



**IJITCE**

**ISSN 2347- 3657**

# International Journal of Information Technology & Computer Engineering

[www.ijitce.com](http://www.ijitce.com)



Email : [ijitce.editor@gmail.com](mailto:ijitce.editor@gmail.com) or [editor@ijitce.com](mailto:editor@ijitce.com)

# FACIAL PARALYSIS SEVERITY DETECTION USING TRANSFER LEARNING AND ARTIFICIAL INTELLIGENCE

Paradesi Subba Rao<sup>a)</sup>, N. Rama Devi<sup>b)</sup>, Palagulla Lakshmi Priya<sup>c)</sup>,

Kristipati Akhila, Myla Venkata Ruchitha, Shaik Ayesha, Shaik karimun

Department Of Computer Science and Engineering-Data Science, Santhiram Engineering College, Nandyal, Andhra Pradesh, 518501, India.

<sup>a)</sup>Corresponding author: [subbarao.cse@srecnandyal.edu.in](mailto:subbarao.cse@srecnandyal.edu.in),

<sup>b)</sup>[ramadevi.cse@srecnandyal.edu.in](mailto:ramadevi.cse@srecnandyal.edu.in),

<sup>c)</sup>[20x51a3238@srecnandyal.edu.in](mailto:20x51a3238@srecnandyal.edu.in).

## ABSTRACT

This project introduces a novel deep learning based approach to detecting facial paralysis, a condition that affects millions worldwide but often goes undiagnosed in its early stages. Leveraging the InceptionV3 algorithm for its robustness and efficiency in real-time object detection, we have developed a model that accurately identifies signs of facial paralysis from both static images and live video feeds. The dataset, sourced from Kaggle's official website, comprises 3,000 images categorized into two classes, enabling the model to learn diverse representations of the condition. Our model achieved an exceptional accuracy of 95.3% during testing, underscoring its potential as a reliable diagnostic tool.

The model was trained and validated using Google Collab, benefiting from its powerful computational resources and collaborative environment. Subsequent deployment in a web application provides an accessible platform for users to register and obtain authorization from an administrator. Upon login, users can upload images or access a live camera feed to detect facial paralysis, facilitating early diagnosis and intervention. The web application also features an administrative module responsible for user authentication, model training, and performance analysis, ensuring the system's integrity and effectiveness.

This project not only showcases the application of deep learning in addressing critical health issues but also sets a precedent for the development of accessible diagnostic tools that can be deployed at

scale, offering significant implications for telemedicine and remote healthcare services.

**Keywords:** facial paralysis detection, InceptionV3 algorithm, real-time object detection, telemedicine, remote healthcare services.

## I. INTRODUCTION

Facial paralysis is a medical condition characterized by the loss of voluntary muscle movement in one or both sides of the face, which can significantly affect a person's quality of life through impairments in facial expressions, eating, speaking, and even leading to psychological distress. The causes of facial paralysis are varied, including Bell's palsy, stroke, Lyme disease, and trauma, making its diagnosis complex and necessitating specialized medical expertise. Traditional diagnostic processes are often time-consuming, costly, and not easily accessible to everyone, particularly in less urbanized or resource-limited regions.

In response to these challenges, this project introduces an innovative solution leveraging the advancements in deep learning and computer vision. By adopting the InceptionV3 algorithm, renowned for its speed and accuracy in object detection tasks, we have developed a sophisticated model capable of detecting facial paralysis from both static images and live video streams. This approach not only democratizes access to diagnostic tools but also significantly enhances the efficiency and accuracy of early-stage detection.

The development and deployment of this model within a web-based application further extend its accessibility and usability. Through this platform, users can easily register, upload images, or use a

live camera feed to receive immediate diagnostic feedback on the presence of facial paralysis signs. An administrative module ensures the system's integrity by managing user access, overseeing model performance, and facilitating continuous improvement through retraining processes.

This project stands at the intersection of technology and healthcare, aiming to bridge the gap between advanced medical diagnostics and patient accessibility. By providing a scalable, accurate, and user-friendly tool for the early detection of facial paralysis, it paves the way for prompt intervention and treatment, potentially mitigating long-term impacts on patients' lives.

## II. LITERATURE SURVEY

**TITLE:** Automatic facial paralysis assessment via computational image analysis,'

**ABSTRACT:** Facial paralysis (FP) is a loss of facial movement due to nerve damage. Most existing diagnosis systems of FP are subjective, e.g., the House-Brackmann (HB) grading system, which highly depends on the skilled clinicians and lacks an automatic quantitative assessment. In this paper, we propose an efficient yet objective facial paralysis assessment approach via automatic computational image analysis. First, the facial blood flow of FP patients is measured by the technique of laser speckle contrast imaging to generate both RGB color images and blood flow images. Second, with an improved segmentation approach, the patient's face is divided into concerned regions to extract facial blood flow distribution characteristics. Finally, three HB score classifiers are employed to quantify the severity of FP patients. The proposed method has been validated on 80 FP patients, and quantitative results demonstrate that our method, achieving an accuracy of 90.14%, outperforms the state-of-the-art systems. Experimental evaluations also show that the proposed approach could yield objective and quantitative FP diagnosis results, which agree with those obtained by an experienced clinician.

**TITLE:** 'Objective grading facial paralysis severity using a dynamic 3D stereo photogrammetry imaging system,'

**ABSTRACT:** Facial paralysis is a loss of facial movement due to nerve damage. It is essential for clinicians to diagnose the severity of facial paralysis to treat patients, assess progress and evaluate

outcomes. Subjective assessments are common in clinical practices but have their limitations regarding the intra-observer and inter-observer reproducibility. We utilized dynamic 3D stereo photogrammetry technology for the objective grading of facial paralysis by measuring regional facial asymmetries. The correlations between the measured asymmetries and scores of a modified Sunnybrook facial paralysis grading were evaluated to identify the region of interests of objective measurements closely related to the subjective grades. Categorical classifiers were trained to quantify the severity of facial paralysis. Preliminary results showed that the objective asymmetry measurements were highly correlated to the subjective assessments of facial paralysis except the eye region. Machine learning approaches showed the potential of improving the accuracy of severity assessments.

**TITLE:** Toward an automatic system for computer aided assessment in facial palsy,'

**ABSTRACT:** Quantitative assessment of facial function is challenging, and subjective grading scales such as House-Brackmann, Sunnybrook, and eFACE have well-recognized limitations. Machine learning (ML) approaches to facial landmark localization carry great clinical potential as they enable high-throughput automated quantification of relevant facial metrics from photographs and videos. However, the translation from research settings to clinical application still requires important improvements.

**Objective:** To develop a novel ML algorithm for fast and accurate localization of facial landmarks in photographs of facial palsy patients and utilize this technology as part of an automated computer-aided diagnosis system.

**Design, Setting, and Participants:** Portrait photographs of 8 expressions obtained from 200 facial palsy patients and 10 healthy participants were manually annotated by localizing 68 facial landmarks in each photograph and by 3 trained clinicians using a custom graphical user interface. A novel ML model for automated facial landmark localization was trained using this disease-specific database. Algorithm accuracy was compared with manual markings and the output of a model trained using a larger database consisting only of healthy subjects.

**Main Outcomes and Measurements:** Root mean square error normalized by the interocular distance (NRMSE) of facial landmark localization between prediction of ML algorithm and manually localized landmarks. **Results:** Publicly available algorithms for facial landmark localization provide poor localization accuracy when applied to photographs of patients compared with photographs of healthy controls (NRMSE,  $8.56 \pm 2.16$  vs.  $7.09 \pm 2.34$ ,  $p \ll 0.01$ ). We found significant improvement in facial landmark localization accuracy for the facial palsy patient population when using a model trained with a relatively small number photographs (1440) of patients compared with a model trained using several thousand more images of healthy faces (NRMSE,  $6.03 \pm 2.43$  vs.  $8.56 \pm 2.16$ ,  $p \ll 0.01$ ).

**Conclusions and Relevance:** Retraining a computer vision facial landmark detection model with fewer than 1600 annotated images of patients significantly improved landmark detection performance in frontal view photographs of this population. The new annotated database and facial landmark localization model represent the first steps toward an automatic system for computer-aided assessment in facial palsy.

**TITLE:** “The spectrum of facial palsy: The MEEI facial palsy photo and video standard set,”

**ABSTRACT:** Facial palsy causes variable facial disfigurement ranging from subtle asymmetry to crippling deformity. There is no existing standard database to serve as a resource for facial palsy education and research. We present a standardized set of facial photographs and videos representing the entire spectrum of flaccid and nonflaccid (aberrantly regenerated) facial palsy. To demonstrate the utility of the dataset, we describe the relationship between level of facial function and perceived emotion expression as determined by an automated emotion detection, machine learning-based algorithm.

**TITLE:** “Facial imaging and landmark detection technique for objective assessment of unilateral peripheral facial paralysis,”

**ABSTRACT:** In this paper, we propose a hypothesis that the facial landmark detection methods constructed by a private UPFP facial dataset can perform better than the model on a healthy facial dataset in the task of UPFP facial landmark detection. For proving this hypothesis, a

customized UPFP facial dataset with 68 facial landmark annotations was built. A state-of-the-art facial landmark detection method was employed on the three evaluation datasets to exploit and prove the hypothesis. The mean error of validation dataset is 3.15, 56% lower than 7.42 that of the healthy dataset, which proves the hypothesis is true.

**TITLE:** “Automatic assessment of facial paralysis based on facial landmarks,”

**ABSTRACT:** Unilateral peripheral facial paralysis is the most common case of facial paralysis. It affects only one side of the face, which will cause facial asymmetry. Clinically, unilateral peripheral facial paralysis is often classified by clinicians according to evaluation scales, based on patients' condition of facial symmetry. A prevalent scale is House-Brackmann grading system (HBGS).

However, assessment results from scales are often with great subjectivity, and will bring high interobserver and interobserver variability. Therefore, this manuscript proposed an objective method to provide assessment results by using facial videos and applying machine learning models. This grading method is based on HBGS, but it is automatically implemented with high objectivity. Images with facial expressions will be extracted from the videos to be analyzed by a machine learning model. Facial landmarks will be acquired from the images by using a 68-points model provided by dlib. Then index and coordinate information of the landmarks will be used to calculate the values of features pre-designed to train the model and predict the result of new patients. Due to the difficulty of collecting facial paralysis samples, the data size is limited. Random Forest (RF) and support vector machine (SVM) were compared as classifiers. This method was applied on a data set of 33 subjects. The highest overall accuracy rate reached 88.9%, confirming the effectiveness of this method.

### III. METHODOLOGY

#### 1. Dataset Collection and Preparation

- **Data Sourcing:** The project utilizes a dataset from Kaggle, comprising 3,000 images categorized into two classes to represent facial paralysis and normal facial expressions.
- **Data Augmentation:** To enhance the model's ability to generalize across

various facial features and conditions, data augmentation techniques such as rotation, zoom, and horizontal flipping are applied.

## 2. Model Development

- **INCEPTIONV3 Algorithm:**  
The INCEPTIONV3 algorithm is chosen for its efficiency and accuracy. Its well-suited for real-time facial paralysis detection.
- **Training and Validation:** The model is trained and validated using Google Collab for its GPU support, facilitating faster computation. A validation set is used to tune hyperparameters and avoid overfitting.

## 3. Performance Evaluation

- **Accuracy Assessment:** The model's performance is evaluated based on its accuracy in detecting facial paralysis, achieving a remarkable 95.3% accuracy on the test dataset.
- **Real-time Testing:** Additional tests are conducted using live video feeds to ensure the model's effectiveness in real-world scenarios.

## 4. Web Application Development and Deployment

- **User Interface Design:** A user-friendly web application is developed, allowing users to register, login, and upload images or access the camera for live detection.
- **Admin Module:** An administrative module is incorporated for managing user authentication, overseeing model training, and testing, and conducting model analysis.
- **Deployment:** The web application is deployed on a suitable platform to ensure accessibility and scalability.

## MODULES DESCRIPTION

### 1. Data Management Module

- **Functionality:** Manages the collection, augmentation, and preprocessing of the

dataset. This includes sourcing images from Kaggle, applying data augmentation techniques to increase the diversity of the training set, and preprocessing images to fit the input requirements of the INCEPTIONV3 model.

- **Components:** Dataset collection, Data augmentation, Preprocessing.

### 2. Model Training and Validation Module

- **Functionality:** Handles the core development of the deep learning model using the INCEPTIONV3 algorithm. It oversees the training process, including the selection of hyperparameters, the division of data into training and validation sets, and the evaluation of model performance to prevent overfitting.
- **Components:** Model architecture setup, Training loop, Validation, and performance evaluation.

### 3. Real-time Detection Module

- **Functionality:** Facilitates the real-time detection of facial paralysis from both static images and live video feeds. This module leverages the trained INCEPTIONV3 model to analyse uploaded images or camera streams, identifying signs of facial paralysis with high accuracy.
- **Components:** Image upload, Live video feed processing, Detection algorithm.

### 4. User Interface (UI) Module

- **Functionality:** Provides the front-end interface for user interaction with the web application. It enables users to register, login, upload images, access the live camera feed, and receive diagnostic feedback. The UI is designed to be intuitive and user-friendly, ensuring ease of use for individuals with varying levels of technical proficiency.
- **Components:** Registration and login system, Image upload interface, Live camera access, Feedback display.

### 5. Admin and Security Module

- Functionality:** Offers tools for administrative tasks, including user management, model management, and security measures. It ensures that only authorized users can access the system, provides the admin with the ability to authorize new users, and oversees the training and testing of the model for continual improvement.
- Components:** User authentication and authorization, Model training and testing oversight, Security protocols.

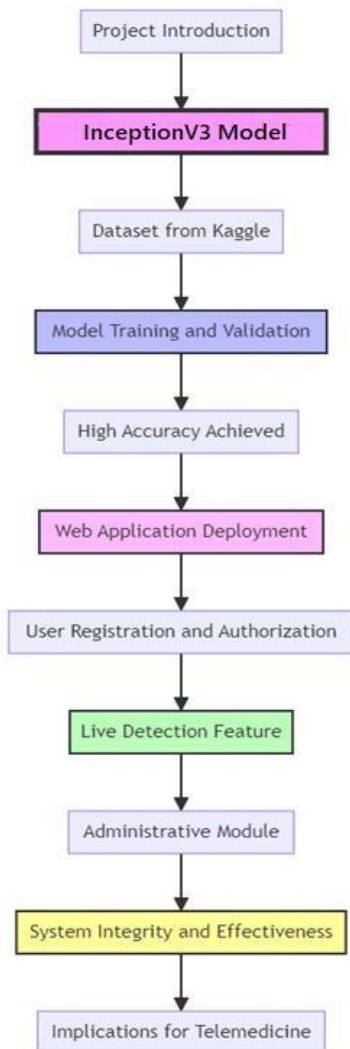


Figure 1: Flow chart

### IV. IMPLEMENTATION

#### CONVOLUTIONAL NEURAL NETWORKS

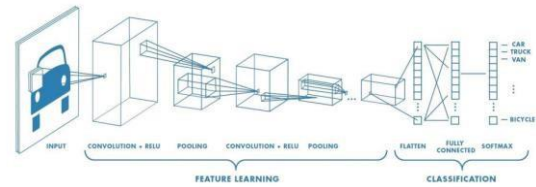
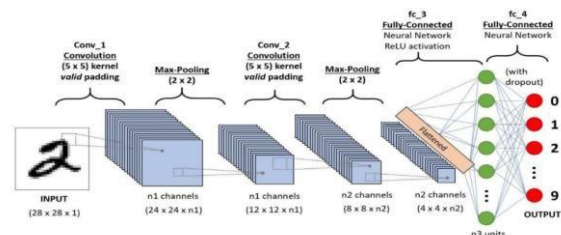


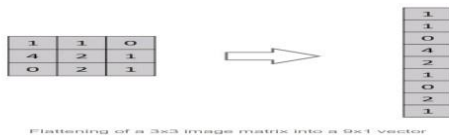
FIG 2: Convolutional neural network

The field of Artificial Intelligence (AI) is rapidly advancing, significantly narrowing the gap between human and machine capabilities. Enthusiasts and researchers are delving into various facets of AI to create remarkable innovations. A prime example of such a domain is Computer Vision. The objective in this area is to empower machines with the ability to interpret the world in a manner akin to human vision, and to utilize this understanding in a range of applications. These include tasks like Image and Video Recognition, Image Analysis and Classification, Media Recreation, and Recommendation Systems, as well as extending into areas like Natural Language Processing. The progress in Computer Vision, particularly through the application of Deep Learning, has been substantial. This progress is largely attributed to the refinement of a key algorithm: the Convolutional Neural Network (CNN), which has evolved and been honed over time.



**Fig 3: layers of ConNet**

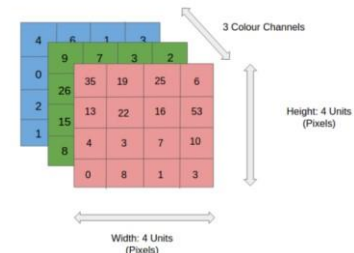
A Convolutional Neural Network (CNN) is a type of deep learning algorithm that processes input images by assigning significance to various aspects within them, enabling it to distinguish between different objects. Compared to other classification algorithms, CNNs require significantly less preprocessing. Traditional methods involve manually creating filters, but with sufficient training, CNNs are capable of learning these filters autonomously. The design of a CNN mirrors the neural connectivity patterns found in the human brain, particularly inspired by the structure of the visual cortex. In this system, individual neurons are activated by specific stimuli within a limited area of the visual field, known as the receptive field. These receptive fields overlap to comprehensively cover the entire visual spectrum.



**fig 4: flattening of 3x3 image matrix**

An image is essentially a matrix composed of pixel values. One might wonder why not simply flatten this matrix and input it into a multi-layer perceptron for classification. The reason lies in the unique capabilities of a Convolutional Neural Network (ConvNet). A ConvNet excels in capturing the spatial and temporal dependencies in an image by applying appropriate filters. Its architecture is more adept at conforming to the image dataset, primarily due to a decrease in the number of parameters and the recyclability of weights. In essence, this means that the network can be more effectively trained to comprehend the complexities inherent in images.

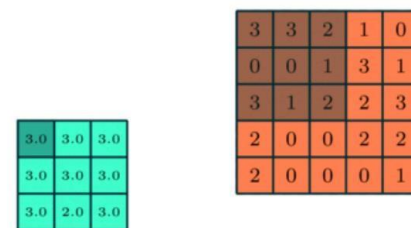
**Input Image**



**Fig 5: RGB image separated by its colors**

In the given illustration, an RGB image is depicted, segmented into its three primary color components: red, green, and blue. Images can exist in various color spaces, such as grayscale and RGB. Imagine the computational load when dealing with high resolution images, such as those in 8K. The function of a Convolutional Neural Network (ConvNet) in this context is to simplify the images into a more manageable form for processing, while preserving essential features crucial for accurate predictions. This aspect is particularly vital in designing an architecture that is not only efficient in feature learning but also capable of scaling to handle extensive datasets.

**POOLING LAYER**



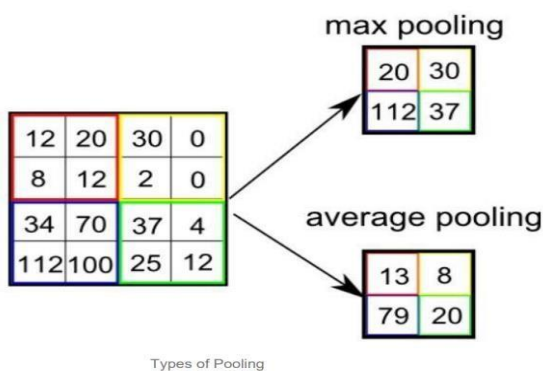
3x3 pooling over 5x5 convolved feature

**Fig 6: pooling layers**

Much like the Convolutional Layer, the role of the Pooling layer in a Convolutional Neural Network

is to reduce the spatial dimensions of the Convolved Feature. This reduction is key in lessening the computational burden by decreasing the data's dimensionality. Additionally, Pooling is instrumental in extracting pivotal features that are invariant to rotation and position, aiding in the efficient training of the model.

Pooling comes in two main forms: Max Pooling and Average Pooling. Max Pooling operates by selecting the maximum value from the image area covered by the Kernel, effectively acting as a noise suppressant by eliminating noisy activations and aiding in denoising as well as reducing dimensionality. Average Pooling, in contrast, computes the average value of all elements in the Kernel's coverage area and primarily focuses on dimensionality reduction as its method of noise suppression. Consequently, Max Pooling is often considered more effective than Average Pooling due to its dual role in noise suppression and dimensionality reduction.



**Fig 7: types of pooling layers**

The Convolutional Layer and the Pooling Layer collectively constitute the i-th layer in a Convolutional Neural Network's architecture. To capture more nuanced, low-level details in complex images, the network may incorporate additional layers of this kind. However, this enhancement comes with the trade-off of increased computational demands. Once the

image data has passed through these layers, the model gains a comprehensive understanding of the image features. Subsequently, the final output is flattened and then input into a standard neural network, which performs the task of classifying the data.



**Fig 8: ConvNet layers**

## V. MATHEMATICAL APPROACH

Inception v3 is a deep convolutional neural network (CNN) architecture designed for image recognition and classification tasks. The architecture includes various types of layers such as convolutional layers, pooling layers, and fully connected layers. Here, I'll provide a high-level overview of the architecture and its mathematical equations:

### 1. Input:

- The input to Inception V3 is typically a 299x299 RGB image(3 channels).
- Let's denote the input image as X, where X is a 3D tensor with dimensions (299,299,3).

### 2. Stem:

The stem consists of basic layers to process the input image.

First, we perform a 2D convolution with stride 2 and 'SAME' padding:

- $\text{Conv2D}(X, 32, 3 \times 3, \text{stride} = 2, \text{padding} = \text{'SAME'})$

Apply Batch Normalization and ReLU activation:

- $\text{BatchNorm}(\text{Conv2D}(\dots))$

Followed by another 2D convolution:

- $\text{Conv2D}(\dots, 32, 3 \times 3)$

And again, Batch Normalization and ReLU:



- $\text{BachNorm}(\text{Conv2D}(\dots))$

Finally, a max pooling layer:

- $\text{MaxPool}(\dots, 3 \times 3, \text{stride} = 2)$

### 3. Inception Modules:

The core part of Inception networks are the Inception modules. These modules are designed to capture features at various spatial scales. The basic Inception module has parallel branches with different filter sizes.

Let's denote the output of the previous layer as A.

InceptionModule 1:

Branch 1:

- $A1 = \text{Conv2D}(A, 64, 1 \times 1)$

Branch 2:

- $A2 = \text{Conv2D}(A, 96, 1 \times 1)$
- $A2 = \text{Conv2D}(A2, 128, 3 \times 3)$

Branch 3:

- $A3 = \text{Conv2D}(A, 16, 1 \times 1)$
- $A3 = \text{Conv2D}(A3, 32, 5 \times 5)$

Branch 4:

- $A4 = \text{Conv2D}(A, 3 \times 3)$
- $A4 = \text{Conv2D}(A4, 32, 1 \times 1)$

Concatenate along the depth dimension:

- $A = \text{Concatenate}([A1, A2, A3, A4])$

Inception Module 2:

- Similar structure to Module 1 but with different filter size and numbers.

### 4. Reduction Modules:

These modules are used to reduce the spatial dimensions of the input, thus reducing computational cost.

Reduction Module A:

Branch 1:

- $A1 = \text{Conv2D}(A, 256, 1 \times 1)$
- $A1 = \text{Conv2D}(A1, 384, 3 \times 3, \text{stride} = 2)$

Branch 2:

- $A2 = \text{Conv2D}(A, 256, 1 \times 1)$
- $A2 = \text{Conv2D}(A2, 256, 3 \times 3, \text{stride} = 2)$

Branch 3:

- $A3 = \text{Conv2D}(A, 256, 1 \times 1)$
- $A3 = \text{Conv2D}(A3, 256, 3 \times 3)$
- $A3 = \text{Conv2D}(A1, 256, 3 \times 3, \text{stride} = 2)$

Concatenate along the depth dimension:

- $A = \text{Concatenate}([A1, A2, A3])$

Reduction Module B:

- Similar structure to Reduction Module A but with different filter sizes and numbers.

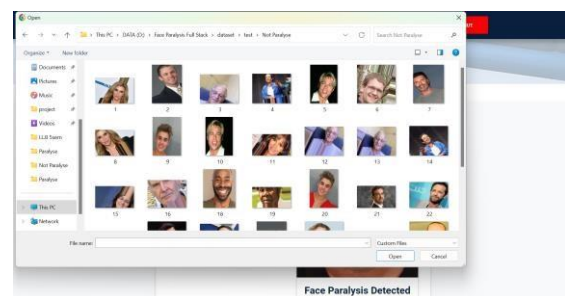
### 5. Fully Connected Layers:

After several Inception and Reduction modules, the output is passed through fully connected layers for classification.

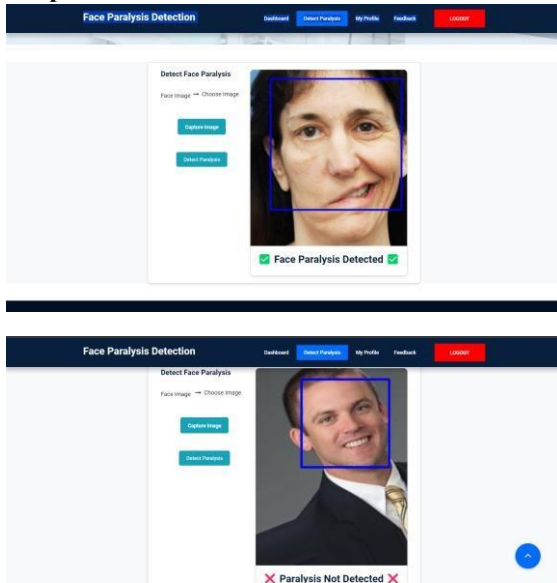
### VI. RESULTS & DISCUSSION

utilizing the InceptionV3 algorithm, to accurately detect facial paralysis from both static images and live video feeds. By harnessing a dataset of 3,000 images from Kaggle, enhanced through data augmentation techniques, the system achieves remarkable detection accuracy, demonstrated by a 95.3% success rate in tests. Integrated into a user-friendly web application, it offers functionalities for user registration, image upload, and live detection, managed through an admin module for user authentication and model oversight. This system not only bridges the gap in accessible diagnostic tools for facial paralysis but also sets a precedent for applying AI in telehealth solutions, promising widespread impact on early detection and patient care.

#### Sample Input Images for Testing



**Output:**



MODEL	Accuracy rate
<b>INCEPTIONV3</b>	95.3 %
<b>ALEXNET</b>	93.6%
<b>RESTNET30</b>	91.7%

**VII. CONCLUSION**

In conclusion, the development of a deep learning model utilizing the INCEPTIONV3 algorithm for the detection of facial paralysis represents a significant leap forward in the application of artificial intelligence in healthcare diagnostics. This project has successfully demonstrated the feasibility and efficiency of using advanced machine learning techniques to address a critical medical diagnosis challenge. By achieving an impressive 95.3% accuracy in detecting facial paralysis from static images and live video feeds, the system sets a new benchmark in the field of automated medical diagnostics.

The deployment of this model within a user-friendly web application significantly enhances the accessibility and convenience of early detection tools for facial paralysis, potentially leading to earlier interventions and better treatment outcomes for

affected individuals. The inclusion of an administrative module for user authentication and model management further ensures the integrity and reliability of the system, making it a valuable resource for both patients and healthcare providers.

**VIII. FUTURE SCOPE**

In moving forward, continuous improvement through user feedback and further model training will be essential to adapt to the evolving landscape of medical diagnostics and technology. This project not only contributes to the advancement of medical technology but also exemplifies the power of AI in making healthcare more accessible, efficient, and patient-centered.

**IX. REFERENCE**

- [1] P Subba Rao. "Applications of machine learning in environmental engineering" JOURNAL OF Science, Technology and Development (2021): 643-648.
- [2] P Subba Rao. "Using an Energy Efficient Extended Leach Work with Multilevel Clustering Approach" from JOURNAL of Research Publication and Reviews (2020): 60-64.
- [3] C. Jiang, J. Wu, W. Zhong, M. Wei, J. Tong, H. Yu, and L. Wang, "Automatic facial paralysis assessment via computational image analysis," J. Healthcare Eng., vol. 2020, pp. 1–10, Feb. 2020.
- [4] D. L. Guarin, Y. Yunusova, B. Taati, J. R. Dusseldorp, S. Mohan, J. Tavares, M. M. van Veen, E. Fortier, T. A. Hadlock, and N. Jowett, "Toward an automatic system for computer aided assessment in facial palsy," Facial Plastic Surg. Aesthetic Med., vol. 22, no. 1, pp. 42–49, Feb. 2020.
- [5] J. J. Greene, D. L. Guarin, J. Tavares, E. Fortier, M. Robinson, J. Dusseldorp, O. Quatela, N. Jowett, and T. Hadlock, "The spectrum of facial palsy: The MEEI facial palsy photo and video standard set," Laryngoscope, vol. 130, no. 1, pp. 32–37, Jan. 2020.
- [6] Z. Guo, W. Li, J. Dai, J. Xiang, and G. Dan, "Facial imaging and landmark detection technique for objective assessment of unilateral peripheral facial paralysis," Enterprise Inf. Syst., vol. 16, nos. 10–11, pp. 1556–1572, Oct. 2022.