Journal of Theoretical and Applied Information Technology
15th May 2025. Vol.103. No.9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PREDICTION OF GEOMETRIC FEATURES IN 3D FACIAL SOFT TISSUE IMAGES USING FULLY CONNECTED NEURAL NETWORK

[1]CHANDRA SEKHAR KOPPIREDDY , [2]SIVA NAGESWARA RAO

[1]Research Scholar,Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522302 Andhra Pradesh, India
[2]Professor,Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522302, Andhra Pradesh, India
chandrasekhar.koppireddy@gmail.com[1], sivanags@kluniversity.in[2]

ABSTRACT

The relevant geometric prominence of the three-dimensional facial soft tissue images is a crucial factor in forensic reconstruction, craniofacial surgery, and biometric identification. The image processing technique is one of the most critical steps in analysing 3D facial images. Traditional image processing methods often face problems while capturing the soft tissue image's non-linear deformations and complex anatomical variation. So, this paper presented a solution to this challenge by formulating an accurate geometric feature prediction model using Fully Connected Neural Network (FCNN) for 3D soft tissue facial images. The important feature of the model is that it was trained on high-definition 3D face scans and leveraged deep learning-driven feature extraction to increase the model's accuracy. Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Structural Similarity Index (SSIM) are the metrics used for evaluation. The presented model performed better than traditional deep learning models based on the results. Finally, the proposed model has provided far more accurate results in shape prediction, which makes this model highly suitable for application in personalized facial reconstruction, clinical diagnostics, and aesthetic surgery planning. These enhancements in the technique contribute to the development of an improved automated 3D facial analysis, making a better way of modelling the soft tissue area and the computational facial anatomy.

Keywords: *3D Facial Soft Tissue, Geometric Feature Prediction, Fully Connected Neural Network, Deep Learning, Craniofacial Analysis.*

1. INTRODUCTION

Interpreting nonverbal cues and reading another person's emotions, thoughts, and intentions is about understanding and responding appropriately to other's emotions. At the same time, humans have various cues such as word choice, voice, facial expression, and intonation to share and interpret emotional states. People who are blind or visually impaired cannot perceive non-verbal cues, such as facial expressions and tone of voice, through touch. Recognizing their accurate emotion is one of the most essential things in social interaction because the function of emotion recognition is to help people communicate more effectively by analyzing people's emotions [2]. Claiming emotional mimicry, or mirroring an individual's non-verbal behaviors that underlie an individual emotional expression [3, 4], has been helping to improve a person's likeability and, subsequently, the probability that they are more likely to develop a connection with the person. People may use various methods to convey their emotions to other people, which include facial expressions [5].

Orthognathic surgery is for dental function, enhancing facial features, and correcting the symmetry and coordination of the face [6]. Orthognathic surgery has gradually been treating patients' esthetical needs with an increasing demand, and concepts based on esthetics have become more common. Every day, social communication happens based on the normal facial shape. A history of social problems, feelings

of inferiority, and depression in patients with dentofacial deformities (DFDs) are not uncommon [7, 8].

rays or CT scans of the head. [13]. Unfortunately, the face width-to-face height ratio can be obtained only from soft tissue images, but orthognathic surgeons can analyse that ratio. Doctors measure facial structures using a scale or a face arch, depending on their level of experience.

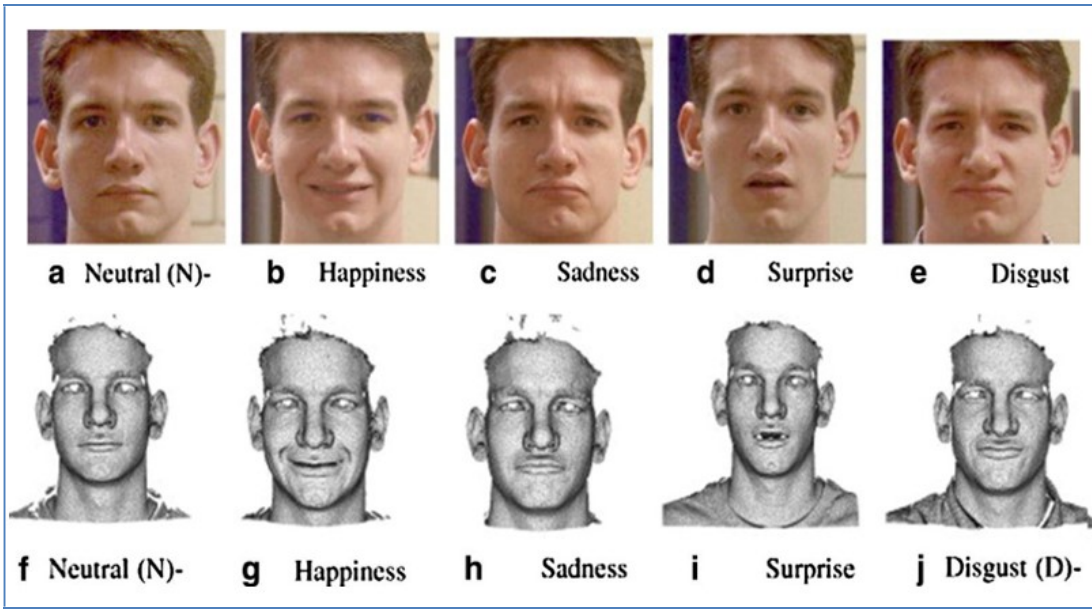


Fig.1. Facial Expression Identification [9]

The need for a facial recognition method was increased to identify people's emotions for better communication. So, it uses imaging technology to recognize facial expressions. different fields, such as orthodontics, genetics, plastic surgery, and craniomaxillo facial surgery, have provided excellent results and multiple benefits from imaging technology, especially in 3D images.

An advanced Cone-beam computed tomography (CBCT) has grown within those fields. Now, CBCT has become one of the most commonly used imaging techniques and is considered to replace Computed Tomography (CT) imaging, a typical imaging technique used before for hard tissues [10]. However, 3D stereophotogrammetry can provide a more detailed and accurate representation of craniofacial soft tissue without ionizing radiation. So, it has become very popular in soft tissue studies. The diagnosis of DFD is mainly done by looking at side-view X-

This method takes a lot of time, depends on personal judgment, and requires a skilled professional. As a result, there is still no quick and objective way to evaluate facial soft tissue in clinical work. [14].

Due to the development of AI within the medical field, there is a possibility of solving this problem. ML has been used in research very recently. It is applied to other applications, like mining texts, media personalisation, classification of images, and multimedia indexing and retrieval [15, 16, 17]. Many other ML algorithms exist, but DL is widely used in almost all applications [18]. Another shared term for DL is RL. This is partly due to its potential in data acquisition and partly due to the fantastic progress made in hardware technologies, such as HPC. Compared to the DL models, CNN provided a better result in face recognition when used on the LFW database and attained an accuracy of 97.45% for facial [19]. VGG19's CNN has an accuracy rate of 89.3% in

predicting the need for orthognathic surgery in a patient. Overall, the CNN accurately understood the outline of soft tissue by analysing the images required for orthognathic surgery [20]. Compared with the other methods, this paper aims to enhance the efficiency of landmark point detection accurately and more. Thus, this paper contributes the following:

- In this paper, an FCNN architecture is used to predict the geometrical characteristics and parameters of 3D soft tissue facial images.
- The model has been trained using a high-definition 3D face scan, enhancing its ability to learn more about facial features.
- Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Structural Similarity Index (SSIM) are the metrics used to evaluate the model's overall performance.
- The results indicate that the proposed FCNN model has outperformed conventional DL models in predicting geometric features using 3D soft tissues. It also improves reliability and accuracy.
- The model was found to be very effective in applications such as clinical diagnostics, facial reconstruction and aesthetic surgery

2. LITERATURE SURVEY

The The problem statement and the challenges and issues the earlier methods face must be identified to understand the proposed model's design pathway. This section briefly explains various earlier methods used for 3D facial image processing. Behzad Hasani et al. (2017) have presented an investigation on using 3D Facial Expression Recognition (FER) in videos. The proposed network design incorporates several 3D Inception-ResNet layers and an LSTM unit. It captures spatial details in face images and tracks changes across video frames over time. Our network uses facial landmark points to highlight the key features that matter most for expressions instead of considering facial areas that don't play a significant role. This network focuses on specific points on the face rather than general face areas. It helps to capture essential details needed for facial expressions while ignoring parts that don't contribute much. By doing this, the system better understands and generates facial expressions.

We have tested our proposed method on 4 publicly available databases to evaluate the model's efficiency. In recent days, various ECG-based face emotion identification techniques have been developed. So, Jianhua Zhang et al. (2020) investigated various emotion-based recognition techniques using multi-modal physiological signals and multi-channel EEG signals. For emotion recognition, it uses feature extraction, feature reduction and various ML classifier methods such as naïve Bayes (NB), KNN, support vector machine (SVM), and random forest (RF) for classification. The proposed models were compared with various conventional models to evaluate the performance of emotion recognition. The result shows that these proposed models also have some problems while recognising the emotion of the image.

Hong Yuan Cao and Chao Qi (2021) have proposed a more efficient method that takes a single 2D image as input and automatically creates a 3D face model as output. This approach uses 3D facial details to predict continuous emotions in a dimensional space. So, this approach can be used for an online learning system. The simulation result shows that the proposed algorithm has a strong ability to recognise facial expressions using 3D input images. Xiaoguang Tu et al. (2021) studied the problem and predicted a 3D moving sequence from a single-face image using a 3D dynamic prediction network. They focused on understanding the issue and creating a method to predict how a face would move in 3D based on just one image. The sparse texture mapping algorithm is then used to render the 3D dynamics further to simulate such details and sparse textures on rendering face frames. Due to its openness, our model is versatile to multiple AR/VR and entertainment problems like face video retargeting and facial video prediction. The model effectively generates high-quality, consistent facial retaining,

and visually appealing video clips that obtain the face from a single source image, which has been shown in superior experimental results.

Gangothri Sanil et al. (2023) studied how to turn two-dimensional images into three-dimensional ones. They created high-quality 3D face recognition models without using complex 3D morphable models. To evaluate and identify the best model, it compares all the model's accuracy in predicting facial expressions. The simulation result shows that the extreme gradient boosting model was the best classification model in terms of accuracy; it attained a 78% accuracy rate in classifying the images. Duc-Phong Nguyen et al. (2021) have presented an investigation of the new DL model called the geometric deep learning model that operates on 3D point cloud data to recognise facial expressions by analysing the 3D image. The presented model is tested on two datasets, SIAT-3DFE and Bosphorus, to estimate the overall efficiency. The simulation result attained a high accuracy rate of 85.5% in the Bosphorus dataset while recognising five expressions: anger, surprise, disgust, neutral and happiness.

In recent days, facial recognition applications have become essential in various fields, such as security, criminal investigation, innovative card applications, and database management systems (DBMS). So, S. Gokulasrishnan et al. (2023) have proposed a DL model called a strawberry-based CNN model for facial expression recognition. The overall efficiency of the presented SBCNN model was tested on the Kaggle face expression dataset to evaluate its overall efficiency. The experimental result shows that this proposed model has performed better in recognising facial expressions. It is necessary to help the students engage in a collaborative learning environment, which helps improve the quality of learning. So, Yi Chen et al. (2022) have proposed a multi-modal deep neural network (MDNN) model to automatically analyse the engagement level of students in a classroom by analysing their facial expressions. The direction of gazing and facial expression are the two components used to predict student engagement. It was tested in a real-time environment to evaluate the model's efficiency. The experimental result shows that this proposed MDNN model predicts student engagement in class better.

.

3. LIMITATION AND MOTIVATION

In forensics detection, cosmetic surgery, facial contour remodelling and biometrics, the prognosis of 3D cephalic soft tissue imaging plays a vital role. For better results from the operation effects, facial contour index and finding facial tissues malformed through the facts of these assumptions. Conventional photo processing and the facial regression models approach have hurdles in detaining the unsystematic, multifaceted, progressive modification that appears in muscle formations. AI and ML brought the most recent developments, such as adapting the Neural Networks to apprehend the multi-faceted correlation between multiple cutaneous tissue layers. For the assessment of 3D anatomical images, this AI and ML provide the FCNN (fully connected neural networks), which is known for its efficacy in examining dimensional attributes. The FCNN is used explicitly for the adversity function, which is accurately used for mathematical outcomes, including tissue density, deflections, and mass deformity.

4. CHALLENGES IN 3D FACIAL SOFT TISSUE FEATURE PREDICTION

Apart from the latest technological developments in medical image processing and analytical modelling, geometrical feature capturing by 3D facial tissue imaging has a challenging side. The characteristic resemblance is due to ageing factors, gender differences, nativity, and external factors such as weight imbalance and surgical treatment history. Due to the facial tissue malformations, it is highly uncoordinated for proposing conventional linear recurrence or by SSM (statistical Shape Model) ineffective. By using the Deep learning framework, this 3D image processing generates massive data for extracting and processing suggestive curvilinear structures. Although this hinders pattern recognition, it is now enabled in fields like CT, MRI AND 3D image scanning. Facial tissue annotation needs

substantive inputs and the restrictive accessibility of the training data comparatively with the cartilaginous structure.

5. PROBLEM STATEMENT

Fully connected networks (FCNs) can generate 3D facial images by constructing layers of hierarchical spatial features. Each layer extracts specific details, enabling the network to predict the facial landmark points very accurately and in detail. FCN present in each layer can obtain various features, enhancing its ability to pinpoint a higher number of facial landmark points accurately. A 3D facial image can be illustrated; it consists of numerous sets of (x, y, z) coordinates captured from the 3D scan. It is a volumetric representation where each voxel holds intensity data. It is a 2D matrix where pixel values can demonstrate the depth information. It models the face surface using interconnected polygons, capturing its contours and details for accurate 3D modelling in facial recognition. The input can be preprocessed before being fed into the FCN using normalisation, alignment, and feature vectorisation techniques. The proposed FCN comprises a set of layers with defined functions. For example, the **i**nput layer, which can do preprocessing normalization from the flattened 3D face data, can be acquired as an input vector:

$$\boldsymbol{X} = [\boldsymbol{x_1}, \boldsymbol{y_1}, \boldsymbol{z_1}, \dots, \boldsymbol{x_i}, \boldsymbol{y_i}, \boldsymbol{z_i}, \dots, \boldsymbol{x_n}, \boldsymbol{y_n}, \boldsymbol{z_n}]$$

To assure uniformity across various subjects, the data have been normalized:

$$\boldsymbol{X_{norm}} = \frac{\boldsymbol{X} - \mu}{\boldsymbol{\sigma}}$$

where μ is the mean and σ is the standard deviation.

Further detailed spatial features can be obtained from each hidden layer, progressively enhancing its understanding; the network can accurately predict more facial landmarks by increasing the output dimensions.

The first level hidden layer performs low lever feature extraction by learning edges and contours, basic geometric structures that can be recognized in 3D data. Linear transformation has been implemented, followed by an activation function:

$$\boldsymbol{z^{(1)}} - \boldsymbol{W^{(1)}} \boldsymbol{X_{norm}} + \boldsymbol{b^{(1)}}$$

$$\boldsymbol{a^{(1)}} = \boldsymbol{f(z^{(1)})}$$

Activation function ReLU performs,

$$\boldsymbol{f(x)} = \boldsymbol{max(0, x)}$$

Output a preliminary computation of landmark regions on the face. The second hidden layer performs spatial feature extraction and expansion. In this layer, the deformation patterns and facial symmetry have been recorded. By understanding the spatial transformation, the number of predicted points can be expanded:

$$\boldsymbol{z^{(2)}} = \boldsymbol{W^{(2)}} \boldsymbol{a^{(1)}} + \boldsymbol{b^{(2)}}$$

$$\boldsymbol{a^{(2)}} = \boldsymbol{f(z^{(2)})}$$

It outputs the detected facial points' accuracy, which additional localized features can enhance. This layer processes the input data to identify and extract more detailed and varied features from the face, where the number of output facial landmarks has been expanded. The system can predict many specific points on the face beyond standard landmark sets, such as 68 or 194 points, to accurately map facial features. It utilizes complex techniques to accurately identify and predict fine details such as wrinkles, skin folds, and gentle facial curves. Output: It will identify and predict more specific facial points (e.g., 200–300 points). The output layer provides high-resolution-based landmark predicted. The concluding outcome contains the 3D coordinates of facial landmark points:

$$\boldsymbol{Y}$$

Here, the total number of predicted facial landmark points is m. Loss function: Mean Squared Error (MSE) to reduce the prediction errors:

$$\boldsymbol{L} = \frac{1}{\boldsymbol{m}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{m}} (\boldsymbol{y_i} - \widehat{\boldsymbol{y_i}})^2$$

The training process involves optimisation algorithms like gradient descent and backpropagation to adjust the network's weights continuously, minimising errors and enhancing the accuracy of facial landmark predictions. At the

same time, the proposed FCN increases the total number of facial landmark points detected. Feature Expansion Across Layers: Further fine-grained features have been studied by each layer, enhancing the landmark precision. Hierarchical Representation Learning: The system has captured the overall structure of the face and localized geometric variations. Dense Regression of Landmark Points: Unlike traditional methods, which predict a fixed number of points, FCNs use interpolation to estimate additional points, resulting in a more detailed and higher-resolution mapping of facial features.

FCNs are capable of identifying and predicting a large number of facial landmarks, typically ranging from 200 to 300+ points. The system can work with various 3D face representations, such as depth maps, meshes, and point clouds, allowing for flexible analysis of facial features. For accurate predictions, the system can continuously learn geometric transformations. The system can efficiently process and analyse high-resolution 3D face images.

6. PROPOSED APPROACH

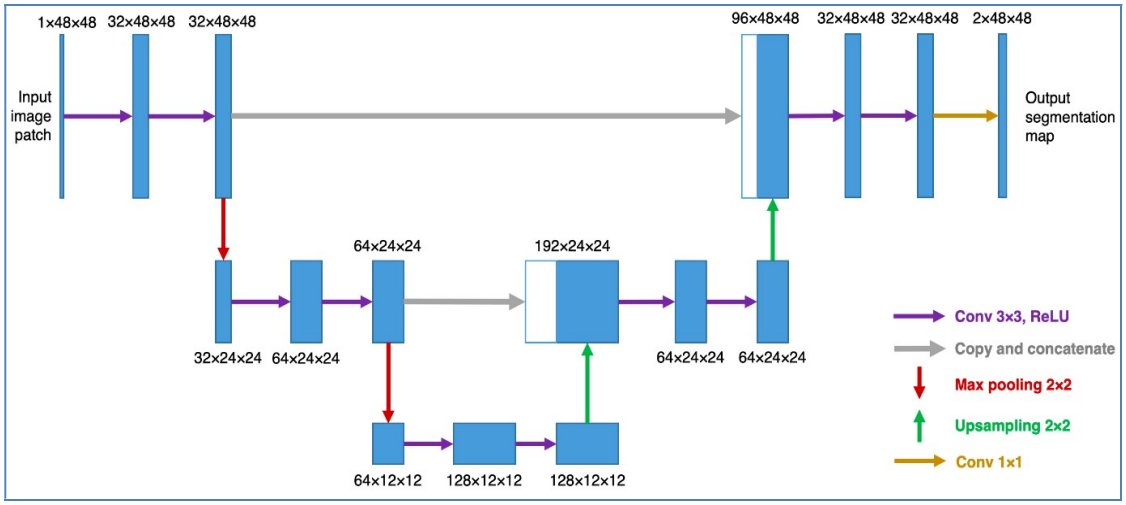


Fig.2. FCN Architecture

This study proposes an FCNN (Fully Connected Neural Network) to balance these difficulties for the assumption of 3D face tissue extractions. This model encourages the acquisition of techniques like finding the correlation between the multiple facial identification done by deep pattern recognition and its progressive competence. FCN uses various layers to process individual images and classifications, as shown in Fig.2. The overall process of FCN involved in this paper is explained well below.

Load 3D Facial Image Data: This is the initial step where the system acquires 3D facial data in various formats, such as Point Cloud, Depth Map, or Voxel Grid, which holds the 3D coordinates (x, y, z) of facial features.

Normalize Input Data: In the network architecture, the mean (μ) represents the average values of the data points, while the σ measures the variation of dispersion from the mean.

$$X_{norm} = \frac{X - \mu}{\sigma}$$

After normalization, the data will be reformed into a One-Dimensional feature vector for input to the Fully Convolutional Network (FCN).

6.1 Fully Connected Neural Network (FCN)

Initialize FCN Architecture: The input layer will take flattened 3D facial data as vectorized representation, with three hidden layers. The first hidden layer will retrieve basic geometric structures, such as edges and contours. The second hidden layer will acquire mid-level spatial transformations, such as facial symmetry and local features. The third hidden layer will detect fine-

granted features and high-resolution facial landmark points.

The FCN architecture is one of the foundational concepts in deep learning. In this architecture every neuron in a layer is interconnected to other neurons in the subsequent layer. The prediction of geometric features in 3D facial soft tissue images, pattern recognition, classification tasks and feature extraction.

6.2 Architecture of FCN

A FCN includes of three types of layers: input, hidden, and output.

Input layer

Unprocessed data, such as sensor readings, is approved. Each node denotes a unique feature, like pixel intensity.

Hidden Layers

Activation functions like ReLU, Sigmoid, and Tanh can introduce non-linearity. Various neuron layers can process the data via weighted connections. Moreover, dropout layers handle problems like overfitting.

Output layer

This layer utilizes activation functions like Linear, which is used for regression, and Softmax for classification. Final predictions are developed in this layer. The determination of a neuron's output in a neural network layer is mathematically represented as:

$$z_j^{(l)} = \sum_{i=1}^{n} w_{ij}^{(l)} a_i^{(l-1)} + b_j^{(l)}$$

$$a_j^{(l)} = f\left(z_j^{(l)}\right)$$

From the given equation, $w_{ij}^{(l)}$ denotes the synaptic weight between neurons i and j. The neuron i is present in the previous layer and j in the current layer. Activation of neuron i of the prior layer is represented by $a_i^{(l-1)}$. $b_j^{(l)}$ indicates the bias term of the neurons. The outcomes of neuron j in layer l is identified as $a_j^{(l)}$. Activation functions like Softmax and Linear can be denoted as $f(x)$ for example ReLu:

$$f(x) = max(0, x).$$

FCN for 3D Facial Feature Prediction

This approach commences by inputting a 3D facial image dataset, which a voxel or point cloud can represent. Next, features in the neurons are extracted, like landmarks, contour points, and surface geometry. These extracted features passed through FCN layers with non-linear activations. The geometric feature points can be predicted with the help of regression-based output. Finally, this model can be optimised with the help of Mean Squared Error (MSE) loss:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2$$

Where the Predicted 3D facial landmark is denoted as $\hat{y}_i$. Ground-truth 3D landmarks can be identified by y_i and N is used to express the total number of samples.

This FCN illustrates relationships between complex geometric and facial landmarks. It is reliable for versions of facial expressions and head pose. FCN is versatile in terms of expanding to high-dimensional 3D data representations. It can easily combine with CNNs or transformers for hybrid models. This FCN model is technically expensive for the 3D data with high resolution. It also trains small datasets due to its risk of overfitting. There is also a need for spatial awareness, which has been developed in CNN-FCN hybrids. Combining spatial feature extraction will improve it. In the future, transformer models will be utilized for advanced landmark predictions. Pruning and quantization techniques will be introduced, which will help reduce model size.

In deep learning, a FCN is one of the essential models in which each neuron in a layer is connected with every neuron in its following layer. Classification, regression, and attribute extraction were the tasks performed by the FCNs. The functional flow of FCN consists of three main phases: Forward Propagation, Loss Calculation and Backpropagation. The process of feature learning is done through forward propagation. In this phase, the input data passes through multiple

Journal of Theoretical and Applied Information Technology
15th May 2025. Vol.103. No.9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

layers, where it gets transformed and abstracted. When the previous layer sends the input data to the following layer of neurons, weight and bias are applied. The activation function transfers the result. This approach is continued until the final output is achieved. For a given neuron j in layer l, the output is:

$$z_j^{(l)} = \sum_{i=1}^{n} w_{ij}^{(l)} a_i^{(l-1)} + b_j^{(l)}$$

$$a_j^{(l)} = f\left(z_j^{(l)}\right)$$

Where $w_{ij}^{(l)}$ denotes the neuron which connects the weight of neuron i to j. Activation of neuron i from the prior layer can be represented by $a_i^{(l-1)}$. In layer l, the output of neuron j is identified as $a_j^{(l)}$. Activation functions like Softmax and ReLU Sigmoid can be denoted as $f(x)$. The output prediction synthesized by this model is compared with ground truth using a loss function. Regression tasks the MSE is commonly used:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2$$

Prediction of 3D landmark coordinate is expressed as $\hat{y}_i$, where the ground truth landmark coordinate is denoted by y_i. N represents the total number of samples.

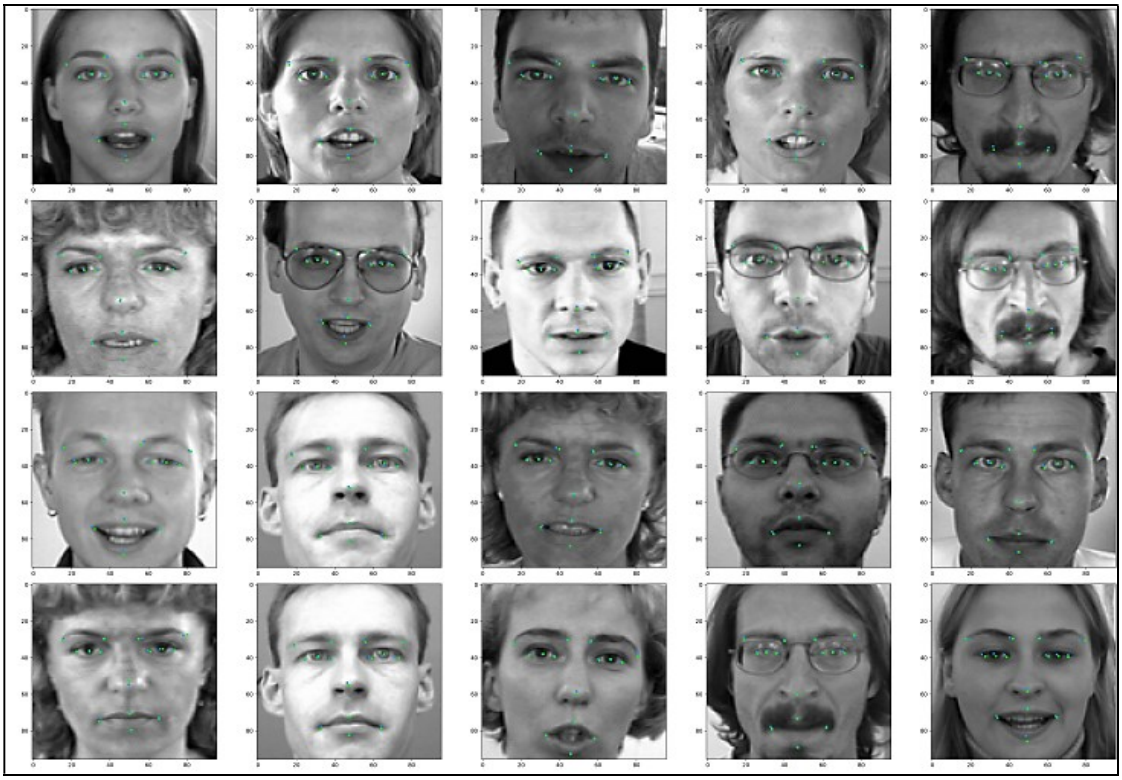


Fig.3. RGB To Gray Scale Conversion

For classification tasks, Cross-Entropy Loss is used:

$$L = -\sum_{i=1}^{N} y_i log(\hat{y}_i)$$

In this equation, y_i represents the True label of (0 or 1), and the predicted probability is denoted by $\hat{y}_i$. Gradient Descent updates the weights throughout the network whenever an error is computed. The updated rule for weights is:

$$w_{ij}^{(l)} = w_{ij}^{(l)} - \eta \frac{\partial L}{\partial w_{ij}^{(l)}}$$

Where, η represent the learning rate, and the Gradient of the loss function is denoted $\frac{\partial L}{\partial w_{ij}^{(l)}}$

concerning the weight. In a neural network, the input layer receives unprocessed data. This data is then passed through the hidden layers, where these layers don't just process the input; they refine it, extract meaningful patterns, and transform it into a more helpful representation. Activation functions like ReLU for non-linearity is applied. Dropout layers are sometimes used to prevent overfitting, randomly ignoring specific neurons during training to make the network more robust. Finally, the output layer generates the final predictions—such as landmark coordinates for facial recognition. It uses Softmax activation for classification, and for regression, it uses Linear activation.

7. RESULTS AND DISCUSSION

Here, the data collection encompasses a High-dimensional 3D facial anatomy examination through preliminary processing approaches like image correction and enlargement used to improve the model resilience. FCN framework includes diverse hidden layers leveraged for spatial feature extraction, ensuring cephalic soft tissue image handling. This RD gives an extensive valuation of the execution of the model by applying different assessment metrics, advancement in coaching and data collection deviations.

The input data are experimented with using the simulation software installed in the system with Windows OS, 11th Gen Intel i5-processor, and 800 GB RAM. The input data are collected from the publicly available Kaggle dataset [29], which includes 7049 rows with 31 columns. The overall input data is split into train and testing. For training, 7049 images are used, and 1783 images are used for testing.

Fig.3. represents a dataset of grayscale facial figures, likely used for facial recognition, expression analysis, or landmark detection. Each face has key landmarks marked with small green dots, indicating crucial points such as eyes, nose, mouth, and jawline. The dataset includes diverse individuals with facial features, expressions, and head positions. Some subjects wear glasses, while others have varying levels of facial hair, which introduces variations in the dataset. The figures appear to be captured under controlled lighting conditions, ensuring consistency for model training. Such datasets are commonly used in DL applications for facial feature detection, emotion recognition, or biometric verification.

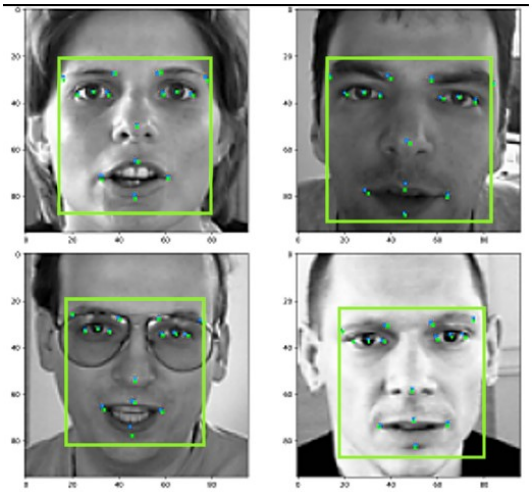


Fig.4. Face Detection

Fig.4. indicates the output of a face detection system. Each face is marked with a green box, showing that the system has detected the faces correctly. There are also blue dots on key facial features like the eyes, nose, and mouth, known as facial landmarks. This dataset is often used to train and test models for facial recognition, emotion detection, or video conferencing tools. The different backgrounds, lighting, and poses help the system learn to detect faces in various real-world situations.

Fig.5 Facial key point plotting

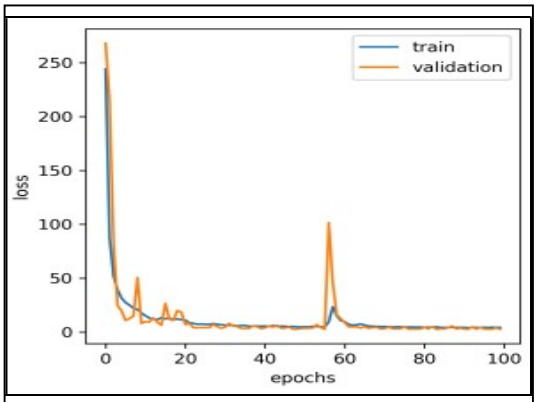
Once the input face is detected, the facial image's key point region is plotted, as shown in Fig.5. The

Journal of Theoretical and Applied Information Technology
15th May 2025. Vol.103. No.9

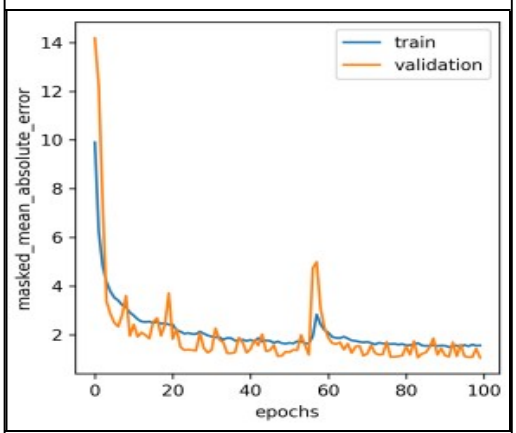
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

proposed model analyses the facial features based on the number of points plotted in the different areas of input facial images. the plotted regions are compared with the stored geometrical feature points value. If both predicted and actual values are the same, then the proposed model segmented the predicted region into multiple colours.

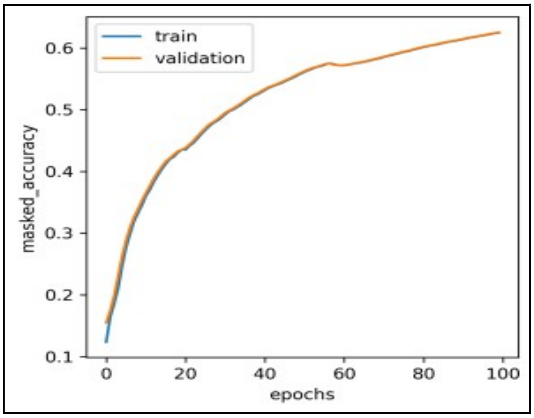
Fig.6 depicts the predicted multiple facial features in different colours. Based on the colour code, the proposed model analysed and identified landmarks like green (noses), red (eyes), yellow (mouths), and blue (eyebrows). The proposed approach can easily recognize the different types of facial images, poses, and signs by tracking these key points from the input facial images.



(a). Epochs Vs. Loss [Training / Validation]



(b) Epochs Vs. MMA Error [Training / Validation]



(c). Epochs Vs. Accuracy [Training / Validation]

Fig.7. Training Output

The loss graph shown in Fig.7(a) indicates the effectiveness of the presented model in analysing the actual values from the input dataset. Initially, the loss is very high, but it quickly drops as the model learns. Around epoch 60, there is a small spike, possibly due to learning rate changes or fluctuations in training data. After that, the loss stabilizes. Figure-7(b) shows masked mean absolute error (MMAE) measure Vs different epochs. Like the loss graph, the error starts high, decreases over time, and has a small spike around epoch 60 before stabilizing. Fig.7(c) depicts the accuracy measures and how well the model makes correct predictions. It starts very low but steadily increases as the model learns. By the end of training, accuracy was much higher, showing that the model had improved significantly.

Table-1. Performance Comparison – Different Epochs

Epochs	Total Images	Time	MMSE	MMAE	Accuracy
20	7049	43min	-	-	-
100	7049	15 min	4.3127	1.584	0.5411
100	11329	25 min	3.7041	1.4212	0.6283
100	17749	40 min	4.2459	1.5622	0.6248

Table 1 compares the model's performance on working with different numbers of epochs based on training time, dataset size, error values, and accuracy. For 20 epochs, 7,049 images were analysed within 43 minutes, but no accuracy or error values were recorded. For 100 epochs, in 15 minutes, reaching 0.5411 accuracies with error values of 4.3127 (MMSE) and 1.584 (more). Then, using the same 100 epochs, 11,329 images were applied and evaluated; it took 25 minutes to train, improving accuracy to 0.6283 while lowering errors to 3.7041 MMSE and 1.4212 more. Similarly, 17,749 images are applied and evaluated with 100 epochs; it takes 40 minutes but slightly drops in accuracy to 0.6248, with error values of 4.2459 mm and 1.5622 more. Overall, training longer and using more data helped improve accuracy.

During the training, it indicates a primary high damage considerably reduced in due course, reflected in the enhanced learning. Around 60 epochs were taken as a limited growth, which recommends instability because of the learning charge regulations or difference in data collection; however, the loss was substantiated. The precision intensifies constantly by affirming the cognitive effectiveness of the model. MAE (Mean Absolute Error) estimates the total variations between forecasted and true values to ensure a persistent assumption. RMSE (Root Mean Squared Error) estimate the medium volume of the uncertainty with smaller values, which show perfect accuracy. SSIM (Structural Similarity Index) evaluates the integrity of modified images, managing the anatomic specifics and assuring elevated constructive authenticity. In addition, the graphic display of the face portrait emphasizes facial points mapped with green dots, which are utilized for face prediction, phrase evaluation and biometrics. The facial recognition system successfully determines the facial anatomy beneath contrasting conditions, which exhibits flexibility in a real-time application and includes forensic analysis, medical reasoning, and video-centric techniques.

8. CONCLUSION

The Fully Connected Neural Network (FCNN) model is a powerful tool for predicting the geometric features of 3D soft tissue facial images. Unlike traditional image processing techniques that struggle with complex anatomical variations, this model leverages deep learning and high-quality 3D face scans to improve accuracy in applications like forensic reconstruction, craniofacial surgery, and biometric identification. Performance evaluations using RMSE, MAE, and SSIM confirm that the FCNN model outperforms other deep-learning approaches by minimizing errors and preserving fine facial details. Its ability to automatically extract features without relying on handcrafted methods makes it highly effective for real-world applications, including medical diagnostics and aesthetic surgery planning. However, its dependence on high-resolution scans and significant computational power poses challenges for resource-limited environments. Future advancements could optimise the model for real-time performance, expand datasets for better generalization, and integrate it with other deep learning architectures to enhance accuracy. Despite its limitations, the FCNN model represents a significant step forward in precise and reliable 3D facial geometry prediction.

9. FUTURE WORK

Data collection extension to incorporate the various cephalic structures and the native mutations. Reduction in quantitative sophistication entitles realistic time refinement for medical and forensic implementation. Merging the FCNN with transformer models or with CNN to upgrade pattern recognition. Scrutinizing its utility in gaming activity, artificial reality, and man-machine interfacing. This model is applied directly to the patient's data in the healthcare industry to test and ensure its sensible relevancy.

REFERENCES

[1] Kunz, A., Miesenberger, K., Mühlhäuser, M., Alavi, A., Pölzer, S., Pöll, D., ... & Schnelle-Walka, D. (2014). Accessibility of brainstorming sessions for blind people. In *Computers Helping People with Special*

Needs: 14th International Conference, ICCHP 2014, Paris, France, July 9-11, 2014, Proceedings, Part I 14 (pp. 237-244). Springer International Publishing.

[2] Van Kleef, G. A. (2009). How emotions regulate social life: The emotions as social information (EASI) model. *Current directions in psychological science*, *18*(3), 184-188.

[3] Hess, U. (2021). Who to whom and why: The social nature of emotional mimicry. *Psychophysiology*, *58*(1), e13675.

[4] Mukhamadiyev, A., Khujayarov, I., Djuraev, O., & Cho, J. (2022). Automatic speech recognition method based on deep learning approaches for Uzbek language. *Sensors*, *22*(10), 3683.

[5] Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional expression: Advances in basic emotion theory. *Journal of nonverbal behavior*, *43*, 133-160.

[6] Park, J. C., Lee, J., Lim, H. J., & Kim, B. C. (2018). Rotation tendency of the posteriorly displaced proximal segment after vertical ramus osteotomy. *Journal of Cranio-Maxillofacial Surgery*, *46*(12), 2096-2102.

[7] Perillo, L., Esposito, M., Caprioglio, A., Attanasio, S., Santini, A. C., & Carotenuto, M. (2014). Orthodontic treatment need for adolescents in the Campania region: the malocclusion impact on self-concept. *Patient preference and adherence*, 353-359.

[8] Mun, S. H., Park, M., Lee, J., Lim, H. J., & Kim, B. C. (2019). Volumetric characteristics of prognathic mandible revealed by skeletal unit analysis. *Annals of Anatomy-Anatomischer Anzeiger*, *226*, 3-9.

[9] Zhou, S., & Xiao, S. (2018). 3D face recognition: a survey. *Human-centric Computing and Information Sciences*, *8*(1), 35.

[10] Ludlow, J. B., & Ivanovic, M. (2008). Comparative dosimetry of dental CBCT devices and 64-slice CT for oral and maxillofacial radiology. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, *106*(1), 106-114.

[11] Dindaroğlu, F., Kutlu, P., Duran, G. S., Görgülü, S., & Aslan, E. (2016). Accuracy and reliability of 3D stereophotogrammetry: a comparison to direct anthropometry and 2D photogrammetry. *The Angle Orthodontist*, *86*(3), 487-494.

[12] Heike, C. L., Upson, K., Stuhaug, E., & Weinberg, S. M. (2010). 3D digital stereophotogrammetry: a practical guide to facial image acquisition. *Head & face medicine*, *6*, 1-11.

[13] Lee, S. H., Kil, T. J., Park, K. R., Kim, B. C., Kim, J. G., Piao, Z., & Corre, P. (2014). Three-dimensional architectural and structural analysis—a transition in concept and design from Delaire's cephalometric analysis. *International journal of oral and maxillofacial surgery*, *43*(9), 1154-1160.

[14] Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, *18*(8), 500-510.

[15] Rozenwald, M. B., Galitsyna, A. A., Sapunov, G. V., Khrameeva, E. E., & Gelfand, M. S. (2020). A machine learning framework for the prediction of chromatin folding in Drosophila using epigenetic features. *PeerJ Computer Science*, *6*, e307.

[16] Amrit, C., Paauw, T., Aly, R., & Lavric, M. (2017). Identifying child abuse through text mining and machine learning. *Expert systems with applications*, *88*, 402-418.

[17] Hossain, E., Khan, I., Un-Noor, F., Sikander, S. S., & Sunny, M. S. H. (2019). Application of big data and machine learning in smart grid, and associated security concerns: A review. *Ieee Access*, *7*, 13960-13988.

[18] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*, 11-26.

[19] Sun, Y., Wang, X., & Tang, X. (2015). Deeply learned face representations are sparse, selective, and robust. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2892-2900).

[20] Jeong, S. H., Yun, J. P., Yeom, H. G., Lim, H. J., Lee, J., & Kim, B. C. (2020). Deep Learning Based Discrimination Of Soft Tissue Profiles Requiring Orthognathic Surgery By Facial Photographs. *Scientific reports*, *10*(1), 16235.

[21] Hasani, B., & Mahoor, M. H. (2017). Facial expression recognition using enhanced deep 3D convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 30-40).

[22] Zhang, J., Yin, Z., Chen, P., & Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion*, *59*, 103-126.

[23] Cao, H., & Qi, C. (2021, December). Facial Expression Study Based on 3D Facial Emotion Recognition. In *2021 20th International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/SmartCNS)* (pp. 375-381). IEEE.

[24] Tu, X., Zou, Y., Zhao, J., Ai, W., Dong, J., Yao, Y., ... & Feng, J. (2021). Image-to-video generation via 3D facial dynamics. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*(4), 1805-1819.

[25] Sanil, G., Prakash, K., Prabhu, S., Nayak, V. C., & Sengupta, S. (2023). 2d-3d facial image analysis for identification of facial features using machine learning algorithms with hyper-parameter optimization for forensics applications. *IEEE Access*, *11*, 82521-82538.

[26] Nguyen, D. P., Ho Ba Tho, M. C., & Dao, T. T. (2021). Enhanced facial expression recognition using 3D point sets and geometric deep learning. *Medical & Biological Engineering & Computing*, *59*(6), 1235-1244.

[27] Gokulakrishnan, S., Chakrabarti, P., Hung, B. T., & Shankar, S. S. (2023). An optimized facial recognition model for identifying criminal activities using deep learning strategy. *International Journal of Information Technology*, *15*(7), 3907-3921.

[28] Chen, Y., Zhou, J., Gao, Q., Gao, J., & Zhang, W. (2023). MDNN: Predicting Student Engagement via Gaze Direction and Facial Expression in Collaborative Learning. *CMES-Computer Modeling in Engineering & Sciences*, *136*(1).

[29] https://www.kaggle.com/competitions/facial-keypoints-detection