Weighted Fuzzy ARTMAP for Osteoporosis Detection

Darlis HERUMURTI [1], Agus Z ARIFIN [1], Rully SULAEMAN [1],
Akira ASANO [2], Akira TAGUCHI[3], Takashi NAKAMOTO[4] , Keiichi UCHIMURA[5]

[1]: Vision and Image Processing Laboratory, Department of Informatics Faculty of Information Technology, ITS, Surabaya, 60111, Indonesia, darlis@cs.its.ac.id, agusza@cs.its.ac.id, rully@is.its.ac.id
[2]: Graduate School of Engineering, Hiroshima University, Japan, asano@mis.hiroshima-u.ac.jp
[3]: Matsumoto Dental University, Nagano, Japan, akiratag@nifty.com
[4]: Hiroshima University Hospital, Hiroshima, Japan, tnk@hiroshima-u.ac.jp
5: Computer Science Department, Kumamoto University, Japan, uchimura@cs.kumamoto-u.ac.jp

Abstract Osteoporotic is a health burden worldwide, resulting in reduction of physical activity, increased risk of mortality, and incremental medical cost. Mandibular trabecular patterns analyzed on dental panoramic radiographs have been widely studied for identifying postmenopausal women with low skeletal bone mineral density (BMD). In this paper we proposed a new method for detecting osteoporosis using Weighted Fuzzy ARTMAP from the features measured in dental panoramic radiographs. The method developed an activation match function by integrating Simplified fuzzy ARTMAP and symmetric Fuzzy ART. Fourier method and segmentation processing were applied for obtaining features of a radiograph in frequency and spatial domain. We also introduced an additional weighted parameter based on pheromone to distinguish clusters based on the amount of its member. The experimental results for osteoporosis detection show that the new method achieved accuracy of 87.88%, sensitivity of 93.33%, and specificity of 83.33%.

1 Introduction

Osteoporosis is a condition in which the bones become fragile and brittle, leading to a higher risk of fractures (breaks or cracks) than in normal bone. Osteoporotic fractures known as a health burden worldwide, resulting in reduction of physical activity, increased risk of mortality, and incremental medical cost. Since incidence rates of osteoporotic hip fracture increase exponentially with aging, this demographic change alone should cause the number of hip fractures worldwide to rise.

The only sure way to determine bone density and fracture risk for osteoporosis is to have a bone mass measurement (also called bone mineral density or BMD test). But this test is very expensive and not so many hospitals can do it. The recently studies proposed computer-aided diagnosis systems based on dental panoramic radiographs which may be cheap and easy to get [1], because people may have greater opportunity to visit dentists for treatment of dental caries and periodontal disease than to visit medical professionals for diagnosis of osteoporosis.

Feature extraction from the dental radiograph is done in two domain space, frequency domain and spatial domain. From the extracted features, we can proceed to the next step that is learning and osteoporosis detection/classification. Kasuba in 1993 propose the simplified Fuzzy ARTMAP (SFAM), which is equivalent to the sequential implementation of ARTMAP [2]. Fast Fourier transformation is capable to get the features from frequency domain [3] [4] and in addition, the spatial domain the features of radiograph can be obtained from segmentation processing using JSEG method [16][17].

In this paper we proposed a new method for detecting osteoporosis using the Weighted Fuzzy ARTMAP, which is based on Adaptive Resonance Theory (ART), the new type of neural network. The ART is designed by Grossberg to solve plasticity-stability dilemma.

There are two main differences between Weighted Fuzzy ARTMAP and ART. The first one is on the activation-match function. The function in the Weighted Fuzzy ARTMAP combines between Fuzzy ARTMAP and Symmetrical Fuzzy ARTMAP by giving symmetrical weighted parameter. The second is the cluster node in the ART method does not watch the number of its member. Somehow, it is important to see the amount of members in the cluster node, especially if there are two or more nearest cluster with same distance which needs to be chosen. So, we need weighted the cluster node based on the number of its member.

2 Global System Description

Generally, the system is divided into three processes, Features extraction process, training process and testing process. Features extraction process transforms the radiograph into numerical information we called the features. In the training process, the features of individual known to have osteoporosis and not are analyzed by doing the learning to get the model for prediction of a new data. During the learning, there are some parameters that require to be filled for the models. The result model of training process will be measured in testing process by accuracy, sensitivity and specificity.

The model with the largest accuracy will be selected to use in detection phase. In the detection phase, the features of new data will be extracted and predicted using the model. Figure 1 shows the flow of the global system.

3 Data

Of 531 women who visited our clinic for DXA measurement between 1996 and 2001, 100 postmenopausal women aged 50 years or older with no previous osteoporosis diagnosis (mean 59.6 years; range 50-84 years) were randomly recruited for this

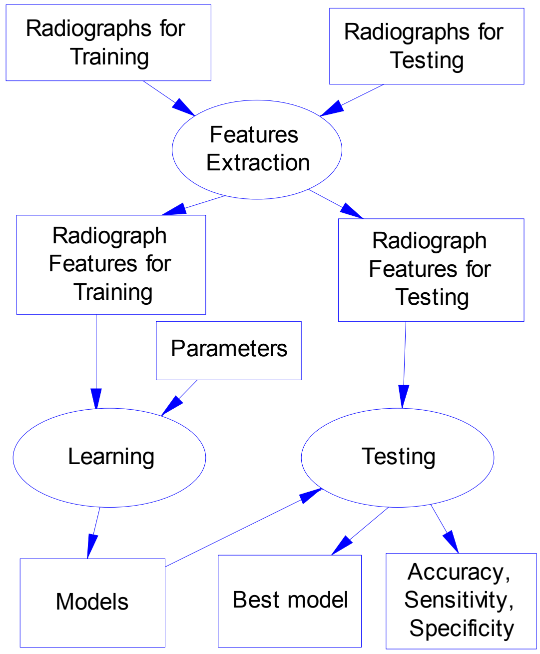
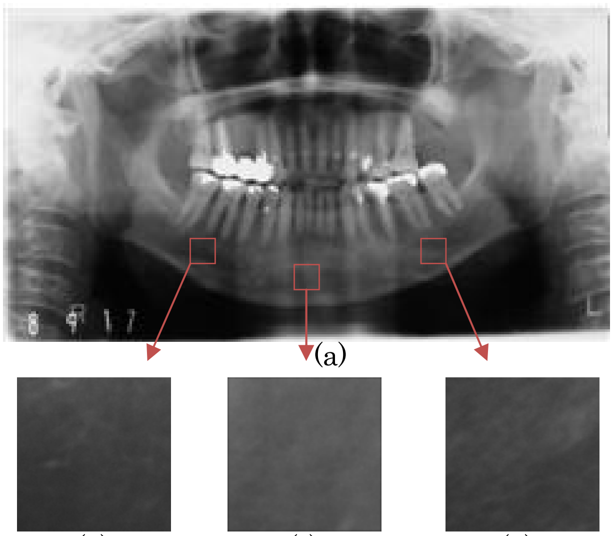


Fig 1: Global system flows

study.

Panoramic radiography was taken for all subjects with informed consent at the time of DXA measurements of the lumbar spine (L2-L4). The dental panoramic radiographs were obtained from 25 osteoporotic women and 75 normal women. From each radiograph, we select three of Regions of Interest (RoI) which are: right, middle and left mandible with size 128x128 for each. Figure 2 shows the radiograph and its 3 RoI's..



(b) (c) (d)

Fig 2: (a) Dental panoramic Radiograph
(b) The RoI of right mandible, (c) The RoI of middle mandible, (d) The RoI of left mandible

4 Features Extraction

Features extraction of radiograph uses one dimensional Fast Fourier Transform (FFT) method and JSEG Segmentation method. Preprocessing of the RoI consisted of top-hat filtering, contrast stretching and noise removal using wiener method. Figure 3 shows the RoI and the result of preprocess, one dimensional FFT and segmentation.

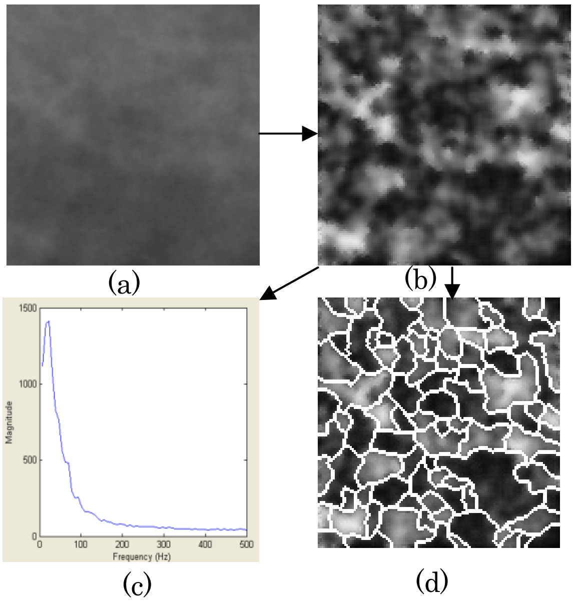


Fig 3: Features Extraction process

(a) RoI of mandible. (b) Preprocessing result. (c) FFT Result. (d) Segmentation Result

One dimensional FFT applied for each row and col of the RoI image. From the Fourier method, we can get the magnitude of frequency and divide into several section of the frequency. For each section, we can find the highest, the lowest, and the average of magnitude. This information from each section is used as the features from frequency domain.

In the other side, the features of Radiographs from spatial domain can be obtained after the segmentation process. In this step, the image RoI is segmented based on grayscale intensity using JSEG segmentation method. For each region of the segmentation result, we can obtain the information about, the average of intensity and size of area. Then we make a group based on range of intensity to get the information of the number of region, the average of intensity and the average of size. We consider this information as the features from spatial domain. The algorithm for features extraction is shown in figure 4.

```
// frequency features extraction
for each row
  mesio=(fft(img(row),128));
magMesio = mean(mesio);
for each col
  apico=(fft(img(col),128));
magApico =mean(apico);
for i=1 to freq_range_cnt
  mesio_min(i)=min(magMesio,freq_range(i))
  mesio_max(i)=max(magMesio,freq_range(i))
  mesio_avg(i)=avg(magMesio,freq_range(i))
  mesio_ratio(i)=log(mesio_max/mesio_min)
  apico_min(i)=min(magMesio,freq_range(i))
  apico_max(i)=max(magMesio,freq_range(i))
  apico_avg(i)=avg(magMesio,freq_range(i))
  apico_ratio(i)=log(apico_max/apico_min)
freqFeatures = all_mesio_apico()
// spatial features extraction
[nReg,regMap,ImgSeg] = JSEG(Img);
SpatFeatures = feature(nReg,regMap,ImgSeg)
```

Fig 4: Features extraction pseudo code

5 JSEG Segmentation

The JSEG method is introduced by Deng and Manjunath[16]. This method separates the segmentation process into two stages, color quantization and spatial segmentation. In the first stage, colors in the image are quantized to several representative classes that can be used to differentiate regions in the image. This quantization is performed in the color space without considering the spatial distributions of the colors. Then the image pixel values are replaced by their corresponding color class labels, thus forming a class-map of the image. The class-map can be viewed as a special kind of texture composition. In the second stage, spatial segmentation is performed directly on this class-map without considering the corresponding pixel color similarity.

Deng and Manjunath make a criterion for "good" segmentation using the class-map, Applying the criterion to local windows in the class-map results in the "J-image," in which high and low values correspond to possible boundaries and interiors of color texture regions. A region growing method is then used to segment the image based on the multiscale J-images.

6 ART and ARTMAP Learning

Adaptive resonance theory (ART) is a new type of neural network that designated by Grossberg to solve plasticity/stability dilemma that is the network can learn from new pattern without forgetting the old knowledge [7].

The first version of ART, ART-1, proposed by Carpenter and Grossberg in 1987, is used to cluster binary data [8][9][10]. Since then several variations of ART have been developed. The most important ones are: ART-2, an extension of ART-1, used to cluster analog data, ARTMAP, a supervised learning mechanism for binary data, and Fuzzy ARTMAP, a supervised learning algorithm for analog data [9][10][11]. Kasuba in 1993 propose the simplified Fuzzy ARTMAP (SFAM), it is equivalent to the sequential implementation of ARTMAP [2]. The algorithm of SFAM is shown in figure 5.

T(x,w) is called the choice function or activation function, which is used to measure the degree of the resemblance of x with w_j , where α is a choice parameter, α > 0.

$$T(x,w_j)=\frac{\|x\Lambda w_j\|}{\alpha+\|w_j\|} \qquad (1)$$

M(x, w_j) is called the match function, which is used to qualify how good is the likeness of w_j to x.

$$M(x,w_j)=\frac{\|x\Lambda w_j\|}{\|x\|} \qquad (2)$$

The function is used in conjunction with the vigilance parameter ρ: 0 < ρ < 1, where M(x, w_j) > ρ means a good match (resonance).

U (x,w) is called the update function, which is used to update a template weight after it resonances with a pattern:

```
while (X not empty)
  get x;
  new = true;
  Loop j =1,C
         Tj = T(x,wj )
  J= sort_desc T
  loop  j = 1,C
    if M(x, wj)>= ρ
      if   L(J)=label(x)
         wj  = U(x, wj );
         new = false;
         break;
     else
          ρ  = M(x,wj ) +ε ;
  If  new= true
    C=C+1;
    wc=x;
    L(C)=label(x)
```

Fig 5: SFAM learning pseudo code

$$U(x,w_j)=(1-\beta)w_j+\beta(x\Lambda w_j) \qquad (3)$$

where β is learning rate, 0<β ≤1. It is called the fast learning in ART when β =1.

The fuzzy AND operator ∧ :

$\mathbf{a}\wedge\mathbf{b}$ = (min(a_1,b_1), min(a_2,b_2) … min(a_D,b_D)) (4)

and the norm operator ||.|| is defined by $\|a\|=\sum_{i=1}^{D}|a|$. Fuzzy ARTMAP requires input pattern to be normalized to prevent category proliferation. The normalization is done by using complement coding, which is

$\mathbf{x}$ = ($\mathbf{x}$,$\mathbf{x}^c$) = (x_1,x_2,…,x_D,1-x_1,1-x_2,…,1-x_D) (5)

Fuzzy ARTMAP use hyper rectangles to represent category weights in a supervised learning paradigm. In general, if x has dimension D, the hyper rectangle R_j includes the two vertices $\wedge_j$ x and $\vee_j$ x , where the i-th component of each vector is defined by the equations:

$(\wedge_j x)_i$ = min {x_i : **x** *has been coded by category j*}

$(\vee_j x)_i$ = max {x_i : **x** *has been coded by category j*}