Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Biomedical Informatics 37 (2004) 99–107

Journal of Biomedical Informatics

www.elsevier.com/locate/yjbin

A computational TW3 classifier for skeletal maturity assessment. A *Computing with Words* approach

Santiago Aja-Fernández, Rodrigo de Luis-García, Miguel Ángel Martín-Fernández, and Carlos Alberola-López*

Laboratorio de Procesado de Imagen, ETSI Telecomunicación, Universidad de Valladolid, Valladolid, Spain

Received 22 December 2003

Abstract

This paper proposes a fuzzy methodology to translate the natural language descriptions of the TW3 method for bone age assessment into an automatic classifier. The classifier is built upon a modified version of a fuzzy ID3 decision tree. No large data records are needed to train the classifier, i.e., to find out the classification rules, since the classifier is built upon rules given by the TW3 method. Only small data records are needed to fine-tune the fuzzy sets used to implement the rulebase.

Keywords: Bone age assessment; TW3; Fuzzy granules; Fuzzy decision trees

1. Introduction

Bone age assessment is a frequently employed procedure in pediatric radiology, as many diseases and syndromes affecting growth result in a significant discrepancy between bone age and chronological age. A quantitative assessment of skeletal maturity is also useful for predicting adult height.

Two major methods are used for bone age assessment in children: the Greulich-Pyle (GP) method [1] and the Tanner-Whitehouse (TW3) method [2]. The former is an atlas-driven method and is based on visually comparing a non-dominant hand-wrist radiograph with a number of atlas patterns. Bone age is assessed on the basis of the pattern which most accurately resembles the clinical image according to the physician's perception. The TW3 method uses a detailed shape analysis of several bones of interest, leading to their individual classification into one of several stages. Scores are derived from each bone stage and summed to compute the assessment.

The subjective nature of the GP method and the considerable complexity of TW3 method, makes the automation of bone age assessment a highly desirable goal, in order to assist the radiologist in performing a more objective, fast and accurate analysis without the intrinsic variability of human activities [3]. Some proposals have been described in the literature.

1.1. State of art

The first attempts made to achieve an automatic method to support the radiologists' work have been reported in the early 1980s. Pathak, Pal, and King posed the problem as a classification problem and they developed some radiograph analysis procedures [4–7] and proposed a syntactic fuzzy classifier for bone age assessment [8,9], based on fuzzy grammars. The input to this classifier is a set of primitives (as points, line segments, and curves) which have been previously extracted from the hand radiograph and the output is the maturity state of each bone. Although the classification is done using the stages of the TW3 method (actually, the authors use the former TW2 version), this proposal is not a direct implementation of the method itself; the use of a set of primitives based on boundaries—far away from the language descriptions of the method—makes it a

* Corresponding author. Fax: +34-983-423-667.

E-mail addresses: sanaja@tel.uva.es (S. Aja-Fernández), rluigar@neptuno.lpi.tel.uva.es (R. de Luis-García), migmar@tel.uva.es (M.A. Martín-Fernández), caralb@tel.uva.es (C. Alberola-López).

1532-0464/$ - see front matter
doi:10.1016/j.jbi.2004.01.002

"new skeletal assessment method". A high amount of examples seems to be needed to train the method properly. In addition, the use of primitives such as "point" or "segment" makes the method hard to understand by a radiologist, so it does not seem straightforward how to use the experience of the experts to improve the accuracy of the classifier. Finally, the fixed structure of the 3-stage hierarchical system proposed by the authors seems a bit rigid. In particular, it is not clear how the authors may deal with the existing overlap between features and output stages. Anyway, the philosophy used to design the classifier is, from our point of view, extremely interesting, but surprisingly, it has not been taken into account—to our knowledge–by any other author since then.

Pietka et al. [10] described a method based on an independent analysis of the phalangeal [11] and the carpal [12] regions of a radiograph. A fuzzy classifier is then developed [10]; since both regions are analyzed independently, two bone age assessments are obtained, one for the carpal region and one for the phalangeal one. The final decision must be made by a radiologist. This fuzzy classifier is built using a great deal of heuristics. In addition, some fuzzy operations are methodologically incorrect. In particular, a *t*-norm—a "min" operator [13]—is replaced by a summation which, as it is well known [14–16], plays the role of a *t*-conorm in standard additive fuzzy systems. We understand this replacement obeys some heuristic rule, but it should not be acceptable from an scholarly point of view. In [17] a real implementation of this system is presented, with a correct classification ratio of 75% for the phalangeal region and 63% for the carpal one. However, when compared with the TW3 method, this approach discards a great amount of useful information, negatively affecting the final result.

Recently, Pietka et al. [18–21] proposed a more refined approach to perform skeletal age assessment, focusing on the phalanges. In a first stage, an epiphyseal/metaphyseal region of interest is found for each bone [18,19] making use of a priori knowledge about the general structure of the hand. Next, for each region of interest, two different cases are distinguished [20]: if the epiphyses are separated from the metaphyses,the bone is decided to be at an early stage of development; if the gap between epiphysis and metaphysis has started to disappear, the bone is assigned to a later stage of development. In the first case, epiphysis and metaphysis are segmented separately, and several distance-related features are extracted. In the latter case, a wavelet decomposition analysis is performed to evaluate the state of epiphyseal fusion [21], also yielding several features. To complete the bone age assessment, a fuzzy classifier has been employed using the referred features [20]. This new classifier now uses a methodologically correct max-min operative, as opposed to their previous work. The classifier seems to give skeletal ages associated with a membership grade out of every bone, and then decisions are aggregated to end up with a final unique solution in a discrete age space. Although the authors use a classifier for each bone, an idea to which we adhere, the procedure to obtain the classifiers rules is not clearly described, so we understand it is inferred from examples. The overall system has been finally integrated into a clinical PACS.

Efford proposed a direct automation of the former TW2 method [22] by performing a shape analysis of certain bones of interest. Most of the work is dedicated to the design of a model-based segmentation of the bones from radiographs, but a skeletal maturity assessment method built upon this segmentation was never actually implemented. So, this work cannot be considered as a complete bone age assessment method.

Other contributions in this field have used neural networks with a backpropagation algorithm in the training phase [23], and a neural architecture called Generalized Softmax Perceptron which estimates posterior probabilities at the output of the network [24] to give a measure of confidence in the final decision.

1.2. A new approach

Technical solutions described so far have in common that they first extract a set of features from the bones and then a classifier is created according to these features, following a classical learning scheme (Fig. 1). So, the implementation of these classifiers is constrained by the selected features.

To achieve a correct learning a great amount of data are needed; however, in a real clinical setting, this amount of data may not always be available for each of the bone ages. In addition, a high accuracy is needed in the age estimation process, a fact which may lead to too high a number of outputs in the classifier, provided that outputs are a set of discrete bone ages [24]. An alternative procedure is to have maturity stages for each bone, as opposed to overall ages, as outputs of *micro-classifiers* (Fig. 2). But to proceed this way is to implicitly follow the same philosophy as the TW3 method, but discarding the rules given by the method.

The TW3 method is a commonly accepted procedure in which the guidelines to analyze each bone are described using words (natural language descriptions), sometimes in a vague way. In addition, one particular bone may show features belonging to different stages or a particular bone shape could be classifiable into two possible predefined labels of the same feature. A flexible classification method is needed to manage these sources of ambiguity.

Fuzzy logic is known to be a very flexible tool in classification problems where imprecise knowledge or not-well-defined features have to be used. Some of the

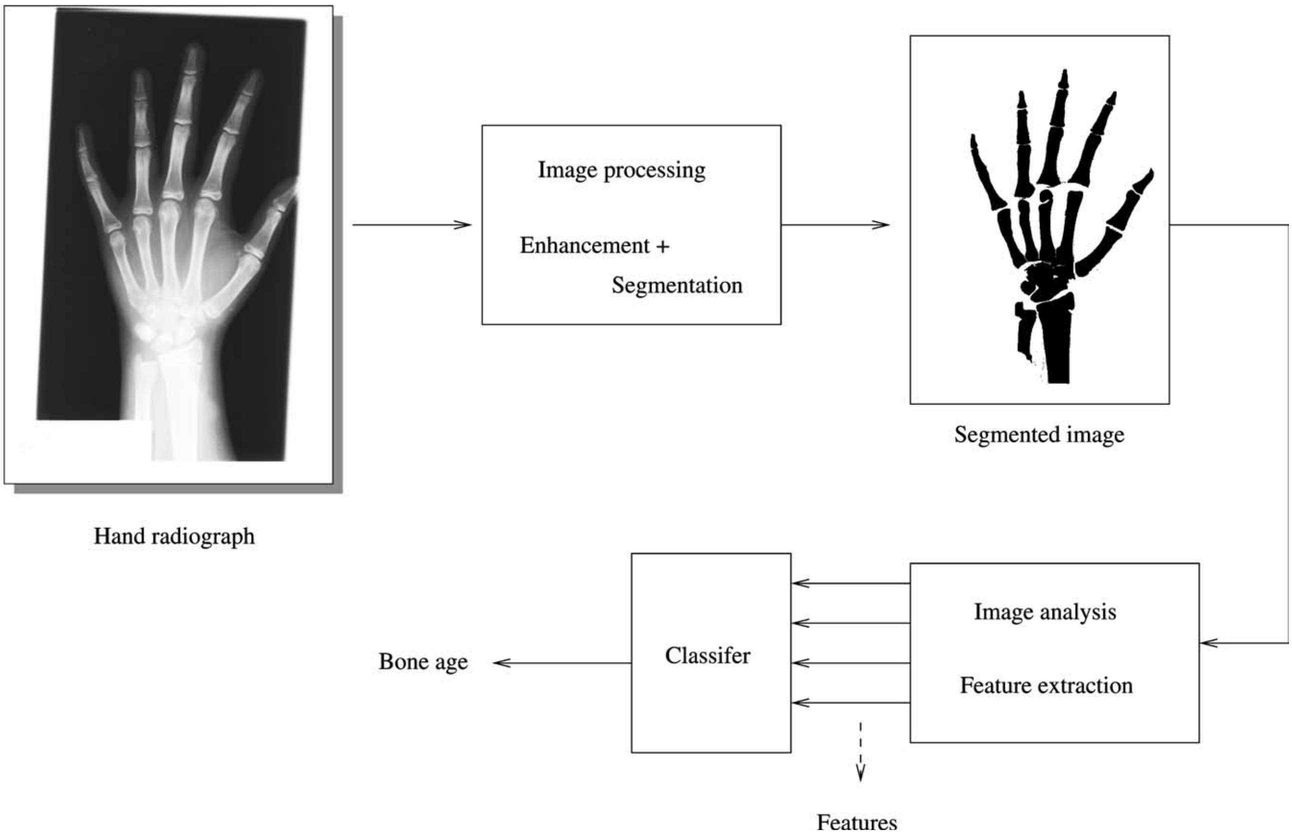

Fig. 1. Bone age assessment automatic process.

authors we have mentioned before [9,10] have been aware of this. But, in addition, fuzzy logic is also a natural selection when information has to be retrieved from linguistic statements; such a methodology, as we will show in the paper, is able to bring the TW3 method "as is" into a computer, using a *Computing with words* (CWW) [25] paradigm. In addition, using the TW3 statements directly to build the computational classifier allows us to benefit from the expertise of the authors of the method. Hence, the lack of lots of training data is no longer a problem since rules are known beforehand, so only a few labeled radiographs are needed to fine tune the rules and to test the classifier. To be specific, in our case 57 radiographs from boys have been used to fine tune the rulebase and 85 from girls to test it. This division between boys for training and girls for testing is arbitrary since there is no morphological difference in bones in terms of gender in children, and the distinction is only of importance for the prediction of adult height (which is the final part of the method, located after the classifier, as indicated in Fig. 2).

Two final remarks to conclude this section are mandatory: first, this paper does not describe an end-to-end system for bone age assessment, since the feature extractor needed to feed the classifier we will propose is still under construction. What we describe is, as stated in the previous paragraph, a CWW methodology to bring human knowledge, expressed in natural language, into a computational scheme. We have used as testbed the TW3 method, under the premise that features used by the method are available. Second, one might argue that if both enough radiographs and the feature extractor were available, an expert system properly trained with these data should be able to perform better than a classifier that directly implements the TW3 method. This statement would probably be true. However, we depart from a different starting point: enough radiographs are hardly available for the whole range of possible ages,[1] so if a classification methodology has been designed (aimed at experts in the field) using many radiographs, it is sensible to bring it into a computer. This method is the TW3, which even though it might not be entirely satisfactory for the whole medical community (for instance, for supporters of the GP method) it is however very frequently used in pediatric radiology; consequently, we have taken it as ground truth for our research, and we have filled in gaps of the method with the help of experts in the field.

[1] Babies are seldom radiographically inspected, unless a bone injury is suspected; however, parents of kids are frequently interested in knowing the prediction of their children's adult height, so they may be willing to get a radiograph of their children to that end. This is the common behavior we have found in our region.

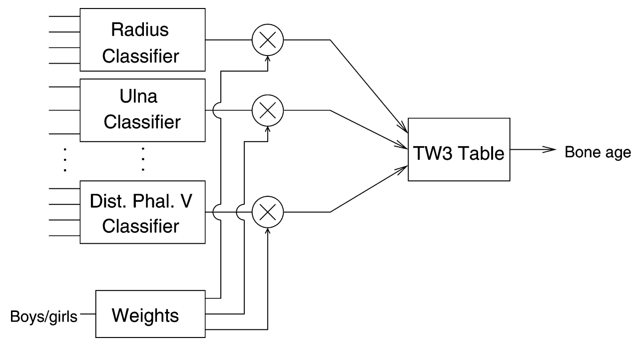

Fig. 2. Global classifier done by one micro-classifier for each bone.

2. Definition of computational features for TW3

The TW3 methodology for bone age assessment consists of a set of rules, expressed in natural language, to describe the prototypical characteristics of the bones of a hand radiograph as they evolve in time. In order for the explanation to be clearer, a graphical sketch (similar to that shown in Fig. 3) with several maturity patterns accompanies the explanation in the original method [2].

Due to the vagueness of the explanations, classical approaches have resorted—as stated in Section 1—to learning paradigms in which rules are inferred from data. However, our understanding is that the natural solution is to use the method itself, i.e., to build one classifier for each bone, with 9 outputs (the possible classification stages for each bone—A,B,…,I—), except for the ulna, which only has 8 stages. Each classifier will be fed with a computational approximation to the input features requested by the method, which, as previously stated, are expressed in natural language.

Two possible analysis schemes are defined for the TW3 method. The first one, named RUS, uses 13 bones (the phalanges, radius, and ulna). The other one, uses 20 bones (the 13 bones previously defined and the 7 bones of the carpal region). We have chosen to build a method based on the 13 RUS bones, because the carpal region is not valid after a certain age.

In what follows we will propose a CWW approach to the problem described. We will only describe the process for the radius, since the process for the rest of the bones is quite similar.

2.1. Feature definition for the radius

Maturity stage for each bone in TW3 is calculated from linguistic statements. For the case of the radius the following two example statements (literally extracted from [2]) give an idea about the sentences given by the method (Fig. 3).

Stage D. The maximum diameter is half or more the width of the metaphysis. The epiphysis has broadened chiefly at its lateral side, so that this portion is thicker and more rounded, the medial portion more tapering. The center third of the proximal surface is flat and slightly thickened and the gap between it and the radial metaphysis has narrowed to about a millimeter.

Stage G. The dorsal surface now has distinct lunate and scaphoid articular edges joined at a small hump. The medial border of the epiphysis has developed palmar and dorsal surfaces for articulation with the ulnar epiphysis; either the palmar or the dorsal surface may be the one that projects medially, depending on the position of the wrist. The proximal border of the epiphysis is now slightly concave.

Overall, six features can be defined that capture all the text information, so they are sufficient to define each possible state.

Presence. Epiphysis is absent or present. If it is absent, the output stage is A. If it is present but is small and hardly visible, the output stage is B. If it is present and well-visible, the output stages are from C to I.

Separation. Relative position of epiphysis and metaphysis: separated (stages B, C, D, E, F, and G), capping (stage H), or fusion has begun (stage I).

Shape of epiphysis I. Oval (stage C) or sharp (stages D–I).

Diameters. Ratio between diameters of metaphysis and epiphysis.

Shape II. A "sharp" epiphysis can have a regular outline (stages D and E), can be adapted to the metaphysis shape (stage F), or can have the articulations form (stages G, H, and I).

Surfaces. Representation of inner surfaces. They can be absent (stages B, C, and D) or present as a white line (stage E), two white lines (stage F) or a c-shaped surface (stages G, H, and I).

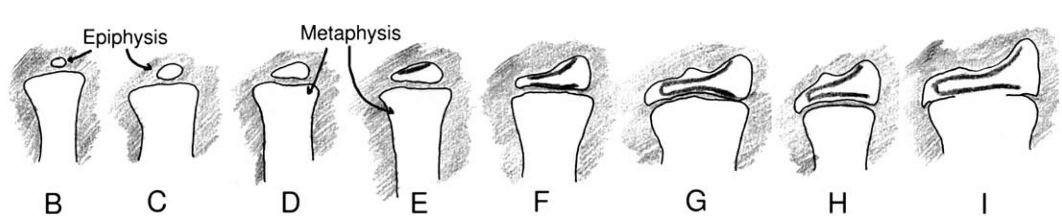

Fig. 3. TW3 stages for the radius.

The feature values for each stage are shown in the next table:

Stage	Presence	Separ.	Shape I	Diam.	Shape II	Surf.
A	No					
B	Small	(Yes)	(oval)	> 2		(no)
C	Yes	Yes	oval	> 2		no
D	Yes	Yes	sharp	$\leqslant 2$	regular	no
E	Yes	Yes	sharp	$\leqslant 2$	regular	1 line
F	Yes	Yes	sharp	$\leqslant 2$	adapted	2 lines
G	Yes	Yes	sharp	$\leqslant 2$	articulation	c-shape
H	Yes	capping	sharp	$\leqslant 2$	articulation	c-shape
I	Yes	fusion	sharp	$\leqslant 2$	articulation	c-shape

However, these features are not independent. As a matter of fact, some of the features are self-excluding: *Shape II* only takes on values when *Shape I* is sharp and *Separation* is only defined for a present epiphysis. Consequently, these features can be merged, a fact which contributes to simplify the classifier. After the fusion process, the resulting feature set is:

Epiphysis. Absent or small and, otherwise, what matters is its relative position with respect to the metaphysis (separated, capping, fusion).

Shape. Outline shape of the epiphysis (oval, regular-sharp, adapted-sharp, articulated-sharp).

Diameters. Ratio between metaphysis and epiphysis diameters.

Surfaces. Inner surfaces (absent, 1-line, 2-lines, c-shape).

The new features values for each stage are now:

Stage	Epiphysis	Diameters	Shape	Surfaces
A	Absent	∞	(oval)	(absent)
B	Small	$\gg 2$	(oval)	(absent)
C	Separated	> 2	oval	absent
D	Separated	$\leqslant 2$	regular sharp	absent
E	Separated	< 2	regular sharp	1 line
F	Separated	< 2	adapted sharp	2 lines
G	Separated	≈ 1	articulation-sharp	c-shape
H	Capping	≈ 1	articulation-sharp	c-shape
I	Fusion	≈ 1	articulation-sharp	c-shape

The same feature definition process has been done for the other bones (see appendix A).

2.2. Feature Fuzzy modeling

In order to build an automatic classifier, the previous features must be modeled more formally to make them suitable for computer processing.

Some studies show that a discrepancy may exist between the results of different radiologists classifications, or even between the results of the classifications done by the same radiologist in different moments [3]. The reason for it is that some stages are easily mistaken with the nearby stages, and sometimes there are bones that show features belonging to different stages. So, features are not strictly well-defined. Instead, they have a certain degree of overlap between them. In practice, there is not a clear bound between a *sharp* epiphysis and an *oval* one; as it evolves in time the transition between shapes is not a step, but a gradual process.

The intrinsic nature of the process and the features themselves have led us to model them by means of *linguistic variables* [13] (the values of which are words instead of numbers). They are characterized by several linguistic terms that can be modeled by fuzzy sets. In a traditional way, these sets are defined over a *base variable* [13] but they can also be described using the degree of overlap between the different terms [26]. In our approach there are features for which a base variable is clearly associated. This is the case for the diameters ratio. But, on the other hand, for features such as *Epiphysis* or *Surfaces* a uniquely defined physical meaning does not exist so as to associate it a base variable;[2] consequently they will be modeled using *fuzzy granules* [26].

Four linguistic variables are created, one for each of the features described in Section 2.1.

Epiphysis. Linguistic variable defined by 5 fuzzy granules named *Absent*, *Small*, *Separated*, *Capping*, and *Fusion*. The degree of overlap between them has been estimated studying some discrepancies between observation of different radiologists over the same radiographs and the discrepancies between the observations of one radiologist over the same set of radiographs in different instants of time. The variable is shown in Fig. 4A.

Shape. Linguistic variable defined by 4 fuzzy granules named *Oval*, *Sharp I*, *Sharp II*, and *Sharp III* (Fig. 4B).

Surfaces. Linguistic variable defined by 4 fuzzy granules named *Absent*, *1-line*, *2-lines*. and *c-shape* (Fig. 4C).

Diameters. The ratio between the diameters of the metaphysis (D) and the epiphysis (d), i.e., $R = (D_{\text{metaphysis}}/d_{\text{epiphysis}})$. Though the input is crisp the classification process will be fuzzy, so 4 fuzzy sets are defined: "*Greater than 2*," "*about 2*," "*about 1*," and "*less than 1*," as shown in Fig. 4D.

3. TW3 classifier

We have observed that the previously defined features may be mutually exclusive: if the *epiphysis* is *absent*, the other features cannot take on any value; if the *epiphysis* is capping the metaphysis it will be very rare for the

[2] A fuzzy set must be defined on a numerical base variable, and each value of the variable receives a membership degree; for instance, the words *young* and *old* have a clear association to a natural number, the meaning of which is *age*. However, cases can be found in which such an association is very artificial, and it is done only for operative, not conceptual, reasons. This would be the case, for instance, for vehicle distinction, say, for a car and a van [15]. Fuzzy granules [26], however, avoid this cumbersome need.

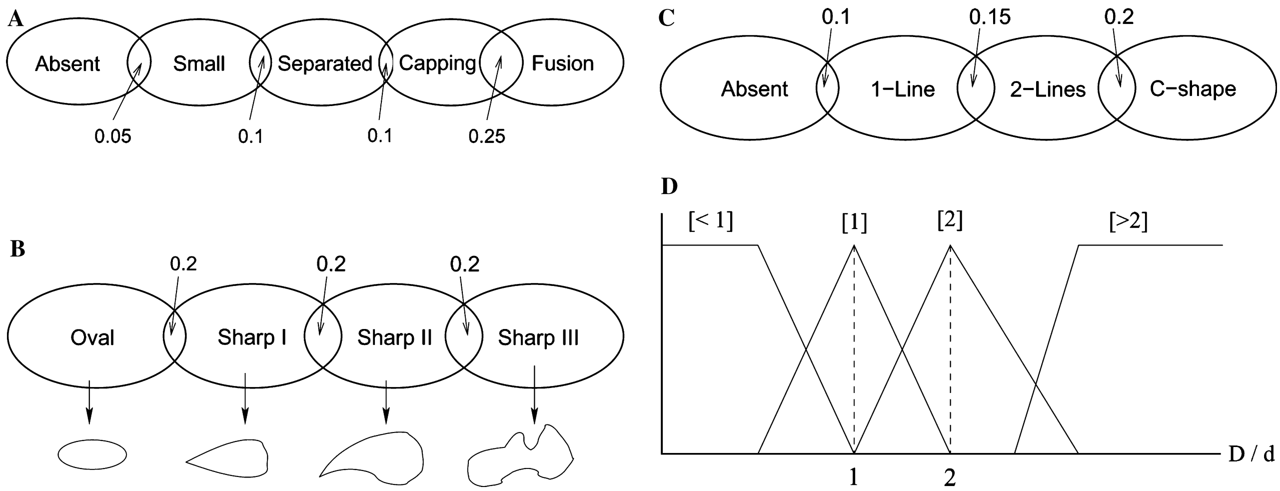

Fig. 4. Linguistic Variables: (A) Epiphysis, (B) shape, (C) surface, (D) diameters.

shape to be oval. This fact has led us to choose a tree-classifier, specifically, a fuzzy decision tree, in order to manage the linguistic variables defined.

As we will be working with discrete domains for the features, an ID3 algorithm is chosen [27]. Following the TW3 method, a different tree will be built for each of the 13 bones defined for the RUS method. Each bone will be finally classified into one stage (A,B,...,I). This scheme will make the classifier suitable for boys and girls.

3.1. Fuzzy decision trees. Fuzzy ID3

ID3 is a classification method based on trees which assumes low-cardinality discrete domains. It requires an a priori partition of the different domains. Umano et al. [27] introduced the fuzzy version of this algorithm. Fuzzy ID3 is an extension of ID3 to be applied over an input fuzzy data set and generates a fuzzy decision tree using the fuzzy sets defined a priori by the user for every attribute. The ID3 algorithm generates a minimum-size decision tree out of a number of classification parameters by choosing the optimum ordering—in terms of maximum information gain—of these parameters along the tree.

A fuzzy decision tree consists of *nodes* for testing attributes, *edges* for branching by values of symbols and *leaves* for deciding class names. To generate the tree, a training data set is needed. The output of the tree will not just be one class, but the membership grade associated to every class. This algorithm will be slightly modified to accept fuzzy granules as input data.

3.2. Training sets

The number of labeled radiographs needed to train a decision tree or any other classifier is very high, and these radiographs must cover all the possible states of each bone. This is equivalent to implementing a new bone age assessment method from scratch. As it has been mentioned in Section 1, an alternative way is to make use of the experience accumulated in the TW3 method (which was created upon 3000 boy and 3000 girl hand radiographs); ignoring the rules derived from the radiologists that had all that information is not advisable, as common sense dictates.

The training set will consist of all the information that will be used to give actual values to the fuzzy sets involved in the fuzzy decision tree. This information is the following:

1. First, 9 *prototypes* are directly extracted from the method; they constitute the "canonical" cases since they reflect the expected feature layout for the 9 possible stages for each bone. These data produced rule consequents (i.e., stages) with membership values equal to 1.
2. Afterwards, several "non-canonical" cases are studied. A template was constructed with all the combinations of all the features involved (all the words in the linguistic variables shown in Fig. 4). Then, the templates were given to expert radiologists, and they were kindly asked to assign a stage and a membership grade (between 0 and 1) to each of these combinations, according to their degree of confidence on the stage given to that particular combination of features. Table 1 shows the radiologists assessments for the radius (membership values are the average of the values given by different experts).
3. The labeled data available are used a posteriori to fine tune the classification tree. Specifically, 57 radiographs were used to modify the membership values of some possible output classes in the nodes of the tree [27] to maximize classification performance.

Notice that no rule is inferred from data nor cases are generated from rules. Rules are known beforehand and

Table 1
Training data set for the radius

Stage	Membership	Epiphysis	Shape	Surface	Diameters
A	1	1	1	1	10
B	1	2	1	1	5
B	1	2	1	1	3
C	1	3	1	1	2.5
C	0.8	3	1	1	2
D	0.2	3	1	2	2.5
C	0.2	3	1	2	1.9
C	0.5	3	2	1	2.5
C	0.5	3	2	1	2.2
D	0.9	3	2	1	2
D	1	3	2	1	1.9
E	0.9	3	2	2	2
⋮	⋮	⋮	⋮	⋮	⋮
I	1	5	4	4	0.8

we assume they are widely accepted; cases not explicitly addressed by the method have been found by creating an exhaustive table with all the feature combinations. These cases have been rated by experts. Consequently, data here are only used to fine tune these rules.

3.3. Fuzzy trees for TW3: construction and test

The fuzzy decision tree for each bone can be easily built using as training set the templates provided by the experts. Fig. 5 shows the decision tree generated for the radius. For simplicity, only the stage with maximum membership grade has been shown in each leaf (but actually the 9 possible output stages are there, each one with a different membership value).

Following the scheme of Fig. 1, the inputs to each tree will be the features of its corresponding bone, and the output will be its skeletal stage. The stages will be numerically weighed following the TW3 method (a different weight is applied for boys and girls, see Fig. 2). The weighed summation is mapped onto the bone age, as described by the method. Our classification procedure therefore has the same age resolution as the original method.

After the decision trees have been built and fine-tuned then 85 diagnosed radiographs from girls have been used to test them. The ulna results, for example, showed that 83 out of 85 radiographs were correctly classified (97.6%). For the proximal phalange I, 81 out of 85 were correctly classified (95.3%). In every case, the classification error was to assign a consecutive stage to the right one.

4. Conclusions and future work

In this paper we have proposed a CWW approach to get a computerized version of the commonly accepted TW3 method for bone age assessment. One of the advantages of the method is the small amount of data needed to fine tune the rulebase used by the classifier. Results have shown that the method's performance is fairly high; in any case, this performance is expected to increase when a systematic rule fine-tune method is

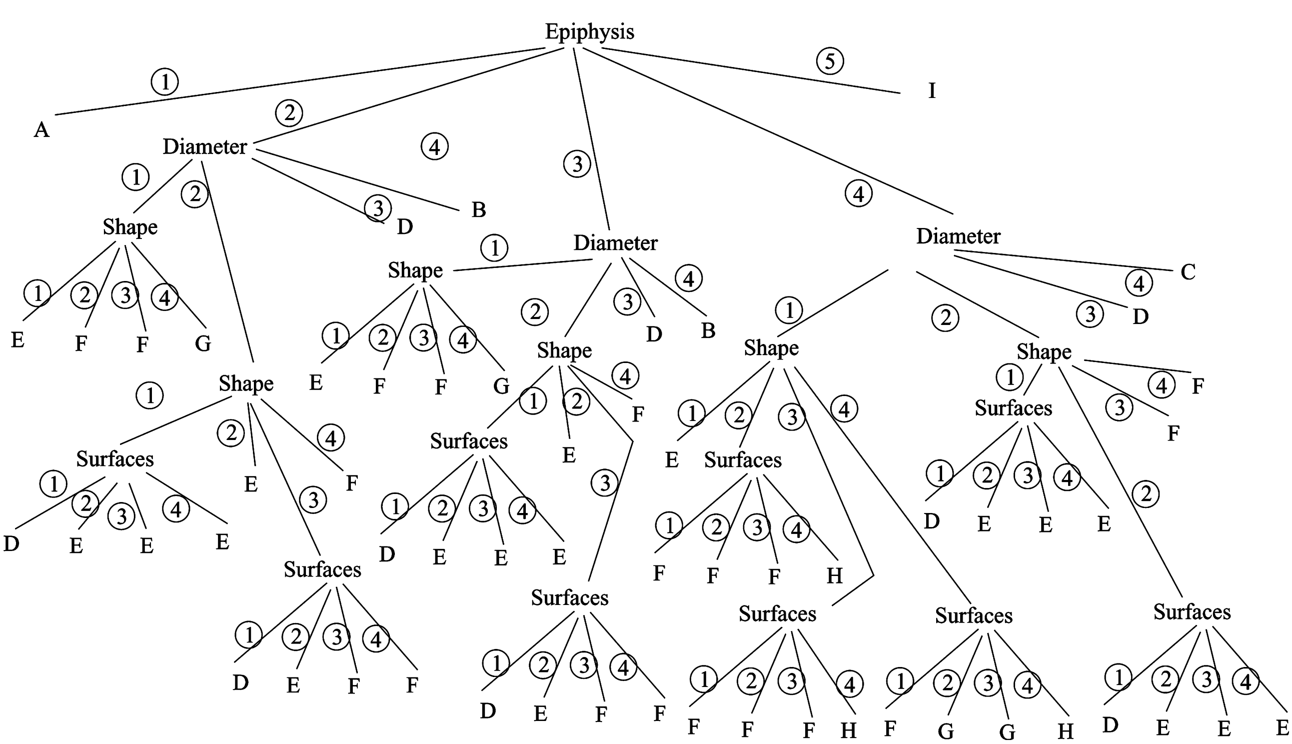

Fig. 5. Decision tree for radius.

implemented—the one used in the paper is based on manual adjustment to maximize performance. In addition, the fuzzy sets used as input features could also be fined-tuned; fuzzy granules are particularly appropriate to that end, since only one parameter, namely, the overlap degree between fuzzy sets, has to be modified.

The paper has focused on the classifier itself, taking for granted that actual features can be extracted from the radiographs; therefore the methodology described in the paper does not constitute an end–to–end classification system. As for the feature extractor, it should not be an issue as soon as a complete and precise bone segmentation is achieved; we have preliminary results of this [24,28,29]. Alternatively, a classical learning scheme may be defined to get linguistic labels out of the image data. If this were carried out it is undoubtable that a smaller number of examples would be needed to learn features than those needed to infer rules from examples.

Acknowledgments

The authors acknowledge Alberto Suárez (UA de Madrid) for his help in the bibliography review about decision trees, the Spanish CICyT for research grant TIC2001-3808-C02 and the J. Castilla y León for research grant VA91/01. Thanks also go to Dr. Viñuela Rueda and Dr. Andrés de Llano, Hospital Río Carrión, Palencia, and Dr. S. Alberola-López, Centro de Salud Jardinillos, Palencia, for their help with the templates are radiograph labelling.

Appendix A. Features of each bone for the RUS skeletal maturity method

We present the different selected features used to classify every bone.

(1) **Radius**
- *Epiphysis.* Absent (1), present and small (2), present and separated (3), epiphysis caps metaphysis (4), fusion has begun (5).
- *Diameters.* Ratio between metaphysis diameter (D) and epiphysis diameter (d).
- *Shape.* Oval shape (1), regular sharp (2), regular sharp with proximal border adapted to metaphysis (3), sharp adapted to articulation with concave proximal border (4).
- *Surfaces.* Absent (1), 1 white line (2), 2 white lines (3), c-shape surface (4).

(2) **Ulna**
- *Epiphysis.* Absent (1), present and small (2), present and separated (3), epiphysis caps metaphysis (4), fusion has begun (5).
- *Diameters.* Ratio between metaphysis diameter (D) and epiphysis diameter (d).
- *Shape.* Oval shape (1), elongated (2), styloid process visible (3), head of the ulna defined (4).

(3) **First metacarpal**
- *Epiphysis.* Absent (1), present and small (2), present and separated (3), epiphysis caps metaphysis (4), fusion has begun (5), fusion is completed (6).
- *Diameters.* Ratio between metaphysis diameter (D) and epiphysis diameter (d).
- *Shape.* Oval shape (1), concavity in the proximal border (2), saddle shape (3).
- *Surfaces.* Absent (1), proximal surface (2).

(4) **Metacarpals III and V**
- *Epiphysis.* Absent (1), present and small (2), present and separated (3), epiphysis caps metaphysis (4), fusion has begun (5), fusion is completed (6).
- *Diameters.* Ratio between metaphysis diameter (D) and epiphysis diameter (d).
- *Shape.* Oval shape (1), fingernail shape (2).
- *Surfaces.* Absent (1), palmar and dorsal surfaces (2).

(5) **Proximal phalanx of the thumb**
- *Epiphysis:* Absent (1), present and small (2), present and separated (3), epiphysis caps metaphysis (4), fusion has begun (5), fusion is completed (6).
- *Diameters:* Ratio between metaphysis diameter (D) and epiphysis diameter (d).
- *Shape:* Oval shape (1), concave proximal border (2), epiphysis follows metaphysis shape (3).

(6) **Proximal phalanges III and V**
- *Epiphysis.* Absent (1), present and small (2), present and separated (3), epiphysis caps metaphysis (4), fusion has begun (5), fusion is completed (6).
- *Diameters.* Ratio between metaphysis diameter (D) and epiphysis diameter (d).
- *Shape.* Oval shape (1), concave proximal border (2), epiphysis follows metaphysis shape (3).

(7) **Middle phalanges III and V**
- *Epiphysis.* Absent (1), present and small (2), present and separated (3), epiphysis caps metaphysis (4), fusion has begun (5), fusion is completed (6).
- *Diameters.* Ratio between metaphysis diameter (D) and epiphysis diameter (d).
- *Shape.* Oval shape (1), triangular shape (2).
- *Surfaces.* Absent (1), proximal surface (2).

(8) **Distal phalanx of the thumb**
- *Epiphysis.* Absent (1), present and small (2), present and separated (3), epiphysis caps metaphysis (4), fusion has begun (5), fusion is completed (6).
- *Diameters.* Ratio between metaphysis diameter (D) and epiphysis diameter (d).
- *Shape.* Oval shape (1), triangular shape (2), saddle shape (3).

(9) **Distal phalanges III and V**
- *Epiphysis.* Absent (1), present and small (2), present and separated (3), epiphysis caps metaphysis

(4), fusion has begun (5), fusion is completed (6).

- *Diameters.* Ratio between metaphysis diameter (D) and epiphysis diameter (d).
- *Shape.* Oval shape (1), triangular shape (2).
- *Surfaces.* Absent (1), palmar and dorsal surfaces (2).

References

[1] Greulich WW, Pyle SI. Radiographic atlas of skeletal development of hand wrist. second ed. Stanford, CA: Standford University Press; 1971.

[2] Tanner JM, Goldstein H, Cameron N, Saunders WB, Healy MJR. Assessment of skeletal maturity and prediction of adult height. third ed. WB Saunders; 2001.

[3] Viñuela-Rueda B, Alberola-López S. Inter and intra observer variability in bone age assessment using TW2-RUS (in spanish), In: 26 Congreso Nacional Soc. Esp. Rad. Médica, Maspalomas, Gran Canaria, Spain; 2002.

[4] Kwabwe S, Pal SK, King RA. Recognition of bones from X-ray of the hand and wrist. Int J Syst Sci 1985;16(4):403–13.

[5] Pal SK, King RA. On edge detection of X-ray images using fuzzy sets. IEEE Trans Pattern Anal Mach Intell 1983;5(1):69–77.

[6] Pal SK, King RA, Hashim AA. Automatic grey level thresholding using index of fuzziness and entropy. Patt Recog Lett 1983;1:141–6.

[7] Pal SK, King RA, Hashim AA. Image description and primitive extraction using fuzzy sets. IEEE Trans System, Man Cybernet SMC 1983;13(1):94–100.

[8] Pathak A, Pal S, King R. Syntactic recognition of skeletal maturity. Patt Recog Lett 1983;1:141–6.

[9] Pathak A, Pal S. Fuzzy grammars in syntactic recognition of skeletal maturity from X-rays. IEEE Trans Syst Man Cybern SMC 1986;16(5):657–66.

[10] Pietka E. Computer-assisted bone age assessment based on features automatically extracted from a hand radiograph. Comput Med Imag Graph 1995;19(3):251–9.

[11] Pietka E, McNitt-Gray MF, Kuo ML, Huang HK. Computer-assisted phalangeal analysis in skeletal age assessment. IEEE Trans Med Imag 1991;10(4):616–20.

[12] Pietka E, Kaabi L, Kuo ML, Huang HK. Feature extraction in carpal-bone analysis. IEEE Trans Med Imag 1993;12(1):44–9.

[13] Klir BYG. Fuzzy sets and fuzzy logic. New Jersey: Prentice-Hall International; 1995.

[14] Kosko B. Fuzzy engineering. New Jersey: Prentice-Hall International; 1997.

[15] Aja-Fernández S, Alberola-López C, Cybenko G. A fuzzy MHT algorithm applied to text-based information tracking. IEEE Trans Fuzzy Syst 2002;10(3):360–74.

[16] Aja-Fernández S, Alberola-López C. Fast inference in SAM fuzzy systems using transition matrices. IEEE Trans Fuzzy Syst [To appear].

[17] Cao FEI, Huang HK, Pietka E, Gilsanz V, Omisky S. Diagnostic workstation for digital hand atlas in bone age assessment. Proc SPIE Med Imag 1998:608–14.

[18] Pietka E, Gertych A, Pospiech-Kurkowska S, Cao F, Huang HK. Computer-assisted bone age assesment: image preprocessing and epiphyseal/metaphyseal roi extraction. IEEE Trans Med Imag 2001;20(8):715–29.

[19] Pietka E, Pospiech-Kurkowska S, Gertych A, Cao F, Huang HK, Gilsanz V. Computer automated approach to the extraction of epiphyseal regions in hand radiographs. J Digital Imag 2001;14(4):165–72.

[20] Pietka E, Pospiech-Kurkowska S, Gertych A, Cao F. Integration of computer assisted bone age assessment with clinical pacs. Comput Med Imag Graph 2003;27:217–28.

[21] Pospiech-Kurkowska S, Pietka E, Cao F, Huang HK, Fuzzy system for the estimation of bone age from wavelet features. In: Proc. of the 16th Biennial International Eurasip Conference Biosignal 2002, Brno, Czech Republic; 2002. p. 441–3.

[22] Efford ND, Knowledge-based segmentation and feature analysis of hand-wrist radiographs, Technical Report, Report 94.31, School of Computer Studies, University of Leeds; 1994.

[23] Gross GW, Boon JM, Bishop DM. Pediatric skeletal age: determination with neural networks. IEEE Trans Med Imag 1995;14(3):689–95.

[24] de Luis-García R, Arribas JI, Aja-Fernández S, Alberola-López C. A neural architecture for bone age assessment. In: Proc. of the IASTED International Conference on Signal Processing, Pattern Recognition and Applications, Creta, Greece; 2002. p. 161–6.

[25] Wang PP, editor. Computing with words. New York: Wiley; 2001.

[26] Aja-Fernández S, Alberola-López C. Inference with fuzzy granules for computing with words: a practical viewpoint. In: Proc. of FUZZ-IEEE 2003, St. Louis, MO; 2003. p. 566–71.

[27] Umano M, Okamoto H, Hatono I, Tamura H. Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems. In: Proc. of the third IEEE conference on fuzzy systems, FUZZ-IEEE'94, Orlando, FL; 1994. p. 2113–8.

[28] de Luis-García R, Martín-Fernández M, Arribas JI, Alberola-López C. A fully automatic algorithm for contour detection of bones in hand radiographs using active contours. In: Proc. of the ICIP 2003, Barcelona, Spain; 2003. p. III:421–4.

[29] Martín-Fernández MA, Martín-Fernández M, Alberola-López C. Automatic bone age assessment: a registration approach. In: Proceedings on Medical Imaging, SPIE, San Diego, CA; 2003. p. 1765–76.