DOI: 10.1016/j.future.2016.08.001

A collaborative clinical analysis service based on theory of evidence, fuzzy linguistic sets and prospect theory and its application to craniofacial disorders in infants

Arcangelo Castiglione *, Raffaele Pizzolante, Christian Esposito, Alfredo De Santis, Francesco Palmieri, Aniello Castiglione

Department of Computer Science, University of Salerno, Via Giovanni Paolo II, 132, I-84084, Fisciano (SA), Italy

HIGHLIGHTS

- A collaborative clinical analysis framework that can be implemented as a Grid or Cloud Service.
- Performs detection and characterization of pathologies by combining the knowledge of a group of experts.
- Combines prospect theory, theory of evidence and fuzzy linguistic sets for developing a computer aided diagnosis tool.
- We considered craniofacial pathologies in infant population as a practical example for better explaining the proposed solution.

Keywords:
Collaborative clinical analysis
Theory of evidence
Fuzzy linguistic sets
Prospect theory
Services management
Pervasive computing

ABSTRACT

Nowadays, it is more and more important to diagnose several kinds of pathologies at their early stage, in order to take the necessary countermeasures before having permanent consequences. Unfortunately, though many pathologies are widespread, there does not exist a unique standardized reference or *Gold Standard* according to which it is possible to evaluate the patients, mainly when the pathology is in the early stages or is not very noticeable, and the doctor is not sufficiently expert in the problem domain. In this work, we deal with this problem, by envisioning new healthcare services supporting a collaborative clinical analysis of symptoms collected from the patients and forwarded to a group of experts, which are geographically distributed. The experts return back their assessment and diagnosis and the system combines these by means of the Theory of the Evidence, in order to provide a single response. The above services can be easily implemented on top of state-of-the-art distributed computing facilities such as Grids or Clouds, providing a connected environment for medical data distributed over different sites and allowing medical experts to collaborate without being co-located, thereby providing transparent access to data and computing resources. Additionally, such services can provide feedbacks to each expert, in order to improve its own knowledge and experience in the case of divergence between the expert response and the global combined diagnosis in recognizing and classifying the received symptomatic indexes from the patient. We have considered the craniofacial pathologies in infant population as a practical example for better explaining the proposed solution.

1. Introduction

Currently, the research community in health informatics is focusing its attention on approaches to support doctors and general practitioners in their job of recognizing pathologies from a series of evidences collected from the patients. The intent is to reduce the probability of misdiagnosis, which is currently quite high *i.e.*, it has been estimated that between 10% and 15% of medical diagnoses are incorrect, and implies high costs for the healthcare systems, *i.e.,* a recent analysis published in *BMJ Quality and Safety* estimated that malpractice claims in the last 25 years had an average price of $386,849 per claim and caused 40,000 to 80,000 hospital deaths yearly in the United States [1]. The problem of misdiagnosis is more serious if we consider that some pathologies can affect infants and leave them permanent deformations and/or

* Corresponding author.
E-mail addresses: arcastiglione@unisa.it (A. Castiglione), rpizzolante@unisa.it (R. Pizzolante), christian.esposito@dia.unisa.it (C. Esposito), ads@unisa.it (A. De Santis), fpalmieri@unisa.it (F. Palmieri), castiglione@ieee.org (A. Castiglione)

health problems, which can be avoided with the early diagnosis of the pathologies and by assuming proper corrective, even simple, actions at the early stage of the pathology. A practical example of this is represented by craniofacial pathologies, such as *plagiocephaly*, *brachycephaly* and *scaphocephaly*, which are very common and widespread among infants, due to either wrong sleeping positions or to congenital problems. The deformation caused by such pathologies, if recognized early, can be smoothed out by simply placing the baby prone while awake or alternating the point of contact between the occiput and bed during sleep [2].

To avoid potential patient harm, reduce the costs caused by misdiagnosis and improve the probability of early diagnosis of pathologies in infants, hospitals and medical practices are turning their attention to clinical *Decision Support Systems (DSS)*, in order to assist them in diagnosing pathologies based on the symptoms manifesting in the patients. Generally speaking, DSS operations are based on the theory of the pattern recognition: the practitioner provides a series of inputs (acquired from patients with different methods), which are compared in a proper manner with a set of labeled data stored in the system, in order to provide as an output the label of the patterns more similar to the inserted ones. Therefore, the effectiveness of the DSS, *that is to say*, the response is exactly the ones we were expecting based on the provided inputs, strongly depends on the quality of the labeled patterns internally stored in the DSS, which is commonly denoted as *Gold Standard* (GS). However, a perfect GS, *i.e.*, the one that allows the DSS to always provide the correct answer based on the provided input, is impossible to be built, since in the healthcare practice there are no strict rules for the disease diagnosis. For example, in craniofacial pathologies, a GS would be a shape which best characterizes the cranial morphology of clinically healthy patients. In addition, the adopted diagnosis methodology substantially relies on subjective assessments, performed according to the skills and experience of the single doctor who carries them out. Specifically, craniofacial pathologies are recognized based on some indexes, such as *Cephalic Index (CI)*, *Oblique Cranial Length Ratio (OCLR)*, *Standard Deviation (SD)*, etc. Such indexes are not very indicative, especially for some mild cases, thus they do not provide the doctor with an immediate and significant support to detect and quantify the pathology, especially in the case of brachycephaly and positional plagiocephaly [3].

In order to address the above issues, we define a methodological framework for the collaborative clinical analysis of a patient, which involves the acquisition of data, along with the detection and characterization of pathologies by a group of experts, geographically scattered over the world. Such a framework, due to its multi-disciplinary and inherently distributed nature, can be easily implemented on top of state-of-the-art distributed computing facilities such as Grids or Clouds, providing a connected environment for medical data distributed over different sites and allowing medical experts to collaborate without being co-located, thus ensuring transparent access to patient records and biomedical knowledge and to computing resources needed for analyzing them. The resulting architecture, which natively provides services dealing with virtualization of distributed data regardless of their location, is able to support geographical transparency by allowing seamless access to the highly heterogeneous resources available in the biomedical domain, as well as to allow the distributed execution of complex analysis tasks, for the development of collaborative diagnoses and for sharing the knowledge between multiple medical centers.

Specifically, we have applied such a collaborative analysis methodology to the problem of early recognition of craniofacial pathologies in infants, but the methodology is generic enough for being applied to several types of disease. In detail, the first stage is the acquisition of data upon with the collaborative decision making works, in order to reach a diagnosis. A set of significant cranial features can be obtained by using thermoplastic material strips to acquire, in a very accurate manner, the contour of the skull. It is important to point out that such acquisition method turns out to be low cost, portable, deployable and non invasive. Subsequently, the patient's clinical data need to be digitalized and characterized by proper sets of points. Such digital inputs can be finally distributed within the identified group of experts by using proper interfaces and communication facilities.

The focus of this work is not on the means for distributing or accessing such data, which can be accomplished by using many mature interfaces, protocols and solutions available on modern grid or cloud platforms, but on the collaborative analysis and decision making methodologies used to implement an effective diagnostic service. More precisely, we combined Prospect Theory, Theory of Evidence and Fuzzy Linguistic Sets for developing a computer-aided diagnosis tool available over a Grid or Cloud through specific service interfaces (e.g., those provided within the Service Oriented Architecture framework). It is important to point out that the Theory of Evidence can be employed even in conflict resolution and context reasoning scenarios [4,5]. In detail, when an expert, which can be a doctor or a healthcare provider, receives the digital inputs of the cranial feature measurements, a proper recognition process is carried out according to the GS locally created by the expert from a training set of clinically healthy patients. We remark that the shape of such a local GS is conform with the one reported in the medical literature [6], as well as with doctor visual assessment and CI checking [7]. Any significant deviation with respect to the GS most likely denotes an anomaly in a given anatomical region of interest. In this work, we propose the method running at each expert side, which, besides being able to detect and quantify local deformations of the cranial shape (e.g., kinking, flattening and bumps), also detects other types of malformations, e.g., bilateral asymmetry, ear deviation and cranial elongation. It is important to emphasize that the method we propose also detects malformations that are highly symmetrical and in a mild form, such as the brachycephaly. Furthermore, we remark that our method may extend and improve the conventional analysis of cranial malformations based on the *Cephalic Index* (*CI*), which can take place in a collaborative manner. We also propose a method for the clinical characterization of eventually detected anomalies. In detail, our approach is based on morphological deformations which characterize a given pathology. We use this approach since it may happen that some pathologies do not occur in the canonical form defined in the medical literature, and a given patient may be simultaneously affected by several different pathologies, eventually mixed among them. Therefore, limiting the classification to the pathologies commonly defined in the literature is not very useful for analyzing malformations which do not fall into the common case histories. After running our detection approach, each expert has performs a multi-criteria assessment of the received input, where a digit and a severity degree have been associated to each assessment criterion for the considered pathology. Then, the multi-criteria assessment is returned back at the patient side and properly combined by using the Theory of Evidence, in order to obtain a single diagnosis concerning the kind of pathology affecting the patient. Such an outcome may be send back to the expert, in case of its wrong assessment, in order to provide a feedback for improving the expert GS (*i.e.*, the cranial features wrongly recognized as abnormal are stored in the GS so as to never repeat the misdiagnosis again in the future). We have used prospect theory in order to assess the consensus degree of a decision maker with the majority and to adjust its weight when participating in future collaborative decision making. Finally, we have used Linguistic Fuzzy Set to model the qualitative assessment of human experts and to combine them with the

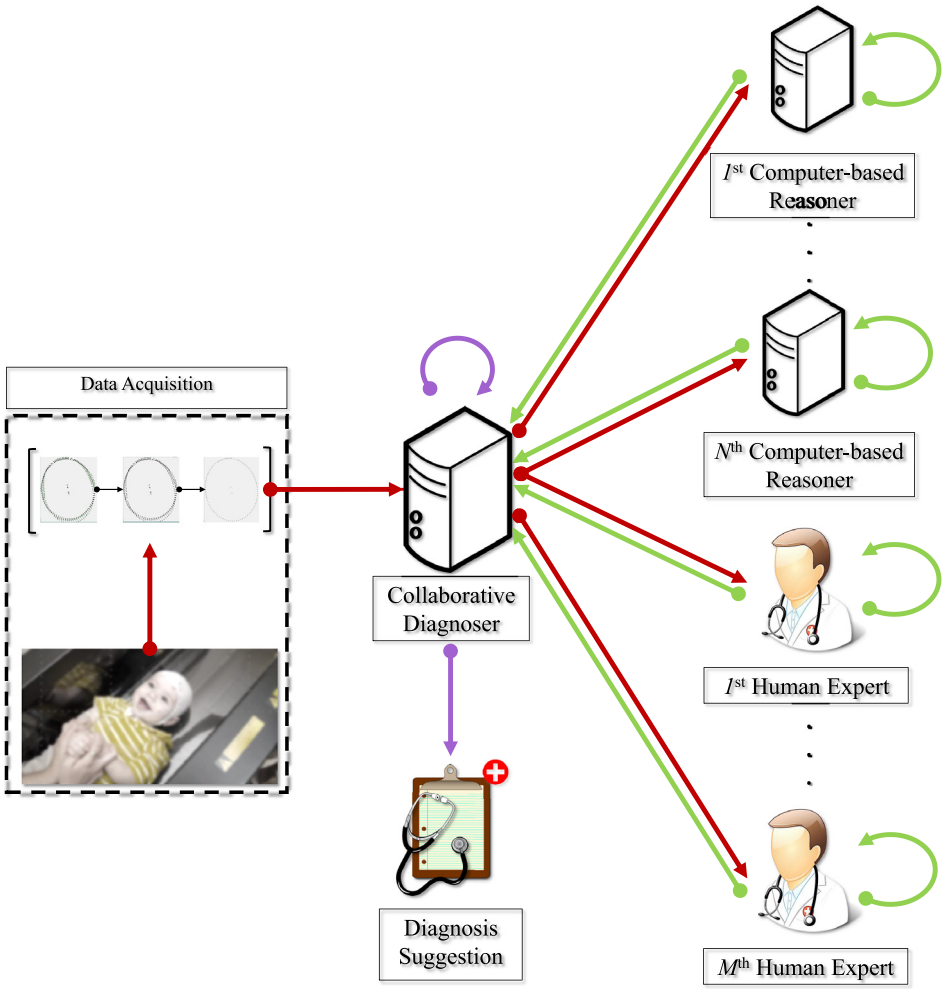


Fig. 1. An overview of the proposed approach for collaborative diagnosis, with application to craniofacial pathologies.

quantitative measures of automated applications for diagnosis analysis. Notice that for what concerns uncertain data analysis, also some other effective methods based on cognitive approaches have been proposed [8,9].

This paper is organized as follows. In Section 2 we describe our proposed methodology for the collaborative clinical analysis of a patient. In Section 3 we propose an application of our collaborative method to the problem of early recognition of craniofacial pathologies in infants. Finally, in Section 4 we draw the conclusions.

2. Collaborative diagnosis

Our driving idea is to adopt the widely known approach of group decision making, in order to support a collaborative diagnosis based on the input healthcare characteristics and a group of experts. Our approach is depicted in Fig. 1, which has been applied to the case of craniofacial pathologies.

Specifically, the first step concerns the acquisition of data by means of a proper equipment used on the patient of interest. The collected data represents an overall picture of the key health state measures of the patient that may support a clinical diagnosis of the possible pathology. Despite the possible measurement approach or equipment used for the data acquisition, at the end of this stage we have a vector $I = \{i_0, i_1, \ldots, i_\iota\}$, with the numerical representation of the measures for the ι measures conducted on the patient. Such a vector is provided to a service, which may be hosted on a cloud platform, implementing our collaborative approach. Such a service forwards the I vector to the set of experts forming the collaborative group, which is composed of N distributed automated reasoners and M geographically-sparse human experts. The size of the group, *i.e.*, $s = N + M$, must be greater than 2. When each expert receives the I vector, a diagnosis decision is determined and returned to the service, which properly aggregate the received decisions in order to formulate a diagnosis suggestion to be presented to the service user. In addition, the service provides also a feedback to each expert, in order to improve its accuracy in the further interactions. Our collaborative diagnosis approach has to deal with three key challenges:

- how to design the group decision making process when experts return their decisions back to the collaborative *diagnoser*?
- how to handle the possibility to have decisions determined in a qualitative manner by an automated application and the ones in a qualitative manner by a human expert?
- how to assess trustability of experts when reporting back their decisions?

In the following subsections we present a solution for each of the above-mentioned challenges, leaving in the following section the determination of the qualitative diagnosis for the particular case of craniofacial pathologies.

2.1. Group-based diagnosis

Let $E = \{e_1, \ldots, e_N, e_{N+1}, \ldots, e_{N+M}\}$ be the set of decision makers, where the first N are the automated applications, while the

remaining ones are the human experts. Each of the $N + M$ experts receives the I vector and has to return a decision. Specifically, we have a set of alternatives $X = \{x_1, x_2, \ldots, x_\kappa\}$, with $\kappa \geq 2$. The decision is expressed as a preference ordering from the best alternative to the worst one, based on some calculations on the provided inputs. Such a preference information can be represented in many ways, as illustrated in [10], but in our approach we have preferred the utility functions. Specifically, a vector of real numbers within the unit interval, namely $\bar{u^i}$, is associated by the decision maker to the ith alternative. In fact, a series of criteria must be assessed in order to reach a diagnosis, by determining the best pathology that may affect the patient under the assessed symptoms. Within the current health practice a concrete example is provided by the Bradford Hill criteria [11], otherwise known as Hill's criteria for causation, which consists of a group of minimal conditions necessary to provide adequate evidence of a causal relationship between an incidence and a possible consequence. Such a set is composed by several criteria, among which the most considered ones are the following:

Strength —a measure of the likelihood of the association between the observed symptoms and the pathology;

Consistency —a measure of the reproducibility of the association between the observed symptoms and the pathology, as observed by different persons in different places;

Specificity —a quantification that the association between the observed symptoms and the pathology is likely to be observed only within a very specific population at a specific site, based on racial and environmental causes;

Plausibility —a plausible mechanism between the observed symptoms and the pathology exists to motivate their association.

Our approach is generic and not bounded to any specific criteria used in the diagnosis approach by each decision maker. We only assume that the group of decision makers have beforehand agreed on the set of criteria to be assessed by each possible alternative pathology that a patient may be affected, so that the best one is the alternative receiving the highest averaged value of utility, while the worst one is that with the lowest value:

$$e_i \succeq e_j \rightarrow OWA(\bar{u}^i) \geq OWA(\bar{u}^j), \quad \forall i \neq j. \tag{1}$$

Therefore, our collaborative diagnosis can be solved as the problem of selecting the alternative e^* with the highest aggregated utility, among the ones received by the group members. The aggregation of the received utilities is determined by means of the *Ordered Weighted Averaging (OWA)* operator. Formally, an OWA operator of dimension n is a mapping $OWA : \psi^n \rightarrow \psi$ that has an associated collection of weights $W = [w_1, \ldots, w_n]$ lying in the unit interval and summing to one, *i.e.*, $\sum_{1=1}^{n} w_i = 1$, such that

$$OWA(a_1, \ldots, a_n) = \sum_{i=1}^{n} w_i b_i, \tag{2}$$

where b_i is the ith largest element of the vector of number provided as the input of the operator. The weighting strategy is a key element for the effective aggregation of the utilities and we will see in the last subsection how it is possible to set it in order to promote truth-telling among the decision makers.

In our collaborative diagnoser, the ordering of the alternatives is based on the *Dempster–Shafer theory*, which represents a generalization of the *Bayesian theory of subjective probability* (*i.e.*, the numeric measure of chance reflecting the degree of a personal belief in the likelihood of an occurrence for a given event) of representing epistemic plausibility, but yielding answers that contradict those achieved by using the classic probability theory.

Indeed, the traditional probability theory is not able to cope with particular situations involving epistemic uncertainty, due to the axiom of additivity and the *Principle of Insufficient Reason* [12]. On the contrary, the Dempster–Shafer theory does not impose these two assumptions so as to handle conflicting evidence, to account for evidence that can be assigned to multiple possible events and to manage evidence that, if combined, bring more or less credence to the likelihood of an event. Specifically, the Dempster–Shafer theory supposes the definition of a set of hypotheses Θ, called the frame of discernment:

$$\Theta = \{H_1, H_2, \ldots, H_N\}, \tag{3}$$

where N is the total number of hypotheses in the frame. In this set Θ, each hypothesis H_i is mutually exclusive (*i.e.*, at most one has to be true), while the set Θ is exhaustive (*i.e.*, at least one H_i has to be true). From Θ, it is possible to build the power set for the frame of discernment, which consists of all the possible 2^N combinations of the N hypotheses, or possible subsets of Θ:

$$P(\Theta) = \{\emptyset, \{H_1\}, \{H_2\}, \ldots, \{H_N\}, \{H_1 \cup H_2\}, \{H_1 \cup H_3\}, \ldots, \Theta\}. \tag{4}$$

Each hypothesis is supported by a given number of evidence, and such a degree of support assumes a value between 0 (*i.e.*, no evidence for the given hypothesis) and 1 (*i.e.*, all available evidence support the given hypothesis). Such a value does not have to be considered as a probability, since the belief in a hypothesis and its negation do not make 1, and both beliefs can be 0 (*i.e.*, no evidence for or against the given hypothesis). Such a degree of support is measured by means of the mass function $m(A)$ (where A is a member of the power set), distributed among the elements of the power set. So, the mass function is defined as:

$$m : P(\Theta) \rightarrow [0, 1] \tag{5}$$

satisfying the following two conditions:

$$\sum_{A \in P(\Theta)} m(A) = 1, \quad m(\emptyset) = 0. \tag{6}$$

The evidence supporting the hypotheses in the frame of discernment comes from a given subjective source of information, whose trustability is measured by means of a given real number α, belonging to the unity interval between 0 (indicating a source with no trustworthiness) and 1 (indicating a source with complete trustworthiness). Such a discounting coefficient can be inserted in the mass function, in order to reflect in such a function the trustability of the source of information upon which the function is built:

$$\exists \alpha \in [0, 1] \Longrightarrow \begin{cases} m^\alpha(\Theta) = (1 - \alpha) + \alpha \cdot m(\Theta) \\ m^\alpha(A) = \alpha \cdot m(A), \\ \forall A \subset \Theta \quad \text{and } A \neq \emptyset. \end{cases} \tag{7}$$

It is possible to encompass more than one source of information providing the evidence supporting the elements in the frame of discernment. The Dempster–Shafer theory encompasses the means to jointly consider the heterogeneity in this sources and any possible conflicts within the information they provide. It is possible to obtain a mass function by combining the mass functions relative to each source of information by means of the Dempster's rule of combination (also called the orthogonal sum) [12]: given two sources of information and their relative mass functions, namely m_1 and m_2, the overall mass function is computed as follows:

$$m(A) = m_1 \oplus m_2 = \frac{\sum_{B \cap C = A} m_1(B) \cdot m_2(C)}{1 - k} \quad \text{with} \quad k = \sum_{B \cap C = \emptyset} m_1(B) \cdot m_2(C) \tag{8}$$

where k quantifies the degree of conflict between m_1 and m_2, spanning from 0 (indicating no conflicts) to 1 (indicating no agreement). The comparison of the discounting coefficients of two different source of information is able to quantify possible conflicts among them: if $\alpha_i > \alpha_j$, the mass function from the ith source is more reliable and trustworthy than the one from the jth source. The mass function can be used to quantify the belief about the hypotheses in the power set, which is represented as intervals, bounded by two values: belief (or support) and plausibility. More precisely, the belief of a hypothesis can be defined as follows:

$$bel(A) = \sum_{\emptyset=B\subseteq A} m(B), \quad \forall A \subseteq \Omega. \tag{9}$$

On the other hand, the plausibility is defined as follows:

$$Pl(A) = bel(\Omega) - bel(\bar{A}) = \sum_{B:A\cup B=\emptyset} m(B), \quad \forall A \subseteq \Omega. \tag{10}$$

The probability of a hypothesis lies within the lower and upper bounds of belief and plausibility:

$$Bel(A) \leq P(A) \leq Pl(A), \quad \forall A \subseteq \Omega. \tag{11}$$

Such a probability is uniquely determined if $Bel(A) = Pl(A)$, otherwise, $Bel(A)$ and $Pl(A)$ provide a fuzzy representation of the probability. The belief interval expressed in Eq. (11) plays a key role when decisions need to be taken. It is possible to derive a crisp number from the belief interval by means of the so called *pignistic probability transformation* [13], which is built based on the expected utility theory to represent beliefs when taking the optimal decision that maximizes the expected utility within the context of a decision making process [14]. Such a transformation determines the pignistic probability as follows:

$$BetP(A) = \sum_{W\subseteq\Omega, A\in W} \frac{m(W)}{|W|}, \quad \forall A \subseteq \Omega, \tag{12}$$

where $|W|$ is the number of entities contained in W.

As above mentioned, our collaborative diagnoser is based on determining a proper ranking of the possible alternative solutions of a diagnosing process, resulting in the choice of the best solution as the one with the highest ranking score. Each of the candidates to motivate the symptoms affecting the patient in the X vector can be considered as the basic information providing evidence to the two elements that constitute the frame of discernment of our selection process: $\Theta = \{IS, NS\}$, where IS indicates the ideal solution, while NS the negative ideal solution. Then, each decision maker supports with proper evidence the hypothesis that the given solution s is the ideal solution, the negative ideal solution or none of them, and proper mass functions and discounting coefficients can be determined. More precisely, let us consider a particular criteria, to which each decision maker has assigned a utility value for each alternative pathology, so as that the utility values of the jth decision maker to the ith criterion assume the following values $\{0.6, 0.65, 0.7, 0.5, 0.9, 0.8\}$, then the maximum distance $d_{\max}$ is 0.9, while the minimum distance $d_{\min}$ is 0.5. For each alternative pathology, it is possible to compute the distance from the ideal solution (*i.e.*, the one with maximum distance), the one from the negative ideal solution (*i.e.*, the one with minimum distance) and the distance from both:

$$\begin{aligned} d_{i,s}(IS) &= |d_i(s) - d_{\max,i}|, \\ d_{i,s}(NS) &= |d_i(s) - d_{\min,i}|, \\ d_{i,s}(IS, NS) &= \left| d_i(s) - \frac{d_{\max,i} - d_{\min,i}}{2} \right|, \end{aligned} \tag{13}$$

where $d_{i,s}$ indicates the distance of the alternative pathology s according to the ith criterion. Knowing such distances, it is possible to compute the mass functions of the three hypotheses:

$$\begin{aligned} m_{i,s}(IS) &= \frac{d_{i,s}(IS)}{d_{i,s}(IS) + d_{i,s}(NS) + d_{i,s}(IS, NS)}, \\ m_{i,s}(NS) &= \frac{d_{i,s}(NS)}{d_{i,s}(IS) + d_{i,s}(NS) + d_{i,s}(IS, NS)}, \\ m_{i,s}(IS, NS) &= \frac{d_{i,s}(IS, NS)}{d_{i,s}(IS) + d_{i,s}(NS) + d_{i,s}(IS, NS)}. \end{aligned} \tag{14}$$

The decision makers have agreed upon which criteria to be used for the collaborative diagnosis, but they have also assigned a weight to each criterion, where the naive case is an equal distribution of the weights among the criteria, *i.e.*, if we have c criteria then $w_i = 1/c$. Given those weights, it is possible to compute from them the discounting coefficients by finding the weight with the highest value and dividing each weight by the highest one:

$$\alpha_i = \frac{w_i}{w_{\max}}, \quad \text{with } w_{\max} = \min_{i\in[0,N+1]} w_i, \tag{15}$$

where α_i is the coefficient for the ith criterion. Given such coefficients, it is possible to discount according to Eq. (7) the mass functions obtained by applying Eq. (14):

$$\begin{aligned} m_{i,s}^{\alpha_i}(IS) &= \alpha_i \cdot m_{i,s}(IS), \\ m_{i,s}^{\alpha_i}(NS) &= \alpha_i \cdot m_{i,s}(NS), \\ m_{i,s}^{\alpha_i}(IS, NS) &= \alpha_i \cdot m_{i,s}(IS, NS) + (1 - \alpha_i). \end{aligned} \tag{16}$$

As soon as the mass functions with respect to the number of diagnosis criteria have been determined, they can be combined by using the Dempster's combination rule in Eq. (8):

$$\begin{aligned} m_{i\oplus j,s}(IS) &= \frac{m_{i,s}^{\alpha_i}(IS) \cdot m_{j,s}^{\alpha_j}(IS, NS) + m_{i,s}^{\alpha_i}(IS, NS) \cdot m_{j,s}^{\alpha_j}(IS)}{1 - [m_{i,s}^{\alpha_i}(IS) \cdot m_{j,s}^{\alpha_j}(NS) + m_{i,s}^{\alpha_i}(NS) \cdot m_{j,s}^{\alpha_j}(IS)]}, \\ m_{i\oplus j,s}(NS) &= \frac{m_{i,s}^{\alpha_i}(NS) \cdot m_{j,s}^{\alpha_j}(IS, NS) + m_{i,s}^{\alpha_i}(IS, NS) \cdot m_{j,s}^{\alpha_j}(NS)}{1 - [m_{i,s}^{\alpha_i}(IS) \cdot m_{j,s}^{\alpha_j}(NS) + m_{i,s}^{\alpha_i}(NS) \cdot m_{j,s}^{\alpha_j}(IS)]}, \\ m_{i\oplus j,s}(IS, NS) &= \frac{m_{i,s}^{\alpha_i}(IS, NS) \cdot m_{j,s}^{\alpha_j}(IS, NS)}{1 - [m_{i,s}^{\alpha_i}(IS) \cdot m_{j,s}^{\alpha_j}(NS) + m_{i,s}^{\alpha_i}(NS) \cdot m_{j,s}^{\alpha_j}(IS)]}. \end{aligned} \tag{17}$$

Eq. (17) is iteratively used until obtaining a unique mass function, namely m_s, that has combined all the functions related to the overall criteria. From such a function, it is possible to adopt the pignistic probability transformation of Eq. (12) and define a ranking among the alternative pathologies, by ordering them according to the computed $BetP^{(s)}(IS)$:

$$BetP^{(s)}(IS) = m_s(IS) + \frac{m_s(IS, NS)}{2}. \tag{18}$$

Based on the result of Eq. (18), the pathology exhibiting the highest value for $BetP^{(s)}(IS)$ is the one to return as a solution to our selection problem.

2.2. Combined qualitative and quantitative representation of expert decisions

In our collaborative diagnosis approach, we have assumed that the group of decision makers can be composed of human experts and automated applications, so our approach have encompassed qualitative and quantitative representation of expert decisions according to the several criteria for a diagnosis. The typical way of indicating assessment of a criteria by means of a number is not viable for human experts, but is more useful if fuzzy variables accepting a number of fuzzy linguistic terms [15] are used to express such an assessment. Specifically, to each criteria the expert can assign a linguistic term; for simplicity, we consider that

only three linguistic terms are sufficient [16]: *LOW*, *MEDIUM* and *HIGH*. To each linguistic term it is possible to associate a fuzzy set characterized by a proper membership function $\mu_{\tilde{A}}$ with a triangular, trapezoidal or other shapes. The membership functions of the three linguistic terms used in this work are respectively a R-function, a trapezoidal function and a L-Function, formalized as follows:

$$\mu_{\tilde{L}_1}(x) = \begin{cases} 1 & B_{lower} \leq x < a_1 \\ \dfrac{b_1 - x}{b_1 - a_1} & a_1 \leq x \geq b_1 \\ 0 & b_1 < x \leq B_{upper} \end{cases}$$

$$\mu_{\tilde{L}_2}(x) = \begin{cases} 0 & B_{lower} \leq x < a_2 \\ \dfrac{x - a_2}{b_2 - a_2} & a_2 \leq x \geq b_2 \\ 1 & b_2 \leq x \geq c_2 \\ \dfrac{d_2 - x}{d_2 - c_2} & c_2 \leq x \geq d_2 \\ 0 & d_2 < x \leq B_{upper} \end{cases} \tag{19}$$

$$\mu_{\tilde{L}_3}(x) = \begin{cases} 0 & B_{lower} \leq x < a_3 \\ \dfrac{x - a_3}{b_3 - a_3} & a_3 \leq x \geq b_3 \\ 1 & b_3 < x \leq B_{upper}. \end{cases}$$

Furthermore, using adjective to qualify the related nouns is a common practice in natural language, *e.g.*, the linguistic term *MEDIUM* can be combined with *VERY* or *QUITE* to strengthen or weaken the feature characterization expressed by the term. This can be modeled by using a hedge as a modifier of a fuzzy value and transforming the relative membership function. In the literature, it is possible to find several kinds of hedges, but in this work we will use only two kinds: a concentration hedge, denoted as *VERY*, which implies the membership function to concentrate around the points with higher degree:

$$\mu^{+}(x) = \mu(x)^p, \quad \forall p > 1 \tag{20}$$

and a dilatation hedge, denoted as *QUITE*, which decreases the membership function:

$$\mu^{-}(x) = \sqrt[p]{\mu(x)}, \quad \forall p > 1. \tag{21}$$

In this work, we have assumed $p = 2$. The tuning of the parameters characterizing the membership functions for the linguistic terms within a fuzzy variable can be made by experts and must be tailored to the specific features of the measure it represents. In this work we assume an equal distribution of the fuzzy set across the unit interval.

To deal with the linguistic fuzzy terms in the collaborative diagnosis approach of the previous subsection, we make use of defuzzification of the linguistic terms provided by the human experts, denoted as $\phi(l)$, where l is a fuzzy linguistic label in the label set Λ. Defuzzification of a linguistic label is performed through the centroid method, by considering the membership function of the linguistic term as the input to the operation. Such a method determines the geometric center of the membership function over the x-axis.

2.3. Weighting of experts' decisions

Each expert is characterized by a given Gold Standard and a precision in providing correct decisions back to the collaborative diagnoser. In order to further improve the diagnoser accuracy across successive iterations and use, a proper weighting of the collaborating decision makers is needed, as above mentioned. Such a weighting scheme should promote the deciders exhibiting higher precision in their replies, and have been designed based on the *prospect theory* [10]. Such a theory is descriptive for forecasting individual actual decision behavior under risk, where the outcomes of the decision making is assumed as gains or losses with respect to a reference point. Respectively, a distance between the single decider aggregated utility value, computed as the OWA operator of the $\bar{u}_k^i$ utility vector for the ith decider for the kth alternative, and the utility value of the selected alternative is determined. A proper S-shaped prospect value function $v(x)$ denotes the gain (when its input is non negative) or loss (when its input is negative) for the decider on the kth alternative:

$$v_k(x) = \begin{cases} x^{\alpha}, & x \geq 0 \\ -\lambda(-x)^{\beta}, & x < 0 \end{cases} \tag{22}$$

where x is the distance for the decider to the selected alternative, α and β are two parameters indicating the concavity and convexity of the function, respectively, and λ is the coefficient of loss aversion. As in [10], we assume that $\alpha = \beta = 0.88$ and $\lambda = 2.25$. After computing the prospect values for all the deciders on the kth alternative, namely $v_k^1, v_k^2, \ldots, v_k^{N+M}$, we can determine the collective prospect value:

$$v_k^c = \frac{1}{N+M} \sum_{1=1} N + M v_k^i, \tag{23}$$

obviously the selected alternative has the highest prospect value. From those prospect values, we can compute a consensus degree among the group of the deciders upon the kth alternative, named as *Prospect Value Consensus Degree (PVCD)* [17], as follows:

$$PVCD(e^k) = \sqrt{\frac{1}{N+M} \sum_{1=1}^{N+M} \left(\frac{v_k^i - v_k^c}{v(N+M) - v(-(N+M))} \right)^2}, \tag{24}$$

where e^k indicates the decision maker. If the PVCD is greater than a given threshold, which we have fixed to 0.75, then its weight can be increased by a fixed constant equal to 0.05; whereas, if the PVCD is lower than the fixed constant, its weight is decreased by 0.05. In this manner, we consider in our approach more the deciders that have a consensus with the majority of the other deciders and decrease the consideration in the ones that diverge with respect to the majority.

3. Application to craniofacial pathologies

In our proposed approach, we have defined a way to diagnose a cranial disease based on the theory of evidence, fuzzy linguistic sets and prospect theory. The starting point of this approach is represented by a set of inputs, namely the set I, upon which our solution makes inference and takes a decision. Such inputs are taken from the cranial data acquisition and are a quantitative measure of possible kinds of malformations by which a patient may be affected. Their characterization is taken from the literature and the widely-accepted medical practice and procedures. Another set of interest for our solution is X, which contains all the possible types of disorders whose symptomatic manifestations are contained in I. In addition, the medical practice has defined a set of criteria, namely $c_{i,j}$, in the form of if-then logical rules based on the satisfaction of an activation condition by the inputs in I: if a condition is verified, then the patient may be possibly affected by one or more diseases in X. Those criteria are not disjoint, but may overlap among each others. For example, an input profile obtained by cranial data acquisition may jointly satisfy more than one activation criterion. Therefore, to each criterion it is possible to associate a plausibility measure $u_{i,j}$ that indicates the satisfaction

Table 1
Anthropometric points used by the proposed acquisition method.

Anthropometric point	*Description*
g	Glabella
v	Vertex
c	Center
op	Opisthocranion
ft_r, ft_l	Frontotemporale right and left
ex_r, ex_l	Exocanthion right and left
fz_r, fz_l	Frontozygomaticus right and left
obi_r, obi_l	Otobasion inferius right and left
po_r, po_l	Porion right and left
t_r, t_l	Tragion right and left
obs_r, obs_l	Otobasion superius right and left
fz_{r-b}, fz_{l-b}	Frontozygomaticus back right and back left

degree by the inputs of the ith criterion considering the jth disease kind. Based on the vector of the plausibility measures for each disease, our method is able to determine the most probable diagnosis for the patient upon which a given input has been taken. The plausibility degree can be estimated in a qualitative manner by a specialist, which assigns a linguistic label to each $c_{i,j}$, or in a quantitative way, through input processing by a software, as described in the following.

3.1. Data acquisition

We define a "*view*" as the patient's cranial morphology observed from the top. The acquisition of such a view is performed by using an extended version of the acquisition method proposed in [18], which is based on a strip of thermoplastic material. Such a method is easy to apply, non-invasive and reliable, besides having a good clinical accuracy and low application cost. In particular, in order to get such a view, the thermoplastic material is positioned by the doctor around the patient's head, in correspondence of the widest transverse circumference. Subsequently, the doctor manually sets on the thermoplastic material a marker (landmark) in correspondence of each anthropometric point defined in Table 1, through a standardized protocol. We remark that the method we propose uses a wider set of markers than the one proposed in [18].

Finally, the strip and the relative markers are digitalized to be represented by a proper set of points, through well-known object detection procedures [19,20]. In particular, the acquisition method first performs the digitalization of the landmarks to get the relative point coordinates. We assume that the output of landmarks digitalization is stored into the *LANDMARK* point set. Again, we denote by *anatomical region* the set of points characterizing a clinically relevant portion of the cranial vault. Each region is comprised between the two landmarks which include it. Then, on each anatomical region, a set of equally spaced points is automatically positioned through the approaches proposed in [20]. Such points are denoted as *pseudo-landmarks*. The output of pseudo-landmarks digitalization is stored into the *PSEUDOLANDMARK* set of points. The *LANDMARK* and *PSEUDOLANDMARK* sets, if considered together, define the *CONTOUR* set, which represents the contour characterizing the morphology of a view (see Fig. 2).

3.2. Gold Standard creation

A GS is the shape which best characterizes the cranial morphology of a clinically healthy patient. We remark that the GS is relative to a specific class of patients, which is characterized by *age*, *gender* and *ethnicity* [21]. In detail, the GS is characterized, for each class of patients, by an average shape and some clinically relevant distances. More precisely, to create the GS, our method first performs the shape averaging [22]. In particular, to create the average shape, we adopt a technique based on *Procrustes Superimposition* [23], since it is the one which ensures the estimates with the least error and no pattern of bias [24]. In our method all the distances are calculated on the GS using the *Euclidean metric*.

3.2.1. Shape averaging

Formally, a *shape* is a member of an equivalence class, formed by removing from a given object the components of translation, rotation and scaling. Our method creates the average shape as described in Algorithm 1.

Algorithm 1 Shape Averaging Procedure

Input:

- *TS*: Vector constituting the training set for a given class of patient.
- *Origin*: Origin of the reference system.
- *ReferenceSystem*: Set of points to rotate the first shape.

Output:

- *GS*: Average shape constituting the Gold Standard.

procedure SHAPE_AVERAGING(*TS*, *Origin*, *ReferenceSystem*)
/* Choose at random an element from the training set */
  $element \leftarrow TS$;
  $element^{TR} \leftarrow$ Translate($element$, $origin$);
  $element^{TR-SC} \leftarrow$ Scale($element^{TR}$);
  $element^{TR-SC-RT} \leftarrow$ Rotate($element^{TR-SC}$, $ReferenceSystem$);
  $GS \leftarrow element^{TR-SC-RT}$;
/* Average other shapes */
  for i = 1 **to** size(TS \ {element}) **do**
    $element_i \leftarrow TS$;
    $element^{TR} \leftarrow$ Translate($element_i$, $origin$);
    $element^{TR-SC} \leftarrow$ Scale($element^{TR}$);
    $element^{TR-SC-RT} \leftarrow$ Rotate($element^{TR-SC}$, GS);
    $GS \leftarrow$ computeAverage($element^{TR-SC-RT}$, GS);
  end for
  return GS;
end procedure

In detail, Algorithm 1 performs the centering with respect to the reference system of an element randomly selected from the training set. Subsequently, the selected element is scaled, so that the *Root Mean Square Distance* (*RMSD*) from its points to the translated origin is equal to one (*uniform scaling*). After the removal of translation and scaling components, the selected element is aligned with respect to the reference system. Once the aforementioned three components have been removed from the first chosen element, the shape averaging procedure creates the average shape, by using all the other elements belonging to the training set, as shown in Fig. 3.

3.2.2. Gold Standard morphology

The techniques commonly employed for comparing objects within images are usually based on the Hausdorff distance [25]. However, through such a distance, it is not possible to assess if in a given anatomical region, misshapen with respect to the GS, there is a *flattening* or a *swelling* [26]. Therefore, we use the *euclidean distance*. In detail, the first measure characterizing the GS is calculated from its centroid to each point belonging to the relative *Average* set, which represents the output of the shape averaging procedure. Therefore, the morphology of a given anatomical region, characterized by its constituting points, is estimated as the distance of these points from the centroid. Another important set of measures which characterizes the cranial

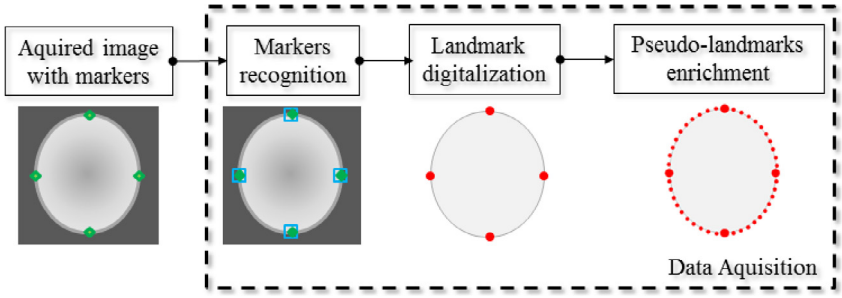


Fig. 2. An overview of the phases for cranial data acquisition.

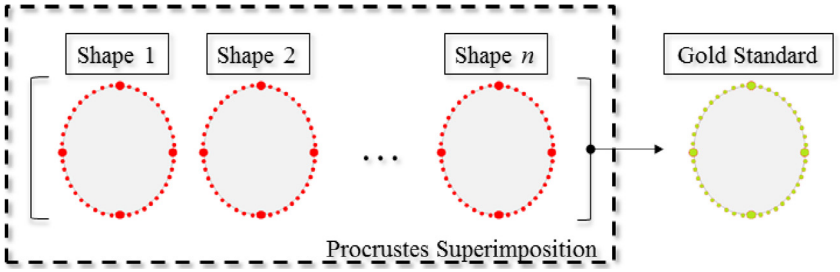


Fig. 3. An overview of the phases for the creation of the Gold Standard.

morphology is that concerning the length and the width of the skull. The former is given by the distance between the landmarks g and op, whereas, the latter is given by the distance between the landmarks t_l and t_r. Let $d^{GS}_{t_l,t_r}$ and $d^{GS}_{g,op}$ be such distances. The distances $d^{GS}_{t_l,t_r}$ and $d^{GS}_{g,op}$ are used to evaluate the cranial elongation, by calculating the ratio between the width and the length of the skull. In addition, our method also calculates the distance from the centroid to the landmarks g and op, respectively. Let $d^{GS}_{g,c}$ and $d^{GS}_{c,op}$ be such distances. In detail, by means of $d^{GS}_{g,c}$ and $d^{GS}_{c,op}$, the information about the lengthening/shortening of both the anterior and posterior hemispheres of the skull can be obtained. Finally, some other distances are calculated in the GS to characterize the *ears displacement*. In detail, we calculate the distances between g and t_l, and g and t_r, respectively, as proposed in [18]. Let d^{GS}_{g,t_l} and d^{GS}_{g,t_r} be such distances. We remark that for each point of the *Average* set we define a tolerance interval (*TolFactor*), representing the range within it can vary, in a healthy patient, the distance between that point and the centroid.

3.3. Clinical analysis

The clinical analysis of a patient, carried out by our method, is performed according to the proper relative GS. In particular, our method first locates and quantifies local malformations related to a single region, then it checks for eventual asymmetries. Finally, some other parameters commonly used for detecting and classifying craniofacial pathologies are also evaluated, namely, *ears deviation* and *cephalic index*.

3.3.1. Per-region checking

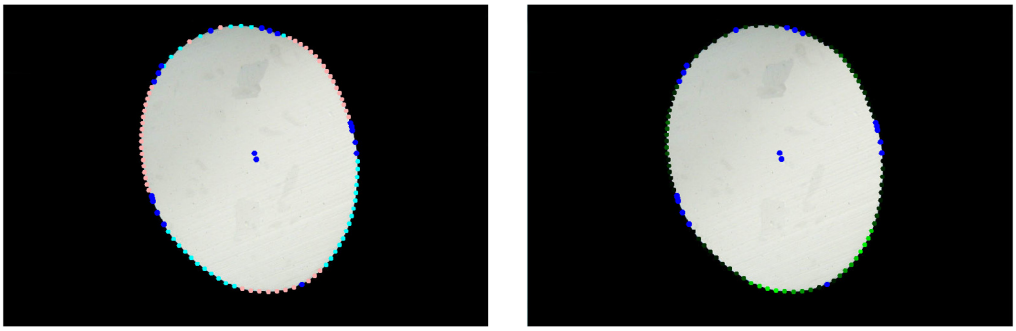
By using Algorithm 2 and Algorithm 3 each anatomical region is evaluated to detect malformations. More precisely, for each anatomical region, it is first checked if the relative points lie within the range defined by GS. Subsequently, for all the points which do not lie within the GS, further checks are performed. In particular, if a subset of those points is contiguous and does not lie within the range defined by the GS, it is checked if such points characterize a swell or a dump. Again, it is checked if a depression denotes a flattening. We remark that for each malformation our method also estimates the severity level.

Let $GS = \{R^G_1, \ldots, R^G_N\}$ be the set of regions belonging to the GS and $S = \{R^S_1, \ldots, R^S_N\}$ the one relative to the patient undergoing examination. Let *AnomalyFactor* be the value characterizing the "*tolerated*" contiguous points which do not lie within the GS range. Let Δ be the *TolFactor*. We remark that for $1 \leq i \leq N$, each R_i is a vector of distances, i.e., one for each point of this region. Notice that the *TolFactor* concerns a single point, while the *AnomalyFactor* is related to a set of points.

In detail, for what concerns the localization of flattening, the method we propose considers the extreme points characterizing a depression, denoted by A and B, respectively, then it draws a segment between them. Finally, our method checks if the points that lie between A and B have a distance from the drawn segment which is comprised between $-$*TolFactor* and $+$*TolFactor*. If this check is satisfied, it is likely that there is a flattening.

3.3.2. Bilateral checking

As it is mentioned in the literature [6], the top view of a healthy patient is highly symmetrical with respect to the imaginary segment passing through the landmarks g and op. We denote the above symmetry as "*bilateral symmetry*". More precisely, the anatomical regions of the view are divided in two sets: the former belonging to the left side of the skull and the latter belonging to the right one. Hence, each point in a side has the corresponding one in the other side. Therefore, to detect the bilateral asymmetry, both the patient being assessed and the relative GS are partitioned into two subsets, which characterize the left and the right hemisphere of the skull, respectively. We remark that each point in a side has the corresponding one in the other side. In detail, for each point, we consider the distance between this point and the centroid. It is important to point out that these distances are calculated for each region of the GS. In this way, our method evaluates the bilateral symmetry for each pair of corresponding points. Finally, in order to detect asymmetries, by using Algorithm 4 our method calculates for each region of the patient undergoing examination the aforementioned distances and compares them with respect to the corresponding ones of the relative GS.



(a) Local malformations. (b) Bilateral symmetry.

Fig. 4. A visual output of the method we propose. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Algorithm 2 Point Anomaly Checking

Input:

- GS: Gold standard.
- S: Shape undergoing examination.

Output:

- *LocalMalformation*: Type of malformation.
- *Severity*: Grade of malformation.

procedure DETECT_POINT_ANOMALY(GS, S)
  for i = 1 **to** N **do**
    $R_i^{GS} \leftarrow \{d_1^{R_i^{GS}}, \cdots, d_z^{R_i^{GS}}\}$;
    $R_i^{S} \leftarrow \{d_1^{R_i^{S}}, \cdots, d_z^{R_i^{S}}\}$;
    $Anomaly_i \leftarrow false$;
    $Counter_i \leftarrow 0$;
    for j = 1 **to** Z **do**
      $Diff_j = |d_j^{R_i^{GS}} - d_j^{R_i^{S}}|$;
      if $Diff_j > TolFactor$ **then**
        $Counter_i + +$;
Note that $|R_i^{DIFF}| \leq z$;
        Add the point $p_j^{R_i}$ and $Diff_j$ into R_i^{DIFF};
      end if
    end for
    if $Counter_i > 0$ **then**
      $OutOfRange_i \leftarrow Counter_i$;
      $Anomaly_i \leftarrow$ true;
      $R^{DIFF} \leftarrow R^{DIFF} \cup R_i^{DIFF}$;
    end if
  end for
  LocalMalformation$\leftarrow$Detect_Local_Malformation(R^{DIFF});
  return *LocalMalformation*, *Severity*;
end procedure

Algorithm 3 Malformations Checking

Input:

- R^{DIFF}: Vector representing regions affected by malformation.

Output:

- *LocalMalformation*: Type of malformation.
- *Severity*: Grade of malformation.

procedure DETECT_LOCAL_MALFORMATION(R^{DIFF})
  for i = 1 **to** sizeof(R^{DIFF}) **do**
    Let $R_{i_{in}}^{DIFF}$ the set of contiguous points in R_i^{DIFF} with distance from the centroid less than the relative ones in the GS;
    Let $R_{i_{out}}^{DIFF}$ the set of contiguous points in R_i^{DIFF} with have distance from the centroid greater than the relative ones in the GS;
    if sizeof($R_{i_{in}}^{DIFF}$)>*AnomalyFactor* **then**
      $LocalMalformation_i \leftarrow$Depression;
      $Severity_i \leftarrow$average of all the distances in $R_{i_{in}}^{DIFF}$;
      $Severity \leftarrow Severity \cup Severity_i$;
      if the majority of the points characterizing the depression lie within the range of the *TolFactor* **then**
        $LocalMalformation_i \leftarrow$ Flattening;
      end if
    end if
    if sizeof($R_{i_{out}}^{DIFF}$)>*AnomalyFactor* **then**
      $LocalMalformation_i \leftarrow$ Swelling;
      $Severity_i \leftarrow$average of all the distances in $R_{i_{out}}^{DIFF}$;
      $Severity \leftarrow Severity \cup Severity_i$;
    end if
  end for
  return *LocalMalformation*, *Severity*;
end procedure

3.3.3. Ears deviation

Several craniofacial pathologies may cause ears deviation. In order to address this problem, we calculate a tolerance threshold on the GS, by means of the distances between g and t_l, and g and t_r, respectively. Finally, in order to detect any deviation of the ears, the same distances are calculated on the patient undergoing examination and compared with those of the relative GS. In detail, ears deviation is assessed by means of Algorithm 5.

3.3.4. Enhanced cephalic index checking

The cephalic index is the ratio between the width and the length of the skull. More precisely, it is an important measure to estimate the proportions of the skull. In detail, the width is given by the distance from t_l to t_r, whereas, the length is given by the distance from g to op. Nowadays, this index is one of the most important parameters used for the evaluation of craniofacial pathologies. However, the evaluation of this index is usually performed in a subjective manner, according to the skills and the experience of the doctor. Instead, by using our method, the analysis based on the CI is enhanced, since it can exploit the GS.

In order to evaluate the skull elongation, we apply the procedure defined in Algorithm 6, which computes the cephalic index on both the GS and the patient undergoing examination, then it compares such two indices to detect any lengthening or shortening. In particular, to detect in which hemisphere of the skull (front/back) the eventual shortening or lengthening is more

Algorithm 4 Bilateral Checking

Input:

- GS: Gold standard.
- S: Shape undergoing examination.

Output:

- *Asymmetry*: Vector constituting the regions affected by asymmetry.
- *Severity*: Vector constituting the grade of severity for each regions affected by asymmetry.

procedure BILATERAL_CHECKING(GS, S)
  for i = 1 **to** k **do**
    $Z_i = \{d_1^{Z_i}, \cdots, d_z^{Z_i}\}, R_i = \{d_1^{R_i}, \cdots, d_z^{R_i}\}$;
    $G_i = \{d_1^{G_i}, \cdots, d_z^{G_i}\}, L_i = \{d_1^{L_i}, \cdots, d_z^{L_i}\}$;
    $Tolerance_i \leftarrow \emptyset$;
    $BS_i \leftarrow \emptyset$;
    for j = 1 **to** z **do**
  /*$Tolerance_i \leftarrow \{t_1, \cdots, t_z\}$*/
      $Tolerance_i \leftarrow Tolerance_i \cup |d_j^{Z_i} - d_j^{R_i}|$;
  /*$BS_i \leftarrow \{v_1, \cdots, v_z\}$*/
      $BS_i \leftarrow BS_i \cup |d_j^{G_i} - d_j^{L_i}|$;
      if $v_j < (t_j$ - $TolFactor)$ **OR** $v_j > (t_j + TolFactor)$ **then**
        $BilateralSC_i \leftarrow BilateralSC_i \cup j$;
        $Severity_i \leftarrow$ average of all the distances in $BilateralSC_i$;
        $Severity \leftarrow Severity \cup Severity_i$;
      end if
    end for
    if $|BilateralSC_i| > AnomalyFactor$ **then**
      $Asymmetry \leftarrow Asymmetry \cup i$;
    end if
  end for
  return *Asymmetry*, *Severity*;
end procedure

Algorithm 5 Ears Deviation Checking

Input:

- S: Shape undergoing examination.
- d_{g,t_l}^{GS}, d_{g,t_r}^{GS}: GS ears distances.

Output:

- *Deviation*: Ears deviation.
- *Severity*: Grade of ears deviation.

procedure EAR_DEVIATION($S, d_{g,t_l}^{GS}, d_{g,t_r}^{GS}$)
  $Tol_{EARS} \leftarrow |d_{g,t_l}^{G} - d_{g,t_r}^{G}|$;
  Compute d_{g,t_l}^{S} and d_{g,t_r}^{S};
  $ED = |d_{g,t_l}^{S} - d_{g,t_r}^{S}|$;
  if $ED < Tol_{EARS} - TolFactor$ **OR** $ED > Tol_{EARS} + TolFactor$ **then**
    $Deviation \leftarrow yes$;
    $Severity \leftarrow EarsDeviation$;
  end if
  return *Deviation*, *Severity*;
end procedure

pronounced, our method calculates the ratio between the width and the length of the front skull hemisphere, that is, the distance from the point g to the centroid. Finally, the ratio between the width and the length of the back skull hemisphere is calculated, i.e., the distance from the centroid to op.

Algorithm 6 Enhanced CI

Input:

- S: Shape undergoing examination.
- d_{t_l,t_r}^{GS}, $d_{g,op}^{GS}$, $d_{g,c}^{GS}$, $d_{c,op}^{GS}$: GS distances used for ears deviation.

Output:

- CI^S: CI of the shape undergoing examination.
- CI^{GS}: CI of the GS.
- CI_f^{GS}, CI_b^{GS}: Front and back CI of the GS.
- CI_f^S, CI_b^S: Front and back CI of the shape undergoing examination.
- *Response*: Clinical response.

procedure ENHANCED_CI($S, d_{t_l,t_r}^{GS}, d_{g,op}^{GS}, d_{g,c}^{GS}, d_{c,op}^{GS}$)
  $CI^{GS} \leftarrow \frac{d_{t_l,t_r}^{GS}}{d_{g,op}^{GS}}$;
  $CI_f^{GS} \leftarrow \frac{d_{t_l,t_r}^{GS}}{d_{g,c}^{GS}}$, $CI_b^{GS} \leftarrow \frac{d_{t_l,t_r}^{GS}}{d_{c,op}^{GS}}$;
  $CI_f^{S} \leftarrow \frac{d_{t_l,t_r}^{S}}{d_{g,c}^{S}}$, $CI_b^{S} \leftarrow \frac{d_{t_l,t_r}^{S}}{d_{c,op}^{S}}$;
  if $CI^S < CI^{GS}$ - $TolFactor$ **then**
    $Response \leftarrow shortening$;
  end if
  if $CI^S > CI^{GS} + TolFactor$ **then**
    $Response \leftarrow elongation$;
  end if
  return $CI^S, CI^{GS}, CI_f^{GS}, CI_b^{GS}, CI_f^S, CI_b^S, Response$;
end procedure

3.4. Clinical response

3.4.1. Characterization and visual assessment of malformations

In detail, in Fig. 4(a) the parts in blue denote "retractions", while the ones in pink denote "bumps". Instead, in Fig. 4(b), the more clear is the level of green the more pronounced the detected bilateral asymmetry. For example, parts in dark green denote mild bilateral asymmetry, whereas, the black ones denote symmetric parts.

4. Conclusions

Nowadays, the research community in health informatics is focusing its attention in approaches to support doctors and general practitioners in their job of recognizing pathologies from a series of evidences collected from the patients, with the aim of reducing the probability of misdiagnosis. The problem of misdiagnosis is more serious if we consider that some pathologies can affect infants and leave them permanent deformations and/or health problems, which can be avoided with the early diagnosis of the pathologies and by assuming proper corrective, even simple, actions at the early stage of the pathology. In order to reduce the costs caused by misdiagnosis and improve the probability of early diagnosis of pathologies in infants, hospitals and medical practices are turning their attention to clinical Decision Support Systems (DSS), to assist them in diagnosing pathologies based on the symptoms manifesting in the patients.

To deal with the above issues, we defined a method for the collaborative clinical analysis of a patient, which involves the acquisition of data, along with the detection and characterization of pathologies by a group of experts. More precisely, we have applied such a methodology to the problem of early recognition of craniofacial pathologies in infants, but the methodology is generic enough for being applied to several types of disease.

After running our detection approach, each expert has performs a multi-criteria assessment of the received input, where a digit and a severity degree have been associated to each assessment criterion for the considered pathology. Then, the multi-criteria assessment is returned back at the patient side and properly combined by using the Theory of Evidence, in order to obtain a single diagnosis concerning the kind of pathology affecting the patient. We have used prospect theory to assess the consensus degree of a decision maker with the majority and to adjust its weight when participating in future collaborative decision making. Finally, we have used Linguistic Fuzzy Set to model the qualitative assessment of human experts and to combine them with the quantitative measures of automated applications for diagnosis analysis.

5. Compliance with ethical standards

All the data and images used in the example application have been taken by resources available in the literature. Furthermore, the work does not contain any study with human participants performed by any of the authors.

References

[1] K.M. Kelly, T.R. Littlefield, J.K. Pomatto, C.E. Ripley, S.P. Beals, E.F. Joganic, Importance of early recognition and treatment of deformational plagiocephaly with orthotic cranioplasty, Cleft Palate Craniofac. J. 36 (2) (1999) 127–130.

[2] P. Cerrato, Reducing the costs of misdiagnosis., Healthcare Finance News, available on line at http://www.healthcarefinancenews.com/news/reducing-costs-misdiagnosis.

[3] C.S. Mendoza, N. Safdar, K. Okada, E. Myers, G.F. Rogers, M.G. Linguraru, Personalized assessment of craniosynostosis via statistical shape modeling, Med. Image Anal. 18 (4) (2014) 635–646.

[4] D. Riboni, L. Pareschi, C. Bettini, Towards the adaptive integration of multiple context reasoners in pervasive computing environments, in: PerCom Workshops, 2010, pp. 25–29.

[5] D. Zhang, M. Guo, J. Zhou, D. Kang, J. Cao, Context reasoning using extended evidence theory in pervasive computing environments, Future Gener. Comput. Syst. 26 (2) (2010) 207–216.

[6] S. Standring, Gray's anatomy, The anatomical basis of clinical practice 3.

[7] W. Likus, G. Bajor, K. Gruszczyńska, J. Baron, J. Markowski, M. Machnikowska-Sokołowska, D. Milka, T. Lepich, Cephalic index in the first three years of life–study of children with normal brain development based on computed tomography, Sci. World J. 2014 (2014) 1–6. http://dx.doi.org/10.1155/2014/502836.

[8] L. Ogiela, Syntactic approach to cognitive interpretation of medical patterns, in: C. Xiong, H. Liu, Y. Huang, Y. Xiong (Eds.), Intelligent Robotics and Applications, First International Conference, ICIRA 2008, Wuhan, China, October 15-17, 2008 Proceedings, Part I, in: Lecture Notes in Computer Science, vol. 5314, Springer, 2008, pp. 456–462. URL http://dx.doi.org/10.1007/978-3-540-88513-9_49.

[9] L. Ogiela, M.R. Ogiela, Semantic analysis processes in advanced pattern understanding systems, in: Advanced Computer Science and Information Technology, Springer, 2011, pp. 26–30.

[10] Y. Dong, N. Luo, H. Liang, Consensus building in multiperson decision making with heterogeneous preference representation structures: A perspective based on prospect theory, Appl. Soft Comput. 35 (2015) 898–910.

[11] A.B. Hill, The environment and disease: Association or causation? Proc. Roy. Soc. Med. 58 (1965) 295–300.

[12] K. Sentz, S. Ferson, Combination of Evidence in Dempster-Shafer Theory, Vol. 4015, Citeseer, 2002.

[13] P. Smets, R. Kennes, The transferable belief model, Artificial Intelligence 66 (2) (1994) 191–234.

[14] P. Smets, Decision making in the tbm: the necessity of the pignistic transformation, Internat. J. Approx. Reason. 38 (2) (2005) 133–147.

[15] T.J. Ross, Fuzzy Logic with Engineering Applications, third ed., Wiley, 2010.

[16] W.-H. Au, K. Chan, Classification with degree of membership: a fuzzy approach, in: Proceedings IEEE International Conference on Data Mining, ICDM, 2001, pp. 35–42.

[17] Y. Dong, J. Xu, Consensus Building in Group Decision Making, Springer, 2016.

[18] L.A. van Vlimmeren, T. Takken, L.N.A. van Adrichem, Y. van der Graaf, P.J.M. Helders, R.H.H. Engelbert, Plagiocephalometry: a non-invasive method to quantify asymmetry of the skull; a reliability study, Eur. J. Pediatr. 165 (3) (2006) 149–157. URL http://www.biomedsearch.com/nih/Plagiocephalometry-non-invasive-method-to/16211401.html.

[19] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, CVPR 2001, Vol. 1, IEEE, 2001, pp. I–511.

[20] B. McCane, Shape variation in outline shapes, Syst. Biol. 62 (1) (2013) 134–146.

[21] M. Hubbe, T. Hanihara, K. Harvati, Climate signatures in the morphological differentiation of worldwide modern human populations, Anat. Rec. 292 (11) (2009) 1720–1733.

[22] C. Goodall, Procrustes methods in the statistical analysis of shape, J. R. Stat. Soc. Ser. B Stat. Methodol. (1991) 285–339.

[23] F. Rohlf, Shape statistics: Procrustes superimpositions and tangent spaces, J. Classification 16 (2) (1999) 197–223.

[24] F. Rohlf, Bias and error in estimates of mean shape in geometric morphometrics, J. Hum. Evol. 44 (6) (2003) 665–683.

[25] D. Huttenlocher, G. Klanderman, W. Rucklidge, Comparing images using the hausdorff distance, IEEE Trans. Pattern Anal. Mach. Intell. 15 (9) (1993) 850–863.

[26] A. Castiglione, R. Pizzolante, A. De Santis, C. D'Ambrosio, A collaborative decision-support system for secure analysis of cranial disorders, in: F. Xhafa, L. Barolli, F. Palmieri, M. Koeppen, V. Loia (Eds.), 2014 International Conference on Intelligent Networking and Collaborative Systems, Salerno, Italy, September 10-12, 2014, IEEE, 2014, pp. 189–196.