Survival Prediction with Clinical and Image Data via Transfer Learning in Head and Neck Squamous Cell Carcinoma

Deepti[1], Goldie Gabrani[2], Nishtha Phutela[3]

[1]Department of Computer Science & Technology, Manav Rachna University, Faridabad, Haryana, India, Email: deeptigupta@mait.ac.in
[2]Department of Computer Science & Technology, Manav Rachna University, Faridabad, Haryana, India, Email: goldiegabrani@mru.edu.in
[3]Department of Computer Science & Technology, BML Munjal University, Gurugram, Haryana, India, Email: nishtha.phutela@bmu.edu.in

Received: 08.11.2023 Revised : 19.12.2023 Accepted: 21.06.2024

ABSTRACT
Head and neck squamous cell carcinoma (HNSCC) is a type of cancer that exclusively affects the mucous membranes in the head and neck area. Accurate prognostication of cancer patients' survival is essential for various therapeutic objectives, including early disease detection, treatment planning, risk assessment, follow-up care provision, patient counselling, and improving the quality of healthcare. The integration of clinical and imaging data to develop predictive models enables the enhancement of accuracy and clinical utility in survival prediction, ultimately resulting in improved patient treatment and outcomes.Our research focuses on the development of a predictive model that leverages transfer learning techniques to utilize clinical and imaging data obtained from TCIA for predicting the survival rates of patients with HNSCC. Transfer learning allows us to harness pre-trained models, significantly enhancing the efficiency and accuracy of feature extraction and prognostic predictions. Specifically, we employed a manual feature selection technique to assess clinical data, while CapsuleNet was utilized to extract features from imaging data. These features were then combined and inputted into DenseNet 121, a deep learning model fine-tuned through transfer learning, to generate prognostic predictions for patients diagnosed with HNSCC.The proposed model's performance was evaluated using metrics such as precision, sensitivity, accuracy, and F1 Score. Our analysis demonstrates that the transfer learning-based model achieved an accuracy rate of 98.1%, surpassing the performance of existing models, including CNN and ResNet50, which were assessed using same dataset. This highlights the significant potential of transfer learning in improving prognostic predictions and ultimately enhancing patient outcomes.

Keywords: Medical Imaging, Head and Neck Squamous Cell Carcinoma, Dense Net121, Capsule Net, transfer learning.

1. INTRODUCTION

Head and Neck squamous cell carcinoma (HNSCC) mainly develops in the oral, throat, and larynx mucosa. HNSCC treatment results depend on clinical characteristics, therapeutic methods, and disease physiology. This substantially complicates survival estimation. Clinical data may be unable to effectively predict HNSCC risk due to its limited core features [1-3]. Clinical data includes demographics, medical history, test findings, and therapy. It also provides patient features, tumour physiology, therapy history, and disease progression, which are crucial for HNSCC survival prediction. A thorough examination of several clinical variables helps doctors identify high-risk patients and improve treatment outcomes. In contrast, PET/CT imaging uses structural and functional imaging to assess tumour size, shape, appearance, and quantity. A patient's clinical outcomes can be predicted by diagnostic imaging data. PET/CT scans may help doctors diagnose HNSCC by revealing the tumour functioning and appearance. Considering their distinct features, integrating clinical and medical imaging data improves survival prediction. A complete patient assessment, better risk classification, earlier identification of high-risk patients, enhanced treatment planning, greater outcome prediction accuracy, easier clinical decision-making, and research and progress are the benefits.

Healthcare standards and results for people being treated for HNSCC are better when machine learning methods are used [6-8]. Deep Learning (DL) algorithms in clinical decision support systems may use health data about patients, like lab test results, background, symptoms, and other relevant data, to make evaluations more accurate and predict how long they will live. Hierarchical features in medical images like

CT scans include many levels of visual details that are needed to explain the main features seen in the image[9–10]. At each level, the basic structures is broken down into more complex and complicated parts. For doctors to correctly understand and identify patients, they need to take hierarchical features. Often, image processing and analysis methods are used to get these features. The Capsule network, which is a deep-learning model, can successfully pull-out hierarchical data from medical images. Because of how it works, CapsuleNet is a good tool for finding tumors[11–12].
Transfer learning is a common technique in which models that have already been trained, such as CapsuleNet, are refined for a particular task or dataset. DenseNet121 can be fine-tuned in the context of HNSCC survival prediction using a dataset comprising medical images of HNSCC patients. Through fine-tuning, the model can adjust its learned features to correspond with the unique features of HNSCC imaging data extracted from CapsuleNet. By capitalizing on the deep learning functionalities of DenseNet121 and integrating imaging characteristics with clinical data provides improved predictive precision, which speeds up treatments and improves it's outcomes. This makes it possible to construct a model that enhances patient outcomes in the management cancer[13-14].
The rest of the study is structured according to the above key aspects of medical imaging. Section 2 of this paper examines the advancement from machine learning to deep learning in the processing of imaging and clinical data and provides a comprehensive overview of the relevant literature. Section 3 provides an overview of the DL models that were used in the development of our suggested model. It outlines the technique used in constructing the suggested model and the dataset utilized along with the data pre-processing procedures carried out for both clinical and imaging data. Section 4 presents the assessment results of the proposed model, CNN model, and ResNet50 model, using performance measures such as precision, sensitivity, accuracy, and F1 Score.

2. LITERATURE REVIEW

This section includes previous work done in the use of machine learning and transfer learning in healthcare and the integration of clinical data and imaging data for better survival prediction.
HNSCC is an acronym for Head and Neck Squamous Cell Carcinoma. Squamous cell carcinoma is a malignancy that arises from the squamous cells that line the mucous membranes in the head and neck area. Squamous cell carcinoma is the predominant form of head and neck cancer, representing over 90% of all occurrences. The treatment for HNSCC often consists of a multimodal approach, which includes surgical intervention, radiation therapy, chemotherapy, targeted therapy, or immunotherapy. The specific treatment modalities are determined based on factors such as the stage and location of the disease, as well as the patient's general condition.
Disease identification and treatment are possible using machine learning in healthcare. Machine learning (ML) improves data prediction and classification, especially in cancer diagnosis. It outperforms biostatistics in the classification, prediction, and grouping of huge amounts of complicated healthcare data. Machine learning has also excelled in detecting biological components in medical images, reconstructing medical images, and segmenting brain tumors [15-18]. Whereas Machine learning is constrained by the current modalities because of the vast quantities of data involved in healthcare. Deep learning can operate effectively regardless of the data volume. Therefore, DL is seen as an advanced method of ML, which overcomes the two main constraints of ML such as dependence on extensive labeled data and training cost. Moreover, Transfer learning has several benefits compared to conventional machine learning methods, such as a reduction in the requirement of data prerequisites, and wider application across various domains and tasks. Transfer learning is a potent method for using existing knowledge and speeding up progress in machine learning research and application development. A study conducted by Ajagbe et al. [20] explored the use of Deep Convolutional Neural Networks (CNN) and transfer learning techniques using VGG-16 and VGG-19 models to diagnose Alzheimer's Disease (AD) based on MRI data. Regarding six performance parameters, namely Area Under the Curve (AUC), accuracy, F-1 score, precision, computational time, and recall, VGG-16 outperformed the others in one statistic, VGG-19 in three metrics, and CNN in two metrics. Iman, M,et.al[21] in their work reviewed the use of Transfer Learning methods over the previous five years. They also discussed deep transfer learning taxonomy and approach. They examined, categorized and analyzed over thirty freshly submitted Deep Transfer Learning (DTL) research articles. They summarized two large DTL studies that demonstrate an effective deep transfer learning technique for various circumstances.
Authors Zhu et al.[24], in their work, concluded that transfer learning can be used to measure lung disease severity, predict its prognosis, and predict therapeutic response and survival. They used transfer learning to make predictions better. Sample weights that had been learned in ImageNet were put into a VGG16 model. With a similar approach authors Srikantamurthy et. al [25], in their work, they developed a hybrid model that utilized Convolutional Neural Network (CNN) and Long Short-Term Memory Recurrent Neural

Network (LSTM RNN) to accurately classify four different kinds of benign and malignant breast cancer subtypes. This CNN-LSTM model utilized transfer learning on ImageNet to classify and forecast four subtypes of each. The recommended transfer learning strategy correctly classified benign and malignant cancer subtypes better than leading machine and deep learning models. The authors, Isaac Shiri et.al. in their work demonstrated that integration of radiomic, clinical, and radiological data for survival prediction of COVID-19 patient gives better results. In a similar work, Hao et. al found that CT imaging and clinical data details help doctors figure out what stage the cancer is in and how well the surgery will go.
We require the medical data of the patients in the form of imaging data, which may include CT scans, MRI scans, or PET scans, as well as clinical data, which may contain features such as demographic information, tumor characteristics, and treatment history, among other things of the patient.

3. Experimental Method

3.1 Background

Use of Multimodal data and Overview of the DL models that were used in the development of our suggested model.

Multi-modal Data

Multimodal data is information that comes from different sources or modes such as clinical and imaging data. Using multi-modal data in medicine has big benefits for taking care of patients, making clinical decisions, advancing research, and improving the efficiency of healthcare systems. In their work, authors, Lobato-Delgado et. al [26], discussed the advanced prediction models for cancer prognosis that use multimodal data, including clinical, molecular (omics and non-omics), and imaging data. Lastly, they also addressed the obstacles and possibilities in the realm of cancer research, which have significant potential to influence the way patients are treated and contribute to the advancement of personalized and precise treatment. In the similar work authors Vollmer, A et.al ,[27], described how the use of multimodal data processing methodologies has the potential to significantly improve the accuracy of predictive algorithms, which in turn leads to more accurate forecasts of long-term survival for patients diagnosed with oesophageal squamous cell carcinoma (OSCC).
In our study we have Combined CT scans with clinical data to provide a comprehensive view of a patient's condition, leading to more accurate diagnosis and personalized treatment plans. It is used to get a full picture of a patient's health, how a disease is progressing, or how well medicine is working. CT scans provide detailed imaging data that can reveal structural abnormalities, tumors, and other critical health indicators. Clinical data includes patient histories, physician notes, lab results, and other narrative information that offers context and details not visible in imaging alone. By leveraging the diverse information from multimodal data, healthcare providers can improve patient outcomes and foster better patient interactions. This holistic approach is paving the way for more precise and personalized medical care.

CapsuleNet for feature extraction from the imaging data of HNSCC patients

Although convolutional neural networks (CNNs) perform very well in a variety of computer vision tasks, they are unable to consider the geometrical relationships between objects. This indicates that CNNs are unable to process training data that includes affine transformations or images that have been rotated. According to the findings of current study, even a little modification to the size or translation of the input picture may have a significant effect on the performance of the network. When it comes to learning visual representations, the most significant issue with standard convolutional neural networks (CNNs) is that they struggle to comprehend input images of varying sizes and orientations. Furthermore, the pooling layer in these networks often disregards positional information. This is the primary difficulty. The representation learning resilience is improved in CapsNet because of the removal of pooling layers and the replacement of such layers with dynamic routing and convolutional strides. This architecture has shown promising results in a variety of popular applications, including classification, identification, segmentation, and natural language processing, among others. Instead of returning scalar outputs, CapsNet returns vector outputs. This is done so that the part-whole relationships may be maintained. One of the fundamental concepts of CapsNet is the use of encryption techniques to encode the link between a wide variety of factors, such as scales, location, posture, and orientation. A non-facial picture that has a mouth, eyes, and nose may nonetheless be classified as a face by convolutional neural networks (CNNs), even though a human eye would easily recognize that such an image does not correspond to a face. Despite this, the capsule network will acquire the knowledge necessary to distinguish non-face photographs as such, as well as the link between facial characteristics such as the eyes and the nose [28]. There is a representation of the structure of CapsNet in Figure 1. CapsNet is comprised of two convolutional layers in its structure. In the first convolutional layer, there are 256 channels. Each channel is equipped with 9 × 9 filters, which are triggered by a RELU function. The stride of each channel is 1. A convolutional capsule layer could be seen in the second layer. This layer had a stride of 2 microns and included capsules that were 6-by-6-by-32. Each of the major capsules is composed of eight convolutional units, each of which has a kernel that is 9 × 9. Through the use of the squashing function, it has been triggered. The completely linked layer, which is the third and final component of the CapsNet architecture, is composed of sixteen D-capsules that are 10 sizes in size. Each of these capsules is referred to as a DigitCap. During the construction of a capsule network, an extra building component, also known as a capsule, is put between a neuron and a layer. The neuronal layers that are layered inside one another are what constitute a capsule. When it comes to data classification, these capsules collect information from each and every other capsule and utilize ten distinct criteria.

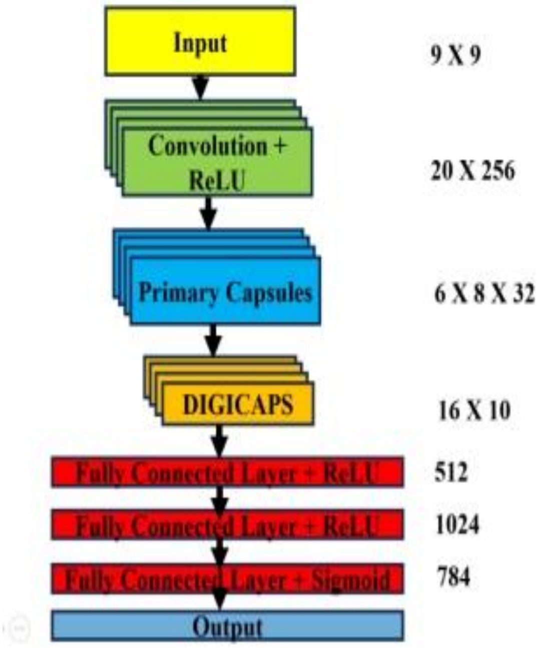


Figure 1. CapsuleNet Architecture

Use of DenseNet121 for survival prediction of HNSCC patients

DenseNet as was created as a subset of CNNs in the original proposal, which was submitted by Huang et al. [29]. The findings achieved using a large number of picture categorization datasets were rather impressive. To construct a feature map that provides data to the layers that are above it, each layer in a DenseNet architecture makes use of inputs from all of the levels that are below it. To do this, thick blocks are used to link the subsequent levels.

As can be seen in equation (1), the nth layer receives all of the inputs from the feature maps that came before it (i.e., i0, i1,..., in respectively).

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

Concerning this particular instance, the concentration of all feature maps that came before it in the n-layer is represented as [n0, n1,..., nn-1]. The nth layer is a composition function that consists of three sequential operations: batch normalization, a ReLU activation function, and convolution. The output of that layer is denoted by the symbol n. ResNet can combine neighboring layers, while DenseNet is only able to concatenate them. Because of this, it is similar to ResNet, although it is not the same. DenseNet can reduce the number of parameters and address the problem of vanishing gradients via the process of repeating characteristic features. Figure 2 serves as an illustration of how DenseNet-121 makes use of four dense blocks. A transition layer is introduced between every pair of blocks. This layer employs down-sampling on the feature maps to construct a convolution layer with dimensions of 1×1 and an average pooling layer with dimensions of 2×2. The dense blocks make use of a series of convolutional layers to bridge the gap between levels that are located at a great distance. The non-linearity of DenseNet-121 is improved using a ReLU activation. Enhancing non-linearity is accomplished using a ReLU activation function in the DenseNet-121 model. In conclusion, a fully connected layer is used for prediction by using a softmax function.

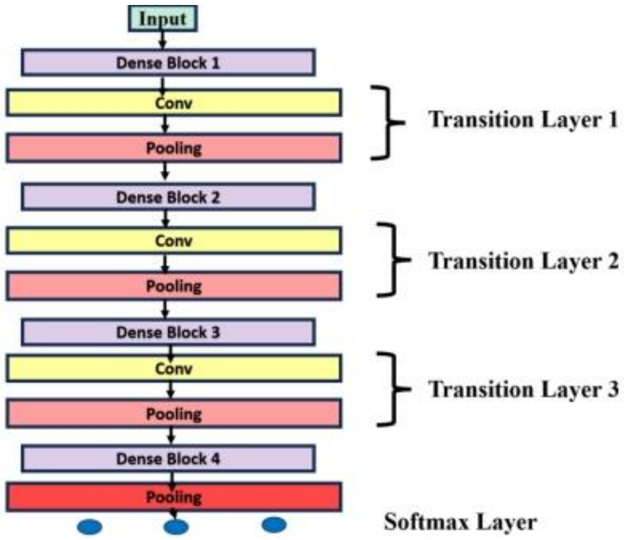


Figure 2. DenseNet121 architecture

3.2 Methodology used for developing Proposed Model

Figure 3 illustrates the workflow of the proposed model. The dataset of clinical data and CT scans of HNSCC patients is downloaded from The Cancer Imaging Archive. The downloaded clinical and imaging data is pre-processed. Imaging data that has been pre-processed is then sent into the capsuleNet model for the feature extraction. On the other hand, the features that are necessary for accurate prediction are selected from clinical data during the pre-processing stage. Further on, the features derived from clinical information are integrated with the imaging data. In the proposed model this integrated data is then sent to DenseNet-121 for survival prediction of the HNSCC patients.

Following the completion of the task, the pre-processed data is fed to CNN and ResNet50 models to predict the survival of HNSCC patients. The proposed model, CNN Model, and ResNet50 models are assessed using metrics such as F1 score, accuracy, sensitivity, specificity, and precision. The performance evaluation of the proposed model is thereafter compared to that of the CNN and ResNet50 models.

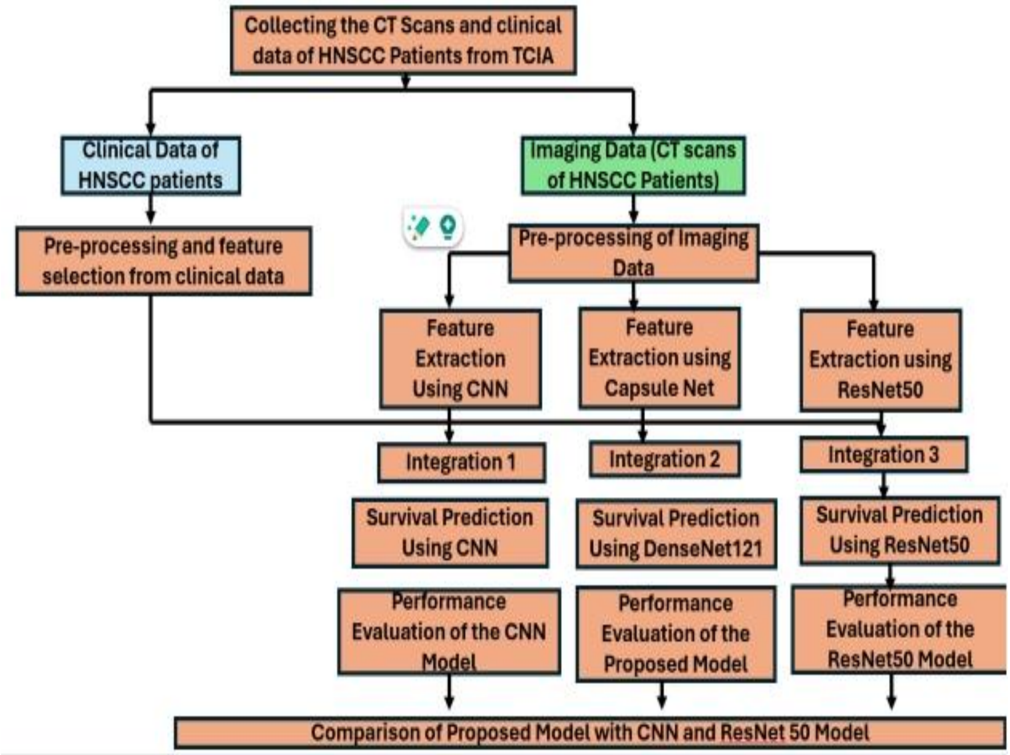


Figure 3. Working Methodology

3.3 Dataset and Data pre-processing

Clinical data comprises of all the information that is gathered while treating people medically, generally in places like hospitals, clinics, or doctors' offices. Demographic information, medical history, vital signs, treatment information, clinical results, and follow-up statistics are all part of the clinical data. To prepare the dataset to be fed into a ML model, pre-processing of the dataset is required. Pre-processing of clinical data includes, selecting the null values or the missing data, removing the duplicate values, encoding the data etc. Data types with both numerical and categorical variables may be encoded using methods like Ordinal Encoding, Onehotencoding, and Feature Hashing, according to research. There may be missing items in the dataset that need to be addressed first. Research has shown that datasets containing many missing elements are amenable to many iterative imputation methods. There can be more columns than rows in the dataset. Executing suitable feature selection algorithms is, hence, crucial. Numerous publications detail feature selection strategies, including PCA, ICA, filter-based methods, wrapper-based approaches, and embedded approaches [30].

We cannot utilize medical images the way we received them. The acquired images need to undergo pre-processing and segmentation. Medical image preprocessing has real limitations due to its complex nature compared to preparing other types of images. Due to limitations in image capture, image division is a tedious process.

Medical images may be provided in either the DICOM, Nifti, or Minc formats in practice. Nuclear medicine, computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, and positron emission tomography (PET) are among the most prevalent imaging modalities.

To get cross-sectional images, computed tomography (CT) scans measure the x-ray attenuation as it passes through a spinning energy source. Variations in tissue composition (density) allow CT imaging to reveal anatomical details. A quantitative measure of tissue density relative to water called the Hounsfield unit (HU), is produced by the nature of CT collection. The voxel HU values in CT scans may be reproduced, with minor variations, across scanners and patients if they adhere to the usual temperature and pressure guidelines. Although computed tomography (CT) offers excellent contrast of important anatomical features, ionizing radiation and poor contrast of soft tissues make it an unpopular choice in many therapeutic contexts. [31].

3.3.1 Data pre-processing of CT scans

The source from The Cancer Imaging Archive consisting of the collection of Head and Neck Squamous Cell Carcinoma (HNSCC) medical images was downloaded. The dataset of 215 patients was categorized into 3,225 series and 765 investigations. We acquired a total of 2593 cancer images from the whole dataset. The original image had dimensions of 512 pixels by 512 pixels.

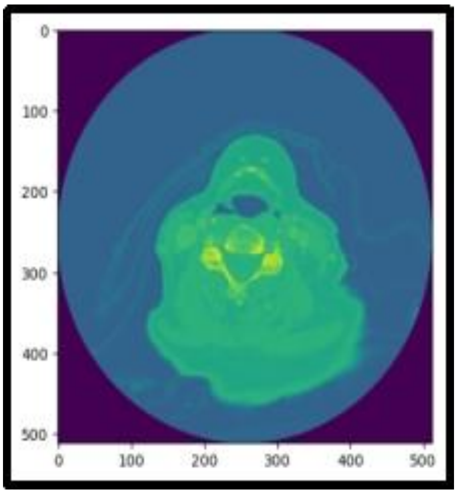
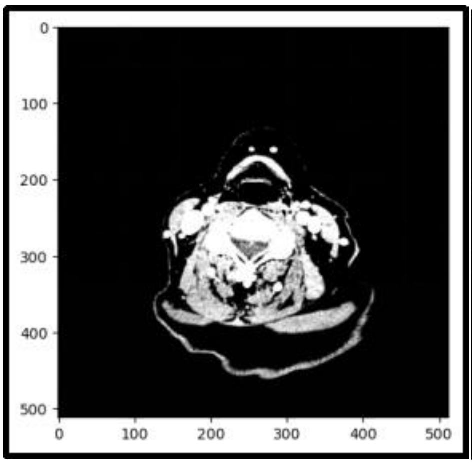


Figure 4. Unprocessed Random CT Scans image of a patient of the Hypopharyngeal region

Medical image preprocessing reduces the distortions in the dataset and increases the accuracy of the model. When DICOM images are converted to grayscale, the spatial organization and linkages between the different anatomical features that are present in the images are preserved inside the images therefore, we converted the images into grayscale. Capsule Networks provide remarkable capabilities in recording spatial hierarchies and linkages. Furthermore, Gray images provide a clear portrayal of the spatial qualities, so reducing the complication that is generated by color channels for visualization purposes. Grayscale image for image in figure 4 is shown in figure 5.



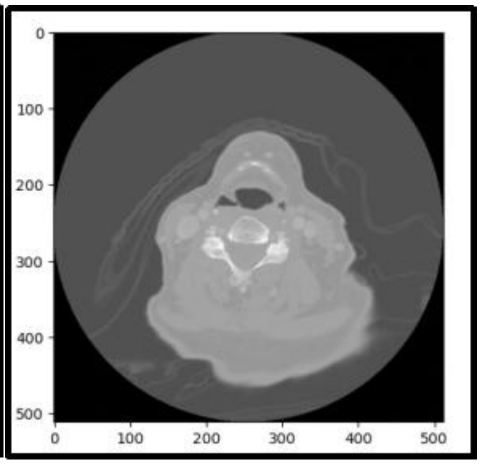


Figure 5. Image converted to Gray Scale Figure 6: Image converted to HU after pre- processing.

When measuring the radiodensity of tissue, CT imaging makes use of the HU. We utilize Hounsfield unit values for grayscale images obtained from CT scans. The inner architecture of the body may be visualized by radiologists via the use of grayscale coloring of tissues based on their HU values. The darker portions represent tissues with lower HU values, whereas the brighter sections suggest tissues with higher HU values because of their higher brightness. Radiologists can differentiate between tissues and identify abnormalities or illnesses based on the density of the CT image with the assistance of this grayscale display. Figure 6 shows a HU image. Next, to improve the quality of the data that is used for model training, it is vital to remove noise during the preprocessing stage. Therefore, the performance of the machine learning model would be increased. An image after noise removal is shown in figure 7.

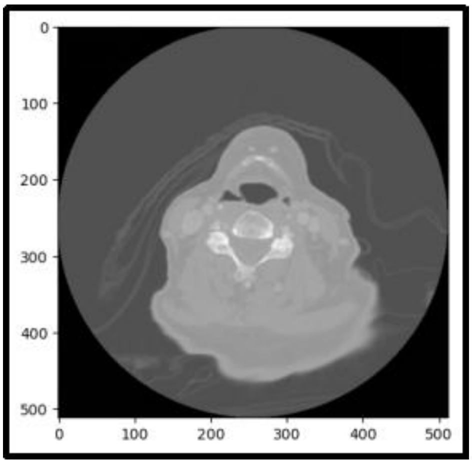


Figure 7. Image after removal of noise

After the images had been pre-processed, the pixel data was extracted and converted into NumPy arrays using Python. The data is randomized to make the training of the model easier. Immediately after the randomization of the data, we separated it into two distinct groups: the training set and the testing set. A total of 1884 images were included in the training set, while 709 images were included in the testing set.

3.3.2 Data pre-processing of clinical data

There are a few distinguishing features of the clinical dataset that was released. For developing our model for survival prediction, we choose certain essential criteria from among the numerous acquired traits. Figure 8 illustrates the clinical characteristics that are included in the patient's clinical dataset. These characteristics include the patient's gender, age, diagnosis, cancer location (site), histology, differentiation grade, tumor, node, metastatic, stage, HPV status, survival, site of recurrence, loco-regional control censor, oncology treatment, smoking history, current smoker status, and whether the patient is active or inactive. We extracted only these features as few of them can be correlated with the imaging data of the same patient and the rest will help the model for better survival prediction.

The features or the characteristics of the patients are represented as columns in the clinical data. Figure 8, displays the features selected such as 'sex, age, site, grade, T, N, M, stage and live or dead' from the complete clinical data.

```
[['Female' 54 'Oropharynx' 'moderately to poorly diff.' 2 '2c' '0' 'IVA'
  'Alive']
 ['Female' 56 'Nasopharynx' 'moderately diff.' 4 2 '0' 'IVA' 'Dead']
 ['Male' 48 'Oropharynx' 'moderately diff.' 2 '2b' '0' 'IVA' 'Alive']
 ['Male' 65 'Oropharynx' 'poorly diff.' '4a' '2c' '0' 'IVA' 'Dead']
 ['Male' 66 'Oropharynx' 'moderately diff.' 2 1 '0' 'III' 'Alive']
 ['Male' 62 'CUP' 'moderately diff.' 0 '2a' '0' 'IVA' 'Dead']
 ['Male' 39 'Oropharynx' 'poorly diff.' 2 3 '0' 'IVB' 'Alive']
 ['Male' 60 'Oropharynx' 'moderately diff.' 2 1 '0' 'III' 'Alive']
 ['Male' 57 'Oropharynx' 'moderately diff.' 3 '2c' '0' 'IVA' 'Alive']]
```

Figure 8. sample clinical data

The data set extracted includes categorial as well as numerical data. We used ColumnTransformer to prepare the data as it can handle numerical and categorical data easily. ML system works with numbers as they cannot handle categorical data, therefore we have used OneHotEncoder which converts the categorical data into a binary vector. The result is stored as numpy array and it is integrated with the features extracted through the imaging data.
Deep learning models like Capsule net, CNN, Resnet50 are used to extract features from image data. Features extracted from models are integrated with clinical data for survival prediction using Densenet121, CNN and Resnet50 respectively.

4. RESULTS

These models ResNet50, CNN, as well as the proposed model can be evaluated using five performance metrics: accuracy, sensitivity, specificity, precision, and F1 score [32-33].**Table 1**, displays the performance metric values for all the models. It can be discerned from the performance metric table that our proposed model is showing better results as compared to other models.

Table 1. Performance Metrices Values for the Three Models

Model Name	**Accuracy**	**sensitivity**	**specificity**	**Precision**	**F1 Score**
CNN	0.964134	0.970497	0.957854	0.957854	0.964134
ResNet 50	0.97069	0.977623	0.963763	0.964231	0.970881
Proposed Model	0.981874	0.981777	0.981975	0.982523	0.98215

As shown in Figure 9, which displays the results of comparing the models' performance indicators, the proposed model outperforms ResNet50 and CNN in terms of accuracy. An outstanding 98.18% accuracy rate defines the proposed model.

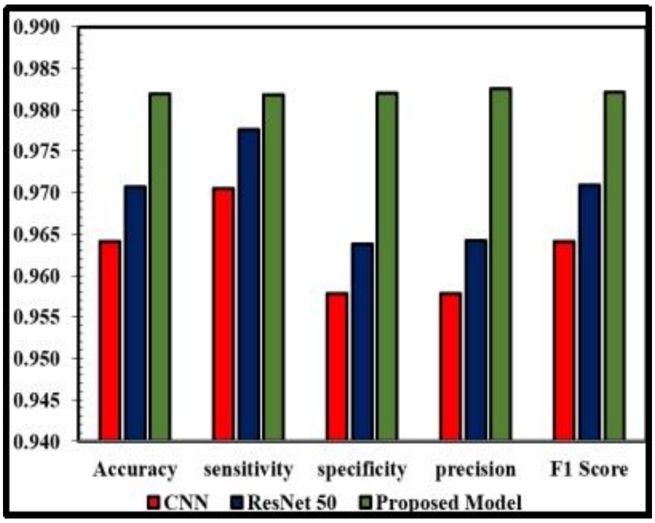


Figure 9. Assessment of the Models using Performance Metrices.

To visualize the performance of model during training and testing accuracy versus epoch graph is shown in figure 10. As illustrated by Figure 10, the training and testing accuracies improved as the number of epochs rose for the proposed model. The training loss as shown in figure 11 is exhibiting a decreasing trend with epochs followed by attaining a steady state, which signifies that the model has effectively acquired the knowledge from the training data and its predictive capabilities has been enhanced.

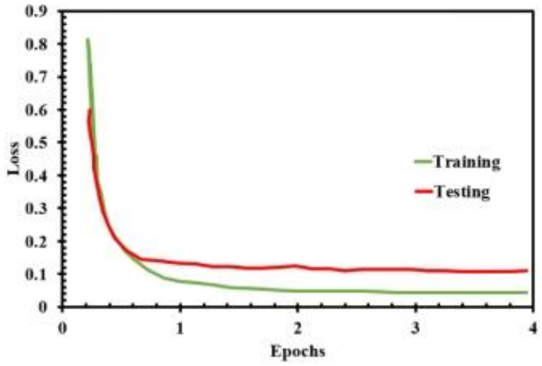


Figure 10. Accuracy Vs Epochs Graph of Proposed Model.

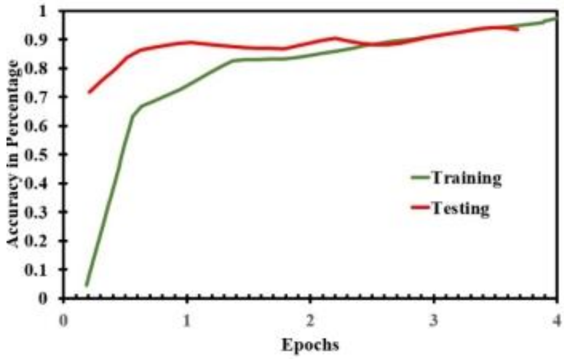


Figure 11. Loss Vs Epochs Graph of Proposed Model.

CONCLUSION

The research suggested a deep learning model that uses clinical and imaging data to predict the survival of patients with Head and Neck Squamous Cell Carcinoma (HNSCC). Transfer learning is used to enhance the precision of the suggested model. CapsuleNet effectively retrieved hierarchical features from medical images for feature selection of imaging data. The integration of clinical data with the attributes derived by capsuleNet has enhanced the possibilities for research and innovation. Transfer learning significantly improved the accuracy of forecasting the survival of HNSCC patients. Conversely, the same approach of integrating data to predict the survival of HNSCC patients is applied where other deep learning models of ResNet50 and CNN are used for feature extraction and survival prediction. The suggested model's performance is compared with CNN and ResNet50, focusing on performance measures like F1 score, accuracy, sensitivity, specificity, and precision. The evaluation metrics have revealed that the suggested model surpasses the performance of the CNN and ResNet50 models.

5. Compliance with Ethical Standards

Disclosure of potential conflicts of interest: Not applicable (Not funding)
Research involving Human Participants and/or Animals: None
Informed consent: None
Funding Declaration: No funding was required

REFERENCES

[1] Johnson, D. E., Burtness, B., Leemans, C. R., Lui, V. W. Y., Bauman, J. E., & Grandis, J. R. (2020, November 26). Head and neck squamous cell carcinoma. Nature Reviews Disease Primers. https://doi.org/10.1038/s41572-020-00224-3.
[2] Barber PR, Mustapha R, Flores-Borja F, Alfano G, Ng K, Weitsman G, Dolcetti L, Suwaidan AA, Wong F, Vicencio JM, Galazi M, Opzoomer JW, Arnold JN, Thavaraj S, Kordasti S, Doyle J, Greenberg J, Dillon MT, Harrington KJ, Forster M, Coolen ACC, Ng T. Predicting progression-free survival after systemic

therapy in advanced head and neck cancer: Bayesian regression and model development. Elife. 2022 Dec 23;11:e73288. doi: 10.7554/eLife.73288. PMID: 36562609; PMCID: PMC9815805.
[3] Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE, Grandis JR: Head and neck squamous cell carcinoma. Nat. Rev. Dis. Prim 6, 1–22 (2020) [PubMed: 31907359].
[4] Haider SP, Burtness B, Yarbrough WG, Payabvash S: Applications of radiomics in precision diagnosis, prognostication and treatment planning of head and neck squamous cell carcinomas. Cancers Head Neck 5, 1–19 (2020) [PubMed: 31938572].
[5] Chinnery T, et al. : Utilizing artificial intelligence for head and neck cancer outcomes prediction from imaging. Can. Assoc. Radiol. J 72, 73–85 (2021) [PubMed: 32735452].
[6] Alkhadar H, Macluskey M, White S, Ellis I, Gardner A. Comparison of machine learning algorithms for the prediction of five-year survival in oral squamous cell carcinoma. J Oral Pathol Med. 2021;50(4):378–84.
[7] Chu CS, Lee NP, Adeoye J, Thomson P, Choi SW. Machine learning and treatment outcome prediction for oral cancer. J Oral Pathol Med. 2020;49(10):977–85.
[8] Mandal S, Gupta A, Chanu WP. Survival prediction of head and neck squamous cell carcinoma using machine learning models. 2021;1–8. Available from: http://arxiv.org/abs/2105.07390.
[9] Tortora, Mario, Laura Gemini, Alessandra Scaravilli, Lorenzo Ugga, Andrea Ponsiglione, Arnaldo Stanzione, Felice D'Arco, Gennaro D'Anna, and Renato Cuocolo. "Radiomics applications in head and neck tumor imaging: a narrative review." Cancers 15, no. 4 (2023): 1174.
[10] Yang, D., Wu, Y., Wan, Z., Xu, Z., Li, W., Yuan, P., Shang, Q., Peng, J., Tao, L., Chen, Q. and Dan, H., 2023. HISMD: a novel immune subtyping system for HNSCC. Journal of Dental Research, 102(3), pp.270-279.
[11] Qin, Chendong, Yongxiong Wang, and Jiapeng Zhang. "CapsANet: Medical image segmentation network based on convolution capsule encoder and attention mechanism." (2023).
[12] Saif, A. F. M., Tamjid Imtiaz, Shahriar Rifat, Celia Shahnaz, Wei-Ping Zhu, and M. Omair Ahmad. "CapsCovNet: A modified capsule network to diagnose Covid-19 from multimodal medical imaging." IEEE Transactions on Artificial Intelligence 2, no. 6 (2021): 608-617.
[13] Salehi, Ahmad Waleed, Shakir Khan, Gaurav Gupta, Bayan Ibrahimm Alabduallah, Abrar Almjally, Hadeel Alsolai, Tamanna Siddiqui, and Adel Mellit. "A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope." Sustainability 15, no. 7 (2023): 5930.
[14] Iman, Mohammadreza, Hamid Reza Arabnia, and Khaled Rasheed. "A review of deep transfer learning and recent advancements." Technologies 11, no. 2 (2023): 40.
[15] Sarker, Mithun. "Revolutionizing Healthcare: The Role of Machine Learning in the Health Sector." Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023 2, no. 1 (2024): 35-48.
[16] An, Qi, Saifur Rahman, Jingwen Zhou, and James Jin Kang. 2023. "A Comprehensive Review on Machine Learning in Healthcare Industry: Classification, Restrictions, Opportunities and Challenges" Sensors 23, no. 9: 4178. https://doi.org/10.3390/s23094178.
[17] Sinha, U.; Singh, A.; Sharma, D.K. Machine learning in the medical industry. In Handbook of Research on Emerging Trends and Applications of Machine Learning; IGI Global: Hershey, PA, USA, 2020; pp. 403–424.
[18] Yan, Z.; Zhan, Y.; Peng, Z.; Liao, S.; Shinagawa, Y.; Zhang, S.; Metaxas, D.N.; Zhou, X.S. Multi-instance deep learning: Discover discriminative local anatomies for bodypart recognition. IEEE Trans. Med. Imaging 2016, 35, 1332–1343.
[19] Rana, M., Bhushan, M. Machine learning and deep learning approach for medical image analysis: diagnosis to detection. Multimed Tools Appl 82, 26731–26769 (2023). https://doi.org/10.1007/s11042-022-14305-w.
[20] Priyatama, Adhigana & Sari, Zamah & Azhar, Yufis. (2023). Deep Learning Implementation using Convolutional Neural Network for Alzheimer's Classification. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi). 7. 310-217. 10.29207/resti.v7i2.4707.
[21] Iman, M.; Arabnia, H.R.; Rasheed, K. A Review of Deep Transfer Learning and Recent Advancements. Technologies 2023, 11, 40. https://doi.org/10.3390/technologies11020040.
[22] Zhu, J., Shen, B., Abbasi, A., Hoshmand-Kochi, M., Li, H., & Duong, T. Q. (2020, July 28). Deep transfer learning artificial intelligence accurately stages COVID-19 lung disease severity on portable chest radiographs. PLOS ONE. https://doi.org/10.1371/journal.pone.0236621..
[23] Srikantamurthy, M.M., Rallabandi, V.P.S., Dudekula, D.B. et al. Classification of benign and malignant subtypes of breast cancer histopathology imaging using hybrid CNN-LSTM based transfer learning. BMC Med Imaging 23, 19 (2023). https://doi.org/10.1186/s12880-023-00964-0.
[24] Isaac Shiri, Majid Sorouri, Parham Geramifar, Mostafa Nazari, Mohammad Abdollahi, Yazdan Salimi, Bardia Khosravi, Dariush Askari, Leila Aghaghazvini, Ghasem Hajianfar, Amir Kasaeian, Hamid

Abdollahi, Hossein Arabi, Arman Rahmim, Amir Reza Radmard, Habib Zaidi, Machine learning-based prognostic modeling using clinical data and quantitative radiomic features from chest CT images in COVID-19 patients, Computers in Biology and Medicine, Volume 132, 2021, 104304, ISSN 0010 4825, https://doi.org/10.1016/j.compbiomed.2021.104304.

[25] Hao, D., Li, Q., Feng, Q. S., Qi, L., Liu, X., Arefan, D., Zhang, Y., & Wu, S. (2022, February 2). Identifying Prognostic Markers From Clinical, Radiomics, and Deep Learning Imaging Features for Gastric Cancer Survival Prediction. Frontiers in Oncology. https://doi.org/10.3389/fonc.2021.725889.

[26] Lobato-Delgado B, Priego-Torres B, Sanchez-Morillo D. Combining Molecular, Imaging, and Clinical Data Analysis for Predicting Cancer Prognosis. Cancers (Basel). 2022 Jun 30;14(13):3215. doi: 10.3390/cancers14133215. PMID: 35804988; PMCID: PMC9265023.

[27] Vollmer, A., Hartmann, S., Vollmer, M. et al. Multimodal artificial intelligence-based pathogenomics improves survival prediction in oral squamous cell carcinoma. Sci Rep 14, 5687 (2024). https://doi.org/10.1038/s41598-024-56172-5

[28] 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.

[29] 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.

[30] Moncada-Torres A, van Maaren MC, Hendriks MP, Siesling S, Geleijnse G. Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival. Sci Rep. 2021;11(1). https://doi.org/10.1038/s41598-021-86327-7.

[31] Masoudi S, Harmon SA, Mehralivand S, Walker SM, Raviprakash H, Bagci U, Choyke PL, Turkbey B. Quick guide on radiology image pre-processing for deep learning applications in prostate cancer research. J Med Imaging (Bellingham). 2021 Jan;8(1):010901. doi: 10.1117/1.JMI.8.1.010901. Epub 2021 Jan 6. PMID: 33426151; PMCID: PMC7790158.

[32] 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.

[33] 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.