Abstract—Breast Cancer Conservative Treatment (BCCT) is considered the gold standard of breast cancer treatment. However, the aesthetic outcome is diverse and very difficult to evaluate in a consistent way partly due to the weak reproducibility of the subjective methods in use. This motivated the research on the objective methods. BCCT.core is a very recent software that objectively and automatically evaluates the aesthetic outcome of BCCT. However, as in other approaches, the system only uses frontal patient information, disregarding volumetric perception on lateral measurements.

In the current work we investigate the improvement of the BCCT.core model by introducing lateral information extracted from patients images. We compare the performance of the model currently used on BCCT.core with the model developed in this study. Experimental results suggest that with lateral measurements the model presents better performance, however improvements are not significant. We can conclude that is essential to use robust models on the BCCT, and the input of 3D models will probably help to obtain better results.

I. INTRODUCTION

Breast cancer is the commonest cancer among women in most countries. According to the American Cancer Society, about 1.3 million of women will be diagnosed with breast cancer annually worldwide and about 465 thousand will die from the disease. Nowadays, after the diagnosis, the number of survival cases that exceeds 10 years is around 80%. As a result many women must live several years with the aesthetical consequences related to surgical procedure and posterior treatment. Womens breasts are a significant part of her body and relate directly with her femininity, and, for that reason, a good aesthetic outcome is extremely important. In the conservative approach however approximately one third of the patients will have a fair or poor aesthetic outcome [1], [2], affecting directly the psychosocial recover and, consequently, patients quality of life.

The absence of a standard method for measuring the aesthetic outcome has been considered an obstacle in the assessment and evaluation of the techniques applied. Until recently most used methods were based on a subjective evaluation, made by one or more observers by visual inspection, although, its lack of reproducibility, precipitate the introduction of objective methods. They are based on measurements taken directly from the patient or from patient photographs, that are essentially based on asymmetries between treated and non-treated breasts [3], [4].

In order to overcome the idea that objective asymmetry measurements were insufficient, other groups introduced a method based on the sum of the individual scores of subjective and objective indices [5]. All these additions improved the aesthetic evaluation methods but were considered insufficient due to a great intra- and inter-observer variability. This fact corroborates the idea that it is necessary to replace or enhance human expert evaluation of the aesthetic result of BCCT. A possible solution should be supported by an objective tool, which has to be easy to use, highly reproducible and acceptable to those who would be evaluated. To overcome the absence of objectivity of the evaluation of the aesthetic outcome of BCCT, a computer-aided medical system named BCCT.core [6] was recently developed. This system is a tool very easy to operate that automatically evaluates the overall aesthetic result of BCCT.

Almost all methodologies reported until now, namely the used on BCCT.core, only use patient frontal images. The only use of face-views may discard some important information, such as volume perception and inexistence of lateral measurements. This can be improved with the introduction of lateral information extracted from patients side-views, providing the model with more information about breast shape. This approach was never addressed on the aesthetic outcome of BCCT and only a small number of researchers used lateral features on their work, such as Kim et al. [7] that compare objective measurements based on breast ptosis with ratings on a subjective scale made by experienced clinical observers. This can be related with the difficulty of extraction of robust features from the side-views.

The objective of the current work is to investigate improvements on the BCCT.core software by the introduction of lateral information, which has to be easy to use, highly reproducible and acceptable to those who would be evaluated. To overcome the absence of objectivity of the evaluation of the aesthetic outcome of BCCT, a computer-aided medical system named BCCT.core [6] was recently developed. This system is a tool very easy to operate that automatically evaluates the overall aesthetic result of BCCT.

II. BCCT.core DESIGN

The development of BCCT.core entailed the automatic extraction of several features from patient’s frontal images (Fig. 1), selecting the aspects with most impact to the overall cosmetic result: breast asymmetry, skin colour changes due to the radioactive treatment and surgical scar appearance. Posteriorly, a Support Vector Machine (SVM) classifier was trained to predict the overall cosmetic result from the recorded features [6], [8].
BCCT.core classifies the aesthetic outcome of BCCT into excellent, good, fair, and poor classes. To achieve the classification, first, a concise representation of a BCCT image is obtained based on the aspects mentioned before. These measurements are preceded by the semi-automatic localization of fiducial points (nipple complex, breast contour and jugular notch of sternum) on the digital photographs [9], [10]; measures are then supported on these fiducial points. After this, all the measures are automatically converted onto an overall objective classification of the aesthetical outcome, using the SVM classifier, trained to predict the overall aesthetical classification on the aforementioned scale of four classes.

In a previous work [8], an accurate and interpretable model for the BCCT.core was found. The objective of that work was to compare between the previous model used on the evaluation of the overall aesthetic result of the BCCT [6], with other methods in order to overcome the non-interpretability of the older one. A set of 143 photographs, recorded at different breast centers, was used on that study. A simplified feature selection was also an aim of that work. The results prove that interpretable models show a similar accuracy to the non-linear models previously used, without sacrificing the performance of the BCCT.core.

### III. BCCT.core Model Improvement Using Lateral Information

Although the current model presents satisfactory results, it is possible that with only frontal views we do not have sufficient information about breast aesthetic, such as, breast volume perception. By using lateral photographs (Fig. 2) we increase the quantity of data related with the breast, augmenting its quality.

![Typical lateral photographs.](image)

Features were defined from the comparison between the treated with the non-treated breast. This represents an additional problem, because we have different images for each breast, unlike frontal analysis. In order to overcome this problem we needed to find a normalization factor between the images making all measurements in the same scale.

#### A. Study population and a gold standard

This study uses a portion of the dataset of 143 photographs with the ground truth used in [8]. The gold standard is used to train the classifier in order to perform an automated analysis and was obtained from the evaluation of patients by an international panel of experts following a Delphi methodology. It only used 63 photographs from the initial data due to the following reasons: definition of the normalization factor between images; fiducial points annotation; poor quality of some of the lateral photographs. Any little moving of the patient making its posture not truly lateral, may influence the measurements taken from the photographs. The distribution of the patients over the four different classes is summarized in TABLE I.

**TABLE I**

<table>
<thead>
<tr>
<th>Class</th>
<th>Excellent</th>
<th>Good</th>
<th>Fair</th>
<th>Poor</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td># cases</td>
<td>10</td>
<td>35</td>
<td>11</td>
<td>7</td>
<td>63</td>
</tr>
</tbody>
</table>

#### B. Feature definition

It is commonly accepted that the cosmetic result after BCCT is mainly determined by visible skin alterations or changes in breast volume or shape. Skin changes can consist of a disturbing surgical scar or radiation-induced pigmentation or telangiectasia. Differently from frontal features, there are not many studies reporting the definition of features in lateral images. As an exception, Kim et al. [7] refer ptosis factor to evaluate the aesthetic results, that is the extent to which the nipple is below the inframammary fold, the crease beneath the breast.

In this phase all fiducial points were manually marked directly on the photographs. In the future, this process will be made automatically. Fig. 3 shows all fiducial points and measurements used for the features definition.

The fiducial points marked on patients photographs were:

- **N**: Nipple;
- **IF**: Lateral terminus of inframammary fold;
- **lvp**: Lowest visible point of the breast contour;
- **tw**: Thoracic wall;
- **U**: Projection of the IF point on the superior breast contour along the tw.

The inframammary fold is a curvilinear structure which is generally hidden behind breast tissue since most women have some degree of ptosis. The lateral terminus of the inframammary fold (IF) is the endpoint of the inframammary fold where intersects the thoracic wall (tw). The lowest visible point (lvp) is located at the most inferior point of...
the breast. The centroid of nipple ($N$) is considered rather than the nipple-areola complex since many women have an irregular shaped areola.

From the fiducial points, the following measurements were defined (Fig. 3):

- **Breast contour (bc):** Breast exterior curve between $IF$ and $U$ point;
- **a:** $\min(dist(N, bc))$;
- **b:** $\min(dist(N, tw))$;
- **c:** $\min(dist(N, lvp))$;
- **Breast size (bs):** $a + b$;
- **Ptosis (p):** $N - lvp$.

As mentioned before, a normalization factor ($nf$) is needed in order to have all measurements of each patient, from two photographs, on same scale, and was defined manually on body patients in each picture. Features were assigned using previous measurements related with the asymmetry between treated and non-treated breast. The indices recorded to assess breast asymmetry were the following (note that all measurements are normalized by $nf$):

- **bs difference**
  \[ \frac{|bs_1 - bs_2|}{(bs_1 + bs_2)/2} \]
- **Nipple distance to bc**
  \[ \frac{|a_1 - a_2|}{(a_1 + a_2)/2} \]
- **Nipple distance to tw**
  \[ \frac{|b_1 - b_2|}{(b_1 + b_2)/2} \]
- **Nipple distances evaluation**
  \[ \frac{|c_1 - c_2|}{(c_1 + c_2)/2} \]
- **Breast compliance evaluation**
  \[ \frac{|c_1 - c_2|}{(c_1 + c_2)/2} \]
- **Breast ptosis difference**
  \[ \frac{|p_1 - p_2|}{(p_1 + p_2)/2} \]

In order to minimize the influence of the different scales between the images of the same patient, other features based on ratios of the aforementioned measurements were also created.

The ratio indices recorded to assess breast asymmetry were the following:

- **bs difference Ratio**
  \[ r_{LBSD} = \frac{\min(BS_1/BS_2, BS_2/BS_1)}{2} \]
- **Nipple distance to bc Ratio**
  \[ r_{LNDbc} = \frac{\min(a_1/a_2, a_2/a_1)}{2} \]
- **Nipple distance to tw Ratio**
  \[ r_{LNDtw} = \frac{\min(b_1/b_2, b_2/b_1)}{2} \]
- **Nipple distances evaluation Ratio**
  \[ r_{LNDE} = \frac{\min(a_1 * b_2/b_1 + a_2, b_1 * a_2/a_1 * b_2)}{2} \]
- **Breast compliance evaluation Ratio**
  \[ r_{LBCE} = \frac{\min(c_1/c_2, c_2/c_1)}{2} \]
- **Breast ptosis difference Ratio**
  \[ r_{LBPD} = \frac{\min(p_1/p_2, p_2/p_1)}{2} \]

### C. Classifier

In this study we applied a cascade generalization based classifier [11]. Cascade generalization uses a set of classifiers sequentially performing, at each step, an extension of the original data by the insertion of new attributes. In this particular application, the frontal features were initially applied to the model developed in [8], resulting in a classification $C_F$. Afterwards, lateral features were introduced on the system, one by one, together with the result of the first model, resulting in a classification $C_G$. The scheme of the classifier is shown on Fig. 4.

This new model is also a classifier based on machine learning techniques, namely SVMs. SVMs have proved themselves as being capable of representing complex classification or mapping functions. They discover the representations using powerful learning algorithms. We make use of the replication method as a tool for mapping this kind of
data [12], because there is an inherent ordering between the classes. As already made in the preliminary study [8], we did not use the same misclassification costs for all classes. Caregivers consider an error in a true excellent or true poor patient more penalizing than an error in the middle classes (fair or good). Moreover, failure to a contiguous class is not as serious as failure to a non contiguous class. From these considerations, and in collaboration with clinical experts, we empirically defined a cost matrix reflecting the penalty of classifying samples from one class as another:

\[
C = \begin{bmatrix}
0 & 2 & 4 & 6 \\
1 & 0 & 1 & 2 \\
2 & 1 & 0 & 1 \\
6 & 4 & 2 & 0
\end{bmatrix}
\]

where \( C(i, j) \) is the cost of classifying a point into class \( j \) if its true class is \( i \). The cost matrix was taken into consideration during the model building process.

### IV. RESULTS

The SVM classifier was based on the data replication method for data naturally ordered [12], both with linear and RBF kernels. We performed a “grid-search” optimizing each parameter using a cross-validation approach in the training set.

For all the 63 photographs the preliminary misclassification error was computed only with the frontal features using the model developed in [8]. The testing phase was made with the same data used on training phase using a Leave One Out scheme. The test results concerning the misclassification error are summarized in TABLE II. For the first ranked of each lateral feature:

<table>
<thead>
<tr>
<th>Lateral Feature</th>
<th>kernel</th>
<th>C</th>
<th>gamma</th>
<th>Error</th>
<th>Weighted Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBSD</td>
<td>Linear</td>
<td>16</td>
<td>0.37</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>rLBSD</td>
<td>Linear</td>
<td>1</td>
<td>0.37</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>LND6c</td>
<td>RBF</td>
<td>0.125</td>
<td>0.39</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>rLND6c</td>
<td>RBF</td>
<td>0.125</td>
<td>0.39</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>LND6w</td>
<td>Linear</td>
<td>0.37</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rLND6w</td>
<td>Linear</td>
<td>0.38</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LNDE</td>
<td>RBF</td>
<td>0.5</td>
<td>0.40</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>rLNDE</td>
<td>RBF</td>
<td>0.5</td>
<td>0.43</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>LBCE</td>
<td>RBF</td>
<td>0.25</td>
<td>0.32</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>rLBCE</td>
<td>RBF</td>
<td>0.25</td>
<td>0.32</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>LBPD</td>
<td>Linear</td>
<td>0.38</td>
<td>0.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rLBPD</td>
<td>Linear</td>
<td>0.39</td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A first observation from TABLE II is that with the integration of these new lateral features the classification error decreases.

### V. CONCLUSIONS

We have researched a new model for the assessment of the aesthetic outcome of BCCT. The accuracy of the model researched herein compared with the model developed for the classic BCCT.core software, showed a better performance, however it did not represented a significant improvement. The number of photographs used, 63, significantly inferior to the initial dataset of 143, was due to quality and possible definition of fiducial points and features on the lateral photographs. Obtained results might have been influenced by the choice of cases.

Additional tests will be needed to understand the relevance of the defined asymmetric lateral features on BCCT cosmetic results evaluation.

In future work, we intend to increase the robustness of BCCT.core with the definition of new lateral features, namely, related with skin colour changes due to the radiotherapy treatment and surgical scar appearance. We also have the intention to create a robustness model based on 3D information extracted from several single images.

### REFERENCES


