*Nat. Volatiles & Essent. Oils, 2021; 8(5): 4252 - 4267*

# TFSAS2M: A novel transfer learning modelled approach for face-skull overlay via augmented salient-point-based scan mapping

**Sharma Tripti [1]* Dubey Sipi 2[2]**

[1] *Rungta college of Engineering & Technology Bhilai 9907995366*

[2] *Rungta college of Engineering & Technology Bhilai9406350006*

**Abstract:**

Overlaying skull CT (computer tomography) scans over facial data requires effective salient-point estimation & mapping. In order to perform this task, salient-points for both facial data & skull CT data are estimated, and their locations are mapped using correlative matching. This requires efficient modelling of face recognition algorithms, which can estimate facial location from both skull CT & image scans. The result of these algorithms is given to a feature extraction & selection unit, which estimates different facial salient-points via geometrical analysis. Correlation algorithms that match these points, perform mapping tasks between any 2 salient-point pairs without considering their original inter-dependencies. Due to which, CT scan of one person, can be easily mapped with facial image of another person, which limits the system's trustworthiness for real-time deployments. Moreover, most of the currently proposed salient-point mapping algorithms work with frontal facial and CT data, which further limits their deployment capabilities. Thus, in this text, a novel transfer learning model is proposed, which performs face-skull overlay via augmented salient-point-based scan mapping. The proposed model initially uses a deep convolutional neural network (DCNN) based on VGGNet-19 architecture, and trains it for facial & skull data separately. This network is evaluated on query images to validate the CT-to-face mapping, which is followed by augmented fusion. The augmented fusion model is responsible for selecting best matching CT scans from database, for any non-matching facial image data. Due to which, the model is able to achieve high accuracy, with better precision & recall performance when compared with state-of-the-art overlay models. The proposed TFSAS2M model was tested on various facial-CT interlinked datasets, and its performance was evaluated in terms of accuracy of facial-to-CT mapping, accuracy of overlay, precision of overlay, and recall of facial-to-CT mapping. Due to use of transfer learning & augmented salient-point mapping, the proposed model showcased 99.2% accuracy of facial-to-CT mapping, 97.4% accuracy for overlay, 94.8% precision for overlay, and 96.5% recall for facial-to-CT mapping, which makes the model suitable for real-time clinical usage. Moreover, this text also recommends some future research approaches, which can be used in order to improve efficiency of the proposed model.

**Keywords:** Facial, CT, mapping, salient-point, CNN, augmented, transfer, CNN

## 1. Introduction

Facial-skull overlay is a multidomain task, which involves face detection, feature extraction, salient-points extraction, feature selection, classification, and salient-point mapping. In order to perform these tasks, a wide variety of image processing models are proposed by researchers over the years. A review of these models is described in the next section of this, which indicates that deep learning and machine learning approaches outperform others in terms of overall accuracy of overlay. An example of such a system model can be observed from figure 1, wherein 2D landmarks, and 3D landmarks are mapped in order to obtain the final skull overlay.

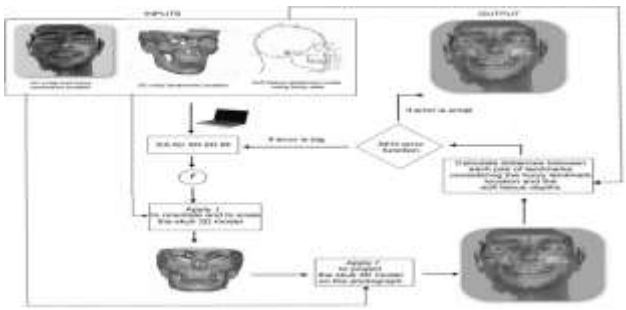

Figure 1. Multiple dimensions for face-skull overlay

From the described model it can be observed that fuzzy landmarks are used for evaluation of distances between each landmark pair. This distance is then compared with list of other distances in order to obtain the final mapping results. But these models do not take into consideration person face-to-person CT mapping before performing landmark matching, which limits their reliability, and deployment capabilities in real-time environment. In order to remove this drawback, section 3 proposes design of TFSAS2M, which is a novel transfer learning modelled approach for face-skull overlay via augmented salient-point-based scan mapping. The proposed TFSAS2M model evaluates salient-points between images that have same test subjects. This similarity is evaluated using a transfer learning model, that compares facial imagery & CT imagery with a database of pre-trained images & scans. Results of this mapping are given to a correlation engine, wherein salient-points are mapped with each other for final overlay. This section is followed by result evaluation, and performance comparison of the proposed model with various state-of-the-art approaches. Finally, this text concludes with some interesting observations about the proposed model, and recommends various ways to improve its performance.

## 2. Literature Review

Skull-face overlay models have matured over the years, and utilize a multitude of image processing operations including classification, post-processing, mapping, etc. The work in [1, 2, 3] proposes use of Artificial immune recognition system (AIRS) based Genetic Algorithm (GA), facial landmarks localization using fuzzy modelling, and recurrent convolutional models for overlays. These models evaluate errors during overlays, and aim at reducing them via feedback learning. Similarly, the models proposed in [4, 5, 6] provide guidelines, and propose methods that utilize distance metrics like Hausdorff distance, Euclidean distance, etc. for reducing this error further. The efficiency of these models is low, but it can be improved via the work in [7], wherein 3D skull model, and 2D face models are described using deep learning features. These features are mapped with facial points like the vertex, Glabella, Zygion, etc. for obtaining high mapping efficiency. Other models proposed in [8, 9, 10] aim at proposing feature extraction & selection models for classification of facial & CT data, and perform their mapping with a wide variety of datasets. Interestingly, the work in [11, 12, 13] propose define relationships between skull & face models, and model use of GA for highly efficient face-skull overlay. Moreover, the work in [14, 15] proposes use of thresholding technologies for high-efficiency face-skull overlay. Thus, it is observed that there is a large scope of research in the field of face-skull overlay, which can be used to improve its overall efficiency. Inspired by this, the next section proposes design of a novel high-efficiency skull-face overlay model, that utilizes augmentation, classification, and registration for improving precision, recall and accuracy of overlay process.

## 3. Design of the proposed novel transfer learning modelled approach for face-skull overlay via augmented salient-point-based scan mapping

From the literature survey, it is observed that a wide variety of models are proposed for skull-face overlap. But most of these models do not verify mapping between skull CT scans, and its corresponding facial imagery. Moreover, none of these algorithms mutate the input facial image to match the corresponding skull scan, and vice versa. Due to this, there is reduction in the efficiency & reliability of mapping. Thus, in this section a novel transfer learning modelled approach for face-skull overlay via augmented salient-point-based scan mapping (TFSAS2M) is proposed. The TFSAS2M model initially utilizes cascade object detection for extracting facial data from both CT scans and images. This facial data is given to a deep learning model for training & validation

purposes. During evaluation, both CT & facial scans are analyzed, and their respective user classes are evaluated. If these user classes are matching, then salient-points from both images are evaluated, and mapped using a correlation-based mapping model. If these classes are not matching, then a stored CT scan for the user is selected, and used for registration of the query CT image. The two images are fused using a registration overlay approach, and final CT image is obtained. The final CT image is then given to a salient-point analyser, and overlay process is performed. The entire process can be observed from figure 2, wherein data flow from different components is visualized.

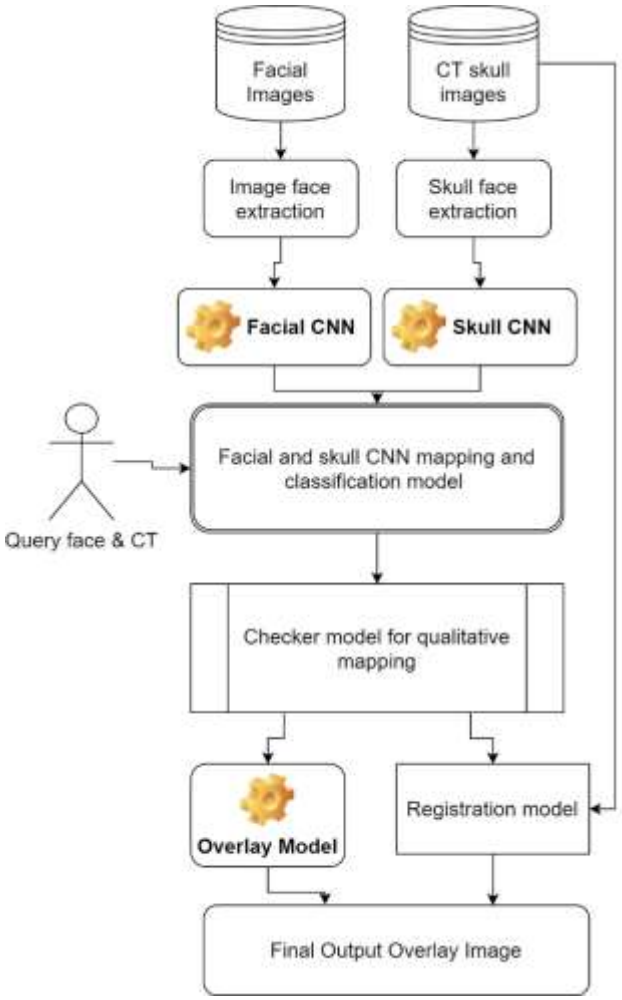


Figure 2. Model for the proposed TFSAS2M method

The proposed TFSAS2M model's description can be divided into 3 different parts, which include, design of facial & skull classification CNN, classification checker & registration model, and final overlay model. Each of these models are described in different sub-sections of this section.

### *3.1. Facial & skull classification CNN model*

From figure 2 it is observed that both facial & skull images are given to a CNN model for classifier training. The trained classifier is then used for validation & classification of query images. But, before going to the CNN model, both the images are given to a customized Haar cascade model for facial data extraction. The architecture for the Haar cascade model which is applied to both facial and CT data can be observed from figure 3, wherein Haar feature pool is combined with multiple stages for obtaining the final output image.

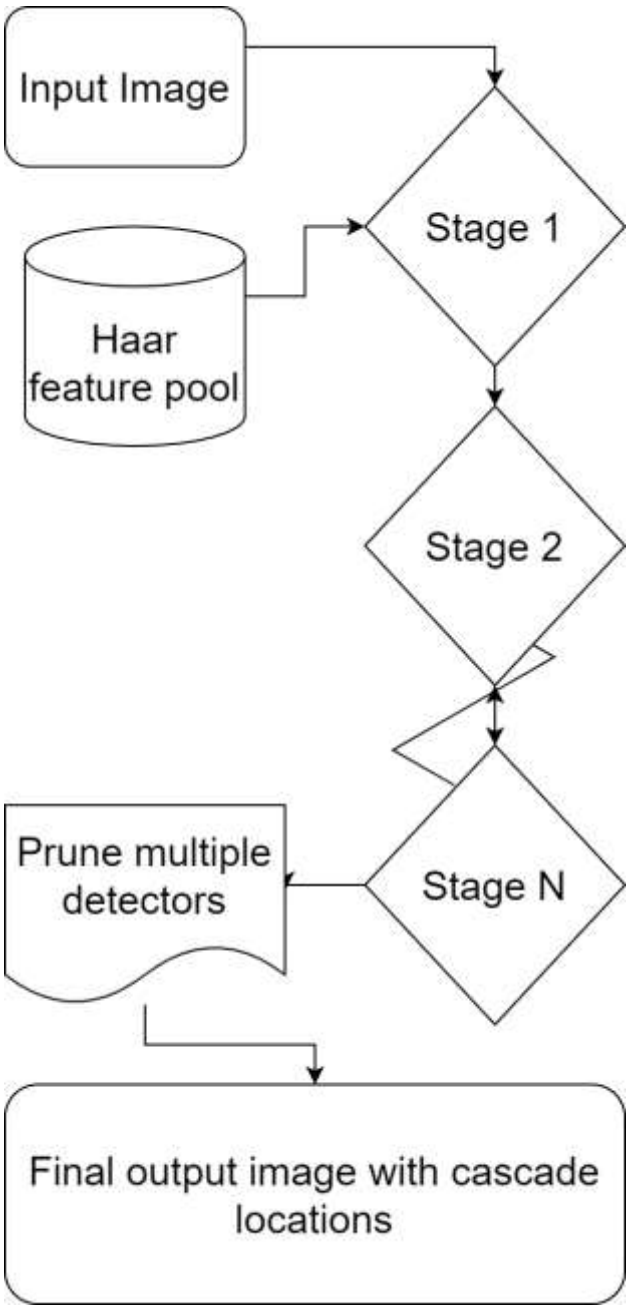


Figure 3. Haar cascade model for facial data extraction

The model utilizes customized Haar cascade configurations in order to estimate facial locations for both the images. In order to perform this task, the following process is followed,

- Skull CT images are manually segmented, and their image data is extracted.
- The extracted data is mapped with original image data.
- This mapping is done via extraction of internal facial components like eyes, mouth, nose, chin, & ears; and then mapping positions of these components with original facial image.
- A difference threshold is evaluated for each of these components using equation 1,

$$DTh_i = \sum_{j=1}^{N_c} sgn([x_i, y_i] - [x_j, y_j]) \frac{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}{N_c} \dots (1)$$

Where, $DTh, x, y, sgn\ and\ N_c$ represents difference threshold, position of the given component, sign of difference and number of facial components detected by the classifier. The difference threshold is evaluated for each facial component, and then Haar features are modified for each component using equation 2,

$$M_{h_i} = S_{h_i} + DT_i \dots (2)$$

Where, $M_h, and\ S_h$ represents modified Haar features, and selected Haar features for the given component. These selected Haar features are referred from the original adaboost model proposed by Rainer Lienhart. This model is used to extract facial data from both image & CT scans, and extracted information is given to a customized CNN model. The proposed CNN model is designed using a customized version of VGGNet-19 architecture, and can be visualized from figure 4 as follows, wherein initially a 1x32 layer convolutional model is combined with 2x2 max pooling model for coarse feature extraction. These features are augmented using the following 32x64, 64x128, 128x256, and 256x512 convolutional layers. These layers assist in extraction of a wide variety of features from both facial and CT scans.

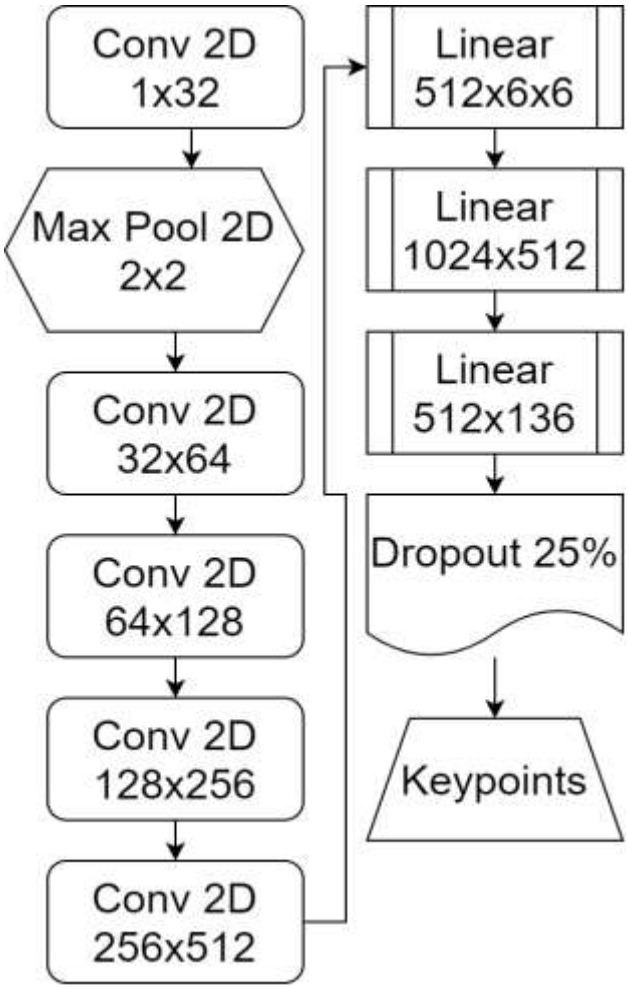


Figure 4. Design of the custom VGGNet-19 CNN model

The extracted features are given to a linear 512x6x6 layer, which performs feature selection using the following equation 3,

$$V_{avg} = \sqrt{\frac{\sum_{a=1}^{m}(x_a - \frac{\sum_{i=1}^{m}\sqrt{\frac{\sum_{j=1}^{n}(x_j - \frac{\sum_{k=1}^{n} x_k}{n})^2}{n-1}}}{m})^2}{m-1}} \dots (3)$$

Where, 'm' is the number of features in the current class, 'n' is number of features in the other class, and 'x' is the feature value. All features with variance less than $V_{avg}$ are removed from the training set, while others are kept for further checking. This process is repeated for 1024x512, and 512x136 sized linear layers, these utilize a combination of max pooling & dropout layers for finer feature selection. The result of this model is used for classification of both facial & CT image, and for extraction of salient-points. Thus, the model inputs facial image data & CT image data, converts them into multiple salient-points, and estimates the user-category of each input image.

### *3.2. Classification checking & CT image registration model*

Upon training the VGGNet-19 based CNN model, it is evaluated for test CT & facial images. This evaluation assists in estimating salient-points, and recognition of the person for which images are being evaluated. Upon evaluations, the system can be in one of two conditions,

- Both facial and CT images belong to the same person.
- Both facial and CT images belong to different persons.

In case 1, both images are directly given to section 3.3 for overlay, but for case 2, CT image of other person is fused with stored CT images of the identified person. This task is performed in order to match the facial and CT images, and thus assists in obtaining better mapping accuracy. To perform this task, the following process is followed,

- Let the CT skull query image be $CT_{query}$
- Let the person class identified by this query image be $P_{query}$
- Extract all CT images from the database, which belong to the $P_{query}$ class
- Find wavelet components for $CT_{query}$, and all the extracted images using equation 5,

$$W(r,c) = \frac{CT(r,c) + CT(r,c+1)}{2} \dots (5)$$

Where, $W, CT, r, and\ c$ represents wavelet components, CT image, row and column of the CT image. Rows and columns are up sampled by a factor of 2, in order to reduce image size by half. Thus, representing wavelet entropy for the input image. This process is repeated until 8x8 sized wavelet components are obtained. Thereby reducing number of features to 64x1 in size.

- The wavelet features of extracted CT images, are compared with query CT image, and correlation is estimated using equation 6 as follows,

$$C_{i,q} = \frac{\sum_{j=1}^{64}(f_{i,j} - f_{q,j})}{\sqrt{\sum_{j=1}^{64}(f_{i,j} - f_{q,j})^2}} \dots (6)$$

Where, $C_{i,q}$ represents correlation of the $i^{th}$ database image with query image, and $f_{i,j}$ represents $j^{th}$ wavelet feature of the $i^{th}$ image.

- Based on this estimation, database image with maximum value of correlation is used for registering with query image.
- Considering that the images do not need rotation, they are registered using equation 7 as follows,

$$\begin{bmatrix} Out_x \\ Out_y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & Q_x \\ 0 & 1 & Q_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} DB_x \\ DB_y \\ 1 \end{bmatrix} \dots (7)$$

where, $Q_x, and\ Q_y$ represents query image pixels in the x & y direction, while $DB_x\ and\ DB_y$ represents matched CT image from the database. The output registered image is given to section 3.3 for salient-point mapping, thereby assisting in improved overlay performance.

### *3.3. Overlay model using salient-point mapping*

Salient-points are mapped using a correlation model, which is activated after CT registration. Salient points are extracted using the customized CNN model described in section 3.1, wherein salient points for both skull CT and facial image are made available for mapping. This mapping is done using equation 8, wherein location of the points, their bounding boxes, and number of points are used.

$$F_s = MAX\left( \left| \frac{\sum_{j=1}^{N_p}(x_i - x_j)(y_i - y_j)}{\sqrt{\sum_{j=1}^{N_p}(x_i - x_j)^2 \sum(y_i - y_j)^2}} + \frac{\min(W_j, W_i)}{\max(W_j, W_i)} + \frac{\min(H_j, H_i)}{\max(H_j, H_i)} \right|_{i \in (1, N_p)} \right) \dots (8)$$

Where, $x, y, W, H, and\ N_p$ represents position of salient-points, bounding box of the points, and number of points respectively. Upon estimating point-wise correlation, its maximum value is used for estimation of mapping facial & CT images. Due to use of maximum value of correlation, the efficiency in terms of accuracy of mapping, accuracy of overlay, precision for overlay, and recall of mapping is improved. This performance is compared with various state-of-the-art models, and is discussed in the next section of this text.

## 4. Results and comparative analysis

The proposed model design utilizes a combination of deep networks, image registration, and correlative mapping. This combination assists in improving feature extraction efficiency, feature selection efficiency, classification performance, and efficiency of salient-point matching. In order to estimate performance of the proposed model, accuracy of mapping ($AMap$), accuracy of face-to-CT matching ($AMatch$), precision of face-to-CT matching ($PMatch$), and recall of mapping ($RMap$) are evaluated using equations 9, 10, 11, and 12 respectively.

$$AMap = \sum_{i=1}^{T_i} \frac{\sqrt{\left(x_i - x_{t_i}\right)^2 + \left(y_i - y_{t_i}\right)^2}}{T_i} \dots (9)$$

Where, $AMap, x_i, x_{t_i}, y_i, and\ y_{t_i}$ represents mapping accuracy, and locations of salient-points for ground truth image and obtained output image. Accuracy of matching is evaluated using equation 10 as follows,

$$AMatch = \frac{1}{N_i} \sum_{i=1}^{N_i} \frac{tp_i + tn_i}{tp_i + fp_i + tn_i + fn_i} \dots (10)$$

Where, $AMatch, N_i, tp, tn, fp\ and\ fn$ represents accuracy of matching, number of test images, true positive, true negative, false positive and false negative rate for the given image set. Here, true positive indicates Num. Images which are correctly classified and belong to the same skull CT category, true negative indicates Num. Images which are correctly classified, but belong to different skull CT category, false positive indicates images which are incorrectly classified, but belong to the correct CT category, and false negative indicates images which are incorrectly classified, and belong to different CT category. Similarly, precision of matching is evaluated using equation 11 as follows,

$$PMatch = \frac{1}{N_i} \sum_{i=1}^{N_i} \frac{tp_i}{tp_i + fp_i} \dots (11)$$

While, recall of mapping is evaluated using equation 12 as follows,

$$RMap = \sum_{i=1}^{T_i} \frac{\sqrt{\left(x_i - x_{other_i}\right)^2 + \left(y_i - y_{other_i}\right)^2}}{T_i} \dots (12)$$

Where, $x_{other}, y_{other}$ represents location of salient-points belonging to another category. Moreover, apart from this performance estimation, face-to-CT mapping was evaluated on both frontal face images & lateral face images. The results for lateral face imagery can be observed from figure 5, wherein lateral face image was mapped with skull CT image. Similarly, frontal face image mapping with its skull is observed from figure 6, wherein the efficiency of mapping is visualized.

Figure 5. Skull-face overlay for lateral images

Similar to figure 5, in figure 6 frontal face with skull overlay can be observed.

Figure 6. Frontal face with face-skull overlay

This visual efficiency was parameterized, and values for precision, recall, and accuracy of mapping & matching were estimated. These parametric results were compared with [2], [3], and [7] for algorithmic validation. During this validation, Num. Images were varied between 20 to 200, and results were tabulated in tables 1, 2, 3, and 4. The mapping accuracy can be observed from table 1 as follows,

| **Num. Images** | $\boldsymbol{AMap}$ **[2]** | $\boldsymbol{AMap}$ **[3]** | $\boldsymbol{AMap}$ **[7]** | $\boldsymbol{AMap}$ **[Proposed]** |
|---|---|---|---|---|
| 20 | 74.90 | 85.60 | 80.25 | 89.16 |
| 30 | 85.60 | 92.74 | 89.17 | 99.07 |
| 40 | 90.95 | 90.95 | 90.95 | 101.05 |
| 50 | 83.81 | 94.16 | 88.99 | 98.88 |
| 60 | 86.79 | 90.86 | 88.83 | 98.70 |
| 70 | 87.18 | 92.18 | 89.68 | 99.65 |
| 80 | 85.93 | 92.04 | 88.99 | 98.87 |
| 90 | 86.64 | 92.31 | 89.47 | 99.41 |
| 100 | 86.58 | 91.85 | 89.22 | 99.12 |
| 110 | 86.38 | 92.09 | 89.24 | 99.16 |
| 120 | 86.53 | 92.07 | 89.30 | 99.22 |

| | | | | |
|---|---|---|---|---|
| 130 | 86.50 | 92.08 | 89.29 | 99.21 |
| 140 | 86.48 | 92.02 | 89.25 | 99.17 |
| 150 | 86.50 | 92.06 | 89.28 | 99.20 |
| 160 | 86.49 | 92.06 | 89.28 | 99.20 |
| 170 | 86.49 | 92.05 | 89.27 | 99.19 |
| 180 | 86.50 | 92.05 | 89.28 | 99.19 |
| 190 | 86.49 | 92.06 | 89.28 | 99.19 |
| 200 | 86.49 | 92.05 | 89.27 | 99.19 |

Table 1. Mapping accuracy for different image sets

From table 1 and figure 7, It is observed that the proposed model is 13% better than [2], 6% better than [3], and 10% better than [7], which makes it suitable for real-time clinical applications. Accuracy of mapping for the proposed model is over 99% due to use of the modified deep learning CNN model.

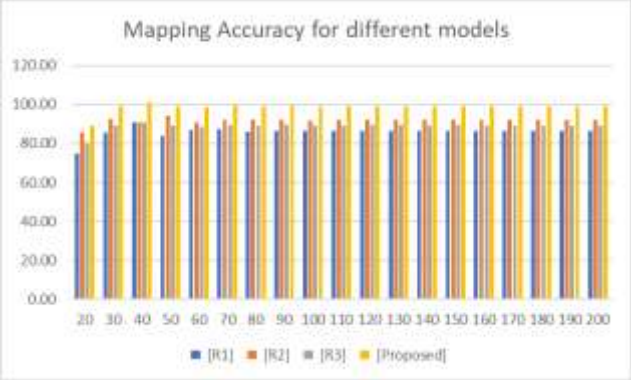


Figure 7. Mapping accuracy of different models

Similarly, the matching accuracy can be observed from table 2 as follows,

| **Num. Images** | $\boldsymbol{AMatch}$ **[2]** | $\boldsymbol{AMatch}$ **[3]** | $\boldsymbol{AMatch}$ **[7]** | $\boldsymbol{AMatch}$ **[Proposed]** |
|---|---|---|---|---|
| 20 | 72.45 | 82.80 | 77.63 | 87.56 |
| 30 | 82.80 | 89.70 | 86.25 | 97.29 |

| | | | | |
|---|---|---|---|---|
| 40 | 87.98 | 87.98 | 87.98 | 99.23 |
| 50 | 81.07 | 91.08 | 86.08 | 97.10 |
| 60 | 83.95 | 87.89 | 85.92 | 96.92 |
| 70 | 84.33 | 89.17 | 86.75 | 97.86 |
| 80 | 83.12 | 89.03 | 86.08 | 97.09 |
| 90 | 83.80 | 89.29 | 86.55 | 97.63 |
| 100 | 83.75 | 88.84 | 86.30 | 97.34 |
| 110 | 83.56 | 89.08 | 86.32 | 97.37 |
| 120 | 83.70 | 89.06 | 86.38 | 97.44 |
| 130 | 83.67 | 89.07 | 86.37 | 97.43 |
| 140 | 83.65 | 89.01 | 86.33 | 97.38 |
| 150 | 83.67 | 89.05 | 86.36 | 97.42 |
| 160 | 83.66 | 89.05 | 86.36 | 97.42 |
| 170 | 83.66 | 89.04 | 86.35 | 97.41 |
| 180 | 83.67 | 89.04 | 86.36 | 97.41 |
| 190 | 83.66 | 89.05 | 86.36 | 97.41 |
| 200 | 83.66 | 89.04 | 86.35 | 97.41 |

Table 2. Matching accuracy for different image sets

From table 2 and figure 8, it is observed that the proposed model is 14% better than [2], 8% better than [3], and 12% better than [7], in terms of matching accuracy, which makes it suitable for real-time clinical applications. Accuracy of matching for the proposed model is over 97% due to use of the modified deep learning CNN model.

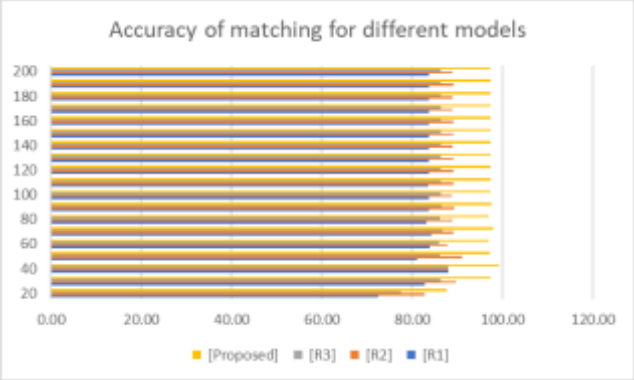


Figure 7. Mapping accuracy of different models

Similarly, the matching precision can be observed from table 3 as follows,

| **Num. Images** | $\boldsymbol{PMatch}$ **[2]** | $\boldsymbol{PMatch}$ **[3]** | $\boldsymbol{PMatch}$ **[7]** | $\boldsymbol{PMatch}$ **[Proposed]** |
|---|---|---|---|---|
| 20 | 72.45 | 82.80 | 77.63 | 85.26 |
| 30 | 82.80 | 89.70 | 86.25 | 94.74 |
| 40 | 87.98 | 87.98 | 87.98 | 96.63 |
| 50 | 81.07 | 91.08 | 86.08 | 94.55 |
| 60 | 83.95 | 87.89 | 85.92 | 94.38 |
| 70 | 84.33 | 89.17 | 86.75 | 95.29 |
| 80 | 83.12 | 89.03 | 86.08 | 94.54 |
| 90 | 83.80 | 89.29 | 86.55 | 95.07 |
| 100 | 83.75 | 88.84 | 86.30 | 94.79 |
| 110 | 83.56 | 89.08 | 86.32 | 94.82 |
| 120 | 83.70 | 89.06 | 86.38 | 94.88 |
| 130 | 83.67 | 89.07 | 86.37 | 94.87 |
| 140 | 83.65 | 89.01 | 86.33 | 94.83 |

| | | | | |
|---|---|---|---|---|
| 150 | 83.67 | 89.05 | 86.36 | 94.86 |
| 160 | 83.66 | 89.05 | 86.36 | 94.86 |
| 170 | 83.66 | 89.04 | 86.35 | 94.85 |
| 180 | 83.67 | 89.04 | 86.36 | 94.85 |
| 190 | 83.66 | 89.05 | 86.36 | 94.85 |
| 200 | 83.66 | 89.04 | 86.35 | 94.85 |

Table 3. Precision of matching for different image sets

From table 3 and figure 9, it is observed that the proposed model is 10% better than [2], 5% better than [3], and 8% better than [7], in terms of matching precision, which makes it suitable for real-time clinical applications. Precision of matching for the proposed model is over 94% due to use of the modified deep learning CNN model.

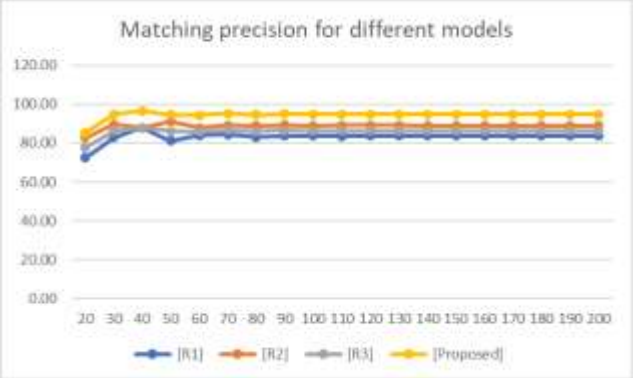


Figure 9. Matching precision of different models

Finally, the mapping recall can be observed from table 4 as follows,

| **Num. Images** | $\boldsymbol{RMap}$ **[2]** | $\boldsymbol{RMap}$ **[3]** | $\boldsymbol{RMap}$ **[7]** | $\boldsymbol{RMap}$ **[Proposed]** |
|---|---|---|---|---|
| 20 | 74.31 | 84.92 | 79.62 | 86.76 |
| 30 | 84.92 | 92.00 | 88.46 | 96.40 |
| 40 | 90.23 | 90.23 | 90.23 | 98.33 |

| | | | | |
|---|---|---|---|---|
| 50 | 83.15 | 93.42 | 88.28 | 96.22 |
| 60 | 86.10 | 90.15 | 88.12 | 96.04 |
| 70 | 86.49 | 91.45 | 88.97 | 96.97 |
| 80 | 85.25 | 91.31 | 88.28 | 96.21 |
| 90 | 85.95 | 91.58 | 88.77 | 96.74 |
| 100 | 85.90 | 91.12 | 88.51 | 96.46 |
| 110 | 85.70 | 91.37 | 88.53 | 96.49 |
| 120 | 85.85 | 91.35 | 88.60 | 96.55 |
| 130 | 85.81 | 91.36 | 88.59 | 96.54 |
| 140 | 85.79 | 91.29 | 88.54 | 96.50 |
| 150 | 85.81 | 91.33 | 88.57 | 96.53 |
| 160 | 85.80 | 91.33 | 88.57 | 96.53 |
| 170 | 85.80 | 91.32 | 88.56 | 96.52 |
| 180 | 85.81 | 91.32 | 88.57 | 96.52 |
| 190 | 85.80 | 91.33 | 88.57 | 96.52 |
| 200 | 85.80 | 91.32 | 88.56 | 96.52 |

Table 4. Recall of mapping for different image sets

From table 4 and figure 10, it is observed that the proposed model is 10% better than [2], 5% better than [3], and 8% better than [7], in terms of mapping recall, which makes it suitable for real-time clinical applications. Recall of mapping for the proposed model is over 96% due to use of the modified deep learning CNN model.

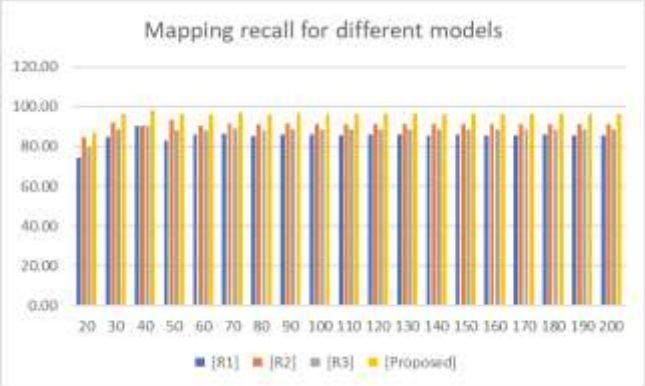


Figure 10. Mapping recall of different models

Thus, it is observed that the proposed model outperforms most of the recently designed models for face-skull overlay, which is mainly due to its efficient CNN design, face-to-CT matching, and image registration properties. These properties make the proposed model applicable for a wide variety of clinical and biomedical applications.

## 5. Conclusion & future scope

Most of the existing approaches used for face-skull mapping utilize deep learning models for salient-point extraction, and mapping. But these approaches do not perform any kind of image registration, due to which their performance in terms of accuracy & reliability is limited. The proposed TFSAS2M model utilizes a combination of feature augmentation, salient-point extraction, face-to-CT matching, CT registration, and correlation-based mapping in order to design a highly effective face-to-CT overlay architecture. The proposed model is capable of achieving 99.2% accuracy of facial-to-CT mapping, 97.4% accuracy for overlay, 94.8% precision for overlay, and 96.5% recall for facial-to-CT mapping, which makes the model suitable for real-time clinical usage. Moreover, the proposed model is 14% better than [2], 8% better than [3], and 12% better than [7], in terms of matching accuracy, and is 13% better than [2], 6% better than [3], and 10% better than [7], which makes it suitable for real-time clinical applications. Furthermore, it is observed that the proposed model is 10% better than [2], 5% better than [3], and 8% better than [7], in terms of matching precision, and is 10% better than [2], 5% better than [3], and 8% better than [7], in terms of mapping recall, which makes it suitable for highly precise applications such as remote imaging. The performance of this model in terms of accuracy, precision & recall can be further improved via use of transfer learning, Q-learning, and reinforcement learning methods, and can be tested on cloud deployments for better mobility.

## 6. References

[1] Natarajan, Yuvaraj & Kousik, N.V & Raja R, Arshath & Saravanan, M.. (2020). Automatic skull-face overlay and mandible articulation in data science by AIRS-Genetic algorithm. International Journal of Intelligent Networks. 1. 9-16. 10.1016/j.ijin.2020.05.003.

[2] B. R. Campomanes-Álvarez, O. Ibáñez, C. Campomanes-Álvarez, S. Damas and O. Cordón, "Modeling Facial Soft Tissue Thickness for Automatic Skull-Face Overlay," in IEEE Transactions on Information Forensics and Security, vol. 10, no. 10, pp. 2057-2070, Oct. 2015, doi: 10.1109/TIFS.2015.2441000.

[3] A. Valsecchi, S. Damas and O. Cordón, "A Robust and Efficient Method for Skull-Face Overlay in Computerized Craniofacial Superimposition," in IEEE Transactions on Information Forensics and Security, vol. 13, no. 8, pp. 1960-1974, Aug. 2018, doi: 10.1109/TIFS.2018.2806939.

[4] Posest-SFO: automatic skull-face overlay, https://skeleton-id.com/posest-sfo-automatic-skull-face-overlay/

[5] V. Mamone, V. Ferrari, S. Condino and F. Cutolo, "Projected Augmented Reality to Drive Osteotomy Surgery: Implementation and Comparison With Video See-Through Technology," in IEEE Access, vol. 8, pp. 169024-169035, 2020, doi: 10.1109/ACCESS.2020.3021940.

[6] Damas S., Cordón O., Ibáñez O. (2020) Craniofacial Superimposition: State of the Art. In: Handbook on Craniofacial Superimposition. Springer, Cham. https://doi.org/10.1007/978-3-319-11137-7_5

[7] An advanced scatter search design for skull-face overlay in craniofacial superimposition, https://www.sciencedirect.com/science/article/abs/pii/S0957417411011456

[8] Ogawa, Tetsuya & Nishimura, Kunihiro & Takahashi, Yasuhiro & Iwami, Kenichiro & Yasumura, Tsuneo & Yo, Kinga & Okamoto, Hiroki & Inukai, Daisuke & Sano, Rui & Watanabe, Tadashi & Kakizaki, Hirohiko. (2020). Incisionless facial resection for Kadish stage C olfactory neuroblastoma: Transcaruncular approach with combined endonasal and skull base surgery. Clinical Case Reports. 8. 10.1002/ccr3.2906.

[9] Ó. Ibañez, Ó. Cordon, S. Damas and J. Santamaria, "Modeling the Skull–Face Overlay Uncertainty Using Fuzzy Sets," in IEEE Transactions on Fuzzy Systems, vol. 19, no. 5, pp. 946-959, Oct. 2011, doi: 10.1109/TFUZZ.2011.2158220.

[10] Ubelaker, Douglas & Wu, Yaohan & Cordero, Quinnlan. (2019). Craniofacial photographic superimposition: New developments. Forensic Science International: Synergy. 1. 10.1016/j.fsisyn.2019.10.002.

[11] Yuadi, Imam & Artaria, Myrtati & Sakina, & Asyhari, Taufiq. (2021). Digital Forensics for Skulls Classification in Physical Anthropology Collection Management. Computers, Materials & Continua. 68. 3979-3995. 10.32604/cmc.2021.015417.

[12] Damas S., Cordón O., Ibáñez O. (2020) Relationships Between the Skull and the Face for Forensic Craniofacial Superimposition. In: Handbook on Craniofacial Superimposition. Springer, Cham. https://doi.org/10.1007/978-3-319-11137-7_3

[13] Bermejo, E., Álvarez, C.C., Valsecchi, A., Ibáñez, Ó., Arroyo, S.D., & García, Ó.C. (2018). Modelling mandible articulation for skull-face overlay in forensic identification.

[14] Bermejo, Enrique & Álvarez, Carmen & Valsecchi, Andrea & ibáñez, Oscar & Damas, S. & Cordon, Oscar. (2017). Genetic algorithms for skull-face overlay including mandible articulation. Information Sciences. 420. 10.1016/j.ins.2017.08.029.

[15] Damas S., Cordón O., Ibáñez O. (2020) Craniofacial Superimposition Techniques. In: Handbook on Craniofacial Superimposition. Springer, Cham. https://doi.org/10.1007/978-3-319-11137-7_4