Title of The Project:
Oral Cancer Detection Using Deep Learning

Supervisor Name and Designation: Dr. Mohammad Monirujjaman Khan

Department Name and University Name: ECE NSU

Group Member Name And ID:

NAME	ID
MD. IMTIAZ HOQUE	1811330042
MD. AZMI SIDDIQUE	1812307042
MD. JAHIRUL ISLAM	1821368642

Submission Date: 03 - April - 2022

1. Background Study of The Research Topic:

Oral cancer is a tumor that develops in a part of the mouth. It may be on the surface of the tongue, the inside of the cheeks, the roof of the mouth (palate), the lips or gums**.** It is difficult, widespread cancer with a high fatality rate. Modern technology and deep learning algorithms make early detection and classification possible. Computer-assisted diagnosis and detection, as well as medical imaging methods, have the potential to change cancer treatment. Deep learning is a machine learning and artificial intelligence (AI) technique that is modeled after how humans learn. Deep learning is a major component of data science, which includes statistics and predictive modeling. Oral cancer has become the most widely reported disease in recent decades, with the World Health Organization recording 4.5 million cases by socioeconomic groups in 2017; around 85 percent of cases result in death. The death rate will be reduced by 70% if the disease is detected early. GLOBOCAN forecasted 354,864 new cases and 177,384 fatalities in 2018 [1]. Oral cancer has been shown to have uncontrolled cell growth that supplies damaged adjacent tissues [2]. Oral cancer is caused by tobacco use in any form, as well as excessive alcohol consumption. In South and Southeast Asia, chewing betel quid, which is made of areca nut, slaked lime, betel leaf, and sometimes tobacco, is a common activity [3]. These quids are now available commercially in sachets and are popular among the general public thanks to aggressive marketing efforts. Late presentation of oral cancer is common, especially in low- and middle-income countries (LMICs), where more than two-thirds of cases present late, resulting in poor survival rates [4]. Oral cancer detection is critical since the sooner it is discovered, the sooner it can be treated and the better the prospects of survival. As a result, we'll use deep learning to detect oral cancer. We'll be looking for lesions to do so. Any damage or aberrant change in an organism's tissue, usually caused by disease or trauma, is referred to as a lesion. It could be benign (not cancerous) or malignant (cancerous) (cancer). We may be able to create an automated approach for diagnosing mouth cancer. As previously stated, the majority of people are poor, making treatment of the terminal stage of cancer unaffordable. However, our system can identify cancer at an early stage and provide patients with useful feedback, and it is also cost effective, so anybody can afford it. Because oral cancer is a global issue, anyone from anywhere in the globe can use our system via an online health platform to stay safe.

2. Description of The Problem Being Solved (Problem Statement):

Oral cancer at early stage if can be detected than the survival chance gets high. Oral cancer being one of the deadliest cancers is costly to detect and if diagnosed then the expense for treatment is high too. In poor countries where betel leaf, tobacco is common has high risk of oral cancer and due to high cost detection test and treatment becomes unaffordable for a large number of people. We’ll be trying to build a system which will use deep learning and will be able detect oral cancer at early stage. Using deep learning we will be using

CNN model, VGG16 architecture Flatten, Dense, Deep pool, fully connected layer. We'll be using large number of dataset and train them for early detection so that oral cancer tissue can be detected.

3. Review of Existing Similar Systems:

In the paper, "Computer-assisted Medical Image Classification for Early Diagnosis of Oral Cancer Employing Deep Learning Algorithms," by Jeyaraj, P. R., and Samuel Nadar, E. R, they conceived and built a deep learning method for an autonomous cancer diagnostic system based on a partitioned convolution neural network. They present the studied elements of hyperspectral medical images from oral cancer case studies, and they employed the stochastic neighbor embedding approach to graphically display the hyperspectral images' evaluated components. They compare built deep CNNs' performance to that of SVM and DBN, as well as other traditional classification algorithms. From a single phase of training, this proposed deep CNN can offer correct categorization. As a result, our deep learning algorithm may be simply implemented on a simple workstation to give an autonomous medical picture classifier without the need for professional experience [8].

In the paper, "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer" by Roshan K, Alex Welikala, Paolo Remagnino, Jian Han Lim, Chee Seng Chan, Senthilmani Rajendran, Thomas George Kallarakkal, Rosnah Binti Zain, Ruwan Duminda Jayasinghe, Jyotsna Rimal, Alexander Ross Kerr, Rahmi Amtha, Karthikeya Patil, and Wanninayake Mudiyanselage Tilakaratne, John Gibson discussed how to take and analyze images of the oral cavity, as well as how to automate oral cancer diagnosis. The contribution of this work is a novel method for merging bounding box annotations from many doctors, followed by a comparison of two deep learning-based automation methods. These encouraging preliminary results demonstrate that deep learning is effective and capable of tackling this demanding challenge. Performance is projected to increase as the dataset develops, which will have a significant impact in low- and middle-income countries with limited health resources [9].

In the paper, "Deep Machine Learning for Oral Cancer: From Precise Diagnosis to Precision Medicine", by Rasheed Omobolaji Alabi, Alhadi Almangush, Mohammed Elmusrati, and Antti A. Mäkitie, They concentrated on deep learning's technical skills and approaches for OSCC. It looks into how deep learning may help identify cancer, classify photos, segment them, synthesize them, and plan treatments. Finally, they talk about how this technology can help with precision medicine and the future of deep learning in the treatment of oral squamous cell carcinoma [12].

In the paper, "Computer-assisted Medical Image Classification for Early Diagnosis of Oral Cancer Employing Deep Learning Algorithm Paper," by Jeyaraj, P. R., and Samuel Nadar, E. R., They created a new learning system that uses weighted majority voting on given criteria to categorize normal, pre-, and post-cancerous areas in hyperspectral image datasets. Models such as SVM, SVM-PCA, and DBM were created and built, and the resulting classifiers were fused. For beginnings, they reduced the hyperspectral image's properties in the SVM and SVM PCA models. The proposed fusion method increases mixed pixel detection sensitivity while also exceeding existing deep learning algorithms in terms of accuracy [13].

4. Objective of The Project:

The objective of this system is developing a system which can detect oral cancer. The percentage of accuracy needs to be good enough to be acceptable. The system needs to be cost friendly as if so it can be widely used and as well as spread for distribution. So, people from poor countries can afford to use the system and medical institutions can widely use it upon need.

5. Feasibility Study Indicating Possible Solutions:

To develop our system, we will be using deep learning. Deep learning is a machine learning and artificial intelligence (AI) technique inspired by human learning. Data science, which covers statistics and predictive modeling, contains deep learning as a major component. Throughout developing our system, we will be using either KNN or CNN model. We will be applying VGG16 architecture in our model. Also, we will use deep pooling, Flatten, Dense, Fully connected layers etc. For performance analysis we will use confusion matrix such as true positive, true negative etc. For datasets we will be using kaggle. Our system will take data as input and then split, process, train the data and lastly provide a result with acceptable accuracy.

6. Output of The Project or Expected Results of The Project:

We'll take a number of different models and put them through their paces one by one. We'll need X-ray photographs of patients with oral cancer because our goal is to develop a system that can detect the condition. Our goal is to find any matches between the collected data and the data that has been submitted recently and to provide a response. Our goal is to reach a 94 to 98 percent accuracy rate. When matching, there's a chance you'll make a mistake or have a percentage error, or the system can contain erroneous data. We'd try to choose a dataset with clear images so that the training data wouldn't be hampered.

7. Detailed Diagrams for The Complete System and All Subsystems:

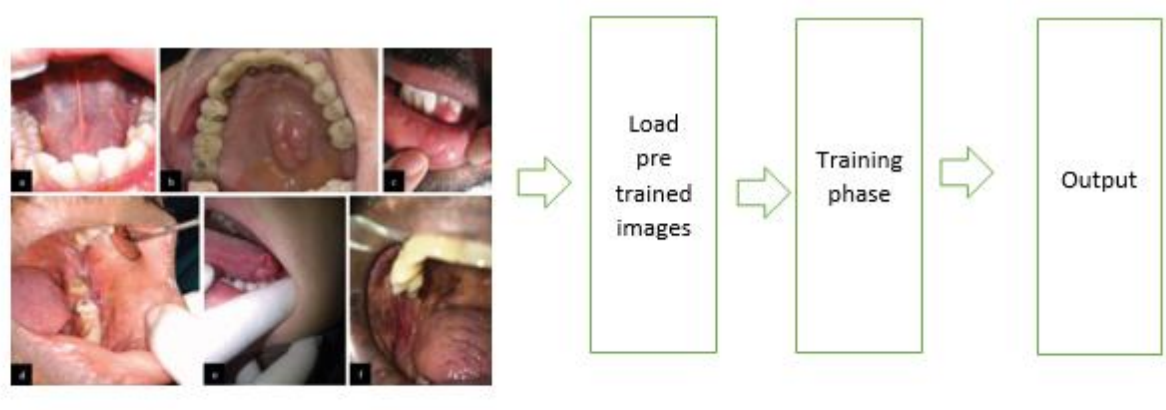


Figure: 1

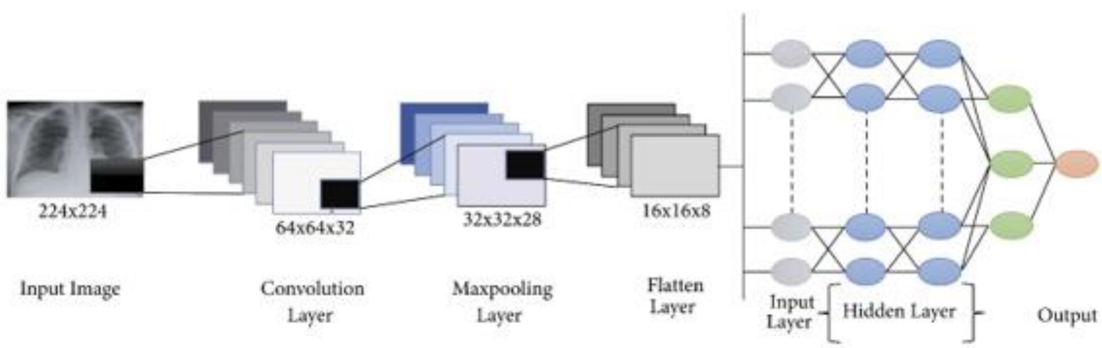


Figure: 2

The Proposed CNN Model:

The CNN model utilized in this work has two important components: feature extractors and classifiers. A CNN model builds a network using a hierarchical technique, resulting in a completely interconnected layer that resembles neurons connected to one another. As a result, this model produces the most accurate and efficient photo categorization results. Figure 3 depicts the general CNN architecture used in this study.

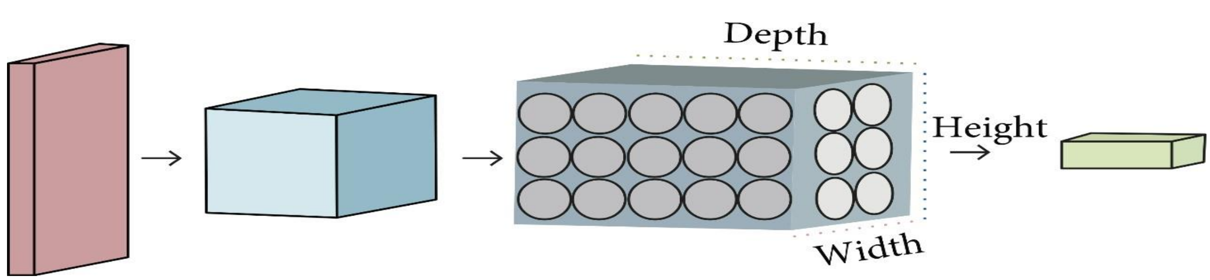


Figure 3: A general 3-dimensional CNN structure [22]

VGG16 is a CNN architecture that can be used to solve image classification problems on a big dataset. The model uses ImageNet to load pre-trained weights while the fully connected (FC) layer head is turned off. Three fully connected layers make up the FC layer, which is followed by a succession of convolutional layers. Each of the first two tiers contained 4096 channels. With 1000 channels, the third layer can do ILSVRC classification in 1000 different ways. The softmax layer, which has the same number of nodes as the output layer, is the final layer. This layer is widely used in multiclass classification situations when more than two labels must belong to the same class.

In deep learning, an epoch is a complete iteration of the samples. The epoch, which is a hyperparameter, determines how many iterations a model is applied to the training dataset. At each epoch during training, the dataset's sample changes the internal model parameters. In an epoch, there may be one or more batches. An epoch with only one batch is referred to by the batch gradient descent learning algorithm [23]. In the coding phase of this investigation, 25 epochs were used. The hyperparameter batch size sets how many data points must be examined before the model parameters are updated, allowing the model to improve over time. Making batch size predictions can be viewed as a series of iterations over one or more variables. The predictions are compared to the predicted output variables at the conclusion of the batch, and the error is calculated. The present model, for example, improves itself from this error by traveling down the error gradient. In the coding phase of this experiment, the batch size was set to 16 and the starting learning rate was set to 13%. Flattening is a method of transforming multidimensional data into one-dimensional data that can then be used in following layers. In this study, flattening was used to turn convolutional layers into a one-dimensional feature vector. The output is then forwarded to the classification layer [24]. We also used a [25] pool size and an average pooling layer (4,4).

Performance Analysis:

The suggested model's performance was assessed using a variety of criteria, including accuracy, recall, sensitivity, specificity, and precision. Various parameters in the confusion matrix, such as true positive (TP), true negative (TN), false positive (FP), and false negative (FN), are used to evaluate the metrics (FN). The following are the metrics:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}, \tag{1}$$

The percentage of real positive cases that are correctly predicted is determined by sensitivity. This statistic assesses the model's ability to anticipate. The following is the equation for calculating sensitivity:

$$\frac{\text{Sensitivity}}{\text{Recall}} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{2}$$

The proportion of actual bad occurrences, which was successfully predicted, was clarified using specificity. The specificity of a model is a measure of its ability to forecast true-negative cases in a given category. These metrics were used to each categorization model to help understand the findings. The formula for calculating specificity is as follows:

$$\text{specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}}. \tag{3}$$

Precision reflects how well the model performed on the test data. It displays the total number of models that were accurately predicted from all positive classes. This should be set to the highest level possible:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{4}$$

The whole code for predicting pneumonia in COVID-19 patients was built using Anaconda's Jupyter notebook. This research can be completed using Jupyter notebook, which is an open-source platform with all of the necessary libraries.

System Architecture:

This system includes a CXR representation, which results in an illustrated forecast. There are three streams in the input form of 224 224. We utilized a filter size of 32 for padding, a kernel size of 3, and a ReLU-based activation function for the first two layers. In the first max-pooling layer, the pool size is 2 and the strides are 2. By integrating all of the pooled properties in a new plain layer, a new cell is generated. In the end, two thick layers were formed. The activation function of the first layer is ReLU, while the activation function of the least thick layer is softmax. The features are introduced to the network after they have been preprocessed.

Convolutional Layer:

Deep transfer learning's foundation is the convolutional layer. The design decisions are made by this group. This layer applies a filter on the image you provide. By integrating the results of the same filters, the function map is formed. A convolution approach is used to multiply the input, which does weight range multiplication. A filter is created by multiplying an array of input data with a two-dimensional collection of weights. When applied to a filter-sized area of the source and filter, a dot product yields a single value. Between the filter-sized patchwork of the input and the filter, this product is employed. The filter's range is narrower than the input, and it's used to multiplex data from many sources using the same filter. Because it thoroughly surrounds the entire frame, the filter is developed as a one-of-a-kind way of distinguishing specific types of features.

Assume the XX input is VRJK, where J denotes the characteristics of the input frequency band and K denotes the whole set of input bandwidths. In this case of filter bank attributes, K refers to the size of the filter bank function vector. Assume that is the band b function's vector. The following formula can be used to compute the activations of the convolution layer:

$$g_{d,e} = \theta\left(\sum_{b=1}^{s} w_{b,d}^{T} v_{b+e-1} + a_d\right), \qquad (1)$$

where is the bias of the feature map, which is the convolution layer's output; s provides the filter scale; is the filter's weight; is the bias of the feature map; The activation function is denoted by (x) [26].

The pooling layer emphasizes the availability of features by allowing feature downsampling. This is mostly due to a convolution layer and has some spatiotemporal invariance. Average pooling and max pooling [27] represent the average presence and maximum active occurrence of a function, respectively.

The pooling layer, in effect, removes extraneous features from the photos and makes them more intelligible. The technique by which the layer averages the relevance of its present attitude is referred to as "average pooling." When max pooling is turned on, the layer selects the greatest value from the current view of the filter each time. The max-pooling approach picks only the highest value using the matrix size specified for each feature map, resulting in fewer output units.

As a result, the image is significantly smaller, but the context remains the same. A dropout layer prevents overfitting, while a pooling layer keeps the number of feature mappings and system parameters to a minimum. The following formula can be used to compute the activation of max pooling:

$$a_{b,c} = \max_{d=1}^{r}\left(h_{b,(c-1)(n+d)}\right), \qquad (2)$$

where is the pooling layer's performance, n is the subsampling factor, r is the pooling scale, and n is the subsampling factor.

Flatten Layer:

In this case, we'll look at how to transform data into a 1D array and build a single long and narrow 1D feature vector using a flattened layer. Vectors can be flattened if needed. The single vector is also connected to the final classification model, resulting in a completely connected layer [28]. All pixel data is included in a single layer that is fully connected. The last steps of deep transfer learning are flattened and completely linked layers. It is converted to a one-dimensional array in order to prepare for the next completely connected layer of photo classification.

Fully Connected Layer:

Deep transfer learning makes considerable use of completely connected layers, which have shown to be quite beneficial in computer vision for image detection and identification. The deep transfer learning approach begins with convolution and pooling, which divides the image into properties and analyzes each one separately [29]. Each input is linked to all of the neurons in a fully connected layer, and the inputs are flattened. As a completely linked layer, the ReLU activation function is commonly employed. The output visuals in the last layer of the totally linked layer were estimated using the softmax activation function. The convolutional neural network architecture employs a fully linked layer. A convolutional neural network's last and most essential layers are these. A fully connected architecture is used in this design.

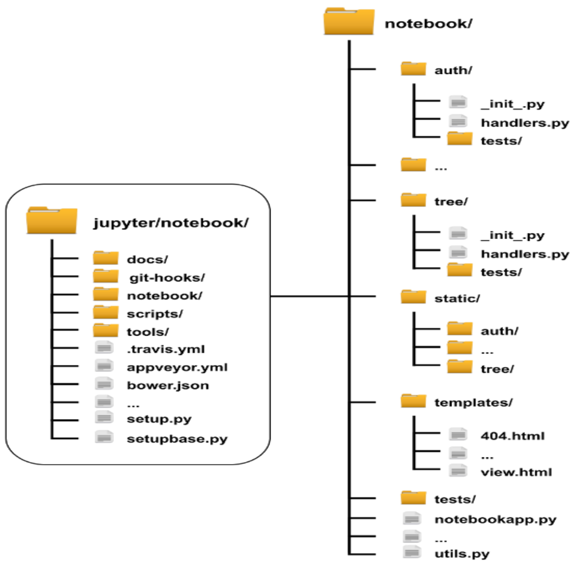
Pretrained Models:

One of the most serious problems that academics engaged in healthcare research face is a lack of medical records or data sets. Both time and money are required for data processing and labeling. The benefit of transfer learning is that it does not require a significant amount of data. Simulators are becoming easier to use. Transfer learning is a method for transferring a previously trained model from a big data set to a new model with less data to train. This process started with deep learning techniques being trained on a small data set for a specific job, then adding a large-scale data set that had already been taught in the training data set models [30].

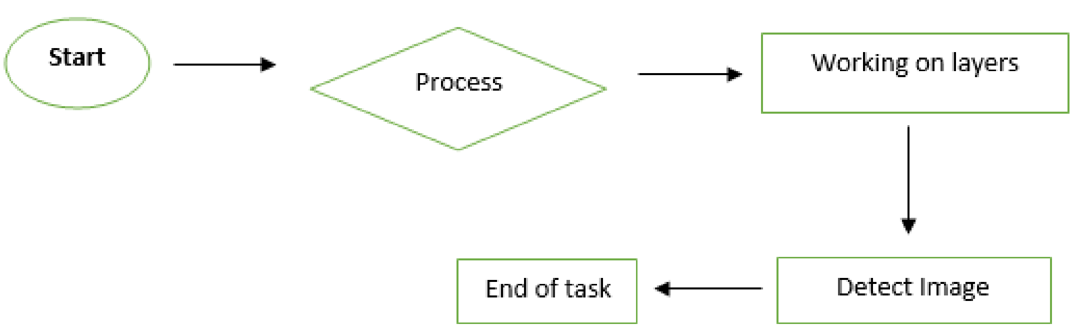
Three deep transfer learning-based pre-trained models were used to categorize CXR pictures in this study. MobileNetV2, VGG-16, and InceptionV3 were the models used. CXR pictures were separated into two groups. One is healthy, whereas the other is infected with SARS-CoV-2 and has pneumonia. This study also used a transfer learning method that uses ImageNet data to perform well with less data and is simple to train.

The architecture of the system is divided into four parts, as indicated in Figure 5. The first two processes are selecting and entering photographs, as well as applying the pre-trained model. Three pre-trained models are loaded in the second stage. In the third segment, the layers shown in Figure 5 were used to alter the pre-trained models. Finally, in the output portion, the findings will be shown as COVID-19 and pneumonia infected and uninfected patients.

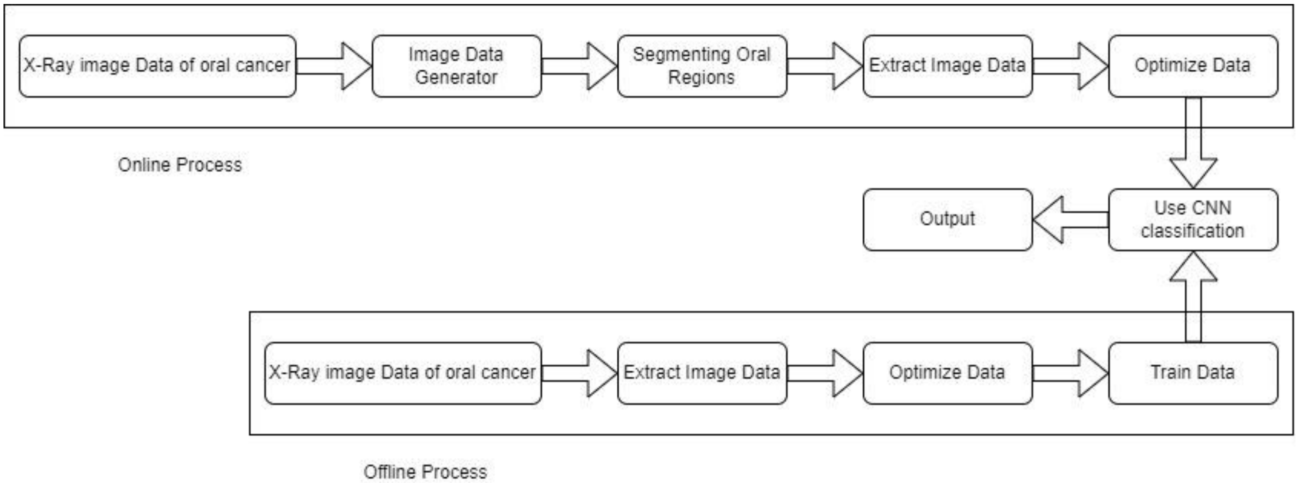
MobileNetV2 improves the cutting-edge efficiency of flexible models on a variety of assignments and seat stamps over a range of model sizes. MobileNetV2 works as a chain of n repeating layers in each line [31]. MobileNet uses depth-wise separable to factorize the regular state into depth-wise convolution. An 11-depth convolution, also known as a point-wise convolution, is also used, according to [32]. InceptionV3 was another pre-trained algorithm that was used. In most circumstances, it has the highest number of pooling layers. The ability of VGG16 to extract features using a little kernel at low concentrations makes it extremely valuable. A small kernel can successfully acquire the properties of CXR pictures.



Jupiter Notebook, a popular web-based IDE for image processing, machine learning, and deep learning, is what we're utilizing. For this project, we primarily use Python.



8. Tables Showing Input and Output Data, As Applicable:



9. Working Steps (Work plan):

A project work plan enables you to define a project's requirements, planning phases, objectives, and team members. This provides visibility to all parties involved, keeps project deliverables in one place, and keeps you on track to reach your objectives. There are some steps:

Background Study: The project background should provide information on why you want to carry out this particular project in this particular way. It must describe the existing situation and its difficulties, as well as the approach you intend to take to resolve them. These assumptions and explanations should be supported by reliable facts.

Proposal Writing: A project proposal is a written document that lays out everything stakeholders need to know about a project, including the schedule, budget, goals, and objectives. The facts of your project should be explained in your project proposal, and your idea should be sold so that stakeholders are interested in joining in the effort.

Present Proposal: The goal of the presentation is to give the evaluator an overview of your project, including both the product and the process. The talk is accompanied with project documents and a demonstration of the product (if any). It enables assessors to clear up any doubts they may have by, for example, asking questions on the spot.

Data Collection: Data collecting allows you to keep track of prior events so that we may look for recurrent trends using data analysis. You may create predictive models that search for trends and anticipate future changes based on those patterns. Because predictive models are only as strong as the data they're built on, good data collecting procedures are essential for creating high-performing models.

Train Data: Training data is the information used to teach an algorithm or machine learning model to predict the outcome you want it to. Test data is used to assess the performance of the algorithm you're using to train the machine, such as accuracy and efficiency.

Develop the System: The process of defining, creating, testing, and implementing a new software application or program is known as systems development. Internal development of custom systems could be part of it.

Testing: Software testing is the process of examining and verifying that a software product or application does what it is supposed to do. Testing has many benefits, including preventing flaws, saving development costs, and improving performance.

Report Writing: The final step is to write a project report. A project report is simply a document that comprises details about the project's overall status as well as specific areas of its progress or performance.

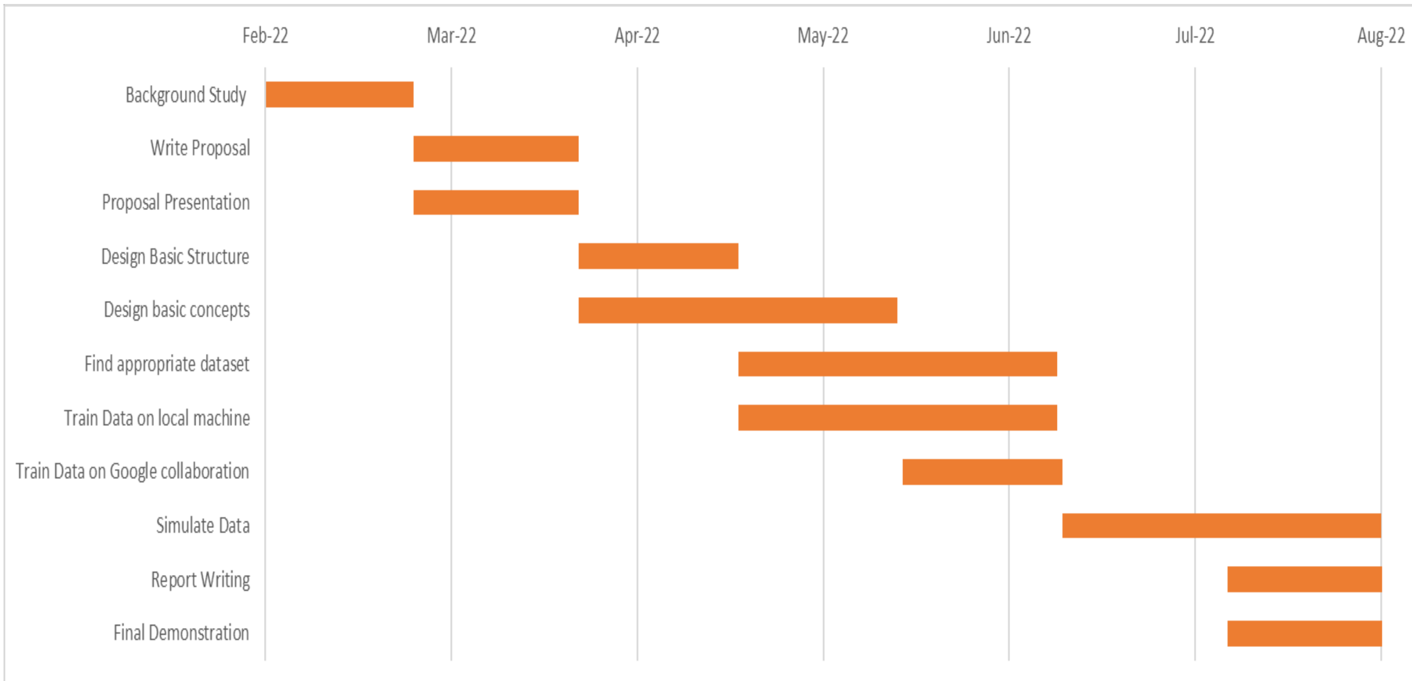
10. Major Milestones:

We've been studying the depths of our system for the past few weeks, learning new ideas and different types of algorithms. We finished our first report and gave our presentation this week.

11. Bill of Materials Required to Build the Circuit or Software System, And the Approximate Cost:

We need extended graphics card GTX2080:	138500 Tk
Need additional ram 16 GB:	22000 Tk
Google collaboration access to unlock new possibilities for **8 months**:	6880 Tk
Total Cost:	**167380 Tk**

12. Gantt Charts Showing the Expected Timeline of Progress or Milestones:



13. Required software tools. Detailed description. What tools and why?

Google collaboration:

The Google Colaboratory, or "Colab" for short, is a product from Google Research. Collab is a Python editor for the web that allows anyone to develop and run Python programs. Machine learning, data analysis, and education are all areas where it comes in handy. [33]

Professors can use Google Collaboration Applications to give new communication and collaboration options for their students.

Several of the many online applications geared toward efficiency and collaboration (Thompson, 2008) are appropriate for higher education, allowing undergraduates to learn how to use cloud computing software and prepare for the workforce. [34]

Jupiter notebook:

Jupiter Notebook is an open-source web tool that allows a user, scientist, scholar, or analyst to create and share a Notebook document that contains live codes, documentation, graphs, plots, and visualizations.

14. Target Population:

The people we're looking for are from the middle class. We all know that any form of oral therapy is expensive all over the world. Oral cancer can bring gloomy days in this condition. So, if we can detect it earlier and avoid it at an early stage, that would be ideal. It may be necessary to obtain x-ray scans of the mouth, therefore it is also cost effective.

15. What Makes the Solution An 'Innovation': (Description of innovative and creative aspects)

This is a breakthrough because it will offer a system that can detect oral cancer at an early stage, allowing patients to receive early treatment and maybe a better chance of survival.

16. How People Will Be Benefit from The Project?

People can be discovered early if they are diagnosed with oral cancer, since the system will allow them to receive an appropriate response.

17. What Are the Risks? How the Risks Will Be Managed?

The hazards include incorrect results or system errors, as well as output delays. To reduce the risk, test the system during the development process and maintain it updated.

18. Unprivileged Women and People Will Benefit?

Yes, unprivileged women and people will be benefited.

19. Disabled Will Benefit?

Yes, Disabled will be benefited.

20. What Is the Positive Impact on Environment, Social, Economic, Design?

We can accomplish a lot of good things with this effort. With early detection, we can give patients new hope for a better life. Early treatment may be able to save his life and riches. It will be beneficial to have a social life.

21. Research Publication on The Problem Statement:

- Deep Machine Learning for Oral Cancer: From Precise Diagnosis to Precision Medicine
  Authors: Rasheed Omobolaji Alabi, Alhadi Almangush, Mohammed Elmusrati and Antti A. Mäkitie

- Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer
  Authors: Roshan Alex Welikala, Paolo Remagnino, Jian Han Lim, Chee Seng Chan, Senthilmani Rajendran, Thomas George Kallarakkal, Rosnah Binti Zain, Ruwan Duminda Jayasinghe, Jyotsna Rimal, Alexander Ross Kerr, Rahmi Amtha, Karthikeya Patil, Wanninayake Mudiyanselage Tilakaratne, John Gibson.

- Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis
  Authors: Huiping Lin, Hanshen Chen, Luxi Weng, Jiaqi Shao, Jun Lin

- Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm
  Authors: Jeyaraj, P. R., & Samuel Nadar, E. R.

- Classifier Feature Fusion Using Deep Learning Model for Non-Invasive Detection of Oral Cancer from Hyperspectral Image
  Authors: Jeyaraj, P. R., Panigrahi, B. K., & Samuel Nadar, E. R.

- Improving Oral Cancer Outcomes with Imaging and Artificial Intelligence. Journal of Dental Research
  Authors: Ilhan, B., Lin, K., Guneri, P., & Wilder-Smith, P. (2020).

- Application and Performance of Artificial Intelligence Technology in Oral Cancer Diagnosis and Prediction of Prognosis: A Systematic Review
  Authors: Khanagar, Sanjeev B., Sachin Naik, Abdulaziz A. Al Kheraif, Satish Vishwanathaiah, Prabhadevi C. Maganur, Yaser Alhazmi, Shazia Mushtaq, Sachin C. Sarode, Gargi S. Sarode, Alessio Zanza, Luca Testarelli, and Shankargouda Patil

- Deep learning-based survival prediction of oral cancer patients

Authors: Kim, D.W., Lee, S., Kwon, S. et al.

- DEEPORCD: Detection of Oral Cancer using Deep Learning
  Authors: R. Dharani and S. Revathy

- Automated Detection and Classification of Oral Lesions Using Deep Learning to Detect Oral Potentially Malignant Disorders
  Authors: Tanriver, Gizem, Merva Soluk Tekkesin, and Onur Ergen

- An Ensemble Deep Neural Network Approach for Oral Cancer Screening
  Authors: B R, N., A, G., H S, C., M S, D. & S, M.

22. Conclusion:

We'll be working on a technique that will help detect mouth cancer early. For this, the CNN model will be employed. Other algorithms, such as pool, flatten, and others, could be used as well. We'll be able to observe if any tissue damage has occurred, as well as whether the tongue has been spotted. For this, we'll need massive datasets that we can segment and train effectively. After successful development, it will produce findings with an accuracy of up to 98 percent.

23. Bibliography:

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries", *CA: A Cancer J. Clinicians*, vol. 68, pp. 394-424, Nov. 2018.

[2] Oral cancer, https://en.wikipedia.org/wiki/Oral_cancer.

[3] J. Rimal, A. Shrestha, I. K. Maharjan, S. Shrestha and P. Shah, "Risk assessment of smokeless tobacco among oral precancer and cancer patients in eastern developmental region of nepal", *Asian Pacific J. Cancer Prevention*, vol. 20, no. 2, pp. 411-415, Feb. 2019.

[4] J. G. Doss, W. M. Thomson, B. K. Drummond and R. J. R. Latifah, "Validity of the FACT-H&N (v 4.0) among Malaysian oral cancer patients", *Oral Oncol.*, vol. 47, no. 7, pp. 648-652, 2011.

[5] B R, N., A, G., H S, C., M S, D. & S, M. (2021). An Ensemble Deep Neural Network Approach for Oral Cancer Screening. International Association of Online Engineering. Retrieved April 1, 2022

[6] Huiping Lin, Hanshen Chen, Luxi Weng, Jiaqi Shao, Jun Lin, Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis, J. Biomed. Opt. 26(8) 086007 (28 August 2021)

[7] Ilhan, B., Lin, K., Guneri, P., & Wilder-Smith, P. (2020). Improving Oral Cancer Outcomes with Imaging and Artificial Intelligence. Journal of Dental Research, 99(3), 241–248. doi10.11770022034520902128

[8] Jeyaraj PR, Samuel Nadar ER. Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm. J Cancer Res Clin Oncol. 2019 Apr;145(4):829-837. doi: 10.1007/s00432-018-02834-7. Epub 2019 Jan 3. PMID: 30603908.

[9] R. A. Welikala et al., "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer," in IEEE Access, vol. 8, pp. 132677-132693, 2020, doi: 10.1109/ACCESS.2020.3010180.

[10] Khanagar, Sanjeev B., Sachin Naik, Abdulaziz A. Al Kheraif, Satish Vishwanathaiah, Prabhadevi C. Maganur, Yaser Alhazmi, Shazia Mushtaq, Sachin C. Sarode, Gargi S. Sarode, Alessio Zanza, Luca Testarelli, and Shankargouda Patil. 2021. "Application and Performance of Artificial Intelligence Technology in Oral Cancer Diagnosis and Prediction of Prognosis: A Systematic Review" Diagnostics 11, no. 6: 1004. https://doi.org/10.3390/diagnostics11061004

[11] Kim, D.W., Lee, S., Kwon, S. et al. Deep learning-based survival prediction of oral cancer patients. Sci Rep 9, 6994 (2019). https://doi.org/10.1038/s41598-019-43372-7

[12] Alabi, R. O., Almangush, A., Elmusrati, M., & Mäkitie, A. A. (2022). Deep Machine Learning for Oral Cancer: From Precise Diagnosis to Precision Medicine. Frontiers in oral health, 2, 794248. https://doi.org/10.3389/froh.2021.794248

[13] Pandia Rajan Jeyaraj, Bijaya Ketan Panigrahi & Edward Rajan Samuel Nadar (2020) Classifier Feature Fusion Using Deep Learning Model for Non-Invasive Detection of Oral Cancer from Hyperspectral Image, IETE Journal of Research, DOI: 10.1080/03772063.2020.1786471

[14] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks", *Proc. Adv. Neural Inf. Process. Syst.*, pp. 91-99, 2015.

[15] K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN", *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 2961-2969, Oct. 2017.

[16] R. Dharani and S. Revathy 2021 J. Phys.: Conf. Ser. 1911 012006

[17] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database", *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 248-255, Jun. 2009.

[18] Kim, K., Li, S. & Cha, I. Nomogram for predicting survival for oral squamous cell carcinoma. *Genomics Inform.* **8**, 212–218 (2010).

[19] Wang, S. J. *et al*. An oral cavity carcinoma nomogram to predict benefit of adjuvant radiotherapy. *JAMA Otolaryngol. - Head Neck Surg.* **139**, 554–559 (2013).

[20] Saintigny, P. *et al*. Gene expression profiling predicts the development of oral cancer. *Cancer Prev. Res.* **4**, 218–229 (2011).

[21] Chang, S.-W., Abdul-Kareem, S., Merican, A. & Zain, R. Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods. *BMC Bioinformatics* **14**, 170 (2013).

[22] Missinglink ,(2019) https://missinglink.ai/guides/neural-network-concepts/convolutiona;-neural-network-build-one-keras-pytorch/.

[23] Ü. H. Ayan, "Diagnosis of pneumonia from chest X-ray images using deep learning," in Proceedings of the 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), pp. 1–5, Istanbul, Turkey, April 2019.

[24] J. Brownlee, Difference between a Batch and an Epoch in a Neural Network, Machine Learning Mastery, San Francisco, CL, USA, 2021, https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/.

[25] Towardsdatascience, "The most intuitive and eauest guide for CNN," 2020, https://towardsdatascience.com/the-most-intuitive-and-easiest-guide-for-convolutionalneuralnetwork3607be47480#:%7E:text=Flattening%20is%20converting%20the%20data, called%20a%20fully%2Dconnected%20layer.

[26] O. A. Hamid, L. Deng, and D. Yu, "Exploring convolutional neural network structures and optimization techniques for speech recognition," ISCA, vol. 11, pp. 73–75, 2013.

[27] J. Brownlee, "A gentle introduction to pooling layers for convolutional neural networks," Machine Learning Mastery, 2021.

[28] J. Jeong, "The Most Intuitive and Easiest Guide for CNN," Medium, 2021.View at: Google Scholar

[29] S. Saha, "A Comprehensive Guide to Convolutional Neural Networks—theELI5 Way," Medium, 2021.View at: Google Scholar

[30] Tanriver G, Soluk Tekkesin M, Ergen O. Automated Detection and Classification of Oral Lesions Using Deep Learning to Detect Oral Potentially Malignant Disorders. Cancers (Basel). 2021 Jun 2;13(11):2766. doi: 10.3390/cancers13112766. PMID: 34199471; PMCID: PMC8199603.

[31] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and C. L. Chieh, "MobileNetV2: inverted residuals and linear bottlenecks," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4510–4520, 2018.View at: Publisher Site | Google Scholar

[32] A. G. Howard, M. Zhu, B. Chen et al., "MobileNets: Efficient Convolutional Neural Networks for mobile Vision Applications," 2017.

[33]"GoogleColab".2022.*Research.Google.Com*.https://research.google.com/colaboratory/faq.html.

[34]Edwards, Jennifer T., and Credence Baker. "A case study: Google collaboration applications as online course teaching tools." *MERLOT Journal of Online Learning and Teaching* 6, no. 4 (2010): 828-838

[35] "Why You Should Be Using A Jupyter Notebook - Elucidata". 2022. *Elucidata*. https://elucidata.io/why-you-should-be-using-a-jupyter-notebook