International Research Journal of Education and Technology

Peer Reviewed Journal

ISSN 2581-7795

Performance of Artificial Intelligence Technology in Oral Cancer Dignosis

Ashvini Mali[1], S.A. Gaikwad[2], Mahesh Waghmare[3]

[1]1M-Tech Student Dept. of Computer Science & Engineering College of Engg. Dharashiv
[2]Professor, Dept. of Computer Science & Engineering, College of Engg. *Dharashiv*
[3]Assi. Professor Dept.of Soil Science & Agricultural Chem. Govt. College of Agricuture, Dharashiv

---***---

Abstract - *Oral cancer is a dangerous and extensive cancer with a high death ratio. Oral cancer is the leading cause of death in the world. In 2022, annual oral cancer cases grew to 377713 people and 177757 deaths in 2020 worldwide. The cancerous tumor appears in the neck, oral glands, face, and mouth.The most common types of oral cancer are lymphoma, Minor salivary gland, mucosal melanoma, sarcomas cancer and squamous cell carcinoma. Early detection of oral cancer plays a vital role in survival of the patient. Imaging tests are used most for identifying the cancer. Imaging tests used in diagnosing oral cancer may include a computerised tomography (CT) scan, bone scan, magnetic resonance imaging (MRI), positron emission tomography (PET) scan, ultrasound and X-ray. AI medical diagnosis is becoming effective day by day. By this approach patients can get their oral cancer scan reports faster. In this project work we are introducing AI assisted web applications which can predict oral cancer. With help of CNN (Convolutional neural networks) deep learning model we are going to detect some types of oral cancer. Apart from deep learning a user friendly interface using react and API using Flask (a python backend framework) is used. The web app is built to add or update existing models for future scenarios with ease. This project is open sourced to help those who are in need.*

Key Words: AI, Oral Cancer, CNN etc.

1. INTRODUCTION

Oral cancer is the most dangerous type of cancer. It is caused in various parts such as lips, tongue, hard and soft palate and floor of mouth. Oral cancer is caused due to tobacco use of any kind, including cigars, pipes, chewing tobacco, snuffs, or alcohol consumption. Its detection and diagnosis is very important or else it can be fatal. Therefore early diagnosis of a cancer type has become a necessity in cancer research. The ability of AI tools to detect key features from composite datasets reveals their significance. The predictive models discussed here are based on various supervised AI techniques as well as on different input features and data sample. Given the rapidly growing trend on the application of AI methods in cancer research, we present here the most recent publications that employ these techniques as an aim to model oral cancer image classification or patient outcomes.

Oral cancer has a high incidence and fatality rate, making it a leading cancer killer. There were an expected 457,000 new instances of lip and oral cavity cancer in 2023, with almost 177,000 fatalities globally, as reported by the International Agency for Research on Cancer.

The most common method currently used to treat oral cancer and its precursor condition is the conventional oral examination (COE), which includes a visual and tactile assessment (and, if necessary, a tissue biopsy). However, general dentists may be unable to distinguish between oral cancer and benign conditions such as aphthous ulcers because of the clinical heterogeneity and subtlety of oral cancer symptoms. Second, biopsy is not always the best screening technique due to its invasive nature and sampling bias, which can result in underdiagnosis or misdiagnosis. In addition, the oral cancer burden disproportionately affects low-resource communities even though specialists can recognize the majority of characteristics that differentiate benign and cancerous lesions. As a result, developing a low-cost screening technique to supplement existing practices is gaining traction.

In recent years, deep learning algorithms have shown superior to feature-based approaches in medical picture analysis. In numerous disease recognition scenarios, deep learning algorithms have been shown to outperform human specialists, as demonstrated by several studies. Automatic analysis of pathology, confocal laser endomicroscopy (CLE), and fluorescence pictures all showed promise for use with deep learning in the detection of oral cancer. To classify oral cancer stages in histo-logical pictures, Kumar et al. presented a two-step process, including a segmentation network and a random forest tree classifier.

The deep convolutional neural network (CNN) performed better than the feature-based classification approaches when Aubreville et al. evaluated it for OSCC diagnosis using CLE images. To detect oral disease, Song et al. created mobile-connected devices to gather fluorescence oral images. These techniques, however, rely upon high-priced equipment or a specialized screening platform that only some have access to. In other words, getting a diagnosis for an illness still necessitates a trip to a doctor's office. As image and sensing technology in cameras continue to advance rapidly, more and more smartphone manufacturers are including high-resolution, low-noise, and speedy camera modules.

1.1 Deep Learning

Deep learning (DL), a branch of Artificial Intelligence, is a family of multi-layer neural network models that excel at the problem of learning from big data. Similar to other machine learning methods, DL consists of the training step where the

ISSN 2581-7795

estimation of network parameters from a given training dataset is carried out, and the testing step that utilises the trained network to predict outputs of new input data. The accumulation of whole transcriptomic profiling of tumour samples enabled the pursuit of the DL model for better accuracy and innovative interpretability for cancer type prediction.

1.2 CNN (Convolutional Neural Network)

CNN also known as ConvNets, consist of multiple layers and are mainly used for image processing and object detection. Yann LeCun developed the first CNN in 1988 when it was called LeNet. It was used for recognizing characters like ZIP codes and digits.

CNNs are widely used to identify satellite images, process medical images, forecast time series, and detect anomalies.A neural network is a mathematical model based on neural units-artificial neurons-similar to biological neural networks. Typically, neurons are organised in layers and the connections are set up between neurons only from adjacent layers. The Input feature vector is put into the first layer and, moving from layer to layer, is transformed to a high level features vector. The output layer neurons finallyleads to classification and their number is equal to the number of classifying classes.Each convolutional layer consists of a set of trainable filters and computes dot production between these filters and layer input to obtain an activation map.Filters are also called kernels and allow detecting same features in different locations. The activation function used is RELU function and sigmoid function.

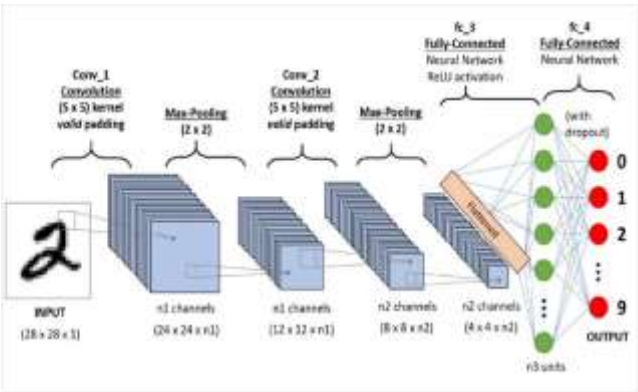

Fig -1: CNN workflow for mnist data

CNN takes an image as an input, distinguishes its objects based on three colour planes, and identifies various colour spaces. It also measures the image dimensions. In order to explain this process, we will give an example of an RGB image given above.

In this image, various colours are based on the three-colour plane that is Red, Green, and Blue, also known as RGB. The various colour spaces are then identified in which images are found, such as RGB, CMYK, Grayscale, and many more. It can become a tedious task while measuring the image dimensions as an example if the image is perse 8k (*7680x4320*). Here comes one of the handy capabilities of CNN: it reduces the image's dimension to the point that it is easier to process, while also maintaining all of its features in one piece. This is done so that a better prediction is obtained. This ability is critical when designing architectures having not only better learning features but also can work on massive datasets of images.

1.2.1 Convolution Layer (Kernel):

The Kernel of CNN works on the basis of the following formula. Image Dimensions = n1 x n2 x 1 where n1 = height, n2 = breadth, and 1 = Number of channels such as RGB. So, as an example, the formula will become I D = 5 x 5 x 1. We will explain this using the image given below.

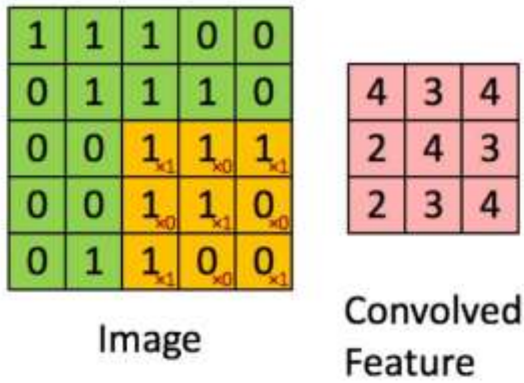

Fig -2: Kernel convolution to find the feature in image

In this image, the green section shows the 5 x 5 x 1 formula. The yellow box evolves from the first box till last, performing the convolutional operation on every 3x3 matrix. This operation is called Kernel (K) and works on the basis of the following binary algorithm.

K =		
1	**0**	**1**
0	**1**	**0**
1	**0**	**1**

Fig -3: Binary algorithm for kernel

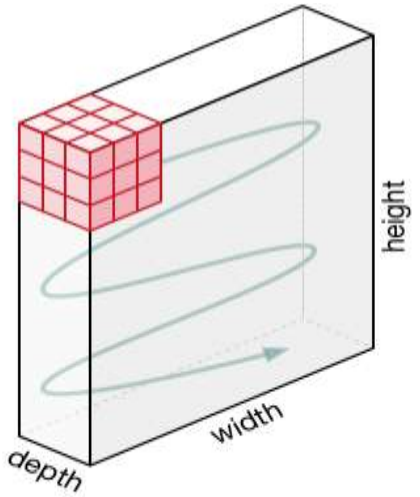

Fig -4: Stride Movements

In the above figure, the Kernel moves to the right with a defined value for "Stride." Along the way, it parses the image objects until it completes the breadth. Then it hops down to the second row on the left and moves just as in the top row till it covers the whole image. The process keeps repeating until every part of the image is parsed. If there are multiple channels such as found in RGB images, then the kernel contains the same depth as found in the input image.

1.2.2 Types of Pooling

There are mainly two different types of Pooling which are as follows:

Max Pooling: The Max Pooling basically provides the maximum value within the covered image by the Kernel.

Average Pooling: The Average Pooling provides and returns the average value within the covered image by the Kernel.

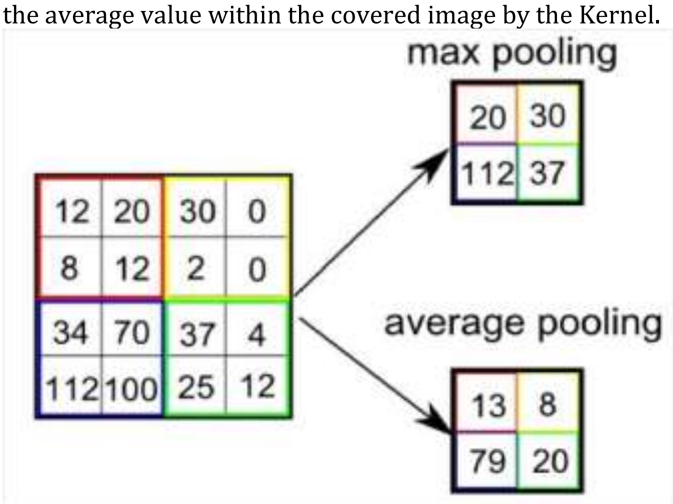

Fig -5: Difference between Max and average pooling

CNN's have multiple layers that process and extract features from data:

Convolution Layer

- CNN has a convolution layer that has several filters to perform the convolution operation. Rectified Linear Unit (ReLU)
  - CNN's have a ReLU layer to perform operations on elements. The output is a rectified feature map.

Pooling Layer

- The rectified feature map next feeds into a pooling layer. Pooling is adown-sampling operation that reduces the dimensions of the feature map.
- The pooling layer then converts the resulting two-dimensional arrays from the pooled feature map into a single, long, continuous, linear vector by flattening it.

Fully Connected Layer

- A fully connected layer forms when the flattened matrix from the pooling layer is fed as an input, which classifies and identifies the images.

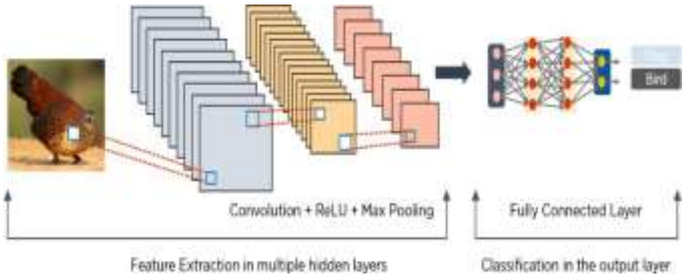

Fig -6: CNN Workflow

2. Website Development

Web development refers to the building, creating, and maintaining of websites. It includes aspects such as web design, web publishing, web programming, and database management. It is the creation of an application that works over the internet i.e. websites.

It can classify into:

- Frontend development
- Backend development
- Deployment

2.1 ReactJS

React is a popular JavaScript library for building user interfaces. Developed by Facebook, React is a powerful and efficient tool for creating fast, interactive and scalable web applications. React is based on the concept of components, which are small, self-contained units of code that can be reused throughout an application. React allows developers to create complex user interfaces by breaking them down into smaller, more manageable components. One of the key features of React is its virtual DOM (Document Object Model). The virtual DOM is a lightweight representation of the actual DOM that React uses to update the user interface. When a user interacts with a React application, React updates the virtual DOM instead of the actual DOM. React then compares the virtual DOM with the actual DOM to determine what needs to be changed, and updates the actual DOM accordingly. Another key feature of React is its focus on declarative programming. With React, developers can define what they want the application to look like, and React takes care of the details of how it should be implemented. This approach makes it easier to reason about the code and to create reusable components. React is also known for its strong community and ecosystem. There are many resources available online for learning React, including documentation,

ISSN 2581-7795

tutorials, and forums. The React community is also constantly creating new libraries and tools to extend the functionality of React and make it easier to build complex applications.

2.2 Firebase

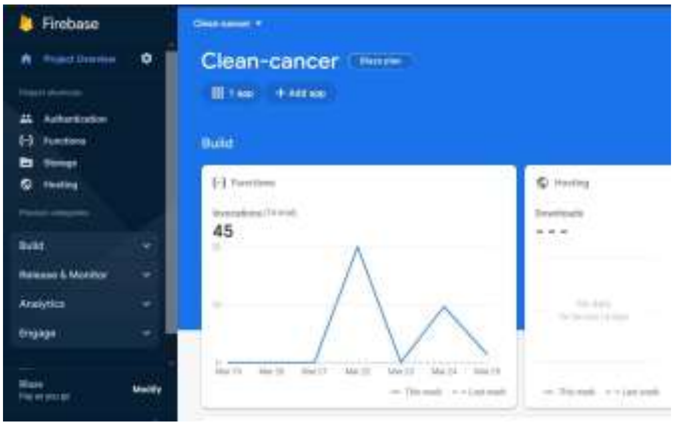

Fig -6: Firebase

Firebase is a comprehensive mobile and web application development platform that provides a wide range of tools and services for building high-quality and feature-rich applications. Developed by Google, Firebase provides a suite of backend services that can be integrated into any application, allowing developers to focus on creating great user experiences without worrying about the infrastructure behind the scenes.

One of the key features of Firebase is its real-time database. This database is a NoSQL cloud-hosted solution that allows developers to store and sync data in real-time. This means that any changes made to the database are immediately propagated to all connected clients, ensuring that all users have access to the most up-to-date information. This real-time feature makes Firebase an excellent choice for building real-time applications such as chat apps, collaboration tools, and multiplayer games. Firebase also includes a comprehensive authentication system that supports various sign-in methods, such as email and password, social media accounts, and phone numbers. This authentication system enables developers to add user authentication to their applications quickly, securely, and with minimal effort. Firebase's cloud messaging service is another powerful feature that enables developers to send push notifications to their users on various platforms, including Android, iOS, and web. The cloud messaging service can be used to send targeted notifications based on user behaviour or to broadcast messages to all users of an application.

Firebase also provides cloud storage that allows developers to store and retrieve user-generated content such as images, videos, and audio files. This feature provides developers with a scalable and secure way to store and serve user-generated content without the need for additional infrastructure.

2.3 Tailwind CSS

Tailwind CSS is a popular utility-first CSS framework for building responsive and customizable user interfaces. It provides a comprehensive set of pre-designed CSS classes that developers can use to style HTML elements quickly and easily. Tailwind is designed to be highly customizable, with an emphasis on utility classes rather than pre-designed components, giving developers complete control over the look and feel of their applications.

2.4 NodeJS

Node.js is an open-source, cross-platform, back-end JavaScript runtime environment built on the V8 engine. It allows developers to use JavaScript to build server-side applications that can be run on a variety of platforms, including Windows, macOS, and Linux. One of the key benefits of Node.js is its performance. Node.js is designed to be fast and efficient, with a non-blocking, event-driven architecture that allows it to handle large numbers of simultaneous connections. This makes it an ideal choice for building real-time applications such as chat applications, multiplayer games, and financial applications.

Node.js also has a large and active community, with a wide range of libraries and modules available for developers to use. These libraries and modules can be easily installed using the Node Package Manager (NPM), which is built into Node.js. This makes it easy for developers to add functionality to their applications without having to write all the code from scratch.

Node.js also includes several built-in modules that make it easy to work with common web technologies such as HTTP, WebSocket, and TCP. These modules can be used to build web servers, web applications, and APIs.

```
const http = require('http');

const hostname = '127.0.0.1';
const port = 3000;

const server = http.createServer((req, res) => {
  res.statusCode = 200;
  res.setHeader('Content-Type', 'text/plain');
  res.end('Hello World');
});

server.listen(port, hostname, () => {
  console.log(`Server running at http://${hostname}:${port}/`);
});
```

Node.js is a powerful and versatile back-end JavaScript runtime environment that is used by many companies and developers to build fast, scalable, and efficient web applications. Its non-blocking, event-driven architecture, scalability, and large community make it an excellent choice for building real-time and high-traffic applications. Additionally, its built-in modules and NPM make it easy to work with common web technologies and add functionality to applications.

3. Emergence of AI in Medical Field

International Research Journal of Education and Technology

Peer Reviewed Journal

ISSN 2581-7795

The first use of AI (artificial intelligence)-based systems in digitalized diagnostic pathology has shown that the diagnostic process and treatment plan can be improved when AI models work with pathologists to find abnormalities more accurately and better manage patients.There is evidence that deep learning models, especially neural network models, can help find oral cancer early. This makes AI a very useful tool for helping doctors get better at diagnosing and reducing mistakes made by humans when analyzing X-rays.

Over the years, researchers have focused on refining AI models for greater accuracy and reliability. This has led to the introduction of more sophisticated ML approaches that can interpret complex medical and dental data to aid in early detection, diagnose more accurately, and predict patient outcomes better. AI-powered tools can now offer a supplementary layer of analysis to traditional examinations, which is increasingly valuable in areas with limited access to specialist care.

One of the key technological advancements in dental AI and ML is the development of Convolutional Neural Networks (CNNs). These deep learning models are particularly well-suited for image recognition and have been repurposed for the analysis of medical and dental images. The ability of CNNs to learn from a large volume of data makes them effective for identifying patterns and anomalies in dental radiographs and visible light images indicative of oral cancer. In addition to CNNs, advances in digital imaging and the application of AI for real-time risk assessment have further elevated the potential for AI in oral cancer detection. Techniques such as computer-aided detection systems have been developed to assist in the analysis of oral lesions, employing algorithms to help delineate suspicious areas from normal tissue.

4. Methodology

4.1 Artificial Intelligence

Artificial intelligence techniques are gaining attention as a means of improving image-based diagnosis. Machine learning and deep learning are 2 subsets of AI, which, although the terms are sometimes used interchangeably, have some important differences. Briefly, machine learning algorithms typically require an accurately categorized data input, whereas deep learning networks rely on layers of the artificial neural networks to generate their own categories based on identifying edges (differences) within layers of neural networks when exposed to a huge number of data points. Therefore, although both of these subsets of AI are "intelligent," deep learning requires much more data than a traditional machine learning algorithm, while machine learning performs better with fewer data sets that are clearly labeled or structured with regard to a gold standard or specific criteria of interest.

Briefly, there are 3 main steps to applying AI to medical imaging: preprocessing, image segmentation, and postprocessing.

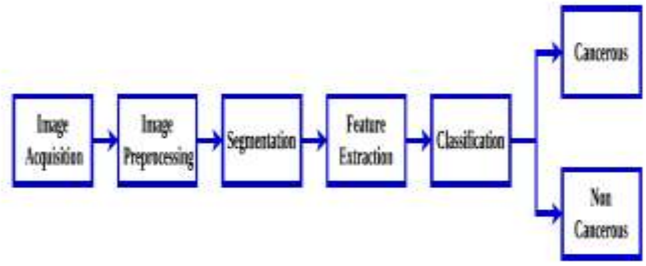

Fig -6: Steps in the Detection of Oral Cancer

1 **Image acquisition**

Image acquisition is the process of capturing visual information from the real world and converting it into a digital image that computers can process.

2 **Image Preprocessing**

Digital Image Processing (DIP) is the process of applying an algorithm on digital images in the computer system. It is a sub field of signals and systems focused on images. DIP is a technique to perform operations on a digital image to get an enhanced image or extract some useful information. Image processing is used in various applications such as processing color, video and so on.

The first step is to collect an image from various sources such as

- Microscopic image (Biopsy sample observed in microscope and connected to a computer to get a digital image)
- X-ray image
- MRI image
- CT Scan image
- PET Scan image
- Color image (captured from mobile phone)

3 **Image Segmentation**

This process recognizes and delineates the region of interest. For cancer imaging, pathological areas of the lesion are distinguished from nonpathological sites. While segmentation can be divided into 4 main classes, there exist a host of different approaches to this process, and often hybrid models combining multiple techniques have been used to improve accuracy.

4 **Postprocessing**

Multiple postprocessing methods exist, whose primary function is to target and extract information on features of interest such as islands, borders, and regions that share the same defined properties). Many different postprocessing techniques are being applied to medical imaging, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), multiscale convolutional neural network (M-CNN), and multiinstance learning convolutional neural network (MIL-CNN). Neural networks can perform complex computational tasks because of the nonlinear processing capabilities of their neurons.

5. Data Collection & Implementation Details

5.1 Data Collection

International Research Journal of Education and Technology

Peer Reviewed Journal

ISSN 2581-7795

Data is collected from the kaggle website.The name of the data is Skin Cancer MNIST: HAM10000

Glossary of Terms:

Nv Melanocytic nevi are benign neoplasms of melanocytes and appear in a myriad of variants, which all are included in our series.

Mel Melanoma is a malignant neoplasm derived from melanocytes that may appear in different variants.

Bkl "Benign keratosis" is a generic class that includes seborrheic keratoses ("senile wart"), solar lentigo - which can be regarded a flat variant of seborrheic keratosis - and lichen-planus like keratosis (LPLK), which corresponds to a seborrheic keratosis or a solar lentigo with inflammation and regression.

Bcc Basal cell carcinoma is a common variant of epithelial skin cancer that rarely metastasizes but grows destructively if untreated.

Akiec Actinic Keratoses (Solar Keratoses) and intraepithelial Carcinoma (Bowen's disease) are common, non-invasive, variants of squamous cell carcinoma that can be treated locally without surgery. Some authors regard them as precursors of squamous cell carcinomas and not as actual carcinomas.

5.2 Implementation Details

EDA (Exploratory data analysis) is important for training of the model

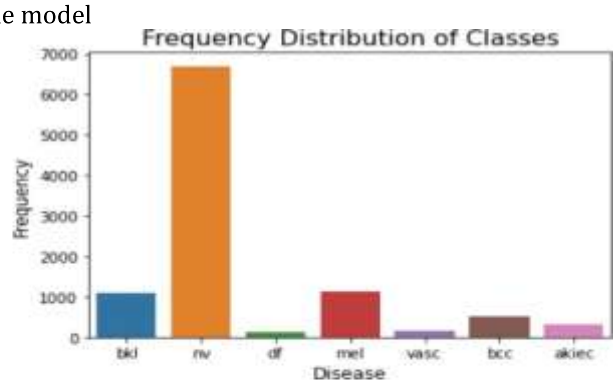

Fig -6: Frequency Distribution of classes

5.3 Data Augmentation

Contemporary advances in deep learning models have been largely associated with large quantities and diversity of data. Having large data helps crucially in improving the performance of machine learning models. Hence, we use the technique of Data Augmentation. It is a technique that enables us to considerably increment the diversity and quantity of data available, without actually aggregating new data.

In this project, the training images are augmented in order to make the model robust to new data which in turn helps in increasing the testing accuracy.

Fig-7 shows the augmentation techniques used on the dataset.

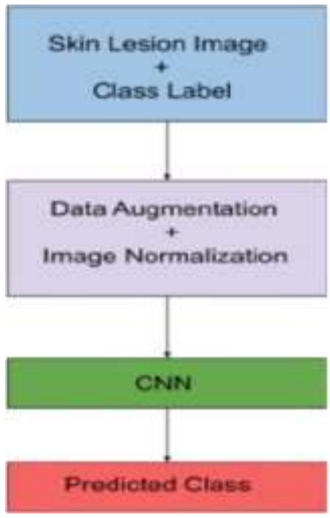

Fig -7: Data Augmentation Steps

5.4 Building an CNN Model

Now we will start developing a convolutional neural network to classify images for correct labels. CNN is best to work with image data.

The architecture of our CNN model

- Conv2D layer – we will add 5 convolutional layers of 32 filters, size of 5*5, and activation as Relu
- Max Pooling – MaxPool2D with 2*2 layers.
- Dropout with a rate of (0.25).
- Flatten layer to squeeze the layers.
- Dense layer(nodes=9, activation="softmax")
- MaxPool2D – Maximum pooling layer is used to reduce the size of images
- Dropout – Dropout is a regularisation technique to reduce overfitting
- Flatten – to convert the parallel layers to squeeze the layers
- Dense – for feed-forward neural network

The last layer will have an activation function as softmax for Multi-class classification.

Fig -8: Builded CNN MODEL (with 5 convolution layer)

3. CONCLUSIONS

1 This study discussed the various techniques to detect an oral cancer at an early stage. It has given an insight of the different types of inputs that are

used for oral cancer detection. It is observed that the histopathology images are much used and also helpful in analyzing and providing the accurate results.

2 A complete collection of different AI techniques are presented with different algorithms like deep learning, machine learning, data mining, genetic algorithms and fuzzy computing. The advantages and disadvantages of all the algorithms used in the study is analyzed.

3 The important AI technique used in recent articles is the deep learning algorithms and it also provides the best detection accuracy that is above 90% and also reduced error rate.

4 An important performance parameter such as accuracy of detecting an oral cancer is specified for all the research articles. Around 60% of the research articles referred in this survey is implemented using MATLAB tool and remaining are implemented using Python, java, data mining tools such as WEKA3.7.9, DTREG. To summarize, artificial intelligence techniques provide a better result as the detection accuracy is above 90% in most of the methods.

5 As the medical procedures like scanning and biopsy involves more cost, the computer algorithms can provide a result with an image taken from a mobile phone. In this way, technology can be considered as a boon in the medical field and can greatly help to save peoples' life.

REFERENCES

1. D. Ilhan B, Guneri P, Wilder-Smith P. The contribution of artificial intelligence to reducing the diagnostic delay in oral cancer. Oral oncol 2021; 116:105254.
2. Khanagar SB, Naik S, Al Kheraif AA, et al. Application and performance of artificial intelligence technology in oral cancer diagnosis and prediction of prognosis: a systematic review. Diagnostics 2021; 11:1004.
3. Kim JS, Kim BG, Hwang SH. Efficacy of artificial intelligence-assisted discrimination of oral cancerous lesions from normal mucosa based on the oral mucosal image: a systematic review and meta-analysis. Cancers 2022; 14:3499.
4. Panigrahi S, Swarnkar T. Machine learning techniques used for the histopathological image analysis of oral cancer-a review. Bioinformatics J 2020; 13.
5. Patil S, Albogami S, Hosmani J, et al. Artificial intelligence in the diagnosis of oral diseases: applications and pitfalls. Diagnostics 2022; 12:1029.
6. Pereira-Prado V, Martins-Silveira F, Sicco E, et al. Artificial intelligence for image analysis in oral squamous cell carcinoma: a review. Diagnostics 2023; 13:2416.
7. Licheng J, Jin Z. A Survey on the New Generation of Deep Learning in Image Processing.IEEE Access 2019; 7: 172231- 172263.
8. Daisuke K, Shumpei I. Machine Learning Methods for Histopathological Image Analysis.Compt and Struct Biotec J, 2020; 16: 34-42.
9. Anne H-H. Texture Feature Extraction Methods: A Survey. Journal of IEEE Access, 2019; 7:8975-8999.
10. Jie C, Jiawei L, Shulin W, Sheng Y. Feature selection in machine learning: A new perspective. Neurocomputing Journal Elsevier, 2018; 300: 70-79.
11. Shruti DP, Wayakule JM, Apurva DK. Skin Segmentation Using YCbCr and RGB Color Models. Int J of Adv Res Comp Sci and Soft Eng, 2014; 4(7).
12. Harikumar Rajaguru, Sunil Kumar Prabhakar, "Performance Comparison of Oral Cancer Classification with Gaussian Mixture Measures and Multi-Layer Perceptron", the 16th International Conference on Biomedical Engineering, 30 June 2017, pp 123-129
13. K.P. Exarchos, Y. Goletsis, D.I. Fotiadis "Multiparametric decision support system for the prediction of oral cancer reoccurrence", IEEE Trans Inf Technol Biomed, 16 (2012), pp. 1127-1134
14. Pandia Rajan Jeyaraj, Edward Rajan Samuel Nadar, "Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm", Journal of Cancer Research and Clinical Oncology 145, pages829–837(2019)
15. R Rajkumar, Dr. M V Sudhamani and Head Department of ISE, RNSIT Bengaluru, Karnataka, India, "Crawler for Image Acquisition from World Wide Web", International Journal of Engineering Trends and Technology (IJETT) –Special Issue-May 2017ISSN: 2231-5381