Key Landmarks Detection of Cleft Lip-Repaired Partially Occluded Facial Images for Aesthetics Outcome Assessment

Paul Bakaki[1,4][0000−0001−8277−2554], Bruce Richard[2][0000−0002−4712−2860], Ella Pereira[1][0000−0002−6013−3935], Aristides Tagalakis[1][0000−0002−4610−0803], Andy Ness[3][0000−0003−3548−9523], Ardhendu Behera[1][0000−0003−0276−9000], and Yonghuai Liu[1][0000−0002−3774−2134]

[1] Faculty of Arts and Science, Edge Hill University, Lancashire L39 4QP, UK
{bakakip, pereirae, Aristides.Tagalakis, Beheraa, yonghuai.liu}@edgehill.ac.uk
[2] Birmingham Children's Hospital, Steelhouse Lane Birmingham B4 6NH, UK
brucerichard@blueyonder.co.uk
[3] British Dental School, University of Bristol, Bristol BS1 2LY, UK
Andy.Ness@bristol.ac.uk
[4] Department of Computer Science, Makerere University, P.O. Box 7062, Kampala, Uganda

Abstract. This paper proposes a novel method for the detection of the symmetrical axis of the cropped face required for the aesthetic outcome estimation from the facial images of patients after their cleft treatment. It firstly applies the Gaussian filter to smooth the images in order to compress noise on the subsequent tasks, then the bilateral semantic segmentation network is applied to segment the facial components out and each region is assigned a distinct colour, thirdly the Canny edge detector is applied to detect the facial feature points and all the contours are further detected and classified into three thirds according to their height. Fourthly, the centres of mass of detected feature points on the contours and the average of all these centres are used to estimate four potential symmetrical axes of the face, the one with minimum Manhattan distance from all the detected feature points is finally selected as the optimal one and used to estimate the aesthetic numerical score through the shape analysis in structural similarity measure. The experimental results based on a publicly accessible dataset shows that it performs well and better than one existing method.

Keywords: Cleft · Facial image · Aesthetic outcome estimation · Symmetrical axis · Shape analysis.

1 Introduction

Cleft lip (CL) is one of the most common maxillofacial congenital malformations with high surgical treatment costs [1]. Computational studies aim to demonstrate the potential for objective outcome of aesthetic assessment following cleft

surgical treatment. The overall intention is to aid audit of the various surgical practices, by encouraging only those with better aesthetic outcome to surgically treat cleft lip. This has the potential to influence most practitioners to adhere to the set standard surgical guidelines for cleft-related treatment [2]. The mouth lip beauty is a targeted outcome measure. The obvious distortion of the lip morphology hinders detection and identification of key features, considered essential for beauty. The features depicted from the facial aesthetic outcome significantly aid towards categorization as success or failure of a cleft repair. Eventually, this aids any audit of different cleft repair practices by assessing the beauty of the mouth lips [3] [4].

Determining facial features in images/videos is predominantly premised on a detected face [5] [6]. Therefore, face detection is a major component of facial feature identification studies. Facial anonymization of aesthetic outcomes is a convention for cleft lip related studies [7]. It is logically commendable and ethically a best practice for unbiased outcome assessment audit of different practices. Anonymization obstructs biased human assessment from any eye colour, ears shape, hair etc, unlike computer-based assessment [7] [8]. Consequently, the images used during outcome assessment bear significant partial occlusion.

Occlusion in computer vision started in the 1960s when Guzman proposed to detect faint lines in polyhedral drawings [9]. Consequently, it has been a subject of study in computer vision for detection of hidden facial features using convolutional neural networks with an attention mechanism [10]; facial appearance and shape learning to robustly detect facial features using an occlusion-adaptive deep network [11]; and cascaded pose regression (CPR) [12] [13].

Some of the facial image features of significant importance include inner eye corners (i.e. inner canthus, lacrimal punctum and inner canthal distance), nose features (tip, ala, root, and nasal base) and mouth features (upper/lower lip vermillion, oral commissure, vermillion border) [14] [15]. Presence of these features in the aesthetics symbolizes beauty. Therefore, computer vision tools aim to detect and locate these features, hence assess beauty using symmetry and other suitable shape defining parameters [16].

Most facial features occur in group or pairwise classification and can be used to determine the symmetric or asymmetric nature of a facial aesthetic. Key mouth features are used in [17] to determine the symmetry axis from which shape analysis is applied for facial aesthetic assessment.

Scars and other skin residues from surgical repair and photography effects can naturally cause features occlusion and influence aesthetic outcome assessment. Given this fact, deep learning techniques and regression studies have registered success regarding feature detection. Our approach disregards face detection because all facial aesthetics in use are anonymized. This study proposes a deep learning-based approach for the detection of facial features from cropped images for the analysis of their aesthetic outcome.

Dollár, Welinder and Perona [13] studied occluded facial landmarks detection using a statistical regression analysis framework. A facial image is partitioned into nine equal portions with anticipated landmarks positions. This study has

been applied to normal facial images in several datasets such as WFLF [18] and COFW [12]. A more robust approach (RCPR) introduced in [12] operates under difficult occlusion with the intention to improve the performance in [13].

So far, all related studies reviewed assume that occlusion is casually created using external objects such as spectacles, caps, hair styling etc. This study introduces and investigates a unique case of CL aesthetics where occlusion originates from the surgical treatment procedure and ethical norms.

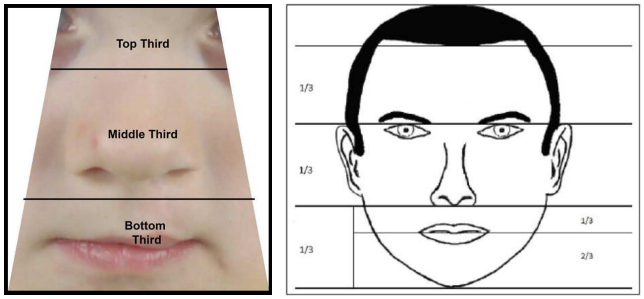

Fig. 1. Illustration and adaptation of horizontal thirds under occlusion of a cropped face (left) and full face (right).

2 Methodology

The approach taken aims to detect as many feature points as possible. Additionally, the objective is to classify and group the detected feature points in the three apparent segments of the facial aesthetics: upper third, middle third and bottom third, represented as the periorbital region, nose region and lips/oral region respectively [3] [19] (Fig. 1).

Successful categorization of these features plays a crucial role in the determination of the most befitting symmetric axis of the face. To this end, a deep learning-based method is proposed to detect the facial feature points of interest in the three regions: pre-processing, feature detection, symmetrical axis estimation and numerical score estimation that are detailed below.

1. Pre-process the aesthetic image using an appropriate filter. Filters have an enhancement and smoothing effect to facilitate generation of better segmentation results [20]. Several filters such as Gaussian, Laplacian of Gaussian, median and the others [20] can be used for this purpose. However, given the nature of our dataset's aesthetics, a 3×3 Gaussian filter was the best choice because image pixels are evenly distributed despite any image degenerative conditions with each element set to 1. Without the Gaussian filtering, less features are detected. Other filters can be designed. Examples of other filters are: mean $f = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ and a vertical Sobel filter: $f = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$ etc. Some of the outputs from these filters are presented in Fig. 2.

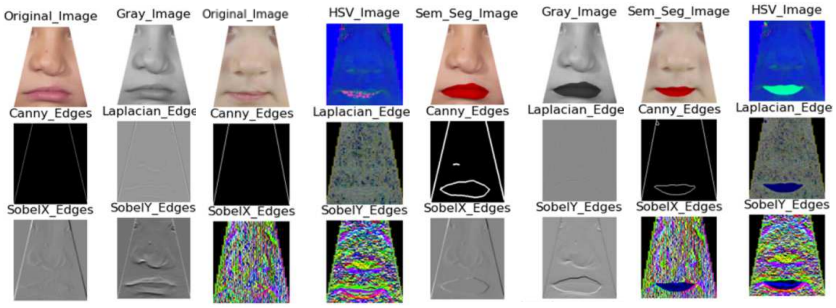

Fig. 2. Different filters used to visualize features for potential classification. Left four columns show that facial features are not clearly localized. Right four columns show clearer features from the same filters after segmentation using an ML approach.

Fig. 3. Segmentation results. Mouth region properly detected in all. From Left to Right: first - right eye corner not detected, fourth - all eye corners not detected.

2. Separate the salient regions through semantic segmentation because facial images present segmentation challenges using ordinary techniques due to low contrast [21]. Inner canthus and oral region features are most salient. Fig. 3 shows that not all the key regions will be detected due to poor anonymization procedures. Further, skin colour tone and scars from surgical treatment complicates the detection of any features [22]. The nose region is not segmented semantically but through any edges that may be detected. Consequently, the head orientation during photo-taking (either looking straight or downwards) influences the detection of the nose region edges and feature points following luminance contrast. Both the bilateral semantic network segmentation algorithm [23] and high-resolution network segmentation [24] produce appropriate segmentation results. The former is a faster and less resource intensive approach. We utilize the detailed module and semantic module of the bilateral segmentation network to acquire the image low/high level features and the semantics of each pixel, respectively. The two modules are combined through a real-time fusion module. The outcome is a clearly segmented mouth region, and the eye canthus, where possible. Fig. 3 shows the red-segmented mouth region and red-segmented eye canthi (*middle two*). For the different scenarios, as discussed in the results section, different features are therefore considered as inputs to the top network layers.
   Segmentation therefore aids the detection of the mouth region and the inner canthi, but it usually completely missed the nose region. The nose was not considered a key CL surgical outcome, hence excluded from training the seg-

mentation network. Besides, the baseline dataset for the bilateral network is annotated to incorporate the nose as a whole, [32], without the two nostrils. Eventually, the two nostrils would be considered significant to determine the symmetry of the partially occluded faces. To this end, we propose to apply Canny edge detection [25]. This further results into more feature points with higher accuracy for detection of the mouth and eye regions, Fig. 4, *middle*. Each of the regions should have feature points to aid with symmetry detection. For eyes, the interest lies with the closest inner canthus distance and the median distance while for the mouth region, the philtrum, vermillion borders, and oral commissures are desirable. Within the nose region, the tip, nostrils and their base are of interest. Fig. 4, *left* shows that classification of feature points can be more complicated before segmentation. To find as many feature points as possible in each region, distinct colours are assigned for improved contrast, as illustrated in Fig. 4, *right*.

Fig. 4. Features identified per horizontal partition. Left: Largely disorderly without segmentation. Middle: Shows improved features mapping and detection after segmentation. Right: Classified per horizontal partition.

3. Because of the partially occluded and anonymised nature of dataset, skin residues etc, some feature points within the key regions are disconnected. Implying there is need to identify connected components for stability to aid the determination of the symmetry. This is aided by filtering the detected feature points using Canny edge detector with the lower and upper thresholds set as decrease and increase by 0.33 of the median of pixel intensities of the whole image. The integral features lie along the eyes corner/inner canthus, nose tips/base/root/nostrils and mouth boundary/ philtrum/oral commissures. Once the feature points are detected, different colours, other than red *(for the predicted set, PS)* and green for the *ground truth sets ($GT1$, $GT2$ and $GT3$)*, are assigned to the feature points per different horizontal third for visualization (Fig. 4, *right*).
4. Link the respective feature points as contours. Successfully determining the feature point's perimeter suggests presence of a closed area/contour. At this stage, all possible contours have been identified as re-sampled contours from fully connected shapes without self-intersection, following the library implementation of [26], Fig. 4, *Left*. Some contours may be very small but necessary for the location of the position of the features of interest. For ex-

ample, a detailed execution shows that the green feature points representing identification of the nose region features (Fig. 4, *right*) are actually more than visibly displayed.

5. Partition the feature points on the detected contours into the three horizontal thirds based on their heights and classify them into the respective horizontal third segment. Fig. 4, *Center and Right* illustrate the features' locations using a colour coding for each of the three horizontal thirds.

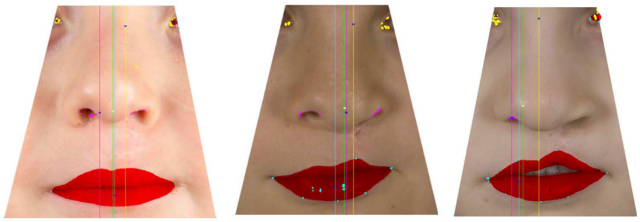

Fig. 5. Potential symmetric axes plotted based on component positions and their averages.

6. Determine the centres of mass of the contours in each of the three horizontal thirds. Draw a vertical line through each of the centres of the mass and also the average of these centres from the top to bottom boundary, yielding four potential symmetrical axes (Fig. 5).
7. Given four plotted potential axes of symmetry (Fig. 5), only one axis should be considered as an optimal one. Determine the Manhattan distance of each of the feature points from each of the potential symmetric axes through *Eq.* 1. Due to aggressive features detection, it has been experimentally proven that there are enough points from which to determine the symmetric axis.

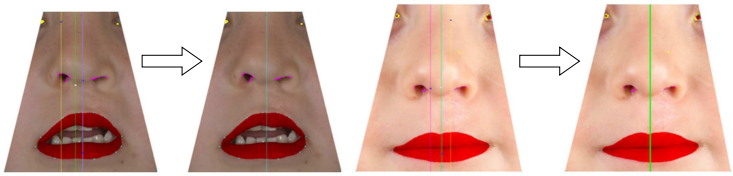

Fig. 6. Most suitable symmetric axis selected using average Manhattan distance.

$$dist(axis_k) = \sum_{j} \sum_{i=0}^{n} \mid axis_k - p_{ij} \mid, k = 1, 2, 3, 4 \tag{1}$$

where n is the number of the detected contours in the image, $axis_k$ is a potential vertical axis of symmetry, p_{ij} is a feature point j in the ith contour. The symmetric axis is finally determined as the one with the minimum distance (please see the green line in Fig. 5 and used in Fig. 6).

8. The determined symmetric axis is the basis for dividing the mouth/lip region into two sections to aid lip shape analysis [17] under different appropriate

scenarios. Lip shape analysis aims to determine how evenly/unevenly shaped the lip region is on either side of the symmetric axis. Generally, shape analysis has been applied to describe human perception features in medical images using contrast improvement ratio (CIR) [27] for example. A study by Loncaric [28] reviews other shape analysis techniques applied to images. In this study, a structural similarity measure [29] is preferred to automatically determine how agreeable the human visual perception is of the mouth lips, nose or a combination of both. Shape analysis is reduced to a structural comparison between the two mouth lip sides using the symmetric axis as a basis.

3 Results and Discussion

The distribution of the number of key feature points per horizontal third of the aesthetics from the public CCUK dataset is shown in Fig. 7. The four subcategories of the public dataset are: (i) Predicted Set (PS) – a set of aesthetics obtained through the proposed algorithm above, and (ii) three expert-generated Ground Truth sets $GT1$, $GT2$, and $GT3$. However, $GT3$ has not been considered in this study because it offers only a single-feature, the mouth/lip boundary. Ground truth sets are generated by manual annotation of the lip region using ImageJ, an open source software.

There exists more features in the upper and bottom horizontal thirds (blue and green, seen in Fig. 7 *left*), probably as expected. Fig. 7 *right* presents a more detailed breakdown of features points distribution per dataset subcategory per horizontal third.

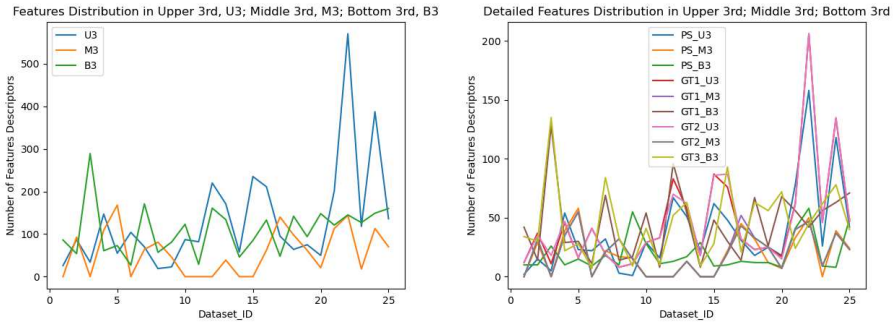

Fig. 7. The number of features detected across the 3 different horizontal upper, middle and bottom thirds: U3, M3 and B3 of each of the 3 considered sub-datasets.

Mouth lip and/or nose shape analysis was performed through three scenarios to assess the beauty of the aesthetic outcome of the surgical repair.
Scenario 1: Mouth region only (Fig. *8, Left*). The physical surgical repair to the cleft on the upper lip is usually taken, in consideration of its alignment with

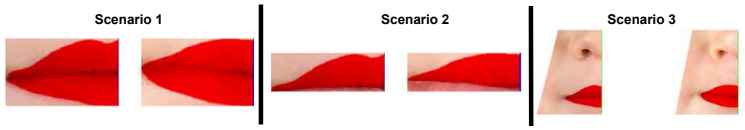

Fig. 8. Visualization of mouth region in Scenario 1 (left), upper lip in Scenario 2 (middle) and both nose and mouth regions in Scenario 3 (right).

the lower lip. Hence, the consideration of the whole mouth region is a natural occurrence when performing aesthetic outcome assessment. Similarity on either side of the symmetric axis through the mouth region is expected for features such as commissures and philtrum.
Scenario 2: The Upper lip, (Fig. *8, Middle*) is the actual region of the mouth that is surgically repaired. Therefore, it is a trivial and fair choice to investigate its structural feature (dis)similarity.
Scenario 3: Combination of the nose region and mouth/lips region (Fig. *8, Right*). Whereas cleft surgical repair is usually performed on the upper lip, human aesthetic outcome assessment naturally occurs with the awareness of other neighbouring features [30]. The closest feature available and applicable to our dataset is the nose region. It is almost trivial to observe any (mal)alignment between the nose and mouth.

These scenarios facilitate shape and structural computation and comparison using colour images, as presented before human assessors [31]. This also implies that the efficacy of our method can be determined by comparing the shape and structural computation with the human-generated numeric score *(HNS)* by human assessors. In [17], binary images for the lip/mouth region were used. After computing the structural similarity measure, s, it is converted to a numeric score between 1 and 5. Three (3) models, designed for conversion of s into a numeric score, are defined as follows: Model 1 ($M1$): $f(s) = 5(1-s^2)+s$, Model 2 ($M2$): $f(s) = \exp\left((1-s)\ln 5\right)$, and Model 3 ($M3$): $f(s) = 5 - \frac{4s}{(1+s)^{\frac{1}{100}}}$ where $0 \leq s \leq 1$ and $1 \leq f(s) \leq 5$.

Therefore, s is computed for each of the three subsets of the main dataset *(PS, GT1, GT2)* and then automatically converted into their respective numeric scores (PS_AENS, $GT1\_AENS$ and $GT2\_AENS$), from which correlation coefficients were calculated against *HNS*. The higher the coefficient, the more accurate the estimated aesthetic numeric score.

Since *PS* is determined automatically, its respective numeric scores PS_AENS are also automatically obtained. The correlation coefficient between *HNS* and PS_AENS is the most significant correlation (MSC) because it compares computer-generated numeric scores with human-generated numeric scores as seen in Table 1. The highest MSC is 23.6% (Model 1, Scenario 1). Overall, Scenario 1 also presents the best MSC. Further reduction of the region of interest (RoI) to study only the upper lip (Scenario 2) produces negative correlation results. It is a potential indication that this is not consistent with the practice and how to determine the region of interest requires further investigation.

Table 1. Correlation coefficients between the aesthetics scores estimated using different methods for different scenarios (S1, S2 and S3) and different models (M1, M2 and M3).

	PS_AENS vs HNS			GT1 vs HNS			GT2 vs HNS		
	S1	S2	S3	S1	S2	S3	S1	S2	S3
M1	0.236	-0.247	0.205	0.079	0.062	0.039	0.055	0.181	0.024
M2	0.200	-0.192	0.151	0.007	0.072	-0.033	-0.056	0.567	-0.102
M3	0.219	-0.226	0.176	0.041	0.052	-0.003	-0.011	0.457	0.056

Table 2. Correlation coefficients between the aesthetics scores estimated using different methods for different scenarios (S1, S2 and S3) and different models (M1, M2 and M3).

	PS_AENS vs GT1			PS_AENS vs GT2			GT1 vs GT2		
	S1	S2	S3	S1	S2	S3	S1	S2	S3
M1	0.809	0.560	0.811	0.854	0.653	0.850	0.903	0.888	0.906
M2	0.827	0.356	0.825	0.834	0.560	0.842	0.940	0.910	0.940
M3	0.836	0.654	0.833	0.856	0.456	0.856	0.924	0.908	0.928

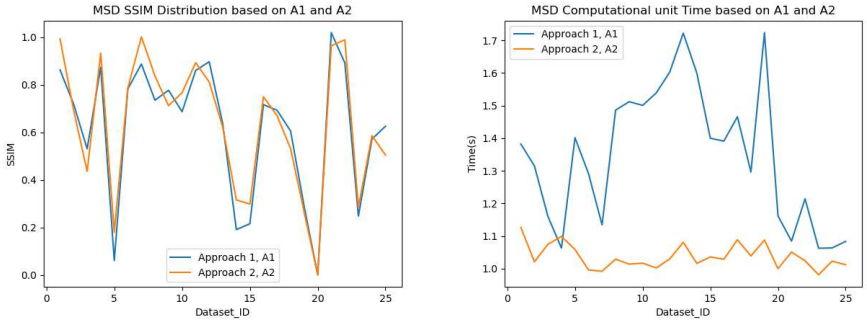

Fig. 9. SSIM (*left*) and computation time (*right*) of different approaches.

Besides, the proposed method has produced a highest correlation result of 94.0% between *GT1_AENS* and *GT2_AENS*, Model 2, Scenarios 1 and 3, as seen in Table 2. This is also a significant outcome, implying potential higher similarity in the datasets generated by the human experts (*GT1* and *GT2*).

Fig. 9 *(right, orange)* presents a comparative study between the proposed method (referred to as approach 2, A2) and the existing one (referred to as approach 1, A1) [17]. It shows that the latter usually generates lower structural similarity across the dataset than the former. This is because it uses more feature points and thus generates more accurate symmetric axes. Fig. 9 *(right, orange)* also shows that A2 takes shorter time than A1. It is an indicator that the whole face feature detection is faster than partial feature detection. Basically, it is easier to perceive and assess a whole face than its portion. Fig.10 compares feature detection using SIFT algorithm and A2, and shows how A1 and A2 map symmetric axes.

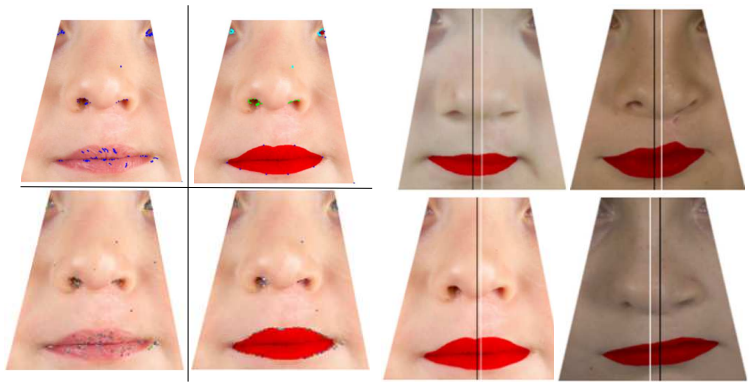

Fig. 10. Features (left two columns) detected using SIFT (Left column) and proposed approach (Second column). Then, symmetrical axis detection from some example cleft images (right two columns) using approach 1 (A1 - Black axis by [17]) and approach 2 (A2 - White axis by the proposed method)

4 Conclusion and Future Works

Detecting key feature positions requires an aggressive approach that combines deep learning and traditional approaches. For instance, a deep learning approach for segmentation combined with traditional edge detection, led to detection of more features (nose region especially) and feature points within the various regions of interest. Automatic structural comparison and analysis of colour aesthetics outcomes is more in harmony with human visual perception and judgement due to inclusion of luminosity and contrast features. This is represented by the consistent MSC results across the scenarios for the three models, M1, M2 and M3. Finally, anonymised and occluded facial aesthetics in our dataset have more features in the upper and lower horizontal third segments, implying that they may provide more potential for the estimation of the aesthetics from the cleft images. Other shape analysis techniques applicable to biomedical images as reviewed in [28] will be tested in future studies. The next step will be investigating purely deep learning techniques to extract, detect and, where necessary predict, facial aesthetic features specific to the public CCUK dataset.

Acknowledgments

The facial images are the cropped and anonymised anteroposterior (A/P) photos of 5-year-old children from the Cleft Care UK (CCUK). This publication presents data derived from the Cleft Care UK Resource (an independent study funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research scheme RP-PG-0707-10034). PB was funded by Graduate Teaching Assistantship, Edge Hill University; YL was partially funded by Shaanxi Province Key Research and Development Plan General Project-Industrial Field (2021GY-171).

References

1. Zhang, Q., Yue, Y., Shi, B., Yuan, Z.: A Bibliometric Analysis of Cleft Lip and Palate-Related Publication Trends From 2000 to 2017. Cleft Palate-Craniofacial J. 56, 658–669 (2019).
2. de Ladeira, P.R.S., Alonso, N.: Protocols in Cleft Lip and Palate Treatment: Systematic Review. Plast. Surg. Int. 2012, 1–9 (2012).
3. Hashim, P.W., Nia, J.K., Taliercio, M., Goldenberg, G.: Ideals of facial beauty. Cutis. 100, 222–224 (2017).
4. Kar, M., Muluk, N.B., Bafaqeeh, S.A., Cingi, C.: È Possibile Definire Le Labbra Ideali? Acta Otorhinolaryngol. Ital. 38, 67–72 (2018).
5. Hassaballah, M., Bekhet, S., Rashed, A.A.M., Zhang, G.: Facial features detection and localization. In: Studies in Computational Intelligence. pp. 33–59. Springer International Publishing (2019).
6. Reisfeld, D., Yeshurun, Y.: Robust Detection of Facial Features by Generalized Symmetry, (1992).
7. Lee, T.V.N., Ireland, A.J., Atack, N.E., Deacon, S.A., Jones, T.E.M., Matharu, J., Wills, A., Al-Ghatam, R., Richard, B.M., Ness, A.R., Sandy, J.R.: Is There a Correlation Between Nasolabial Appearance and Dentoalveolar Relationships in Patients With Repaired Unilateral Cleft Lip and Palate? Cleft Palate-Craniofacial J. 57, 21–28 (2019).
8. Shkoukani, M.A., Chen, M., Vong, A.: Cleft lip - A comprehensive review. Front. Pediatr. 1, 1–10 (2013).
9. Hoiem, D., Efros, A.A., Hebert, M.: Recovering occlusion boundaries from an image. Int. J. Comput. Vis. 91, 328–346 (2011).
10. Li, Y., Zeng, J., Shan, S., Chen, X.: Occlusion Aware Facial Expression Recognition Using CNN With Attention Mechanism. IEEE Trans. Image Process. 28, 2439–2450 (2019).
11. Zhu, M., Shi, D., Zheng, M., Sadiq, M.: Robust facial landmark detection via occlusion-adaptive deep networks. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2019-June, 3481–3491 (2019).
12. Burgos-Artizzu, X.P., Perona, P., Dollar, P.: Robust face landmark estimation under occlusion. Proc. IEEE Int. Conf. Comput. Vis. 1513–1520 (2013).
13. Dollár, P., Welinder, P., Perona, P.: Cascaded pose regression. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 1078–1085 (2010).
14. Hennekam, R.C.M., Cormier-Daire, V., Hall, J.G., Méhes, K., Patton, M., Stevenson, R.E.: Elements of morphology: Standard terminology for the nose and philtrum. Am. J. Med. Genet. Part A. 149, 61–76 (2009).
15. Hall, B.D., Graham, J.M., Cassidy, S.B., Opitz, J.M.: Elements of morphology: Standard terminology for the periorbital region. Am. J. Med. Genet. Part A. 149, 29–39 (2009).
16. Sharma, V.P., Bella, H., Cadier, M.M., Pigott, R.W., Goodacre, T.E.E., Richard, B.M.: Outcomes in facial aesthetics in cleft lip and palate surgery: A systematic review. J. Plast. Reconstr. Aesthetic Surg. 65, 1233–1245 (2012).
17. Bakaki, P., Richard, B., Pereira, E., Tagalakis, A., Ness, A., Liu, Y., Shape Analysis Approach Towards Assessment of Cleft Lip Repair Outcome. In International Conference on Computer Analysis of Images and Patterns, pp. 165-174. Springer, Cham, (2021).
18. Sagonas, C., Tzimiropoulos, G., Zafeiriou, S., Pantic, M.: 300 faces in-the-wild challenge: The first facial landmark Localization Challenge. Proc. IEEE Int. Conf. Comput. Vis. 397–403 (2013).

19. Erian, A., Shiffman, M.A.: Advanced surgical facial rejuvenation: Art and clinical practice. Adv. Surg. Facial Rejuvenation Art Clin. Pract. 1–740 (2010).
20. Frery, A.C.: Image Filtering. In: de Mello, C.A.B. (ed.) Digital Document Analysis and Processing. pp. 55–70. Nova Science Pub Inc, New York (2013). https://doi.org/10.1201/b10797-8.
21. Oliveira, R.B., Filho, M.E., Ma, Z., Papa, J.P., Pereira, A.S., Tavares, J.M.R.S.: Computational methods for the image segmentation of pigmented skin lesions: A review. Comput. Methods Programs Biomed. 131, 127–141 (2016).
22. Sandy, J., Kilpatrick, N., Ireland, A.: Treatment outcome for children born with cleft lip and palate. Front. Oral Biol. 16, 91–100 (2012).
23. Yu, C., Gao, C., Wang, J., Yu, G., Shen, C., Sang, N.: BiSeNet V2: bilateral network with guided aggregation for real-time semantic segmentation. arXiv. (2020).
24. Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., Liu, D., Mu, Y., Tan, M., Wang, X., Liu, W., Xiao, B.: Deep High-Resolution Representation Learning for Visual Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 43, 3349–3364 (2020).
25. Canny, J.: A Computational Approach to Edge Detection. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-8, 679–698 (1986).
26. Wu, S.-T., Silva, A.C.G. da, Márquez, M.R.G.: The Douglas-peucker algorithm: sufficiency conditions for non-self-intersections. J. Brazilian Comput. Soc. 9, 67–84 (2004).
27. Kimori, Y.: Morphological image processing for quantitative shape analysis of biomedical structures: Effective contrast enhancement. J. Synchrotron Radiat. 20, 848–853 (2013).
28. Loncaric, S.: A survey of shape analysis techniques. Pattern Recognition 31(8), 983-1001 (1998).
29. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 13, 600–612 (2004).
30. Deall, C.E., et al: Facial Aesthetic Outcomes of Cleft Surgery: Assessment of Discrete Lip and Nose Images Compared with Digital Symmetry Analysis. Plast. Reconstr. Surg. 138, 855–862 (2016).
31. Mosmuller, D.G.M., et al.: Scoring systems of cleft-related facial deformities: A review of literature. Cleft Palate-Craniofacial J. 50, 286–296 (2013).
32. Liu, Z., Luo, P., Wang, X', Tang, X., Deep Learning Face Attributes in the Wild, in IEEE International Conference on Computer Vision (ICCV), (2015).