ISSN: 2633-4828 Vol. 5 No.2, June, 2023

International Journal of Applied Engineering & Technology

SURVIVAL PREDICTION IN HEAD AND NECK SQUAMOUS CELL CARCINOMA USING TRANSFER LEARNING

Deepti and Susmita Ray
Department of Computer Science & Technology, Manav Rachna University, Faridabad, Haryana, India
deeptigupta@mait.ac.in and susmita@mru.edu.in

ABSTRACT

Head and neck cancer impacts it's patients not only physiologically but psychologically, financially and aesthetically too. Survival prediction gives an insight about life expectancy after treatment. This helps in planning treatment strategies to improve its outcome, making future planning, and enhancing patient care. In this paper, we propose a novel approach for the prediction of survival for head and neck squamous cell carcinoma (HNSCC) patients using transfer learning. Transfer learning in medical imaging can develop an efficient model by utilizing information from per-trained models. In our work, we have employed CapsuleNet for feature extraction from 2593 CT scans of HNSCC patients, as it has the capability of attaining better feature extraction from less training dataset. For HNSCC survival prediction we have used DenseNet121 on the features extracted from CapsuleNet, as it requires fewer features as compared to traditional CNN. Performance metrics such as accuracy, sensitivity, specificity, F1 Score, and precision are used to evaluate the prediction results. To compare the results other deep learning models like CNN and ResNet50 were also employed over same data-set for feature extraction as well as prediction. The proposed model achieved the accuracy of 97.3% which outperformed other deep learning models.

Keywords: Transfer learning, Head and Neck Squamous Cell Carcinoma, Dense Net121, Capsule Net.

1. INTRODUCTION

Head and neck cancers, commonly referred to as malignancies, often originate from the squamous cells that coat the mucosal surfaces of the head and neck, including the mouth, throat, and voice box. The medical term for these malignancies is Head and neck squamous cell carcinoma (HNSCC). Authors Mateus, P., Volmer et al.[1], in their research mentioned that on a global scale, 3% of cancer cases are HNSCC. Oropharyngeal, hypopharyngeal, laryngeal, and oral cavity squamous cell carcinomas are the primary types of HNSCC. HNSCC patients have complicated daignosis options based on where the cancer is located, its stage, and the results of any tests that were done. Head and neck cancer treatment necessitates a multidisciplinary approach involving oncologists, surgeons, radiation oncologists, and other specialists to develop individualized treatment plans that optimize outcomes while minimizing side effects and preserving quality of life. A multidisciplinary approach and early detection are crucial for optimal treatment [18].

Clinical Oncology aims to help medical practitioners make smart decisions about cancer treatment. Radiomics integration in clinical oncology offers precision medicine and improved patient management, providing insights into tumour biology, treatment response, and patient outcomes. Radiomics is the extraction and analysis of quantitative information from medical imaging data, such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), or positron emission tomography (PET) scans. Through the utilization of medical imaging data, radiomics can provide new avenues for precision medicine and better patient management along the cancer care continuum. It is an exciting and relatively new area of study that aims to make predictive and prognostic models by taking semi-quantitative or quantitative traits from medical images[2-5]. Radiomics analysis may use high-resolution structural imaging modalities for extracting radiomic characteristics like shape, intensity and texture. New developments in machine learning and artificial intelligence have made it possible for medical images to be evaluated automatically, which results in developing the identifiers that can be used for diagnosis and prognosis.

Radiomics, paired with machine learning techniques, is a cutting-edge approach to cancer research and clinical oncology. Radiomic properties may be categorized as statistical features, which include histogram-based and texture-based features, model-based features, transform-based features, and shape-based features. Machine learning algorithms can then be used to examine these radiomic properties and create predictive models for a variety of clinical outcomes, such as diagnosis, therapy response, prognosis, and patient survival. The large volumes of data in this field have enabled models that may improve medical decision-making, outcome prediction, and precision healthcare [9-12].

Deep learning (DL), a machine learning approach deviates from the conventional approach by using a complicated framework that has the potential to outperform conventional machine learning models for imaging data. Medical Images include images from magnetic resonance imaging (MRI), digital radiography (RD), or ultrasonography (US), X-ray, computed tomography (CT) scan. CPU and memory limits parameters are responsible for the data pre-processing and building of the algorithms[15]. Based on the patient's particulars, medical image analysis helps in finding out the diagnosis and the future survival prediction of the patient. Analysing methods for medical images have been very helpful in diagnosing head and neck cancers[17]. However, it can be hard to find out what these images mean because of the complicated anatomy, different types of tumor biology, and traits that cross between cancerous and noncancerous spots.

Conventional imaging methods often rely on human skill, which may be uncertain and susceptible to inconsistency. Deep learning methods, particularly convolutional neural networks (CNNs), have the potential to enhance the accuracy, speed, and consistency of medical picture processing. These sophisticated computer systems have the ability to autonomously acquire characteristics from unprocessed picture data. This allows for the extraction of therapeutically meaningful data and has the potential to surpass human performance in some tasks[7-8].

Transfer learning has considerable potential for increasing automated analysis, diagnosis, and treatment planning in clinical practice by using pre-trained models and customizing them for medical imaging applications. Transfer learning using convolutional neural networks attempts to enhance performance on a new task by using the information obtained from comparable tasks learned before. It has significantly advanced medical image analysis by addressing the issue of limited data availability and optimizing time and hardware resources.

CNN Transfer Learning (TL) utilizes parametric data transfer. Trained CNN models use convolutional layer parameters for medical image analysis. Medical image classification may be trained using generic features from natural image classification when the labels are available in both domains in TL with CNN. Deep learning has been extensively employed in medical imaging tasks. One of the challenges that arise when using deep learning to the interpretation of medical images is the limited amount of annotation and labeling that is available. As a result of the restricted labels, there is a limited amount of data that is accessible, which makes it difficult to train deep learning models effectively. Consequently, transfer learning has emerged as a viable alternative to deep learning. TL may use CNN models either by extracting features or by fine-tuning. Feature extractors immobilise the convolutional layers, whereas fine-tuning modifies the parameters throughout the model fitting process [20, 25-26, 28].

In this research, we have proposed a model that utilizes the use of transfer learning to perform better than the conventional machine learning models. We have extracted the radiomic features from the CT scans using the capsule Net. Capsule Net is the pre-trained deep learning model best used for feature extraction. After extracting the features, we have trained the model with DenseNet 121 to predict the survivability of HNSCC patients. This paper includes four sections where **Section 2** is an overview of previously published work on the machine learning and deep learning models used in healthcare, **Section 3** gives the background of the models used for developing the proposed model and the comparison between the conventional methods and radiomics. It also includes the methodology and proposed work, data set pre-processing details. **Section** 4 includes the results and the conclusion.

2. LITERATURE REVIEW

This section includes the approaches used in past years for analyzing medical images for HNSCC. Medical images provide detailed information about tumour characteristics like size, extent and location of tumours.

Researchers are focusing on exploiting the deep learning architecture's capabilities of image recognition, classification, feature extraction, object detection and localization in extracting quantitative features and analyzing medical images.

Tortora M et.al.[6], reviewed in their work, HNC includes a wide range of different types of cancer that can grow in many different locations of the body. The head and neck region, which includes the paranasal sinuses, pharynx, mouth cavity, larynx, thyroid, lymph nodes, soft tissues, and nearby bones, can be affected by various types of cancer. Squamous cell carcinomas (SCCs) are the most common type of head and neck cancer, accounting for 90% of cases based on histology. The primary causes of the increasing number of head and neck squamous cell carcinoma (HNSCC) cases are tobacco use and alcohol consumption. The majority of HNCC cases consist of oropharyngeal squamous cell carcinoma (OPSCC) and nasopharyngeal carcinoma (NPC). The authors , Bray F. et.al. [7], in their work, stated that Head and neck cancer affects more than 830,000 people a year around the world. They are linked to a high death rate because they often grow slowly and without any symptoms.

Pham et al. [14] talked about how medical records can be used to use DL to predict useful results in healthcare. The authors said that deep learning can be used to model the complicated connections between various health problems and guess how a patient's medical care might change over time. The study used information from patients' electronic medical records who had diabetes, high blood pressure, and heart disease, among other illnesses. CNNs and RNNs were used in the suggested DL model to find the links between the data's time and space. The study found that the deep learning model was very good at predicting how patients would get care in the future with a high level of accuracy. In their conclusion, the writers talked about how deep learning could change the way healthcare is provided by making predictions more accurate and allowing for more personalized care. Despite this, the authors agreed that using deep learning in healthcare is still in its early stages and needs more research to fully reach its full potential.

Use of deep learning in head and neck cancer has increased rapidly. Deep learning has improved disease identification, anatomical structure segmentation, and treatment outcome prediction in medical image analysis. Deep learning algorithms can comprehend complex patterns from medical pictures, exhibit efficient generalization to novel data, and perform a multitude of jobs akin to humans. This ability might considerably enhance medical image analysis accuracy, speed, and consistency, improving patient care and results. Illimoottil et.al.[19] in their work reviewed the progress, possibilities, and challenges involved in using deep learning methods in images for head and neck cancer. They also suggested ways to improve patient results and personalized treatment plans for the future. They gave a beneficial review of the current state and possible outcomes in HNSCC imaging by including the latest developments by an in-depth review of both standard and deep learning methods. According to their review DL techniques have significantly improved HNSCC imaging, enhancing tasks like detection of tumor and prediction of outcome. Advanced models like convolutional autoencoders and GANs have improved data compression and feature extraction. Integrating radiogenomics into deep learning models could guide personalized treatment strategies.

Many studies in the field have focused on and validated deep learning in medical image processing. Recent developments in computational pathology and algorithmic computing have made DL a beneficial tool for a variety of cancer diagnostic tasks, including tumor kind, grade, and prognosis prediction. Several studies have demonstrated its efficacy and utility in solving the numerous issues encountered in medical imaging activities.The medical industry has seen a meteoric rise in the use of CNNs in recent years, with many hospitals and clinics collecting massive amounts of imaging data. Among the many uses for these programs are tasks as varied as medical diagnosis and image segmentation. Their intricate structure hinders the inexplicability of neural networks' choices. More data is often needed by CNNs compared to more conventional machine learning methods from a variety of sources. More and more models have been published in this subject, which is encouraging because of its potential. Reasons such as the absence of standards for reproducibility and instructions on how to construct reliable imaging models explain why only a small percentage finds its way into healthcare settings [10-11]. In their study, Liu et al.[12] concluded that the analysis of head and neck pathologic images was first conducted in

2017 utilizing deep learning (DL), which is an advancement of machine learning (ML). This work aims to systematically examine and perform a meta-analysis to summarize and analyze the effectiveness of deep learning algorithms in contouring head and neck Organs at risk. After eliminating duplicate literature, doing primary screening, and re-screening, a total of 149 articles were acquired, out of which 22 studies were selected for the meta-analysis.They concluded that deep learning has great potential for collaborating with institutions to create high-quality medical data sets, promoting precision radiotherapy and providing patients with tailored, standardized treatment plans.

Kanna et al. [13] carried out a study that used transfer learning to build durable deep-learning convolutional neural networks. They developed deep-learning models to detect prostate cancer tissue in whole-slide images. They utilized a large and high-quality training dataset, as well as advanced convolutional network architectures such as MobileNet V2, InceptionResNet V2, DenseNet 169, ResNet101 V2, and NasNetMobile for extensive training. The performance was evaluated using a confusion matrix, which included accuracy, loss, and RMSE values. The findings indicate that Inception ResNet V2 achieved superior performance compared to other models.

3. EXPERIMENTAL METHOD

3.1 Background

In recent years, the intersection of transfer learning with pre-trained architectures has garnered increasing attention from researchers, reflecting a growing recognition of its potential in cancer patient management.

Improvement of Machine Learning for Medical Image Analysis over Traditional Methods

At its core, radiomics is a branch of mathematics that applies algorithms to transform two- or three-dimensional images into tabular data including radiomics properties. Radiomics allows a comprehensive analysis of tumors, which may be used to forecast the effectiveness of the therapy. Radiomics may improve diagnosis and decision-making by collecting image-based metrics. The conventional approach to radiomics is using a pre-existing picture Region of Interest (ROI) to extract characteristics, which are then used to train a classification algorithm for the purpose of predicting a therapeutic objective. The ROI may be characterised by three primary properties: intensity distribution, shape and size, and texture. Contouring of the ROI involves delineating the main tumour and the lymph nodes that are connected to it. There are a lot of options in traditional radiomics that have to do with image discretization and feature extraction, which may lead to thousands of features that might end up being much more in number than the number of patients. Overfitting occurs when a feature set becomes excessively preoccupied with training or validation data and ignores to extrapolate to unidentified test data which is sourced externally. which results the impossibility of exhaustively searching for all possible combinations of salient features [25-26].

Convolutional neural networks overcome the limitations of radiomics as they are used in deep learning radiomics, where images are sent into the network with or without tumor outlines. It is possible to avoid doing things like contouring tumors and nodes and calculating radiomics characteristics. To make predictions, the network learns discriminant characteristics automatically during training. For image pre-processing, a few factors need to be considered such as variances in image collection and reconstruction techniques, uncertainties, and variances caused in ROI by the outside observer. In CNN, a well-designed feature selection algorithm chooses a subset of traits that accurately reflect the whole patient population based on the medical data. This enhances the model's interpretability[27-28].

Capsule Network used for Feature extraction in the Proposed Model

While convolutional neural networks (CNNs) excel at many computer vision tasks, they fail to account for objects' geometrical connections. This means that CNNs can't handle training data that contains the affine transformations or the rotated images. A slight adjustment to the input image's size or translation may have a large impact on the network's performance, according to recent research. The main problem with traditional convolutional neural networks (CNNs) is that when it comes to learning visual representations, they have trouble processing input images of different sizes and orientations, and the pooling layer in these networks generally disregards positional information. The removal of pooling layers and their replacement with dynamic routing and convolutional strides has increased the representation learning of Capsule Network (CapsNet). Popular tasks including as classification, identification, segmentation, and natural language processing have shown encouraging results using this architecture. CapsNet returns vector outputs rather than scalar ones. The core idea of CapsNet is to use encryption methods to encode the connection between many variables, including scales, location, posture, and orientation. Even though a human eye would readily see that a non-facial image with a mouth, eyes, and nose is not a face, convolutional neural networks (CNNs) might still not label it. Nonetheless, the capsule network will learn the relationship between facial features like the eyes and nose, and it will correctly recognize non-face images as such [21–23]. Medical imaging often utilizes three-dimensional data, such as CT scans and MRIs. Capsule Networks can process 3D pictures more efficiently compared to regular Convolutional Neural Networks (CNNs). Caps Nets are intrinsically capable of preserving spatial hierarchies, which makes them well-suited for processing volumetric data without sacrificing important information[29-30].

In Figure 1, the structure of CapsNet shows, a pair of convolutional layers make up CapsNet. There are 256 channels in the first convolutional layer, and each channel has 9 × 9 filters that are activated using a RELU function, and the stride is 1. In the second layer, there was a convolutional capsule layer with 6-by-6-by-32 capsules and a 2-micron stride. Eight convolutional units, each with a 9 × 9 kernel, make up each main capsule. It has been activated using the squashing function. The layer that is fully connected, the final component of the CapsNet design, is made up of sixteen ten-size D-capsules. The name for these pills is DigitCaps. An additional building block, or capsule, is inserted between a neuron and a layer in a capsule network. The neuronal layers that make up a capsule are nested sets. To classify data, these capsules gather data from every other capsule and use 10 different criteria[31-32].

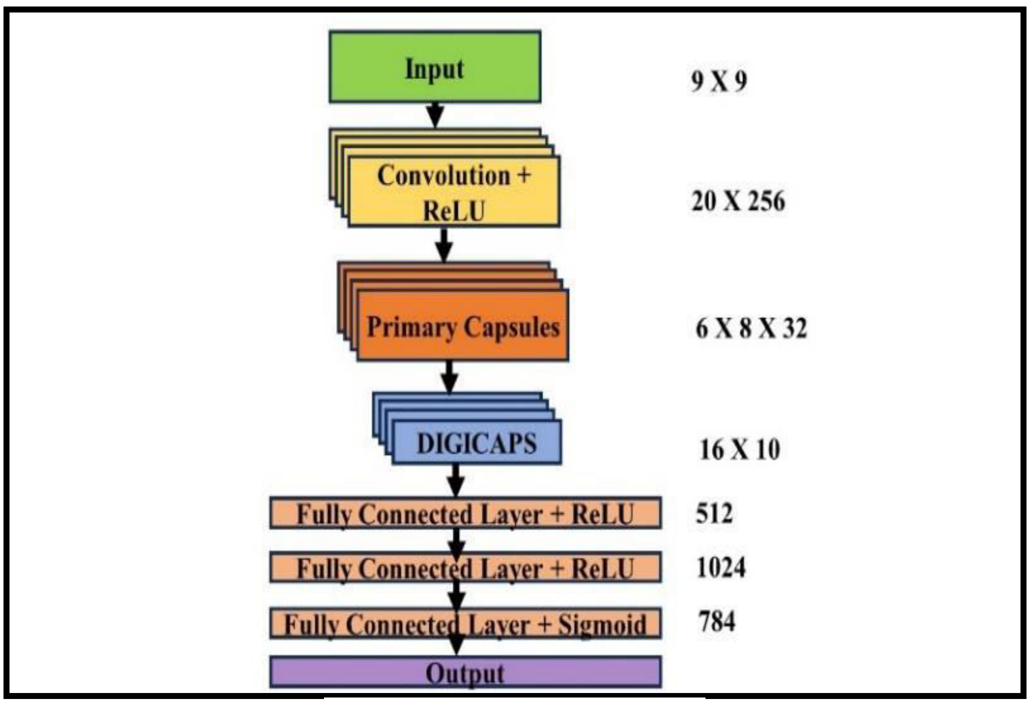

Figure 1: Capsule Net Structure

When it comes to medical image analysis capsule networks have shown encouraging results. Organ segmentation, anomaly detection, and medical scan-based illness diagnosis are all areas that might benefit from CapsNets' ability to record structural connections. Complicated structures are common in medical imaging. CapsNets have the potential to enhance automated system's ability to analyze medical images, leading to more accurate diagnosis [29-30].

DenseNet121 used for Survival Predication of HNSCC by using the features extracted by CapsNet

In their initial proposal, G. Huang et al. [33] established DenseNet as a subset of CNNs. Impressive results were obtained on many image classification datasets. Each layer in a DenseNet design uses inputs from all the layers below it to build a feature map that feeds data to the layers above it. This is achieved by connecting the layers with dense blocks.

The n^{th}- layer gets all the inputs from the preceding feature maps (x_0, x_1, …, x_n) as seen in equation (1):

$$x_n = H_n([x_0, x_1, \dots\dots\dots\dots, x_{n-1}] \qquad (1)$$

In this case, the concentration of all preceding feature maps of the n^{th}-layer is represented as [x_0, x_1, …, x_{n-1}]. The n^{th} layer is a composition function that includes three consecutive operations—batch normalization, a ReLU activation function, and convolution—and x^n is the output of that layer. While ResNet merges adjacent layers, DenseNet just concatenates them. This makes it comparable to ResNet but not quite the same. By repeating features, DenseNet decreases the parameters and tackles the issue of vanishing gradients. In Figure 2, the use of four dense blocks by DenseNet-121 is demonstrated. A transition layer is inserted between each set of blocks; it uses down-sampling on the feature maps to generate a 1×1 convolution and a 2×2 average pooling layer. To bridge the gap between faraway levels, the dense blocks use a succession of convolutional layers. A ReLU activation is used by DenseNet-121 to enhance non-linearity. The DenseNet-121 model makes use of a ReLU activation function to enhance non-linearity. Lastly, a fully connected layer is used for prediction using a softmax function.

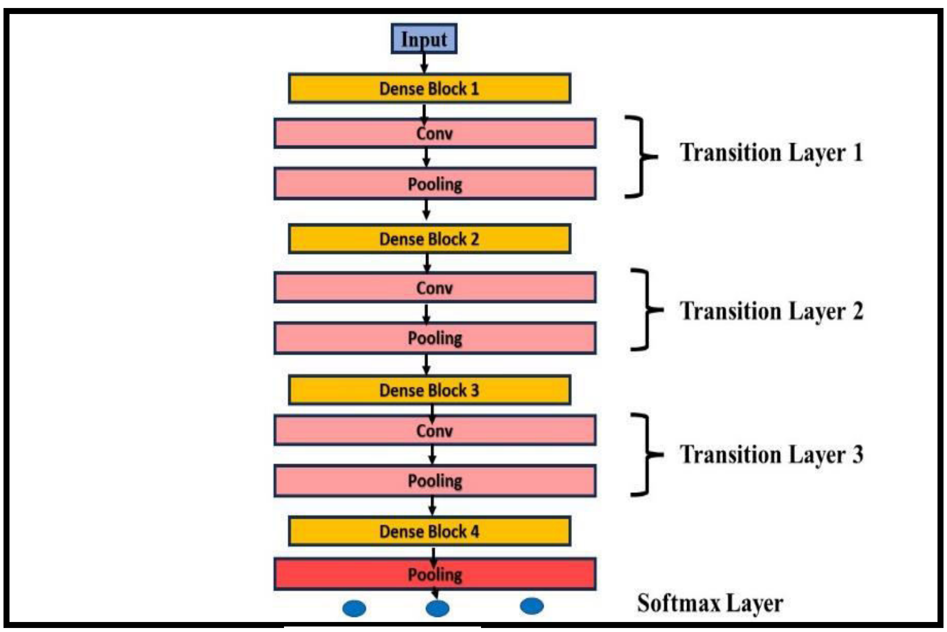

Figure 2: Four dense blocks of DenseNet 121 Structure

3.2 Methodology Adopted to Develop the Proposed Model

In the Figure 3, the workflow of the suggested system is depicted. The CT images of patients with Head and Neck Squamous Cell Carcinoma (HNSCC) are obtained from The Cancer Imaging Archive. Subsequently, the acquired dataset undergoes pre-processing. Subsequently, the data is inputted into the capsuleNet model to extract features. The characteristics obtained from the CapsuleNet are then inputted into the DenseNet-121 model to predict survival. Similarly, the preprocessed information is simultaneously inputted to both the CNN and ResNet50 models to predict the survival of HNSCC patients. The suggested model, CNN Model, and ResNet50 models are assessed using metrics such as F1 score, accuracy, sensitivity, specificity, and precision. The performance evaluation of the suggested model is thereafter compared to that of the CNN and ResNet50 models.

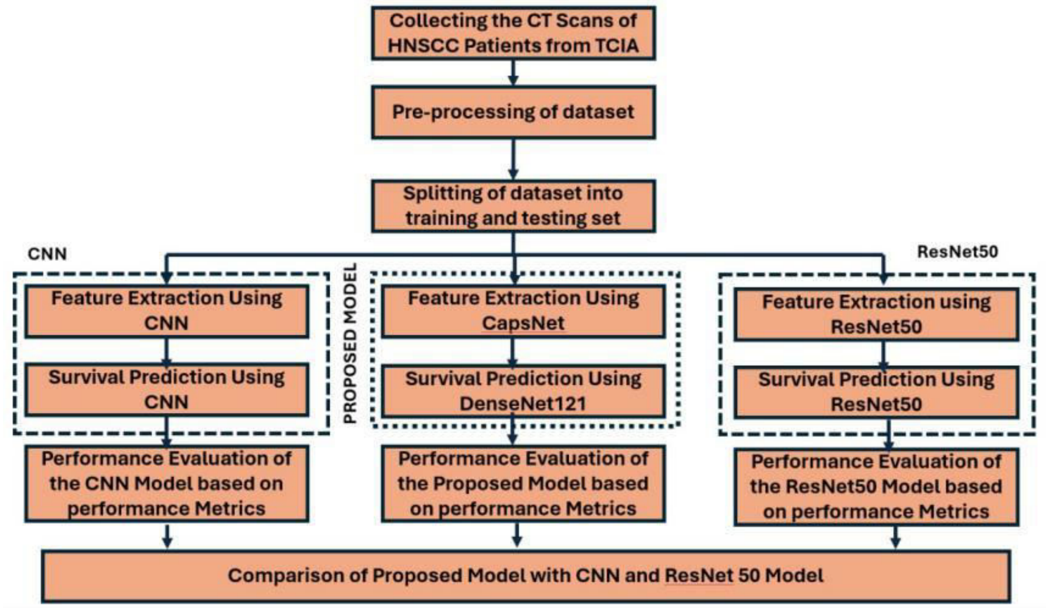

Figure 3: Working methodology

Data Preprocessing of the Data extracted from the TCIA Dataset

The medical imaging dataset of HNSCC comprising 433,384 DICOM files from 3,225 series and 765 investigations from 215 patients is retrieved from The Cancer Imaging Archive. From the complete dataset, we extracted 2593 cancer images. The size of the original image was 512 X 512. An unprocessed CT scan image of a patient's Hypopharyngeal region is shown in figure 5.

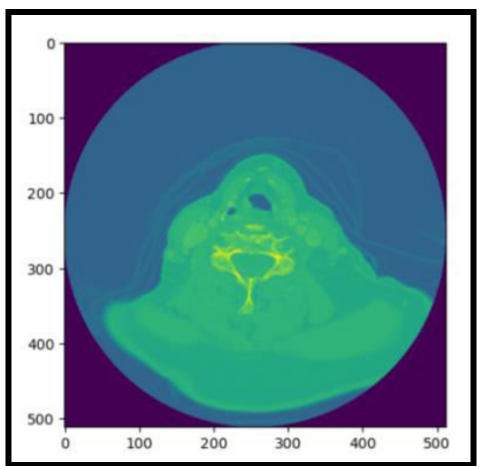

Figure 4: Unprocessed Random CT Scans image of a patient of the Hypopharyngeal region

Medical image preprocessing reduces the distortions in the dataset and increases the accuracy of the model. In our work we have done the pre-processing involves several the steps. Converting DICOM images to grayscale maintains the spatial organization and interconnections among various anatomical elements present in the images. Capsule Networks demonstrate outstanding ability in capturing spatial hierarchies and relationships. Moreover, monochromatic images offer a transparent depiction of the spatial characteristics, avoiding the complexity introduced by color channels. In figure 5, we can see the gray image of the CT scan which was demonstrated in Figure 4.

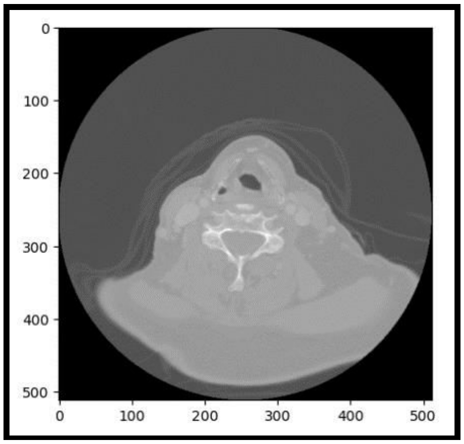

Figure 5: Image converted to Gray Scale

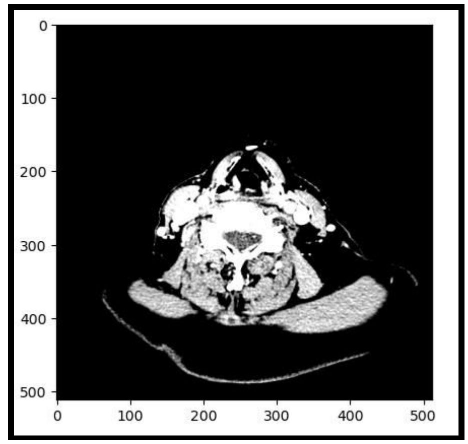

Figure6: Image converted to HU after pre- processing.

CT imaging uses the HU to measure tissue radiodensity. CT scan grayscale images use Hounsfield unit values. Radiologists can see the body's interior architecture by coloring tissues grey depending on their HU values. Brighter regions indicate tissues with greater HU values, whereas darker areas reflect tissues with lower HU values. This grayscale display helps radiologists distinguish tissues and spot anomalies or diseases depending on CT picture density. Figure 6, demonstrates the HU image of the gray image of Figure 5.

Noise removal during preprocessing is essential for enhancing the data quality used for model training. it results in improved performance machine learning model. The figure 6 after noise removal from the HU image is demonstrated in figure 7

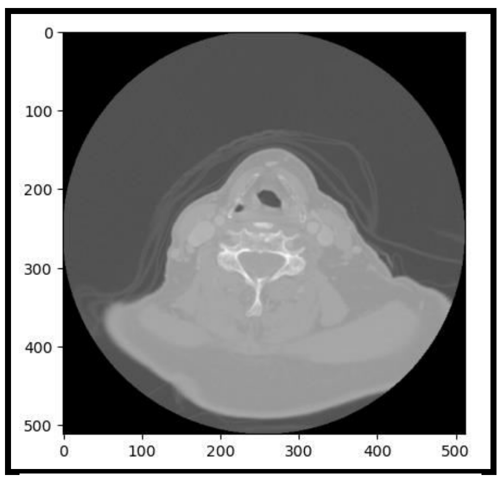

Figure 7: Image after removal of noise

The pixel data of the pre-processed images was taken and transformed into NumPy arrays. The data is randomized to facilitate training of the model utilizing unordered data. Once the data was randomized, we partitioned it into two distinct groups: the training set and the testing set. The training set consisted of 1884 photos, whereas the testing set included 709 images.

4. RESULTS AND CONCLUSION

The sensitivity, accuracy, F1 Score, specificity, and precision are five performance measures that were evaluated for the three models: Proposed Model, ResNet50, and CNN. One way to measure a prediction model's performance in statistics and ML is by using a confusion matrix, which is a tabular representation. For a quick overview of the prediction results, it shows the total number of correct, incorrect, FP, and FN predictions. The confusion matrix allows for the evaluation of all performance parameter indicators. In this research, we have CT scan images of HNSCC patients. The confusion matrix is represented as shown in Table 1. Where True Positive (TP) is present in a dataset, it means that the expected number of survived patients (positive results) match the actual values. True negative values (TN) indicate that the dataset's actual non survived patients values match the negative projected outcomes. When the predicted results are false and the actual values in the dataset are positive, we get false negative values (FN). When the projected results are correct but the actual values in the dataset are negative, this is called false positive values (FN) [34-35].

Table I: Confusion Matrix used for Depicting Predicted Values

Confusion Matrix		**Actual Values**	
		Positive	**Negative**
Predicted Values	**Positive**	TP	FP
	Negative	FN	TN

Table II: Representation of Confusion Matrix Values for calculation performance metrices.

Model Name	**No. of true positive values**	**No. of true negative values**	**No. of false positive values**	**No. of false negative values**
Proposed Model	1275	1250	30	38
CNN	1260	1240	60	33
ResNet50	1278	1230	45	40

Performance Metrics used for the evaluation of the proposed model.

$$Accuracy = \left[\frac{(TN + TP)}{(TN + TP + FP + FN)}\right] \quad (2)$$

$$Sensitivity = \left[\frac{(TP)}{(TP + FN)}\right] \quad (3)$$

$$Specificity = \left[\frac{(TN)}{(TN + FP)}\right] \quad (4)$$

$$Precision = \left[\frac{(TP)}{(TP + FP)}\right] \quad (5)$$

$$F1\,Score = [2 \times \frac{(Precision \times Sensitivity)}{(Precision + Sensitivity)}] \quad (6)$$

It is critical to recall that accuracy is only one efficiency indicator and not the be-all and end-all, especially when dealing with an imbalanced dataset. The machine learning algorithm could develop biases due to the imbalanced dataset's skewed distribution of outputs. Equation 2, specify Accuracy of a test and is determined by its sensitivity and specificity, which indicate the existence or absence of a condition. Equation 3, specify Sensitivity, often called the true positive rate, is a measure of how likely it is that a test would provide a positive result. A low number of false negatives in the confusion matrix is desirable in situations when a high-sensitivity condition is present, such as when the trial calls for the classification of entries that fulfil the given condition. In addition, the probability of correctly detecting a genuine positive is high. This becomes of the utmost importance in cases when the consequences of treatment failure are substantial or when the therapy has remarkable efficacy with few side effects. Because very sensitive tests rarely miss the existence of an illness, a negative result is very important. Equation 4, specifies Specificity, sometimes called the real negative rate which is the measure of the probability of a negative test result. One way to determine specificity is by how well it can differentiate between true negatives. It is common for specificity and sensitivity to go hand in hand; that is, when sensitivity is reduced, specificity is increased, and vice versa. The trade-off condition describes this situation. A class's accuracy is its true positive rate as a percentage of all forecasts for that class. Equation 5, specifies the percentage of correct predictions relative to the total number of correct predictions produced by the model. By finding a median between the two, the F1 score can decrease both false positive and false negative numbers. In terms of reliability, F1 scores are superior. It is not easy to understand the F1 score since it is not clear whether accuracy or recall is being maximized. Improving evaluations is achieved by combining the F1-score with other measures.

Table III: Performance Metrices Values for the Three Models

Model Name	Accuracy	sensitivity	Specificity	precision	F1 Score
CNN	0.964134	0.974478	0.953846	0.954545	0.964409
ResNet 50	0.967219	0.969651	0.964706	0.965986	0.967815
Proposed Model	**0.973776**	**0.971059**	**0.976563**	**0.977011**	**0.974026**

The proposed model demonstrates higher accuracy compared to ResNet50 and CNN, as seen in Figure 8, via the evaluation of performance metrics. The proposed model has a remarkable accuracy rate of 97.3%.

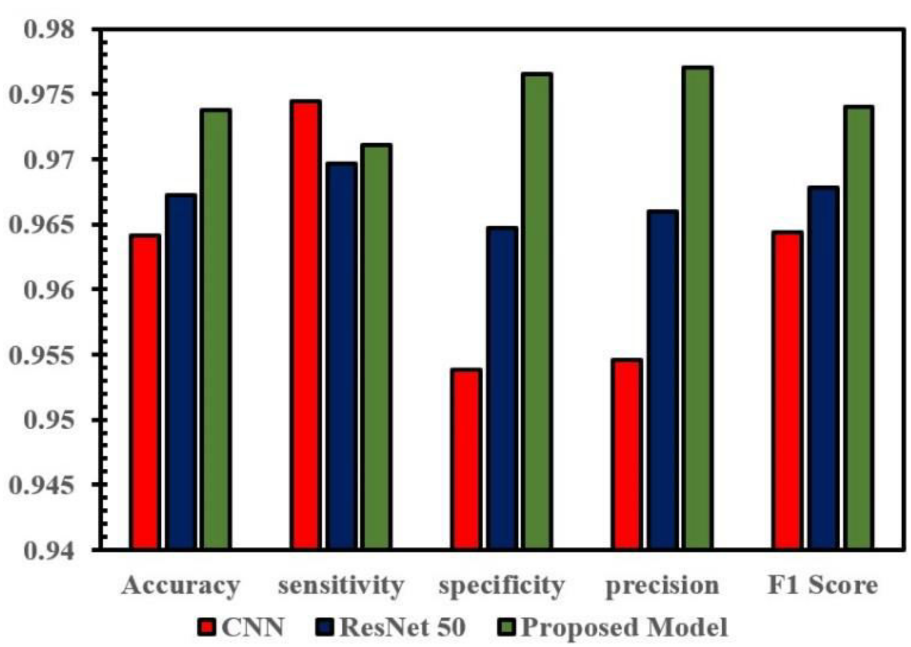

Figure 8: Comparison of the Models

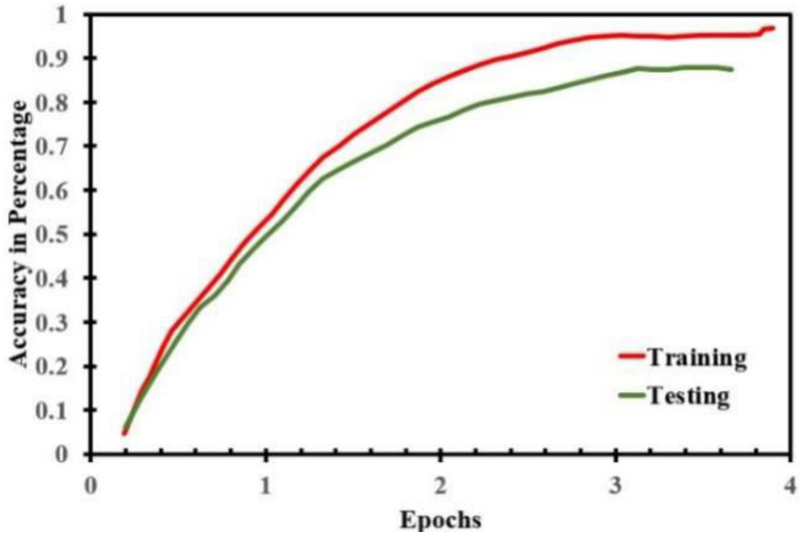

Figure 9: Accuracy Vs Epochs Graph of Proposed Model.

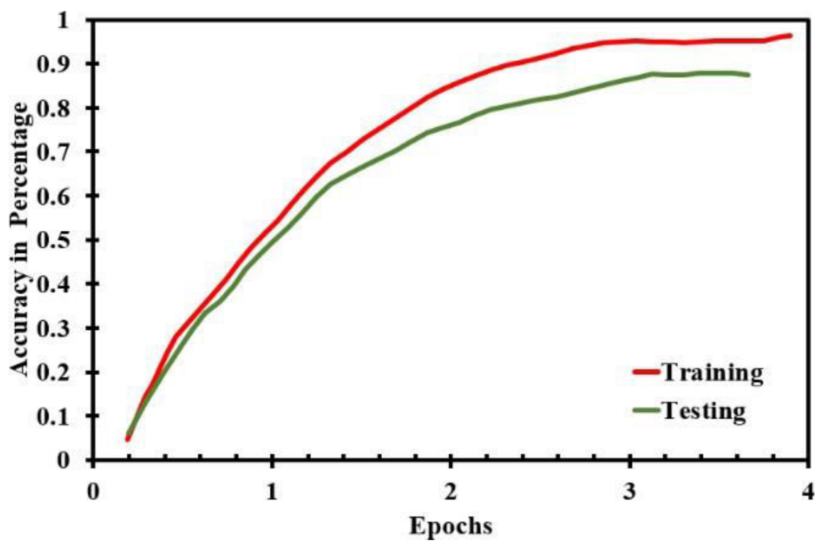

Figure 10: Loss Vs Epochs Graph of ProposedModel.

To understand the statistics, the rate of change of accuracy with change in epochs for all three models are represented graphically in Figures 9,10, and 11. As illustrated by Figure 10, the training and testing accuracy improved as the number of epochs rose for the proposed model. Figure 11, demonstrates the training and testing loss curves of the proposed model. It can be seen that as the number of epochs increased eventually gap between training and testing accuracy decreased. It implies that the model has learned to generalize successfully from training data to new data that is the model can recognize the underlying patterns in the data without just memorizing the training instances.

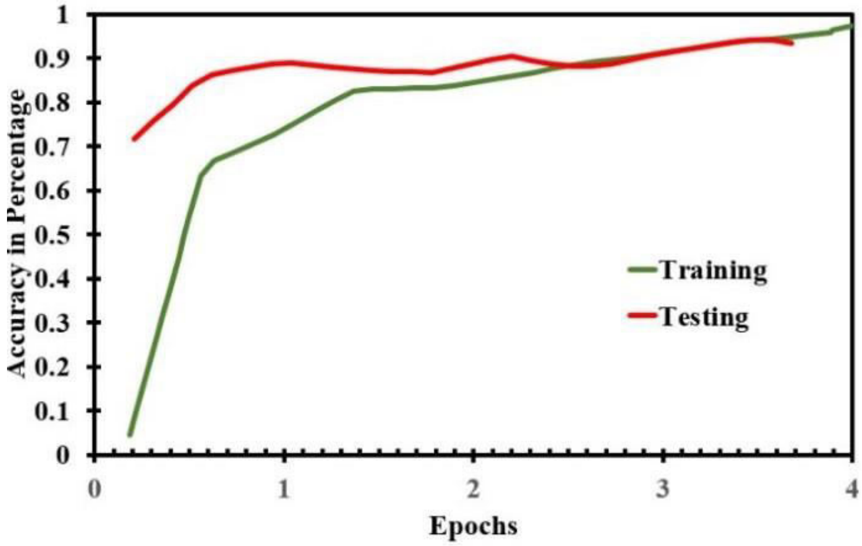

Figure 11(a): Accuracy Vs Epochs of ResNet50 Model CNN

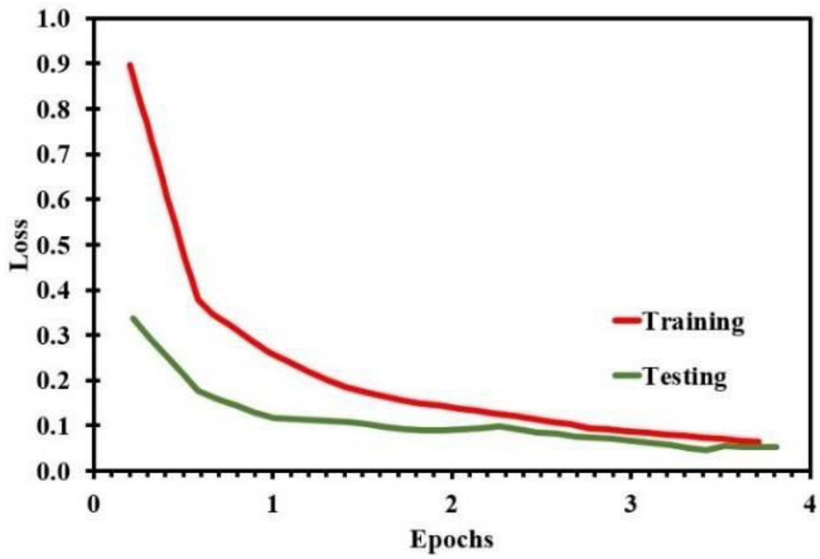

Figure 11(b): Accuracy Vs Epochs Graph of Model.

Figure 11(a) illustrates that the training and testing accuracy increased as the number of epochs rose for the ResNet50 model. Similarly, its depicted in Figure 11(b), that the training and testing accuracy increased as the number of epochs rose for the CNN model. After a threshold the gap in training and testing accuracy start to increase.

The study suggested that deep learning models can predict the survivability of HNSCC with high accuracy. The features recovered by CapsNet provide significant information about the input medical images, including traits that indicate the presence of malignant tissues. DenseNet121 successfully applies the acquired knowledge from CapsNet, which involves feature extraction, to produce predictions in the job of cancer prediction. The differential architectures and training procedures of CapsNet and DenseNet121 may result in the acquisition of different parts of the input medical images. By integrating the hierarchical representations acquired by CapsNet with the generic characteristics acquired by DenseNet121, the model has the potential to attain a more extensive comprehension of the input data, which resulted in enhanced cancer prediction performance. In our work we found that the proposed model shows promising results with greatest levels of accuracy as compared with CNN and ResNet50. Furthermore, the findings show that confusion matrices can serve as performance metrics for cancer prediction using CT scans. Evaluation criteria, such as F1 score, accuracy, sensitivity, specificity, and precision, were used to assess the models' efficacy. Comparative examination of the three models—Proposed Model, ResNet50, and CNN showed encouraging results, when it came to predicting patients' survival from HNSCC. The assessment metrics determined that the proposed model gives the highest accuracy in predicting the survivability of the patients.

REFERENCES

1. Mateus, P., Volmer, L., Wee, L. et al. Image-based prognosis in head and neck cancer using convolutional neural networks: a case study in reproducibility and optimization. Sci Rep 13, 18176 (2023). https://doi.org/10.1038/s41598-023-45486-5.

2. Pfister, DG, Spencer, S, Adelstein, D, Adkins, D, Anzai, Y, Brizel, DM, et al. Head and neck cancers, version 2.2020, NCCN clinical practice guidelines in oncology. J Natl Compr Cancer Netw. (2020) 18:873–98. doi: 10.6004/jnccn.2020.0031.

3. Afshar, P, Mohammadi, A, Plataniotis, KN, Oikonomou, A, and Benali, H. From handcrafted to deep-learning-based cancer radiomics: challenges and opportunities. IEEE Signal Process Mag. (2019) 36:132–60. doi: 10.1109/MSP.2019.2900993.

4. Diamant, A., Chatterjee, A., Vallières, M. et al. Deep learning in head & neck cancer outcome prediction. Sci Rep 9, 2764 (2019). https://doi.org/10.1038/s41598-019-39206-1.

5. Dwdewe Tortora, M, Gemini, L, Scaravilli, A, Ugga, L, Ponsiglione, A, Stanzione, A, et al. Radiomics applications in head and neck tumor imaging: a narrative review. Cancers. (2023) 15:1174. doi: 10.3390/cancers15041174.

6. Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int. J. Cancer 136, E359–E386. doi: 10.1002/ijc.29210.

7. Fidler, M. M., Bray, F., Vaccarella, S., and Soerjomataram, I. (2017). Assessing global transitions in human development and colorectal cancer incidence. Int. J. Cancer 140, 2709–2715. doi: 10.1002/ijc.30686.

8. Koelzer, V. H., Sokol, L., Zahnd, S., Christe, L., Dawson, H., Berger, M. D., et al. (2017). Digital analysis and epigenetic regulation of the signature of rejection in colorectal cancer. Oncoimmunology 6:e1288330. doi: 10.1080/2162402x.2017.1288330.

9. Limkin, E. J., Sun, R., Dercle, L., Zacharaki, E. I., Robert, C., Reuzé, S., et al. (2017). Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Ann. Oncol. 28, 1191–1206. doi: 10.1093/annonc/mdx034.

10. Lu, C., Lewis, J. S., Dupont, W. D., Plummer, W. D., Janowczyk, A., and Madabhushi, A. (2017). An oral cavity squamous cell carcinoma quantitative histomorphometric-based image classifier of nuclear morphology can risk stratify patients for disease-specific survival. Mod. Pathol. 30, 1655–1665. doi: 10.1038/modpathol.2017.98.

11. Ibragimov, B., and Xing, L. (2017). Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. Med. Phys. 44, 547–557. doi: 10.1002/mp.12045.

12. Liu, P., Sun, Y., Zhao, X. et al. Deep learning algorithm performance in contouring head and neck organs at risk: a systematic review and single-arm meta-analysis. BioMed Eng OnLine 22, 104 (2023). https://doi.org/10.1186/s12938-023-01159-.

13. Kanna, G.P., Kumar, S.J., Kumar, Y. et al. Advanced deep learning techniques for early disease prediction in cauliflower plants. Sci Rep 13, 18475 (2023). https://doi.org/10.1038/s41598-023-45403-w.

14. Miotto, R., Wang, F., Wang, S., Jiang, X. & Dudley, J. T. Deep learning for healthcare: Review, opportunities and challenges. Brief. Bioinform. 19, 1236–1246 (2018).

15. Pham T, Tran T, Phung D, Venkatesh S. Predicting healthcare trajectories from medical records: A deep learning approach. J Biomed Inform. 2017 May;69:218-229. doi: 10.1016/j.jbi.2017.04.001. Epub 2017 Apr 12. PMID: 28410981.

16. Ker, J., Wang, L., Rao, J. & Lim, T. Deep learning applications in medical image analysis. IEEE Access 6, 9375–9389 (2018).

17. Gangil, T., Shahabuddin, A.B., Dinesh Rao, B. et al. Predicting clinical outcomes of radiotherapy for head and neck squamous cell carcinoma patients using machine learning algorithms. J Big Data 9, 25 (2022). https://doi.org/10.1186/s40537-022-00578-3.

18. Illimoottil M, Ginat D. Recent Advances in Deep Learning and Medical Imaging for Head and Neck Cancer Treatment: MRI, CT, and PET Scans. Cancers (Basel). 2023 Jun 21;15(13):3267. doi: 10.3390/cancers15133267. PMID: 37444376; PMCID: PMC10339989.

19. Iesmantas, Tomas, Alzbutas, Robertas, 2018. Convolutional capsule network for classification of breast cancer histology images. In: International Conference Image Analysis and Recognition. Springer, Cham, pp. 853–860.

20. Wang, Duolin, Liang, Yanchun, Dong, Xu., 2019. Capsule network for protein posttranslational modification site prediction. Bioinformatics 35 (14), 2386–2394.

21. Panigrahi, Santisudha & Das, Jayshankar & Swarnkar, Tripti. (2020). Capsule network based analysis of histopathological images of oral squamous cell carcinoma. Journal of King Saud University - Computer and Information Sciences. 34. 10.1016/j.jksuci.2020.11.003.

22. Afshar, P., Mohammadi, A., Plataniotis, K.N., 2018. Brain tumor type classification via capsule networks. In: 2018 25th IEEE International Conference on Image Processing (ICIP), IEEE, pp. 3129–3133.

23. Salmanpour, Mohammad R., et al. "Deep versus Handcrafted Tensor Radiomics Features: Prediction of Survival in Head and Neck Cancer Using Machine Learning and Fusion Techniques." Diagnostics 13.10 (2023): 1696.

24. Wang, Zipei, et al. "Radiomics and Deep Learning in Nasopharyngeal Carcinoma: A Review." IEEE Reviews in Biomedical Engineering (2023).

25. Baker, Qanita Bani, et al. "Lung Cancer Survival Time Prediction Using Machine Learning and Deep Learning Techniques." 2023 Fourth International Conference on Intelligent Data Science Technologies and Applications (IDSTA). IEEE, 2023.

26. López-García G, Jerez JM, Franco L, Veredas FJ. Transfer learning with convolutional neural networks for cancer survival prediction using gene-expression data. PLoS One. 2020 Mar 26;15(3):e0230536. doi: 10.1371/journal.pone.0230536. PMID: 32214348; PMCID: PMC7098575.

27. Kim, H.E., Cosa-Linan, A., Santhanam, N. et al. Transfer learning for medical image classification: a literature review. BMC Med Imaging 22, 69 (2022). https://doi.org/10.1186/s12880-022-00793-7.

28. Gu, J.; Wu, B.; Tresp, V. Effective and efficient vote attack on capsule networks. arXiv 2021, arXiv:2102.10055.

29. Guarda, L.; Tapia, J.E.; Droguett, E.L.; Ramos, M. A novel Capsule Neural Network based model for drowsiness detection using electroencephalography signals. Expert Syst. Appl. 2022, 201, 116977.

30. Haq, M.U.; Sethi, M.A.J.; Rehman, A.U. Capsule Network with Its Limitation, Modification, and Applications—A Survey. Mach. Learn. Knowl. Extr. 2023, 5, 891-921. https://doi.org/10.3390/make5030047.

31. Gu, J.; Tresp, V.; Hu, H. Capsule network is not more robust than convolutional network. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 14309–14317.

32. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700- 4708.

33. S. Mascarenhas and M. Agarwal, "A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification," 2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON), Bengaluru, India, 2021, pp. 96-99, doi: 10.1109/CENTCON52345.2021.9687944.

34. S. Namani, L. S. Akkapeddi and S. Bantu, "Performance Analysis of VGG-19 Deep Learning Model for COVID-19 Detection," 2022 9th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2022, pp. 781-787, doi: 10.23919/INDIACom54597.2022.9763177.

35. Dataset Retrieval: - https://wiki.cancerimagingarchive.net/display/DOI/Data+from+Head+and+Neck+Cancer+CT+Atlas#24281354036220c66a5a436f90e4a0b54367bfae