ensemble machine learning (VEML).

Keywords: Convolutional neural network, Head and neck cancer, Hidden markov random field model, Min-max normalization, VGG-16.

1. Introduction

Head and neck cancer (HNC) is a heterogeneous disorder that initiates in the mucosal layer of head and neck regions [1]. The HNSCC is a group of malignant tumors that typically instigates squamous cells lining the mucous membranes of HN regions [2]. The HNSCC is a substantial global health concern, posing challenges in early treatment, detection, and overall patient management [3]. The yearly worldwide occurrence of HNSCC is 500,000 cases approximately and it is the sixth most common type of cancer [4]. This cancer includes various group of tumors that affects the throat, mouth, or other parts of the head and neck [5]. The treatment includes surgery, chemotherapy, radiation and targeted therapy it is essential to have precise information about accurate cancer location [6]. After the treatment, monitoring the patient is also significant because of their high risk of emerging a second primary cancer [7]. Multimodal image methods like PET/CT generate high specificity and sensitivity in which the PET is mainly utilized to evaluate physiology and CT is mainly utilized to evaluate anatomy [8].

The accurate segmentation of HNC from medical image data such as CT and MRI is fundamental for precise delineation of tumor boundaries and monitoring of disease progression [9]. The HNC segmentation from PET/CT is intended to recognize the best method for leveraging rich information in the context of initial tumor segmentation and output calculation [10]. The tumor demarcation techniques are endless, cost-effective and cause errors. The automatic segmentation techniques for HNC were more helpful for optimizing patient treatment strategies [11]. The AI techniques particularly in CL-based techniques are extensively utilized to assist quick segmentation procedures [12]. Machine

learning techniques with high-dimensional personalized datasets have been shown as a prevailing tool for prognostic and predicting therapeutic effectiveness [13]. The radiomic features are utilized in various ways for detecting the HNC that assists in curing the disease symptoms [14]. The radiomics is utilized to extract information from the radiological image by applying hand-crafted filters on preselected regions [15]. The major contribution is given as follows:

- The preprocessing is done by using the min-max normalization method which improves the model performance and feeds into the segmentation process.
- In segmentation, the HMRFM is utilized to effectively identify and isolate segments in sequential data which makes it easier to analyze the accurate data from input sequence.
- The VGG-16 model is fine-tuned by certain convolutional layers which is utilized due to less image dataset and overcome the overfitting issues.

The research paper is described as follows, section 2 presented the relative researches. The section 3 explained about the proposed method, the section 4 illustrated the results and comparative analysis of the proposed method and section 5 is the conclusion of this manuscript.

2. Literature review

Joonas Liedes [16] introduced an instinctive segmentation of HNC through deep learning techniques. This model utilized a CNN for detection and the 2D-Unet was trained and tested on PET-MRI from 44 patients identified by HNSCC. The images are occupied 12 weeks after chemoradiation treatment and these images produced 178 PET-MRI portions with cancer. The preprocessing protocols are essential to verify high-quality training data. However, the developed model has lack of interpretability which reduce the model performance.

Henri Hellström [17] developed a CNN technique for HNC classification through HNACC from PET images. The developed model was trained, and tested on 5fold cross-validation through data of 1990 2D images attained by separating the original 3D images. The deep CNN technique was utilized through U-Net architecture to classify data into dual groups according to whether a given image comprises cancer or not. This model assists in preventing overfitting issues. However, in the developed model the augmentation technique enhances the sample size.

Pavel Nikulin [18] implemented an automated lesion delineation with residual 3D U-Net Convolutional Neural Network integrating a multi-head self-attention block. The network capacity was differentiated among primary and LN metastases which were evaluated by considering voxels subset in both manual and CNN delineations. This model performance was evaluated through a 5-fold cross-validation primary dataset and pooling of results from 5 established models in external data. The developed 3D U-net CNN model has advantages like high accuracy rates and robustness to noise. However, this model has limited memory, false tolerance.

Fuk-Hay Tang [19] presented various tumor volume sizes for the prediction of HNSCC through VEML model. The six various tumor volumes such as gross tumor volume (GTV), planning target volume (PTV), extended GTV and PTV, and diminished GTV and PTV. The extracted features are examined through the random forest, decision tree, support vector machine and extreme boosting. The developed VEML model enhances the accuracy by integrating multiple model strengths and minimizing the impact of individual models. However, this model does not identify the substructure of tumor regions.

Anirudh Chandrashekar [20] introduced a DL technique for HNC through CT images. The random forest classification with CT radiomic features was utilized to discriminate tumor and inactive tissues. Then, the generative adversarial network (GAN) was utilized to train the CT without tracer injection. The developed model efficiently enhances the accuracy due to its capability for capturing data distribution. However, this model has limitations such as inherent patient inconsistency and enhanced radiation exposure.

Nadia Brancati [21] suggested a head and neck cancer staging classification through radiomics features. These features are extricated from PET/CT images then selection was performed to remove redundancy in the dataset. The ML techniques are utilized for classification and diminishing the features. The shape features are considered from PET and CT images which is most relevant in terms of pN-stage classification. This technique was utilized to ensure generalization and dependability and reduce the computational time. However, by utilizing this method the low-grade and high-grade tumors were not classified.

Jianan Chen and Anne L. Martel [22] implemented an HNC segmentation through 3D-Unet with a multi-instance neural network. The 3D-Unet was utilized for segment tumor and lymph nodes according to CT images. Various combinations of loss functions with three sets of features are

ensembled to produce a robust model. The output of the SoftMax ensemble model was fused by PET scans. The location and clinical variables of tumors are encoded with radiomic features as extra inputs. The developed model enables continuous segmentation of 3D volumes with better performance and accuracy which is adopted to resolve various segmentation issues. However, this model has trained only CT data due to its restricted time and resources.

Mark Gardner [23] developed a DL-based segmentation in HNC during radiation therapy. The conditional GAN was utilized to detect and segment cancer during RT. The procedure of producing realistic and synthetic CT deformation was established to augment training data and create segmentation techniques for patients. The training data was augmented through synthetically deformed CT to produce additional DRRs. The developed model enhances the overall performance and minimizes the overfitting issues with a lower learning rate. However, this model was not suitable for detecting a tumor in its early stages.

Qiongwen Zhang [24] implemented a Multi objective, multi classifier, multi-modality radiomics model for head and neck cancer prediction. The developed model with PET/CT radiomic features and clinical data for predicting the results of patient with HNSCC after the radiation treatment. This model presented the potential for detect the cancer in early stages after radiation treatment. However, the developed model was not adopting to handle the long-term dependencies and sequential information in input data.

From the overall analysis, the 2D-Unet has lack of interpretability which reduce the model performance. The deep CNN the augmentation technique enhances the sample size. The 3D-Unet CNN has limited memory, false tolerance. The VEML does not identify the substructure of tumor regions. The conditional GAN was not suitable for detecting a tumor in its early stages. Hence, these limitations are considered and overcomes by proposed CNN classifier in this manuscript.

3. Proposed methodology

The HNSCC-3DCT-RT dataset utilized in this paper has 3D high-resolution fan-beam CT scans. The preprocessing utilizes the min-max normalization method which improves the performance of the proposed model. In segmentation, HMRFM is utilized to effectively identify and isolate segments in sequential data which makes it easier to analyze the accurate data from input sequence.

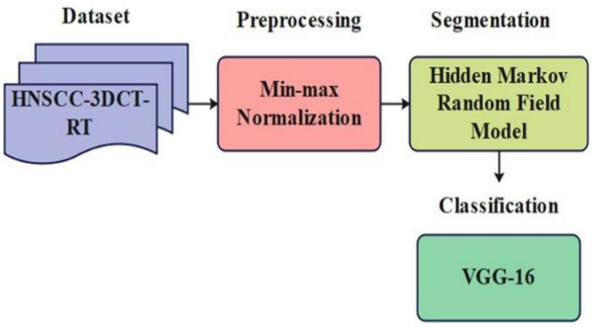

Figure. 1 Block diagram of the proposed methodology

After segmentation, the segmented regions are classified using the VGG-16 model. The HMRFM enhances processing power, minimizing noise and dimensionality problems like model processing time, computational cost, and complexity. The VGG-16 model is fine-tuned by certain convolutional layers which is utilized due to less image dataset and overcome the overfitting issues. Fig. 1 illustrates the block diagram of the proposed methodology.

3.1 Dataset

The dataset used in this analysis is the HNSCC-3DCT-RT dataset from TCIA [25] which comprises 3D high-resolution fan-beam CT scans. This is composed through pre-, mid and post-treatment utilizing Siemens 16-slice CT scanner by medical protocol for 31 HNSCC patients. The pre-treatment was accomplished with a median of 13 days earlier treatment, the mid-treatment was accomplished at fraction 17 and the post-treatment was accomplished at fraction 30. Additionally, this dataset comprises contours of anatomical structures for radiotherapy.

3.2 Preprocessing

The HNSCC-3DCT-RT dataset is preprocessed by the min-max normalization method which improves the model performance. Data preprocessing is the process of converting the raw data into a desired format, the dataset from several resources may consist of incomplete data. So, for further analysis, this data is required to be filtered and normalized. Data normalization is the process of preprocessing the input data. The highest score of the feature is transformed into 1, the smallest score of the feature is transformed into 0 and other values of the feature are transformed into an integer between 0 and 1. The mathematical representation of min-max normalization is shown in Eq. (1),

$$f(x,y) = \frac{f(x,y) - Z_{min}}{Z_{max} - Z_{min}} \quad (1)$$

Where, f is the input image, x and y are pixel positions in the image, Z_{max} and Z_{min} are the maximum and minimum pixel values. After the preprocessing step, the features are segmented by using a multilevel threshold technique.

3.3 Segmentation

The segmentation is performed by utilizing the hidden markov random field model which is a statistical model utilized to define a sequence of observed events where every event is associated with unobservable and underlying states. Segmentation involves dividing a sequence of observations into segments where every segment corresponds to a distinct hidden state. In segmentation, the HMRFM is utilized to effectively identify and isolate segments in sequential data which makes it easier to analyze the accurate data from the input sequence [26]. The complexity is to recognize the process of implicit parameters from the observable parameter and then utilize these parameters for further process. In HMRFM, the state is undistinctive then a few variables affected by the state are observable. Every state has particular distribution possibilities on the possible results so the output sequence symbols can expose few efficient data about the sequence state.

In HMRFM, let us consider the S is utilized to illustrate the state set where $S = \{s_1, s_2, \dots, s_i, \dots, s_n\}$, n in many states. Though these states are hidden, there are few physical data related to events in our set of states. Assume $H = \{h_1, h_2, \dots, h_i\}$ illustrates observation symbols set, where i is a various number of observations which is to be a result from every state. The possibility distribution of state transmission is $A = \{a_{ij}\}$ and the distribution possibility observations of state j is $B = \{b_j(k)\}$ that can illustrate the possibility of state j results from the equivalent observation scores. The primary distribution state is $\psi = \{\psi_i\}$. According to this, the HMRFM can be distinguished by five-tuple $\sigma = (S, 0, A, B, \psi)$. The HMRFM parameter set is illustrated below:

1. Set of states $S = \{s_1, s_2, \dots, s_i, \dots, s_n\}$
2. In the set of observed states for continuous observations, there can be unlimited elements in the set
3. Primary possibility
4. Transfer possibility
5. Possibility of observable symbols in jth state.

The primary HMRFM model transition is from left to right to minimize the HMRFM model parameters. The primary state of the HMRFM model estimates the probability π and transition probability matrix A. When an observation is constant, let $\delta\theta$ illustrate a target range when it continuously estimates target and θi denotes range of state i. Concerning the HMRFM state transition, state i is exposed to $\delta\theta < \theta i$. Let's consider, the $A = \{a_{ij}\}$ illustrates the possibility of state i to the following stage j, then the primary score of A is presented in Eq. (2) and (3),

$$a_{ij-1} = a_{ij+1} = \delta\theta/2\theta_i \tag{2}$$

$$a_{ij} = (\theta_i - \delta\theta)/\theta_i \tag{3}$$

Additionally, the primary positioning of target is commonly measured to be regularly distributed thus, the possibility of primary state is calculated in Eq. (4),

$$\pi_i = \theta_i / \sum_{i=1}^{N} \theta_i \tag{4}$$

As mentioned above equations, number of HMRFM models is established on target type, additionally, it minimizes the identification performance of the HMRFM model. If ith HMRFM produces the highest probability, then the sequential features are integrated by i th target type. Let's consider, attaining the sequence $o = \{o_1, o_2, o_3, \dots, o_M\}$ from unidentified target type T, then the observation sequence probability o is specified by joint probability $P(o/q, T)P(q/T)$ over possible paths q which is shown in Eq. (5),

$$P(o/T) = \sum_{all\ q} P(o/q, T)P(q/T) \tag{5}$$

If target type T_i generates maximum probability for observation sequence o, which is presented in Eq. (6),

$$P(o/T_i) \geq P(o/T_k), \forall T_k \tag{6}$$

At that time, consider the sequence o goes to the target type T_i. The HMRFM model completely learns features to produce identification results. The HMRFM enhances processing power, minimizing noise and dimensionality problems like model processing time, computational cost and complexity.

3.4 Classification

In this research, a pre-trained CNN model is utilized for classification which utilizes various labeled images for model training acquired from the segmentation process. The CNN takes an input and allocates weights to several image features that the image is different from various images which are

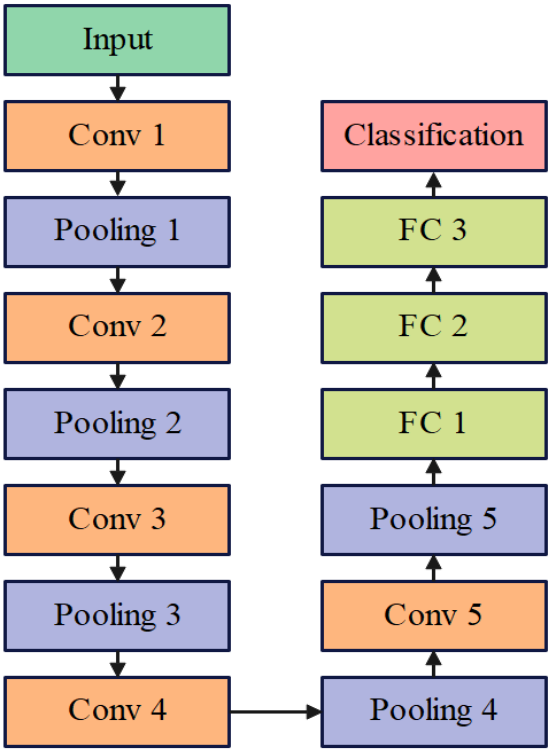

Figure. 2 Structure of VGG-16 model

managed by the same methodology. CNN has a unique capability to capture the spatial dependencies of the images. This CNN converts the image into one form that is geometrically simple while managing the crucial features that are obtainable in the image. The essential block of CNN is convolution operation which is a distinctive feature among regular neural networks and CNN. The main role of the convolution operation is to extract the highest feature from images. Another important operation is the pooling operation which mainly removes image dimensionality in energy to minimize computational complications that need to develop the input information. Both convolutional and pooling operations are presented as layers in CNN. Both are combined and form ith layer of CNN and it is dependable for feature extraction in CNN. The extracted features are stuffed into a regular neural network for classification. The pre-trained CNN model utilized for the classification is VGG-16. The dense layer is an important part of the classification of head and neck tumors which depends on the results from convolution layers. It produces results from the previous layer to every neuron which generates a single result to the next layer. The VGG-16 pre-trained CNN model is utilized due to less image dataset and to overcome the overfitting issues, this model is fine-tuning through certain convolutional layers. The VGG-16 pretrained CNN model contains 16 convolution layers. Fig. 2 represents the structure of CNN based VGG-16 model.

The model takes the head and neck cancer image as input with a dimension of $224 \times 224 \times 3$. The model includes conv layers with a fixed filter size of 3×3 and five max-pooling layers of dimension size of 2×2 within the network. The model also has an activation function of ReLU and two FC layers which have a SoftMax output layer. The VGG-16 model is an extensive network that includes 138 million hyperparameters approximately. The method assembles multi-conv layers to develop deep neural networks which improves the capacity to learn features of handcrafted which is invisible.

4. Experimental result

In this paper, the proposed model is stimulated by utilizing a Python environment with the system configuration: RAM:16GB, processor: intel core i7 and operating system: windows 10. The parameters like dice score coefficient (DSC), intersection over union (IoU), area under curve (AUC), accuracy, sensitivity, specificity and F1-score are utilized to evaluate the model performance. The mathematical representation of these parameters is shown in Eqs. (7), (8), (9), (10), (11), (12) and (13),

$$DSC = \frac{2 \times TP}{(TP+FP)+(TP+FN)} \tag{7}$$

$$IoU = \frac{TP}{(TP+FP+FN)} \tag{8}$$

$$AUC = \frac{\sum R_i(I_I) - I_I(I_I+I)/2}{I_I+I_f} \tag{9}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{11}$$

$$Specificity = \frac{TN}{TN+FP} \tag{12}$$

$$F1-score = 2 \times \frac{Precision \times sensitivity}{Precision+sensitivity} \tag{13}$$

Where, $Precision = \frac{TP}{TP+FP}$, R_i is the rate of ith image, I_I and I_f is the number of positive and negative images, TP, TN, FP and FN illustrate the True Positive, True Negative, False Positive and False Negatives respectively.

Table 1. Performance of HMRFM with other methods

Methods	DSC (%)	IoU (%)	Sensitivity (%)	Specificity (%)	HD (mm)
K-means	88.2	87.4	89.6	85.8	21.2
Fuzzy C-means	89.4	88.6	91.4	86.9	19.8
Otsu's Threshold	91.5	90.7	92.8	88.5	18.7
GMM	93.7	92.2	94.5	89.4	16.4
HMRFM	**94.3**	**94.5**	**96.2**	**91.7**	**13.5**

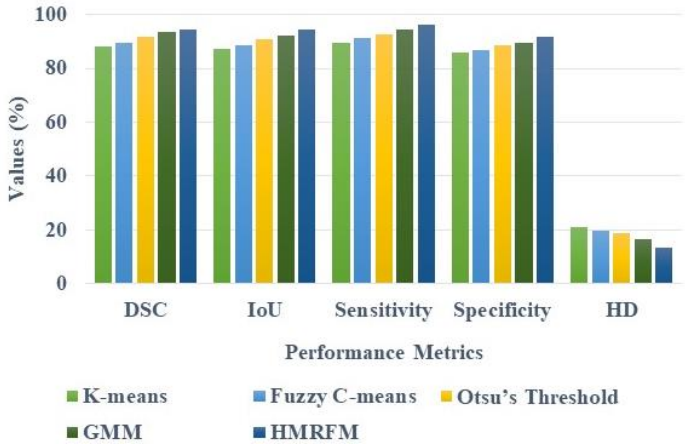

Figure. 3 Performance of HMRFM with other methods

Table 2. Performance of VGG-16 with other methods

Methods	AUC (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
AlexNet	86.7	90.8	85.3	89.4	85.8
EfficientNet	88.4	91.5	86.8	90.6	87.2
DenseNet	90.6	93.2	88.4	92.5	88.7
ResNet-50	92.5	94.6	89.2	94.7	90.3
VGG-16	**93.4**	**96.8**	**89.2**	**95.1**	**90.3**

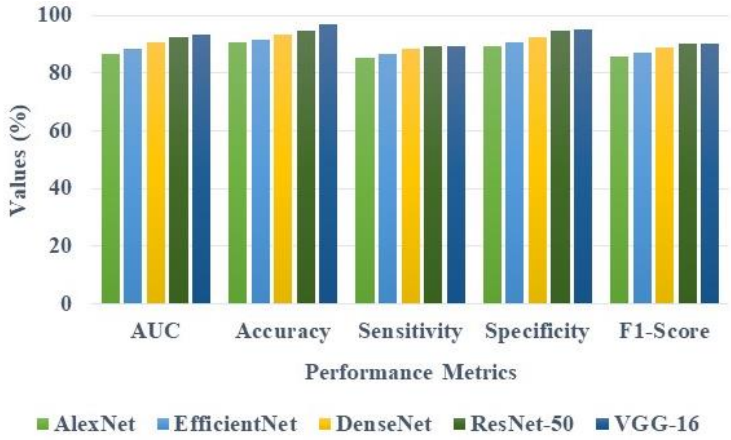

Figure. 4 Performance of VGG-16 with other methods

4.1 Quantitative and qualitative analysis

This section illustrates the quantitative and qualitative analysis of HMRFM with the VGG-16 model with regards to DSC, IoU, Hausdorff Distance (HD), AUC, accuracy, sensitivity, specificity and F1-score are shown in Tables 1, 2 and 3. Table 1 illustrates HMRFM performance with other algorithms, Table 2 illustrates the performance of VGG-16 with other algorithms and Table 3 illustrates the performance of proposed HMRFM with VGG-16 model with other methods. The HMRFM enhances processing power, minimizing noise and dimensionality problems like model processing time, computational cost and complexity. The VGG-16 model is fine-tuned by certain convolutional layers which is utilized due to less image dataset and tackles the overfitting issues.

Table 3. Performance of proposed HMRFM with VGG-16 with other methods

Methods	AUC (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
AlexNet	89.5	91.6	85.5	89.3	86.4
DenseNet	90.3	93.5	87.7	91.6	87.2
ResNet-50	92.8	94.2	88.3	93.4	89.1
VGG-16	93.4	96.8	89.2	95.1	90.3
HMRFM with VGG-16	**95.2**	**98.7**	**91.4**	**97.5**	**92.7**

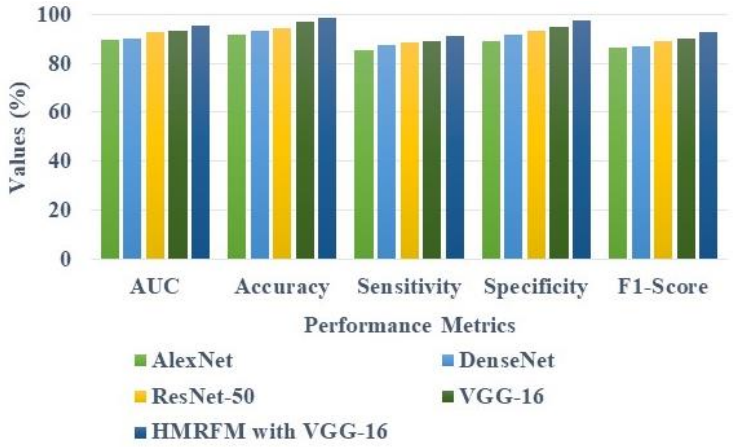

Figure. 5 Performance of proposed HMRFM with VGG-16 with other methods

Table 4. Comparative analysis of the proposed method with existing methods

Method	Dataset	AUC (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
Deep CNN model with data augmentation [17]	HNSCC-3DCT-RT	85.1	78.6	72.4	84.9	77.4
3D U-Net CNN [18]		N/A	98.0	N/A	N/A	N/A
VEML [19]		N/A	88.3	79.9	96.6	N/A
ML based Grid Search optimization [21]		93	N/A	N/A	N/A	N/A
Multi objective, multi classifier, multi-modality radiomics model [24]		N/A	85	93	83	N/A
Proposed HMRFM with VGG-16 model		**95.2**	**98.7**	**91.4**	**97.5**	**92.7**

Table 1 and Fig. 3 represent the performance of the HMRFM model by using performance metrics like DSC, IoU, Sensitivity, specificity, and HD. The performance of k-means, fuzzy c-means, Otsu's threshold, and Gaussian Mixture Model are compared to the proposed HMRFM. The obtained result shows that proposed HMRFM attains a DSC of 94.3%, IoU of 94.5%, sensitivity of 96.2%, specificity of 91.7 and HD of 13.5 mm which is comparatively better than the other existing methods.

Table 2 and Fig. 4 represent the performance of VGG-16 model by using performance metrics like AUC, accuracy, sensitivity, specificity and f1-score. The performance of AlexNet, EfficientNet, DenseNet and ResNet-50 are compared to the VGG-16 method. The attained result shows that proposed VGG-16 attains an AUC of 93.4%, accuracy of 96.8%, sensitivity of 89.2%, specificity of 95.1%, and f1-score of 90.3% which is comparatively higher than the other existing methods.

Table 3 and Fig. 5 represent the performance of HMRFM with the VGG-16 model by using performance metrics like AUC, accuracy, sensitivity, specificity and f1-score. The performance of AlexNet, DenseNet, ResNet-50 and VGG-16 are compared with the proposed HMRFM with the VGG-16 model. The obtained result shows that the proposed HMRFM with VGG-16 model attains an AUC of 95.2%, accuracy of 98.7%, sensitivity of 91.4%, specificity of 97.5%, and f1-score of 92.7% which is comparatively higher than the other existing methods.

4.2 Comparative analysis

This section illustrates the comparative analysis of the proposed HMRFM with the VGG-16 model

with performance metrics like AUC, accuracy, sensitivity, specificity and f1-score as shown in Table 4. The existing result such as [17, 18, 19, 21] and [24] are used for evaluating the ability of the model. The result obtained from Table 4 shows that the proposed HMRFM with the VGG-16 model attains better performance when compared with the existing methods. The AUC was improved to 95.2%, accuracy of 98.7%, sensitivity of 91.4%, specificity of 97.5% and f1-score of 92.7%.

4.2.1. Discussion

In this section, the advantages of the proposed method and the limitations of existing methods are discussed. The existing method has some limitations such as the 2D-Unet [16] model has limited training data and the augmentation technique enhances the sample size. The Deep CNN model with data augmentation [17] enhances the network complexity and computational cost. The 3D-U-net CNN [18] model has huge computational costs and extreme memory usage. The VEML [19] model has high computational cost, time-consuming and memory requirements. The GAN [20] model has limitations such as inherent patient inconsistency and enhanced radiation exposure. The ML-based Grid Search optimization [21] model has limitations such as time-consuming and overfitting issues. The 3D-Unet [22] model has trained only CT data due to its restricted time and resources. The proposed HMRFM with the VGG-16 model overcomes these existing model limitations. The HMRFM is utilized to effectively identify and isolate segments in sequential data which makes it easier to analyze the accurate data from the input sequence. The HMRFM enhances processing power, minimizing noise and dimensionality problems like model processing time, computational cost and complexity. The VGG-16 model is fine-tuned by certain convolutional layers which is utilized due to less image dataset and overcome the overfitting issues.

5. Conclusion

The HNSCC is a group of malignant tumors that typically instigates in the squamous cells lining mucous membranes of HN regions. This paper proposed a HMRFM and VGG-16 for effective segmentation and classification of head and neck cancer. The preprocessing is done by using the min-max normalization method which improves the model performance and feeds into the segmentation process. In segmentation, the HMRFM is utilized to effectively identify and isolate segments in sequential data which makes it easier to analyze the accurate data from input sequence. The VGG-16 model is fine-tuned by certain convolutional layers which is utilized due to less image dataset and overcome the overfitting issues. This method is evaluated on the HNSCC-3DCT-RT dataset and attains better results with regards to AUC, accuracy, sensitivity, specificity, and F1-score values about 95.2%, 98.7%, 91.4%, 97.5% and 91.5% correspondingly. In the future, hyperparameter tuning can be applied to improve the classification performance.

Conflicts of interest

The authors declare that they have no conflict of interest.

Author contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration, have been done by 2nd author.

References

[1] M. Thor, A. Iyer, J. Jiang, A. Apte, H. Veeraraghavan, N. B. Allgood, J. A. Kouri, Y. Zhou, E. LoCastro, S. Elguindi, and L. Hong, "Deep learning auto-segmentation and automated treatment planning for trismus risk reduction in head and neck cancer radiotherapy", *Physics and Imaging in Radiation Oncology*, Vol. 19, pp. 96-101, 2021.

[2] S. Francis, P. B. Jayaraj, P. N. Pournami, M. Thomas, A. T. Jose, A. J. Binu, and N. Puzhakkal, "ThoraxNet: a 3D U-Net based two-stage framework for OAR segmentation on thoracic CT images", *Physical and Engineering Sciences in Medicine*, Vol. 45, No. 1, pp. 189-203, 2022.

[3] V. I. Strijbis, M. Dahele, O. J. G. Champion, G. J. Blom, M. R. Vergeer, B. J. Slotman, and W. F. Verbakel, "Deep learning for automated elective lymph node level segmentation for head and neck cancer radiotherapy", *Cancers*, Vol. 14, No. 22, p. 5501, 2022.

[4] Z. Wei, J. Ren, S. S. Korreman, and J. Nijkamp, "Towards interactive deep-learning for tumour segmentation in head and neck cancer radiotherapy", *Physics and Imaging in Radiation Oncology*, Vol. 25, p. 100408, 2023.

[5] M. R. Salmanpour, G. Hajianfar, M. Hosseinzadeh, S. M. Rezaeijo, M. M. Hosseini,

E. Kalatehjari, A. Harimi, and A. Rahmim, "Deep learning and machine learning techniques for automated PET/CT segmentation and survival prediction in head and neck cancer", In: *3D Head and Neck Tumor Segmentation in PET/CT Challenge*, Cham: Springer Nature, Switzerland, pp. 230-239, 2022.

[6] Y. C. Lin, G. Lin, S. Pandey, C. H. Yeh, J. J. Wang, C. Y. Lin, T. Y. Ho, S. F. Ko, and S. H. Ng, "Fully automated segmentation and radiomics feature extraction of hypopharyngeal cancer on MRI using deep learning", *European Radiology*, pp. 1-9, 2023.

[7] M. A. Naser, K. A. Wahid, A. J. Grossberg, B. Olson, R. Jain, D. E. Habashy, C. Dede, V. Salama, M. Abobakr, A. S. Mohamed, and R. He, "Deep learning auto-segmentation of cervical skeletal muscle for sarcopenia analysis in patients with head and neck cancer", *Frontiers in Oncology*, Vol. 12, p. 930432, 2022.

[8] D. M. Lang, J. C. Peeken, S. E. Combs, J. J. Wilkens, and S. Bartzsch, "Deep learning based HPV status prediction for oropharyngeal cancer patients", *Cancers*, Vol. 13, No. 4, p. 786, 2021.

[9] M. E. Rasmussen, J. A. Nijkamp, J. G. Eriksen, and S. S. Korreman, "A simple single-cycle interactive strategy to improve deep learning-based segmentation of organs-at-risk in head-and-neck cancer", *Physics and Imaging in Radiation Oncology*, Vol. 26, p. 100426, 2023.

[10] I. Shiri, A. V. Sadr, A. Akhavan, Y. Salimi, A. Sanaat, M. Amini, B. Razeghi, A. Saberi, H. Arabi, S. Ferdowsi, and S. Voloshynovskiy, "Decentralized collaborative multi-institutional PET attenuation and scatter correction using federated deep learning", *European Journal of Nuclear Medicine and Molecular Imaging*, Vol. 50, No. 4, pp. 1034-1050, 2023.

[11] J. Y. Ryu, H. K. Hong, H. G. Cho, J. S. Lee, B. C. Yoo, M. H. Choi, and H. Y. Chung, "Deep Learning for the Automatic Segmentation of Extracranial Venous Malformations of the Head and Neck from MR Images Using 3D U-Net", *Journal of Clinical Medicine*, Vol. 11, No. 19, p. 5593, 2022.

[12] T. Weissmann, Y. Huang, S. Fischer, J. Roesch, S. Mansoorian, H. A. Gaona, A. O. Gostian, M. Hecht, S. Lettmaier, L. Deloch, and B. Frey, "Deep learning for automatic head and neck lymph node level delineation provides expert-level accuracy", *Frontiers in Oncology*, Vol. 13, p. 1115258, 2023.

[13] K. Abe, N. Kadoya, K. Ito, S. Tanaka, Y. Nakajima, S. Hashimoto, Y. Suda, T. Uno, and K. Jingu, "Evaluation of the MVCT-based radiomic features as prognostic factor in patients with head and neck squamous cell carcinoma", *BMC Medical Imaging*, Vol. 23, No. 1, p. 102, 2023.

[14] Q. C. Le, H. Arimura, K. Ninomiya, T. Kodama, and T. Moriyama, "Can persistent homology features capture more intrinsic information about tumors from 18F-fluorodeoxyglucose positron emission tomography/computed tomography images of head and neck cancer patients?", *Metabolites*, Vol. 12, No. 10, p. 972, 2022.

[15] B. Huang, Y. Ye, Z. Xu, Z. Cai, Y. He, Z. Zhong, L. Liu, X. Chen, H. Chen, and B. Huang, "3D lightweight network for simultaneous registration and segmentation of organs-at-risk in CT images of head and neck cancer", *IEEE Transactions on Medical Imaging*, Vol. 41, No. 4, pp. 951-964, 2021.

[16] J. Liedes, H. Hellström, O. Rainio, S. Murtojärvi, S. Malaspina, J. Hirvonen, R. Klén, and J. Kemppainen, "Automatic segmentation of head and neck cancer from PET-MRI data using deep learning", *Journal of Medical and Biological Engineering*, pp. 1-9, 2023.

[17] H. Hellström, J. Liedes, O. Rainio, S. Malaspina, J. Kemppainen, and R. Klén, "Classification of head and neck cancer from PET images using convolutional neural networks", *Scientific Reports*, Vol. 13, No. 1, p. 10528, 2023.

[18] P. Nikulin, S. Zschaeck, J. Maus, P. Cegla, E. Lombardo, C. Furth, J. Kaźmierska, J. M. Rogasch, A. Holzgreve, N. L. Albert, and K. Ferentinos, "A convolutional neural network with self-attention for fully automated metabolic tumor volume delineation of head and neck cancer in [^{18}F] FDG PET/CT", *European Journal of Nuclear Medicine and Molecular Imaging*, pp. 1-16, 2023.

[19] F. H. Tang, E. Y. W. Cheung, H. L. Wong, C. M. Yuen, M. H. Yu, and P. C. Ho, "Radiomics from various tumour volume sizes for prognosis prediction of head and neck squamous cell carcinoma: a voted ensemble machine learning approach", *Life*, Vol. 12, No. 9, p. 1380, 2022.

[20] A. Chandrashekar, A. Handa, J. Ward, V. Grau, and R. Lee, "A deep learning pipeline to simulate fluorodeoxyglucose (FDG) uptake in head and neck cancers using non-contrast CT images without the administration of radioactive tracer", *Insights Into Imaging*, Vol. 13, No. 1, p. 45, 2022.

[21] N. Brancati, M. L. Rosa, G. D. Pietro, G. Esposito, M. Valentino, M. Aiello, and M. Salvatore, "An Investigation on Radiomics

Feature Handling for HNSCC Staging Classification", *Applied Sciences*, Vol. 12, No. 15, p. 7826, 2022.

[22] J. Chen and A. L. Martel, "Head and neck tumor segmentation with 3D UNet and survival prediction with multiple instance neural network", In: *3D Head and Neck Tumor Segmentation in PET/CT Challenge*, Cham: Springer Nature, Switzerland, pp. 221-229, 2022.

[23] M. Gardner, Y. B. Bouchta, A. Mylonas, M. Mueller, C. Cheng, P. Chlap, R. Finnegan, J. Sykes, P. J. Keall, and D. T. Nguyen, "Realistic CT data augmentation for accurate deep-learning based segmentation of head and neck tumors in kV images acquired during radiation therapy", *Medical Physics*, 2023.

[24] Q. Zhang, K. Wang, Z. Zhou, G. Qin, L. Wang, P. Li, D. Sher, S. Jiang, and J. Wang, "Predicting local persistence/recurrence after radiation therapy for head and neck cancer from PET/CT using a multi-objective, multi-classifier radiomics model", *Frontiers in Oncology*, Vol. 12, p. 955712, 2022.

[25] Dataset: https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=39879146.

[26] S. Goumiri, D. Benboudjema, and W. Pieczynski, "A new hybrid model of convolutional neural networks and hidden Markov chains for image classification", *Neural Computing and Applications*, pp. 1-16, 2023.