Review

Review: Facial Anthropometric, Landmark Extraction, and Nasal Reconstruction Technology

Nguyen Hoang Vu 1, Nguyen Minh Trieu 2, Ho Nguyen Anh Tuan 3, Tran Dang Khoa 3 and Nguyen Truong Thinh 2,*

1 Human Anatomy Department, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City 700000, Vietnam
2 Department of Mechatronics, Ho Chi Minh City University of Technology and Education, Ho Chi Minh City 700000, Vietnam
3 Human Anatomy Department, Pham Ngoc Thach University of Medicine, Ho Chi Minh City 700000, Vietnam
* Correspondence: thinhnt@hcmute.edu.vn

Abstract: Facial anthropometrics are measurements of human faces and are important figures that are used in many different fields, such as cosmetic surgery, protective gear design, reconstruction, etc. Therefore, the first procedure is to extract facial landmarks, then measurements are carried out by professional devices or based on experience. The aim of this review is to provide an update and review of 3D facial measurements, facial landmarks, and nasal reconstruction literature. The novel methods to detect facial landmarks including non-deep and deep learning are also introduced in this paper. Moreover, the nose is the most attractive part of the face, so nasal reconstruction or rhinoplasty is a matter of concern, and this is a significant challenge. The documents on the use of 3D printing technology as an aid in clinical diagnosis and during rhinoplasty surgery are also surveyed. Although scientific technology development with many algorithms for facial landmarks extraction have been proposed, their application in the medical field is still scarce. Connectivity between studies in different fields is a major challenge today; it opens up opportunities for the development of technology in healthcare. This review consists of the recent literature on 3D measurements, identification of landmarks, particularly in the medical field, and finally, nasal reconstruction technology. It is a helpful reference for researchers in these fields.

Keywords: facial anthropometric; landmark extraction; nasal reconstruction technology

1. Introduction

Anthropometric data is the measurements of human size and shape, which are usually used to reconstruct a digital human morphology. The figures are usually used to identify the gender, ethnicity, or age of humans in different regions; moreover, they are also used in clinical diagnosis or as a reference for patients before and after surgery. Moreover, they are also used as reference parameters in the design of personal protective equipment. Reconstructive surgery is repairing congenital defects, damaged organs, or parts of the body destroyed after a trauma. This helps patients feel more confident in communication or is used in cosmetic surgery to increase facial beauty. In cosmetic surgery, doctors try to create 3D models that are used as a supported tool to determine clinical diagnoses and communicate with patients. The stages of building 3D images are created thanks to the techniques of the staff, which increases their work pressure. The support of machine learning algorithms needs to be applied to automate these processes; recently, the algorithms for identification, classification, or recognition have been introduced in many different fields, such as agriculture, medicine, transportation, etc. [1–3], to improve the quality of human life. Artificial intelligence (AI) was first introduced in 1956.
at Dartmouth College. It is a term for machine learning intelligence through learned weights, it is used in research about nature language networks and classifiers to solve real-world problems. Facial landmark (aka keypoint) detection is a necessary and first task in facial human studies [4]. Further, the information on the facial landmarks is also used in areas such as human–computer interaction, face recognition, and emotion recognition [5–7], and, in particular, it is applied in anatomical problems and reconstructing 3D images for clinical diagnosis [8]. The number of facial landmarks is determined differently based on its application. For example, Irtija et al. [9] detect fatigue based on 68 facial landmark extraction, Fabian Benitez-Quiroz. Ref. [10] identify 66 landmarks to recognize facial emotions, Yashunin et al. [11] determine 19 landmarks for face detection, etc., which is presented in Section 3.

In medical and cosmetic surgery, simulating procedures is necessary for clinical diagnosis, patient education, and medical education; however, the research on reconstructing parts of the human body is a significant challenge. Patient education impacts how patients behave and results in the influence of knowledge, attitudes, and skills required to preserve or enhance health [12]. This means that doctors can teach patients about their faces before and after surgery based on simulation models. The facial landmarks are the key points determined on a human face, which are specific, such as the eyes, nose, mouth, etc. [13]; moreover, they are used to calculate the anthropometric measurement. The difficulty of facial landmark detection is due to the location of the landmark points, which change in various poses combined with the fact that the face surface is not flat, thus, each point is captured by different light at different locations on the face, all of which creates noise that interferes with recognition. Facial landmarks are inputs of the different studies, such as face detection, doze detection, etc. In many different countries and regions, three-dimensional image analysis is a topic that has attracted interest and prompted the creation of anthropometric databases for different purposes; this is presented in Section 2. Doctors can rely on anthropometric data to make clinical decisions before surgery, however, this depends heavily on their experience. With the development of technology, many algorithms were proposed to recognize landmarks automatically, such as Active Appearance Models (AAM), Support Vector Machine (SVM), Regression Trees, Coarse-to-Fine Shape Searching (CFSS) [14], Fast Shape Searching Face Alignment (F-SSFA) [15], etc., or the deep convolutional neural network algorithm, which is in Section 3. Because the nose is one of the most attractive parts of the face [16,17], nasal reconstruction after trauma is extremely important as it helps patients to overcome low self-esteem when communicating. Moreover, rhinoplasty or nasal reconstruction is a fundamental operation, becoming increasingly popular and attractive [17]. The pre-operative clinical assessment depends on the doctor’s experience, and this job puts pressure on doctors and makes patient education difficult. The development of 3D technology reduces the difficulty of the aspects of rhinoplasty surgery as 3D images are able to provide a more objective perspective for patients and reduce the risk of surgery. Furthermore, 3D printed technology is popular in the market, therefore, making them easier to access at a low cost and to be easily invested in by agencies or individuals. The 3D models of the nose or nasal bone are reconstructed for different purposes, such as clinical diagnosis, educational purposes, and guiding doctors during surgery, they are presented in section 4. These studies open up a lot of potential in the field of reconstruction, and the material is the most important factor in creating human prosthetics. Reconstruction helps clinical usage at minimal cost and risk to the patients.

In this work, we review studies related to three main issues: 3D photographic analysis, facial landmark detection, and nasal reconstruction technology. The purpose of the article is to update, survey, and evaluate recent studies on the mentioned issues to provide help future studies with an overview in this field. In the previous reviews, the authors usually only review a certain topic for each application. This makes it difficult for the reader to have an overview of studies in this area, especially in medicine. In fact, there are many studies that achieved a high accuracy of landmark extraction, which is presented in
Section 3, but their applications differ from medical applications. This report discusses all three issues to provide readers with more clarity on the progress made within the field. Surveying 3D photographic analysis in countries or regions is presented in Section 2. Identifying landmarks from images automatically using methods of non-deep and deep convolutional neural networks is presented in Section 3. Section 4 presents nasal reconstructing technology based on 3D-image and 3D printed technology. Finally, the discussion and evaluation are described in Section 5.

2. Anthropometric Measurement Analysis of Different Regions

Anthropometric data are the measurements of a human's size and shape, which are usually used to construct a digital human model. They depend on various factors, such as sex [18–29], race [20,21,23,26,30–32], age [33,31], etc. In the 13th century, Marco Polo first described man during his circumnavigation of the world, from which the description and studies of the human body started to develop. In recent years, studies about people have attracted authors from many different countries. Anthropometric measurements are set up to facilitate the calculation and comparison of the sizes of body parts. In medicine, the determination of the location of landmarks always requires high accuracy for applications in reconstruction and clinical diagnosis to reduce costs and risks for patients or to use in medical education. As a result, most of the research was conducted for the purpose of collecting anthropometric data, landmarks were marked carefully and manually. Anthropometric data are used in clinical diagnosis as a reference to compare changes in patients with normal people. Many studies were conducted to create databases used by doctors in medicine and cosmetic surgery or to simply survey to design protective clothing for workers. For example, in 2005, Farkas et al. [30] conducted a study on facial anthropometric factors of participants from different countries. Four groups from 25 different countries and regions took 14 different facial measurements to determine differences in the anthropometric measurements of the participants. This can be considered the largest survey at that time, which took five years to complete with the support of the medical labs from countries around the world. However, the data used in the study was unbalanced as 53.1% of participants came from Europe, which was two-times that of all three groups of participants from elsewhere.

In another study, Kwon et al. [33] performed facial measurements of Kanwomen to demonstrate age-related facial morphology. In this study, 192 Korean women of different ages were measured with non-invasive facial measurements. Morpheus 3D light-emitting diode-based light scanner was used to collect images, and according to the authors, this is the first study to examine the relationship between age and anthropometric data. However, in 2010, a previous study by Zhuang et al. [31] performed measurements to demonstrate the influencing factors of anthropometric data, such as gender, ethnicity, and age. Their study surveyed 3997 participants with the aim of surveying the head sizes of workers and assisting in the design of protective products to be used in construction. Their results show that the factors of gender, ethnicity, and age group are statistically significant concerning anthropometric data. Because their research was conducted on American workers, the number of participants that were White Americans accounted for more than 47% of the total participants compared to other ethnicities, such as Hispanics, African-Americans, and others. Therefore, the strength of the conclusions about the differences in the measurements related to race is affected. Furthermore, the authors should be commented as this is a study with a large sample; it has great significance in anthropometrics. At the same time, Husein et al. [32] also studied Indian American anthropometric landmarks on 30 manual measurements from 3-direction (frontal, left, right) images of 102 participants. The authors used these anthropometric figures to compare with the data available for North American White women (NAW) surveyed in the study by Farkas et al. [30] and concluded that 25 of the 30 measurements were different. The sample sizes of the two comparisons are different, hence, it may affect the reliability of the conclusions. Anthropometric data of Malaysians were collected and published by Othman et al. [18].
They studied the facial morphology of Malay people between the ages of 20 and 30. They conducted a survey on 109 participants who were all ethnic Malay. The scope of the survey focused on patients who were in the area, which is not random in nature, thus, according to the authors, the study should have been conducted with a group of larger sample size. Twenty-three landmarks on the face were detected to measure the ocular dimension, nasal dimension, orolabial dimension, etc. Similar to this approach, Menéndez López-Mateos et al. [19] analyzed the faces of European adults with 100 participants. The conclusion was that in men the figures were higher than in women in vertical and transversal dimensions, however, the measurements of the upper lip and mandibular prominence were not different in men and women. The survey was conducted with the aid of modern equipment and measured on collected 3D images, however, landmarks and measurements were performed manually.

Furthermore, many surveys were carried out to collect anthropometric databases for regions or countries, including Virdi et al. [20], who surveyed the Kenyan-African anthropometric baseline measurement; anthropometric data of Italians and Egyptians by Celebi et al. [21]; anthropometric data for Chinese [22], Basrah [23], and Greek [24] with 31 landmarks; the study of Staka et al. [25] with 8 facial measurements on Albanians; 21 landmarks and 27 measurements were presented by Miyazato et al. [26], all with applications intended to build anthropometric databases applied in medicine, cosmetic surgery, and identification. In another study by Dong et al. [27], they presented the anthropometric measurements of the Chinese nose, which are useful documents for doctors in cosmetic surgery with 9 linear measurements, 3 angular measurements, and 7 proportions. Moreover, in some studies conducted to survey the aging of the face, such as the anthropometric study of Icelandic children [28], the boys of northeast Iran [29], etc., the purpose was to create an anthropometric database for various fields; studies in different regions and countries are summarized in Table 1. Measurements were performed manually with the aid of commercially available cameras such as the VECTRA-M5 360 camera (Canfield Scientific Inc, Fairfield, NJ, USA), Planmeca ProMax 3D ProFace® (Planmeca USA, Inc.; Roselle, IL, USA), 3D stereophotogrammetry system (3DSS-II; Shanghai Digital Manufacturing, Shanghai, China), or digital camera. Anthropometric data collected by the authors are commendable, however, most of the studies were conducted in a certain area or locality rather than randomly drawn from many places, hence, the other effects of the factors were not considered in these studies. Studies with larger sample sizes are always expected to strengthen the conclusions. In addition, natural or artificial influences (weather, nutrition, etc.) can also be factors affecting anthropometric indices. Most of the studies surveyed the identification of landmarks and anthropometric measurements manually. It increases work pressure for authors and its accuracy cannot be accurately assessed—it only depends on the experience of the operator. Therefore, with the current development of computer vision and machine learning, automatic identifications and measurements are studied, which is reviewed in Section 3. Table 1 summarizes the characteristic parameters for this survey, such as author, location, subjects, number of participants, year in which the study was published, surveyed subjects, and methods. Anthropometry of the Head and Face in Medicine was introduced by LG Farkas in 1981 and was updated in 1994 and consists of 21 facial landmarks [34]. This document has continued to develop and update the atlas as well as released datasets to add new measurements and normative data using a strict measurement technique developed based on the recognition of anthropometric landmarks. To this day, these landmarks and measurements are still used in most of the research in this field, especially in digital human reconstructs. Figure 1 presents facial landmarks of case studies and it is used as a common standard in anthropometric studies.
Table 1. Statistical tables of anthropometric data of the reviewed studies.

<table>
<thead>
<tr>
<th>N/C</th>
<th>Sample Size</th>
<th>Year</th>
<th>Age</th>
<th>Subjects</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farkas et al. [30]</td>
<td>Europe, the Middle East, Asia, Africa, and North America</td>
<td>1470</td>
<td>2005</td>
<td>18–30</td>
<td>Ethnicity, gender</td>
</tr>
<tr>
<td>Kwon et al. [33]</td>
<td>Korea</td>
<td>192</td>
<td>2021</td>
<td>20–79</td>
<td>Age</td>
</tr>
<tr>
<td>Zhuang et al. [31]</td>
<td>African Americans, Hispanic, Caucasians</td>
<td>3997</td>
<td>2010</td>
<td>18–29</td>
<td>Gender, Ethnicity, Age</td>
</tr>
<tr>
<td>Husein et al. [32]</td>
<td>Indian American</td>
<td>102</td>
<td>2010</td>
<td>18–30</td>
<td>Ethnicity</td>
</tr>
<tr>
<td>Othman et al. [18]</td>
<td>Malay</td>
<td>109</td>
<td>2016</td>
<td>20–30</td>
<td>Gender</td>
</tr>
<tr>
<td>Menéndez López-Mateos et al. [19]</td>
<td>Southern Spain</td>
<td>100</td>
<td>2019</td>
<td>20–25</td>
<td>Gender</td>
</tr>
<tr>
<td>Virdi et al. [20]</td>
<td>Kenyan-African</td>
<td>72</td>
<td>2019</td>
<td>18–30</td>
<td>Gender, Ethnicity</td>
</tr>
<tr>
<td>Celebi et al. [21]</td>
<td>Italian, Egyptian</td>
<td>259</td>
<td>2018</td>
<td>18–30</td>
<td>Gender, Ethnicity</td>
</tr>
<tr>
<td>Dong et al. [22]</td>
<td>Chinese</td>
<td>100</td>
<td>2011</td>
<td>20–27</td>
<td>Gender</td>
</tr>
<tr>
<td>Al-Jassim et al. [23]</td>
<td>Arab, Arian, and Mixed</td>
<td>1000</td>
<td>2014</td>
<td>&gt;18</td>
<td>Race, sex</td>
</tr>
<tr>
<td>Zacharopoulos et al. [24]</td>
<td>Greek</td>
<td>152</td>
<td>2016</td>
<td>18–30</td>
<td>Gender</td>
</tr>
<tr>
<td>Staka et al. [25]</td>
<td>Albania</td>
<td>204</td>
<td>2017</td>
<td>18–30</td>
<td>Gender</td>
</tr>
</tbody>
</table>
3. Facial Landmark Extraction

Facial landmarks are defined as key points on the human face, which are detected from 2D or 3D images (video) using different methods. In this work, we review approaches to landmark locations including using deep convolutional neural networks (DCNNs) and non-deep convolutional neural networks. DCNNs are widely used in image or video recognition and classification applications, with kernels used to extract features. Nowadays, many network structures based on DCNNs are introduced to solve the extract features problem, namely YOLO, VGG, Resnet, etc. [35–37]. Moreover, the accuracy and speed of processing are two important factors in most applications, therefore, the hybrid networks are a kind of deep learning for specific research subjects [38]. In order to increase the stability of facial landmark location, several studies have been carried out where the head posture is calibrated before determining the location of the landmarks [39–41]. In this section, facial landmark locations are defined by different methods; we divided them into three parts based on their method and application. As follows, Section 3.1 presents recent studies on landmark locations using DCNNs, with different numbers and locations of landmarks by various network structures and are not intended for medical use because they are not made according to any medical theories. Similar to the purpose of Section 3.1, Section 3.2 presents landmark location studies using specific novel methods for this application. Finally, research in medicine is presented in Section 3.3 based on both methods.

3.1. Deep Convolutional Neural Networks

In this section, some recent studies on facial landmark extraction are reviewed. The purpose of this section is to provide a review of the published studies in this field using DCNNs, but the scope of these studies is not based on theories in cosmetic surgery. They are used in information technology applications, particularly facial analysis. Of course, facial landmark detection is the first step in the applications of the human face or facial attribute analysis. The number of points can differ based on the applications, for example, Duffner et al. [13] detected five key points on the face by using a convolutional neural network that has a structure of six layers, including the input of the model that is the extracted face image, three convolution layers, and two fully connected layers. However,
the performance is not clearly presented, it is only shown by some experimental images, and the accuracy of the model is not optimal, especially the large mean error for occlusion. For the purpose of reducing occlusion errors, a study by Sun et al. [42] proposed a different traditional networks model, called Deep Convolutional Network Cascade, which is divided into three levels. The model has a complex structure, so the speed of the model is not optimized, each image takes 120 ms to extract five points with higher accuracy. The limitation of their study was that it did not detect a large number of landmarks and the speed of the model was very low. However, at this time, the old equipment did not meet the processing speed of the AI models, hence, these studies are highly appreciated and are a premise for further research. Afterward, a novel end-to-end CNN model was introduced by Zhu et al. [43], their architecture consists of shared layers and component-aware branches. The first part extracts the features by using two convolutional layers, and the second part includes six branches to reduce imbalanced errors between parts of the human face. A later study by this same group of authors suggested using a combination of Branched Convolutional Neural Networks (BCNNs) with Jacobian Deep Regression (JDR) to boost the accuracy of landmark locations [44]. By using an additional JDR model, the coordinates of landmarks are refined, increasing model accuracy and managing uncontrolled conditions. Of course, the speed of the model is significantly reduced, an image takes approximately 25 ms, whereas the BCNN model takes 12 ms; this means that by using the BCNN–JDR combination, the time to recognize an image is doubled compared to previous research [43].

The accuracy is affected by occlusion cases, which is a big issue in facial landmark location, with possible reasons being the angle of the face, the pose, light, and defects, thus, there are many studies are carried out to solve this problem. For instance, the occluded locations are found using the locations of nearby visible points, which is proposed by Valle et al. [45]. Two types of loss functions are applied to find the matching heatmap and manage missed points. In their study, Cascaded Heatmaps Regression into 2D Coordinates (CHR2C) appended with two encoder-decoder CNNs similar in structure was introduced to estimate landmark points in different poses and occlusion. However, the model was not accurate for large occlusions and makeup faces. The processing speed of the model is rated at an average of 90 ms per image, but according to the authors this speed can be improved by removing part of the proposed model. A new location is determined thanks, in part, to the location of previous points that was also proposed by Lai et al. [46]. In their study, a new approach was presented, namely the VGG19 network, applied to estimate the initial coordinates, then the LSTM model was used to refine the facial landmark detection. With this approach, the number of parameters is quite large, so processing speed is not the main factor of interest in this study, and it is difficult for real-time applications. In the research by Hoang et al. [47], they proposed the Stacked Hourglass Network to predict the location of the facial landmark point. The Residual Network (Resnet) was applied as the backbone instead of the 7 × 7 convolution layer, furthermore, the 1 × 1 convolution is used to reduce the parameters of the model. To improve the accuracy of their model, Original Residual blocks were proposed instead of the Residual Dense blocks, but the speed of the model is lower. According to the authors, their proposed model takes 60 ms for landmark detection, which is approximately 12 times slower than the 3DDFA [48] method but the accuracy is higher. Therefore, other networks with fewer parameters such as MobileNet can be considered to replace ResNet. Moreover, the authors had not solved the problem of occlusion; the results show that with the occluded cases, the model’s predictions became mixed up. A study by Xiangyu Zhu et al. [48] proposed a combination of three models consisting of 3DMM, Cascaded Regression, and CNN to align the face when changing pose. 3D Dense Face Alignment (3DDFA) was proposed in their work, and a cost function was also introduced to increase the speed of the model. According to the authors, traditional models are not flexible enough to recognize points when appearance changes or major changes in pose. In their work, they reconstruct the
face from the 2D image and the entire process takes 63.9 ms. To optimize the model parameters, the authors propose a novel cost function named Optimized Weighted Parameter Distance Cost (OWPDC) with demonstrated efficiency compared to different cost functions.

In a different approach, the proposed network structures are used and calibrated to suit the intended application, namely using YOLO networks in the study of Rao et al. [49]. They used YOLO to detect the face, then the Active Shape Model (ASM) was used to extract the landmark points. Previously, the HAAR cascade classifier was applied to detect two key points on the human face that was presented by Asi et al. [4]. Their applied YOLO model gave surprising results in terms of model accuracy and speed. Their result showed that landmark detection takes 0.0021 ms per image, and the time of the ASM model to detect landmark points was not presented by the authors. In their study, the YOLO model was presented carefully, but it was only used for face detection, and the facial landmark detection was carried out by using the ASM model. The input of this model is a 2D image with a frontal face and it does not deal with occlusion cases. In addition, the evaluations are only performed on a small sample, which greatly affects the strength of the conclusions about the accuracy of the model. Another approach many authors have used to achieve the advantages of different machine learning models is to introduce hybrid models, which means combining the CNN model with a different model as a classifier layer of the traditional CNN. CNN was used to extract the features, then the Support Vector Machine (SVM) was applied to recognize facial landmark points that were researched by Tao et al. [50]. The classifier of the traditional CNN is usually dense layers with softmax function, and these layers consist of many parameters that reduce the speed of the model. Similarly, Chen et al. [51] also used CNN to extract features of the image, then Conditional Random Field was used in addition rather than just fully connected layers. Furthermore, the CNN–LSTM–RNN model was proposed to detect the facial landmarks by Sivaram et al. [52]. Long Short Term Memory (LSTM) was used to find the initial location, then they applied a model with the same structure as suggested by Chen et al. [53] to refine the results. The result was a high evaluation and the accuracy of their proposed model was impressive, which is shown in Section 4. Previously, the Stacked Hourglass Network (SHN) was applied to detect 68 points on the face that was introduced by Yang et al. [54]. The images were rotated and rescaled by using a Supervised Transformer Network, then four SHN was used to extract the features and locate the key points. To compare and evaluate the proposed deep model, the authors built two other models called LBF_Yang and SDM_Yang, and the accuracy of the proposed deep model was the highest of the three models, but the speed of the models was a few seconds per frame because the two stages consist of many parameters to be calculated.

The paper of Zhu et al. [55] introduced Occlusion-adaptive Deep Networks (called ODN), which included three modules: geometry-aware module, distillation module, and low-rank learning module. The last layer of ResNet–18 was replaced by the occlusion-adaptive framework; their model is effective with occlusion problems. The feature map was extracted and fed into the geometry-aware module and distillation module, then the output of two modules was input into the low-rank learning module. Their model achieved a high accuracy for occlusion problems; however, the authors are not really satisfied with this result so, in a later study [56], they proposed to add the feature of attention to improve performance, namely, to replace the distillation module with the attentioned distillation module, and two factors, channel and spatial, were attentioned. Three models were introduced, including the geometry-aware module, the attentioned distillation module, and the low-rank learning module to achieve impressive results in occlusions and various poses. The aim of this study improved and developed previous research in terms of accuracy and speed.

Another approach to increase model accuracy is to refine the loss function, this was presented in the research of Feng et al. [57]. In their approach, they used ResNet, and they proposed a new loss function called Wingloss function. With this approach, their results
proved the effectiveness of recommending a new loss function, and this function was tested on numerous architectures with impressive speed, namely, for a model with 3.8M parameters, the speed of the model was 2.5 ms per image, and for the Resnet model with 25M parameters, the speed was 33 ms per image; however, accuracy is also affected by different models. In another study, the authors used a set of anchor templates as references to address cases of the large variation in facial posture. A proposed model was introduced by Xu et al. [58] called Anchorface; the model results were evaluated effectively on a 300W dataset. Overall, the studies surveyed show a trade-off between the accuracy and speed of the models. If many parameters are used to build a model, the speed is decreased, and vice versa. Therefore, novel algorithms have been proposed to have both the speed and accuracy of a specific model. For instance, Fard et al. [59] used the MobileNet structure to speed up the model to avoid reducing the accuracy of extracted processing, and two different models were used—two teacher models and a student model. The structures in the training model stage were very complicated, but there are very few parameters while running the model. In their study, the loss functions were used differently in each stage. Assistive Loss (ALoss) was defined for the two teacher models, then the KD loss function was proposed to update the weight of the student model. According to the results of the article, the student model’s parameters were lowered to the number of 2.4M parameters, namely, the MobileNet structure. This result was impressive, achieving quite a good accuracy with the occlusion cases, they called this model mv2KD. Moreover, the self-learning model is an algorithm that has recently attracted interest, being able to reduce the manual labeling process for training. In semi-supervised facial landmark detection, the study by Dong et al. [60] proposed to use a teacher model and two student models, with the teacher model as a detector to check and filter the unsatisfactory samples before the samples are fed into the train for the model. The improved model accuracy after applying two student models is shown in the results section of the paper. The semi-supervised learning approach, proposed in many studies in different fields, is surveyed in the paper of Van Engelen et al. [61]. However, the significant challenge for this method is overfitting because of the accepted pseudo label data used to train the main model and the allowed errors between the pseudo label and the truth label. These are cumulative errors, therefore, this is a challenging issue that needs to be resolved for this method.

3.2. Non-Deep Convolutional Neural Networks

Besides using DCNN, some proposed models, such as the Active Shape Model, Active Appearance Model (AAM), Support Vector Machine (SVM), coarse-to-fine framework, etc. [62–65], are non-deep models that are often used in landmark detection. The advantages of these methods over the DCNN approach are that the number of parameters is less, and the model structure is simpler, however, they often handle tasks with direct computations. For example, the Support Vector Machine (SVM) classifier was used to calculate the score of each point on the image that was introduced by Belhumeur et al. [15]. They built a global detector to handle the case that goes astray from the local detector and their innovation in the study was to combine the Bayesian model with the local detector and nonparametric global models. The local detector process takes up a lot of time because it has to slide and calculate many scores for each pixel, thus, the speed of their model was lower than the current approach, and the time to process an image was 400ms. Face alignment was introduced by Kazemi et al. [66], a new method based on regression trees, the loss functions were optimized to minimize at test time; similarly, Thakur et al. [67] applied the Cascaded Regression tree model to detect the key points. Unlike the different approaches, the idea of finding a preliminary assessment of the locations of landmarks and then using the added model to refine the facial landmark locations was presented by Zhu et al. [14]. To reduce the dependency on the initialization of the model cascades, they proposed a coarse-to-fine framework that could consider and solve the diverse pose. Their result shows that this method was more effective than the cascade model for significant changes in pose. However, the number of computations of this algorithm is enormous
compared to its real-time applications, which are difficult. To solve this problem, another study [64] proposed the F-SSFA model to improve the speed of their model. Their result shows that the processing speed is impressive, reaching 1.43 ms per image, 53 times more than the previously proposed CFFS model. This approach was quite similar to the approach of Zhang et al. [65] about how the model works, however, the authors used different architectures to solve their problems. Coarse-to-Fine Auto-encoder Networks were proposed to locate the landmarks, the points were identified and preliminarily evaluated before being refined by the 2nd Stacked Auto-encoder Networks (called SANs) with high accuracy. The structure of their model was built with the desire to reduce calculations compared to other models, the low-resolution image was used to estimate landmark locations, then the model was used to refine the coordinate positions of facial landmarks. Their result showed that the speed of the model achieved 23 ms per image to detect 68 facial points. In a previous study, a Support Vector Machine (SVM) was used as a local detector to return a score at the focal point, then the global model was applied by using the Bayesian model that was presented by Belhumeur et al. [15]. Their model was the non-parametric model, and the responses of the key points are considered as a hidden variable, and a consensus of exemplars was used to optimize this global model. However, the localizer takes 400 ms per fiducial, and the local detector takes up most of the model time.

The studies of facial landmark extraction are summarized in Table 2, including using the deep CNN model and non-deep CNN model. The factors are considered, such as the number of landmarks, year, and model used that was introduced in each study. Nevertheless, fps is for reference only because it depends on many factors, including the hardware of each study being different.

Table 2. Summarizing the facial landmark extraction using deep CNN and non-deep CNN models, reviewed in Sections 3.1 and 3.2.

<table>
<thead>
<tr>
<th>Amount</th>
<th>Fps *</th>
<th>Year</th>
<th>Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duffner et al. [13]</td>
<td>4</td>
<td>N/A</td>
<td>2005</td>
</tr>
<tr>
<td>Sun et al. [42]</td>
<td>5</td>
<td>8.33</td>
<td>2013</td>
</tr>
<tr>
<td>Zhu et al. [43]</td>
<td>68</td>
<td>83.33</td>
<td>2018</td>
</tr>
<tr>
<td>Valle et al. [45]</td>
<td>68</td>
<td>11.11</td>
<td>2020</td>
</tr>
<tr>
<td>Lai et al. [46]</td>
<td>68</td>
<td>N/A</td>
<td>2016</td>
</tr>
<tr>
<td>Hoang et al. [47]</td>
<td>68</td>
<td>16.67</td>
<td>2020</td>
</tr>
<tr>
<td>Xiangyu Zhu et al. [48]</td>
<td>68</td>
<td>86.20</td>
<td>2017</td>
</tr>
<tr>
<td>Rao et al. [49]</td>
<td>19</td>
<td>N/A</td>
<td>2019</td>
</tr>
<tr>
<td>Asi et al. [4]</td>
<td>4</td>
<td>N/A</td>
<td>2014</td>
</tr>
<tr>
<td>Chen et al. [51]</td>
<td>68</td>
<td>N/A</td>
<td>2019</td>
</tr>
<tr>
<td>Sivaram et al. [52]</td>
<td>49–74</td>
<td>N/A</td>
<td>2019</td>
</tr>
<tr>
<td>Chen et al. [53]</td>
<td>68</td>
<td>10</td>
<td>2015</td>
</tr>
<tr>
<td>Yang et al. [54]</td>
<td>68</td>
<td>N/A</td>
<td>2017</td>
</tr>
<tr>
<td>Zhu et al. [55]</td>
<td>68</td>
<td>N/A</td>
<td>2019</td>
</tr>
<tr>
<td>Sadiq et al. [56]</td>
<td>68</td>
<td>N/A</td>
<td>2019</td>
</tr>
<tr>
<td>Feng et al. [57]</td>
<td>68</td>
<td>8–400 (based on number of parameters)</td>
<td>2018</td>
</tr>
<tr>
<td>Xu et al. [58]</td>
<td>68</td>
<td>45</td>
<td>2021</td>
</tr>
<tr>
<td>Fard et al. [59]</td>
<td>68</td>
<td>253–417</td>
<td>2022</td>
</tr>
<tr>
<td>Dong et al. [60]</td>
<td>68</td>
<td>N/A</td>
<td>2019</td>
</tr>
<tr>
<td>Belhumeur et al. [15]</td>
<td>29</td>
<td>2.5</td>
<td>2013</td>
</tr>
</tbody>
</table>
3.3. Facial Landmark Detection in Medicine

Different from other applications, in the field of medicine, high accuracy and location are factors considered based on the sustainability of medical theories. For example, 18 facial landmarks were extracted automatically using the Anthropometric Face Model introduced by Sohail et al. [68]. They used the distance of the eyes as the main parameter for finding other areas of the face, and the rotation angle of the face was recognized and calibrated to the frontal angle to increase the accuracy of the model. Then, 18 facial landmarks were also located by an Anthropometric Face Model proposed by them in a previous study [69]. However, the study only corrects the face horizontally and is not effective for large vertical rotations; in addition, the accuracy of the model depends on the accuracy of the recognition of the centers of the two introduced eyes, according to the study of Fasel et al. [70]. In another study, Alom et al. [71] established mathematical relationships among horizontal and vertical distances from the 18 landmarks identified to perform 16 Euclidean measurements. The purpose of this study was to use 4 out of 16 measurements to predict age group by using SVM-Sequential Minimal Optimization (SMO) with an accuracy of 96%. With the same method and purpose, Du et al. [72] with 11 horizontal and 5 horizontal measurements were used to predict age group by the SVM-SMO algorithm. Recently, a facial anthropometric collection system for a Vietnamese population was proposed by Tuan et al. [73], and YOLOv4 was applied to detect anthropometric features with 27 anthropometric points. In their other study of the same data set, they proposed Faster R-CNN [74] to detect 14 key points on the human face, with the aim of automatically identifying landmarks, then the 3D model was built for medical applications and pre-diagnosis for cosmetic surgery. They also recorded anthropometric data of 182 Vietnamese aged from 23–46 years old, this is also the anthropometric dataset collected on the Vietnamese. Guarin et al. [75] made an assessment of facial palsy using a machine learning model previously exported to identify 68 landmarks; they hypothesized that previous studies using training datasets collected from normal people had not been demonstrated to be sufficiently accurate when tested in patients. The authors trained and compared the accuracy of the machine learning model on two data sets, one is the available data called 300-W and the data set they collected on the patients; the results were very impressive. The purpose of the study was to identify landmarks in facial paralysis patients, thereby objectively assessing the patient’s disease, and the hypothesis about their training dataset was demonstrated.

In the study of Kong et al. [76], the authors proposed to detect acromegaly from human face images by machine learning methods. The state-of-the-art models were applied in this study, such as Generalized Linear Models (LM), K-nearest neighbors (KNN), Support Vector Machines (SVM), Forests of Randomized Trees (RT), Convolutional Neural Network (CNN), and Ensemble Method (EM), which increased the accuracy of acromegaly detection, however, multiple models were used to increase the model’s calculation parameters. The facial landmarks were identified through three main steps: recognize faces from images using the SVM model, then extract features of the image, namely, landmarks by CNN, and classify as acromegaly or not. The structure of this method was very complicated, making it difficult for real-time applications, yet, in medicine, accuracy is always the most important factor. Twenty-seven key points on the human face were recognized by using the VGG16 model, researched by AbdAlmageed et al. [77]. The purpose

<table>
<thead>
<tr>
<th>Kazemi et al. [66]</th>
<th>194</th>
<th>1000</th>
<th>2014</th>
<th>Regression Trees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thakur et al. [67]</td>
<td>68</td>
<td>N/A</td>
<td>2018</td>
<td>Cascaded Regression Tree</td>
</tr>
<tr>
<td>Wang et al. [64]</td>
<td>68</td>
<td>700</td>
<td>2017</td>
<td>Fast Shape Searching Face Alignment model (F-SSFA)</td>
</tr>
<tr>
<td>Zhang et al. [65]</td>
<td>68</td>
<td>43.48</td>
<td>2014</td>
<td>Coarse-to-Fine Auto-encoder Networks</td>
</tr>
</tbody>
</table>

* Fps is frame per second, nevertheless, it is for reference only because it depends on many factors including the hardware; N/A: Not Available.
of their study was to predict congenital adrenal hyperplasia from the analysis of facial landmarks. The limitation of their study is that the dataset size was limited, and the predictions made only on 2D images. Their results showed that there was a difference in the face of normal people and patients, which had great significance in the field of medicine. Another application of AI called DeepGestalt, was proposed by Gurovich et al. [78]. They studied genetic diseases using anthropometric landmarks on the patients’ faces, they identified 130 landmarks on the patients’ faces by using ten convolutional layers. DCNN cascade was applied to detect the face of the patients, then the landmarks were detected, and the accuracy of prediction achieved was 91%. This is highly regarded research and opens up a lot of potential in the field of medicine.

Furthermore, Williams–Beuren syndrome was researched by Liu et al. [79], who used five CNN architectures to identify five landmarks for WBS facial identification in clinical practice. They collected data to train a model of 340 participants, and train on different models, from their published results, the VGG19 model is rated as the model with the best performance, however, its number of parameters is quite large (144 million) compared to other network architectures, such as ResNet-18, ResNet-34, and MobileNet-V2. Anthropometric data used to train machine learning models is still limited and it is specific for each region and for different purposes. Application in reconstructive and aesthetic cosmetic surgery was suggested in research by Nachmani et al. [80]. They introduced an app on the smartphone to calculate four measurements automatically; a best-in-class machine learning model was developed by Google that was applied in their study. To evaluate the accuracy of the app, the authors compared 15 healthy subjects. Still, the accuracy of the model is considered to be relatively low, so it cannot be applied to replace the manual measurement method for diseases in aesthetics. Improving accuracy is a necessary task of the authors, but this is also a new study on the application of ML in clinical diagnostics used on smartphones. Additionally, in this field, a new method with low equipment cost for tracking landmarks over time was proposed by Gerós et al. [81] in a study in 2016. Facegram was a new standard for analyzing facial movements with patient input and five anatomical landmarks on the face. It was rated as a powerful and accessible face tracking tool by Petrides et al. [82]. The authors used Kinect cameras instead of using three infrared-light cameras, as in the previous study by Hontanilla et al. [83], to collect the videos and images. A study on facial beauty was proposed by Aarabi et al. [84]. Facial landmarks, such as the face, eyes, eyebrows, and mouth, are identified by grayscale transformations and using a variant of the KNN model. Eight element vectors were applied to calculate the score of the human face automatically. In a 2013 study by Zhao et al. [85], they presented a novel method to detect Down Syndrome using only 2D imaging and diagnosis. Seventeen facial landmarks were detected, then SVM was applied as a classifier to decide if it is Down Syndrome or not. The accuracy of the SVM model was more than 97%, but the process of extracting facial landmarks of the patients used manual methods, and their data set was quite small, including only 24 patients. Another study conducted to improve on the unautomated problems of this study was carried out by Qin et al. [86], they used the 2D image to identify Down Syndrome based on DCNN, it was performed through three steps to provide an automated Down Syndrome identification tool.

The process of identifying landmarks in medicine is highly appreciated and is a development direction for the analysis of disease syndromes through the change of facial landmarks. However, to our knowledge, there is relatively little research carried out based on the theories of medicine, which needs to be developed to reduce the work stress for medical staff. The studies of facial landmark extraction in medicine are summarized in Table 3, including the number of landmarks, fps, sample size, and proposed approaches.
Table 3. Summarizing the facial landmark extraction in medicine.

<table>
<thead>
<tr>
<th>Amount</th>
<th>Fps *</th>
<th>Sample Size</th>
<th>Year</th>
<th>Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sohail et al. [68]</td>
<td>18</td>
<td>N/A</td>
<td>150</td>
<td>2008 Anthropometric Face Model</td>
</tr>
<tr>
<td>Alom et al. [71]</td>
<td>18</td>
<td>N/A</td>
<td>50</td>
<td>2012 Support Vector Machine (SVM)–Sequential Minimal Optimization (SMO)</td>
</tr>
<tr>
<td>Du et al. [72]</td>
<td>18</td>
<td>N/A</td>
<td>50</td>
<td>2013 Support Vector Machine (SVM)–Sequential Minimal Optimization (SMO)</td>
</tr>
<tr>
<td>Tuan et al. [73]</td>
<td>27</td>
<td>N/A</td>
<td>182</td>
<td>2021 YOLOv4 model</td>
</tr>
<tr>
<td>Tuan et al. [74]</td>
<td>16</td>
<td>9</td>
<td>182</td>
<td>2022 Faster Region Convolutional Neural Networks (Faster R-CNN)</td>
</tr>
<tr>
<td>Guarin et al. [75]</td>
<td>68</td>
<td>N/A</td>
<td>200</td>
<td>2020 Cascade of Regression Trees</td>
</tr>
<tr>
<td>Kong et al. [76]</td>
<td>68</td>
<td>N/A</td>
<td>1123</td>
<td>2018 Ensemble method</td>
</tr>
<tr>
<td>Gurovich et al. [78]</td>
<td>130</td>
<td>N/A</td>
<td>329</td>
<td>2019 Deep Convolutional Neural Network (DCNN)</td>
</tr>
<tr>
<td>Nachmani et al. [80]</td>
<td>468</td>
<td>N/A</td>
<td>15</td>
<td>2022 Google’s ML Kit algorithm</td>
</tr>
<tr>
<td>Gerós et al. [81]</td>
<td>5</td>
<td>N/A</td>
<td>4</td>
<td>2016 Facegram</td>
</tr>
</tbody>
</table>

* Fps is frame per second, nevertheless, it is for reference only because it depends on many factors including the hardware; N/A: Not Available.

4. Nasal Reconstruction Technology

With the development of science and technology, medical research is always appreciated and has great scientific significance. Cosmetic surgery is increasingly being chosen by more people to enhance the beauty of the face, or reconstruct injured organs after accidents, correct deformities, and disease of advanced nasal vestibular malignancy [87]. The clinical assessment is always appreciated as it helps doctors more objectively assess the condition of each patient and to make communication with the patient easier. This can be achieved by using a 3D model or 3D image to help them get a better overview. Faris et al. [88] studied the satisfaction of nasal enlargement defects and prostheses in reconstructive surgery. According to the authors, this is the first study on the negative effects of rhinectomy nasal defect on health utility; it was performed on 273 adult naïve observers and concluded that age and skin color affect health utility. The reconstruction of organs helps patients restore basic functions and aesthetics, which helps improve health utilities. Rhinoplasty is the oldest form of cosmetic surgery, with forehead flaps to reconstruct the nose first recorded in India millennia ago; this history of rhinoplasty is studied by Shaye et al. [89]. Moreover, the nose is the subject of much research with the aim of increasing the accuracy of the patient’s diagnosis. Authors [90,91] have presented factors affecting nasal airway obstruction, however, machine learning methods have not been used in this study. For clinical evaluation before rhinoplasty, doctors rely on CT (computed tomography) images to diagnose other pathologies based on bone structure and nasal septum, as presented in [92]. Peters et al. [93] studied the changes during rhinoplasty through 3D images collected by the optical three-dimensional scanner. The paramedian forehead flap and the bilobed flap were used to reconstruct the nose, however, aesthetically, the scar needs to be considered to ensure patient satisfaction. Twenty-four measurements were determined automatically by eighteen landmarks previously determined in 30 patients. They used 3D imaging to evaluate the effectiveness of surgery and scar size, which is considered a study that forms the basis for future studies in the evaluation of rhinoplasty.

The nose is the most attractive organ of the face and nasal reconstruction or rhinoplasty is performed for purposes such as revising nose defects after trauma, the removal of tumors, or rhinoplasty to improve the appearance and increase the beauty of the face. In order to easily evaluate and communicate well with patients, 3D technology is now used in the field to create visible 3D models. The research on this issue is also reviewed in this work, it helps readers to see the technologies that have been applied in the nasal reconstruction field. To create a 3D model of the nose that is similar to the patient’s anatom-
ical structure, Baldi et al. [94] attempted to reconstruct the 3D nose model using 3D printing technology using low-dose computed tomography (LDCT) image fusion combined with magnetic resonance imaging (MRI) to reduce the harmful effects of X-rays. Bone and tissue structures were reconstructed on 3D software and printed with PLA and TPU, which were used for clinical diagnosis before surgery. However, using X-rays to collect images of the human face can cause unnecessary side effects. The combination uses images from LDCT and MRI to minimize these effects. The authors also emphasize that the 3D model is meaningful in terms of patient education and education overall, adjusting the aspirations compared with the doctor’s expectations during rhinoplasty surgery.

Although this theory is not entirely new, it opens up opportunities for research to be able to evaluate and control the surgical process. In rhinoplasty, creating visible nasal models with plaster is essential for physicians to diagnose surgical sites, this was studied by Suszynski et al. [95]. They used 3D obtained images to create a nasal model for clinical evaluation before rhinoplasty was performed. This is also the first study to use 3D printing in aesthetic rhinoplasty. Then, Jung et al. [96] proposed a nasal bone model using 3D printing with PLA plastic to be used by doctors in traumatized rhinoplasty. According to the authors, 10 patients were detail planned for rhinoplasty surgery, this study had significant predictive methods that can be applied during surgery. Compared to prototyping with gypsum [92], 3D printing with PLA is more time- and cost-saving. However, in this study, other parts affected by nose surgery were not mentioned, such as tissue, nasal septum, etc. In another study, the figures of 3D measurements were extracted from specialized cameras. Doctors rely on these figures to create 3D models and send them to the MirrorMe3D Corporation to 3D print these faces using plaster, wax, and cyanoacrylate studied by Klosterman et al. [97]. They create 3D models for the clinical diagnosis of rhinoplasty, letting patients know the changes in advance and to have an objective view after surgery; this way, the costs to pay for a model were 45 times higher than the PLA modeling method presented by Bekisz et al. [98]. They used Blender software to build 3D models of the human head from 2D images with different angles, then the face was extracted, and the structure of the nose was adjusted to the desired shape manually and, finally, it was used for 3D printing; however, in these 3D models, only the external structure of the patient’s face can be described. Therefore, the internal bone structures were not noticed in this study, as with these models, there is only the external evaluation view images of after the surgery to help patients have a more general view of predicted nasal shape. Twelve patients were surveyed in this study, and according to the author, no complications occurred in surgeries. With the development of technology, numerous free software used to reconstruct images of objects were developed, such as Blender and RhinOnBlender, which were widely used in cosmetic surgery. Much software used to create SLT/OBJ format files, used for rapid prototyping with 3D printers by using PLA plastic, was introduced by Sobral et al. [99]. In this study, the authors only wanted to reconstruct the shape of the nose after surgery with the main purpose of surgical education and patient communication. 2D images collected from smartphone cameras are used to build virtual facial molding, from which patients can see their own faces after rhinoplasty surgery. Their results showed that the study was conducted on 3 patients with their absolute satisfaction with the results after surgery, the results compared the actual shape of the nose after surgery, and the 3D printed model was similar. However, the small sample size can lead to non-generalized conclusions, with the benefits of this model being that it is low cost and fast.

In 2020, Choi et al. [100] published a study on a 3D printed rhinoplasty guide tool. Three-dimensionally-printed rhinoplasty guides were introduced, and the authors created two models before and after surgery to evaluate the 3D model, which increases the accuracy for clinical diagnoses. They were subtle in checking the satisfaction of the patients after the surgery with specific questions. 3D images are created using the Morpheus three-dimensional scanner, then calibrate the parameters for each patient and get consent from the patients before being sent to a 3D printing company. It is evaluated as an effective 3D model to make surgery easier; it was used as a testing tool and guide for rhinoplasty
in the study of Lee et al. [101]. In a similar purported study published in 2021, Gordon et al. [102] performed a one-year study from 2019 to 2020 in 15 rhinoplasty patients. 3D images were collected from a commercial Vectra H1 camera and processed on specialized software. The aim of this study was to facilitate communication between the patient and the physician in the preoperative period, 3D printing as a guide to be used during surgery for the physician, and it was used to assess the accuracy of actual and simulated surgical procedures. A custom 3D printed profile contour guide was created and the authors were very careful in evaluating the accuracy of the method, specifically the ceramic models of the patient’s nose before and after surgery created by MirrorMe3D. However, the study was performed on a rather small sample, which affects the strength of the overall conclusions. It is necessary to evaluate the process on larger samples to increase the strength of the general conclusions, or other studies applying 3D printing technology to assist doctors in evaluating rhinoplasty surgery and in rhinoplasty education using 3D printing [103,104], or assisting in recovery after rhinoplasty surgery, specifically less edema and ecchymosis [105]. Reconstruction using 3D printing technology was an area of potential for reducing the risks of surgery. In brief, most of the studies were conducted with quite small sample sizes, hence, it affects the strength of the outcome. Moreover, highly commercialized 3D printers require suitable materials, without compromising the quality of the surgery and the patient’s health. The accuracy of the points and 3D models that were built depends entirely on the experience of the staff performing. The facial landmark extraction and necessary measuring angles for fully automatic face reconstruction is also a big challenge. Anthropometric measurements also need to be fully automated to reduce pressure on medical staff and reduce the costs of medical services. The case studies using the 3D printing method are recently surveyed and summarized in Table 4. The criteria evaluated in the version include the software used by the authors in the process of creating 3D printing files, the materials used, the number of patients participating in the study, and the results achieved after the printing process; additionally, the data necessary for the reproduction of the images and files are also presented.

Table 4. Summary of nasal reconstruction studies using 3D printing.

<table>
<thead>
<tr>
<th>Year</th>
<th>Input</th>
<th>Output</th>
<th>Sample Size</th>
<th>Material</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baldi et al. [94] 2021</td>
<td>CT scans and magnetic resonance imaging</td>
<td>Bone tissue model, soft tissue and cartilage models</td>
<td>10</td>
<td>PLA, TPU</td>
<td>3-Matic software, CURA (Materialise, Belgium)</td>
</tr>
<tr>
<td>Suszynski et al. [95] 2018</td>
<td>3D image from Vectra H1</td>
<td>Three-dimensionally gypsum model</td>
<td>–</td>
<td>Gypsum</td>
<td>MirrorMe3D (New York, NY, USA)</td>
</tr>
<tr>
<td>Jung et al. [96] 2020</td>
<td>CT image</td>
<td>3D model for nasal osteotomy</td>
<td>11</td>
<td>PLA</td>
<td>DICOM, InVesalius, Meshmixer</td>
</tr>
<tr>
<td>Klosterman et al. [97] 2018</td>
<td>Image from canfield H1 camera</td>
<td>Two 3D models before and after surgery</td>
<td>6</td>
<td>Gypsum</td>
<td>Vectra Sculptor Software (Canfield Scientific, Fairfield, NJ, USA)</td>
</tr>
<tr>
<td>Bekisz et al. [98] 2019</td>
<td>3D digital photographic images</td>
<td>Predicted postoperative PLA model</td>
<td>12</td>
<td>PLA</td>
<td>Blender (Version 2.78, Amsterdam, The Netherlands)</td>
</tr>
<tr>
<td>Sobral et al. [99] 2021</td>
<td>Image 2D from smartphone</td>
<td>3D surgical guides</td>
<td>3</td>
<td>PLA</td>
<td>Blender (Stichting Blender Foundation, Amsterdam, The Netherlands) and RhinOnBlender (Cicero Moraes, Sinop, Brazil)</td>
</tr>
<tr>
<td>Choi et al. [100] 2018</td>
<td>Morpheus 3D scanner</td>
<td>3D printed rhinoplasty guide</td>
<td>50</td>
<td>PLA</td>
<td>Morpheus (Morpheus Co., Ltd., Seongnam City, Gyeonggido, Republic of Korea)</td>
</tr>
<tr>
<td>Gordon et al. [102] 2021</td>
<td>3D image from Vectra H1</td>
<td>3D models as surgical guides</td>
<td>15</td>
<td>Canfield software (approved by the Institutional Review Board (no. 2020-12420))</td>
<td></td>
</tr>
<tr>
<td>Locketz et al. [106] 2020</td>
<td>3D image from Vectra H1</td>
<td>3D printed modeling to define contours on the nose</td>
<td>5</td>
<td>Rigid plastic</td>
<td>Mirror (Canfield Scientific Inc.)</td>
</tr>
</tbody>
</table>
### 5. Discussions and Evaluations

Medical research is aimed at predicting risks during treatment and the clinical diagnosis of some diseases observed through symptoms, as well as reducing work pressure for medical staff. Humans express their emotions through their face. Thus, it is considered one of the most important parts of the human body [107]. The number of facial landmarks extracted depends on the different applications. This paper reviews studies from around the world that generate anthropometric datasets for each different country and research applying the achievements of AI for landmark recognition and nasal reconstruction technology. Anthropometry plays an important role in medicine, the authors tried to find the relationship of factors affecting anthropometric data, such as ethnicity, sex, and age. These figures are used as a reference for identification, reconstructive surgery, or simply in the design of personal protective equipment for suitable workers. In addition, facial landmarks and measurements were used for clinical diagnosis to predict some disease symptoms, assess facial beauty, predict internal bone structure, or reproduce 3D images. Rhinoplasty is a familiar phrase in cosmetic surgery as the nose is the most attractive part of the human face, it increases the beauty of the face according to feng shui in some Asian cultures [73], and reconstruction of damaged nasal passages is necessary to avoid difficult communication for patients. Facial landmark extraction is a topic that has many different applications, such as emotion recognition, facial beauty adjustment applications, and medical applications, such as depression detection, Down Syndrome, etc. For medicine, facial landmarks are defined differently from other apps in terms of the number, name, and location of the key points, and the accuracy of the recognition process is the top concern. Currently, in hospitals, surveyed landmarks are manually identified with the semi-help of some commercial software. The development of 3D printing technology is also evident where prosthetics are built into 3D models for the purpose of education and clinical prediction to create a more general view for the patient’s changes before and after surgery. 3D models created by commercially available plastics and printers, according to objective assessment, cannot create 3D printed prosthetics to replace damaged human body parts; however, this depends largely on the materials used to reconstruct the human part.

Three-dimensional photographic analyses of the human face were researched by many authors in different countries with different numbers of participants and purposes. The database was created in different regions and countries with the number of measurements taken on the face varying from 8–38 measurements. This is the premise for the development of a system for predicting the internal characteristics of bones and reconstruction automatically. Nowadays, authors perform studies on anthropometric data collection using commercial 3D cameras as presented. Anthropometric databases were concluded based on the results of measurements, but in some studies with small sample sizes where the participants were not randomly selected from different areas, the strength of the conclusions could be affected. A large number of random samples in the different areas were expected to increase the reliability of the figures. In this review, we evaluated a number of well-known studies in different parts of the world on anthropometric measurements. This is also a topic that promotes development in the process of identifying and reconstruction, these make patients more confident in communication. It is also a solid basis for future studies on automatic facial landmark recognition in biomedical research.
With the development of algorithms including both deep and non-convolutional neural networks, many authors proposed different approaches to extracting the facial landmarks. The number of landmarks is determined depending on the application, and the surveyed papers were determined between 5 and 80 key points. Some studies only identify a few areas of the face, such as two eyes, nose, and mouth, or recognize a few facial landmarks that are applied in simple applications, such as determining the rotation angle of the face and recognizing the face, or a large number of landmarks in facials were used in emotional identification and reconstruction. With a non-deep learning method, Milborrow et al. [108] showed that the accuracy of landmarks depends on the number of points identified because the points are determined based on a part of the coordinates of other neighborhood points. They showed a big change in accuracy when changing the number of landmarks from 3 to 68. Factors considered in landmark recognition are the accuracy and speed of the model, and the model’s ability to handle various poses or occlusion. To solve the problem of accuracy in occlusion cases, network structures with a large number of parameters are proposed [109,110]. In these studies, the issues of landmark detection were solved very well, but too many model parameters make it difficult to have real-time applications. Accuracy and processing speed are considered a trade-off, as shown in [111]. The mentioned studies are applied in many cases, but they have not been used in medical studies. In hospitals, the recognition of facial landmarks is now mostly a manual method; it is essential to apply automatic identification and measurement in medicine. Many studies have impressive results but have not yet connected them to studies in the medical field. In some cases, the anthropometric measurements were performed automatically from facial landmarks that were manually marked [93]. Nowadays, 3D printing technology is developed as a rapid prototyping device that is applied in many fields. Recently, it has been applied to medicine, especially for clinical diagnoses, and 3D models are created for medical education and patient education. Furthermore, the studies created guidelines to be used during surgery or fixation and post-surgery shaping. The 3D models are created depending on the following key factors: the skill of the technicians in creating the 3D image file, the accuracy of the recording equipment, and the accuracy of the printer. In our view, the creation of 3D images is very necessary for medicine to facilitate the surgical process, however, at present, this model has not been particularly popular, especially in Asia. The current 3D printed models deployed in studies were largely outsourced, the 3D reconstructing image, landmark extractions, and measurements are all performed manually by professional staff. It is necessary to put new technologies into practice that reduce the amount of work for employees, in addition, to help to normalize measurements to avoid human measurement errors, especially reducing time and cost.

Nowadays, software that reconstructs 3D from 2D images is also becoming more popular, but it often comes with expensive hardware compared to the digital camera. With the development of science and technology, machine learning algorithms gradually replace the hardware of complex collection systems. This also reduces investment costs for 3D devices that address a disadvantage of 3D, which was presented by Lekakis et al. [112]. The 3D printer was applied in medical reconstructions, which was a potential development direction to improve the quality of healthcare, especially in the field of cosmetic surgery. 3D printed models used for medicine as a supported tool in surgery, medicine education, and patient communication. The 3D printed model could be used like a model to clinical diagnosis [98], a guide for surgery (Figure 2), or external nasal splint after rhinoplasty (Figure 3). To brief, in this paper, we surveyed studies that represent approaches from the authors in landmark extraction and their application. Several authors have conducted studies to collect facial data for research in this area, such as 300W [113], CMU Multi-PIE [114], AFLW [115], Helen [116], LFPW [15], FRGC [117], XM2VTSDB [118], BioID [119], PUT [120], AR Purdue [121], MUCT [122]. The Multi-PIE dataset [114] with the number of 755,370 images was also investigated in this work. The number of images in the datasets published by the authors is shown in Figure 4b. Normalized Mean
Error (NME) is an error evaluation method defined by Formula (1). The diagram in Figure 4a shows the NME of the process of identifying landmarks automatically on the 300W dataset [113] in all three parts of the dataset, including 300W-Common, 300W-Full, and 300W-Challenging.

\[
NME = \frac{1}{N} \sum_{n=1}^{N} \left( \frac{y_n - \hat{y}_n}{d_n} \right)^2
\]

where \( N \) is number of sample size; \( y_n \) is the ground truth landmark; \( \hat{y}_n \) is the predicted landmark; \( d_n \) is the square-root of the ground truth bounding box.
Figure 4. (a) NME description of the studies surveyed on a 300W dataset [4,13,15,39–41,43,47,50–55,62], (b) the number of datasets published in the studies.

Sun et al. [123] reviewed 40 publications in the nasal reconstruction field in a complete review published in 2021. The authors have systematized advances in the field with papers on 3D imaging technology, computer-assisted surgery, and 3D printing. However, they only provide a general review of previous studies not in-depth comments on each technology. In another review, Matteo Bodini [124] reviewed studies on facial landmark extraction on 2D images and video. The author updated papers on facial landmark recognition in 2018 and the article provides an overview of the proposed methods, however, the article only focuses on evaluating the structures of models in the field of technology and does not refer to studies based on any sustainable medicine theories. A review with a similar purpose was carried out by Johnston et al. [125] in 2018. Unlike previous reviews, we reviewed previous studies in both technology and medicine. The three issues evaluated included facial anthropometrics, landmark extraction, and nasal reconstruction technology. Anthropometric measurements of the human face are carried out in studies to collect anthropometric data of people in different regions. Facial landmarks are defined according to sustainable medicine theories. Research on landmark extraction has achieved many achievements in recent years, but their applications in medicine are still limited. The purpose of this review is to give readers an overview of both the technique and medical aspects. The nasal reconstruction technology studies are reviewed in Section 4, which updates new research in the field that can serve as a reference for future research. In this work, we only focus on reviewing the reconstruction techniques without evaluating many of the articles on materials; this is a small limitation. Despite what we have analyzed, research in these fields has been continuing. To our knowledge, applications in medicine are expected to be considered more in future studies, such as automatic anthropometric measurement systems, novel materials to replace damaged organs, 3D printing technologies suitable for different materials, and high resolution.

6. Conclusions

This survey provides an overview of current state-of-the-art algorithms in the fields, such as facial landmark detections, anthropometric indexes of human faces in different countries, and, in particular, landmark detection and nasal reconstruction technology. Anthropometric data is important in medicine; it helps in the clinical diagnosis process, reconstructing organs, or is used as data to design clothes for people. The differing nature of each region and the different ethnicities and ages affect the anthropometric indicators of the people. Facial landmark extraction was developed through many studies with different purposes, however, there are fewer applications in the medical field. Making the
connection between the studies in the different fields is extremely necessary. The studies in rhinoplasty, landmarks, and 3D models are created manually depending largely on the medical staff's experience. With the development of science and technology, the automatic process is a necessary and challenging task in this field. In addition, the datasets that provide the field of landmark recognition are well developed and refined, which are presented in Section 5. The medical field is always an attractive choice, but it requires more precision than other applications because it affects human health. Specialized cameras are increasingly commercialized to collect 3D images for these processes, but the cost of investing in this system is very high. Meanwhile, algorithms need to be developed to replace these hardware systems; this is a task and a challenge for future studies. Currently, 3D printed reconstruction is mostly used only as a reference for doctors and patients, yet the development of this field promises prosthetics capable of replacing damaged parts of the patient. This requires a combination of intensive materials research and the accuracy of 3D printing for the medical field.

**Author Contributions:** Conceptualization, N.T.T., N.H.V. and N.M.T.; methodology, N.T.T. and N.M.T.; software, N.M.T.; validation, N.H.V.; formal analysis, H.N.A.T. and T.D.K.; investigation, N.T.T. and N.H.V.; resources, N.M.T.; data curation, N.T.T.; writing—original draft preparation, N.M.T.; writing—review and editing, N.T.T.; visualization, H.N.A.T. and T.D.K.; supervision, N.T.T.; project administration, N.T.T.; funding acquisition, N.T.T. and H.N.A.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not available.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


96. Jung, Y.G.; Park, H.; Seo, J. Patient-Specific 3-Dimensional Printed Models for Planning Nasal Osteotomy to Correct Nasal Deformities Due to Trauma. OTO Open 2020, 4, 2473974X20924342.


