

Automatic whole-body bone scan image segmentation based on constrained local model

Ema Rachmawati¹, Jondri², Kurniawan Nur Ramadhani³, Achmad Hussein Sundawa Kartamihardja⁴, Arifudin Achmad⁵, Rini Shintawati⁶

^{1,2,3}School of Computing, Universitas Telkom, Indonesia

^{4,5,6}Department of Nuclear Medicine and Molecular Theranostic, Dr. Hasan Sadikin General Hospital, Faculty of Medicine, Universitas Padjadjaran, Indonesia

Article Info

Article history:

Received Feb 21, 2020

Revised Apr 27, 2020

Accepted May 9, 2020

Keywords:

Bone scan images

Cancer lesion

Constrained local model

Landmark points

Regularized landmark

mean-shift

ABSTRACT

In Indonesia, cancer is very burdensome financially for sufferers as well as for the country. Increasing the access to early detection of cancer can be a solution to prevent the situation from worsening. Regarding the problem of cancer lesion detection, a whole-body bone scan image is the primary modality of nuclear medicine for the detection of cancer lesions on a bone. Therefore, high segmentation accuracy of the whole-body bone scan image is a crucial step in building the shape model of some predefined regions in the bone scan image where metastasis was predicted to appear frequently. In this article, we proposed an automatic whole-body bone scan image segmentation based on constrained local model (CLM). We determine 111 landmark points on the bone scan image as the input for the model building step. The resulting shape and texture model are further used in the fitting step to estimate the landmark points of predefined regions. We use the CLM-based approach using regularized landmark mean-shift (RLMS) to lessen the effect of ambiguity, which was struggled by the CLM-based approach. From the experimental result, we successfully show that our proposed image segmentation system achieves higher performance than the general CLM-based approach.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Ema Rachmawati

School of Computing

Universitas Telkom

Jalan Telekomunikasi, Bandung, Indonesia

Email: emarachmawati@telkomuniversity.ac.id

1. INTRODUCTION

In many countries, cancer had been a significant cause of morbidity and mortality for several decades. It is estimated that by 2030, the incidence will increase by 70% to 1.3 million cases [1]. For the detection of bone problems such as cancer lesion, a whole-body bone scan is the primary modality of nuclear medicine, to find out whether the cancers have spread to the bones besides finding out how the metastases respond to chemotherapy and radiotherapy treatment. Besides, a bone scan is an essential tool for detecting whether cancer has spread over the skeleton started from different organs [2].

The skeleton segmentation process in the bone scan image is the initial and essential stage of the whole stage of the metastasis detection system in bone scan images [3-6]. For that purpose, the skeleton of the bone must be divided into several predetermined regions, such as the head and spine region, upper arms and collarbone region, chest region, and pelvis and upper thighs region. In general, the bone scan images produced by whole-body scintigraphy with ^{99m}Tc-methylene diphosphonate (^{99m}Tc-MDP) often

degraded by a large amount of noise, which might affect the segmentation process [7]. Tissue inhomogeneity, images with low resolution, noise, and changes in structure are some examples of challenges faced by the segmentation process in this domain. The results of segmentation will significantly affect the calculation of the bone scan index (BSI) value which further will be calculated as a guide in determining whether or not metastatic lesions are detected in a particular region of bone in cancer patients [4, 8-11]. Another challenge of automatic skeleton bone segmentation lies in the fact that the segmented regions of the skeleton might be overlapped. For example, the upper arm region is overlapped with the chest region; the head-spine region is overlapped with the upper arm and chest region [3, 11]. Those conditions might cause considerable difficulties when applying classical segmentation techniques [12] or deep learning methods [5, 13, 14].

Meanwhile, deformable models have been extensively studied and widely used in medical image segmentation, with promising results [12]. Active shape model (ASM) is one kind of extension of deformable models that incorporate prior shape information to offer robustness to image noise and boundary gaps and allow integrating boundary elements into a coherent and consistent mathematical description [12]. ASM is one of the constrained local model (CLM) based approaches which was widely used in bone scan image segmentation [3, 15, 16]. On the other hand, ambiguity is commonly a significant problem in CLM based approach. We have to learn local image patch detectors from labeled training images for each landmark point in the object. However, the local support is small, and the variation of appearance is abundant in the training set, which in turn leads to the problem of ambiguity. From each landmark detector, we can observe this ambiguity problem in the non-parametric distribution of landmark point locations. To solve the problem, we use the regularized landmark mean-shift (RLMS) [17] in our proposed approach.

2. RELATED WORKS

Cristinacce and Cootes [18] initially proposed CLM to describe the modeling of a class with a distinct set of corresponding features. By using the iterative template generation and shape constrained search technique, the model is matched to new instances of an object. The appearance variation of a set of template regions surrounding individual features is learned by CLM. The advantages of CLM influence some researchers in their works, such as [19] and [20]. Baltrusaitis *et al.* [20] develop Conditional Local Neural Fields (CLNF) to detect landmark points in their OpenFace system with very satisfying accuracy. CLM was also used as the basis for the development of discriminative response map fitting (DRMF), which was used in the face fitting scenario [19]. One of the popular CLM based approaches is ASM. The utilization of ASM to localize structures in medical images was first proposed by Cootes *et al.* [21]. The proposed shape model is based on the representation of objects based on a collection of labeled points, with each point is positioned at a particular part of the object. By evaluating the position statistics from the labeled point, a point distribution model can then be generated. This model gives the average position of the points and the main shape descriptions of the variations found in the training data.

Some researchers sought to develop the ASM method namely [22-24]. Van Ginneken *et al.* [24] proposed the development of a new appearance model as a representation of the gray-level variation around the object boundary. If the classic ASM method uses a normalized first derivative profile, this study considers the use of a local image structure descriptor, the local moment histogram, which is extracted from images that have passed the filter stage using a Gaussian derivative filter bank. Furthermore, statistical analysis is carried out to investigate which descriptors are most informative at each resolution and each landmark. Meanwhile, Cristinacce and Cootes [22] used non-linear boosted-features that were trained using GentleBoost [25] to model each feature. In their research, Cristinacce and Cootes [22] investigate local feature detection using boosted features and boosted regression, which aimed to obtain local feature point predictions without having to search for sliding windows in the local neighborhood. The proposed approach can produce satisfactory performance. Ebhotemhen [23] use the k-nearest neighbor technique to adjust the position of each landmark extracted using the Haar filter. The best position is the position where the local edge structure is closest to the corresponding landmark. Several neighboring points from the initial landmarks were evaluated in the process. This stage is carried out on the shape model training to get a better segmentation quality of the shape model results.

Further, Sadik *et al.* [3] implemented the ASM method at the segmentation stage to build a statistical model of the object in the image. Eight shape models (with one model covering four anterior parts and four posterior parts) are produced from the training model, which is further used in the skeleton segmentation process. Continuing their work, Sadik *et al.* [15] surveyed with their Computer-Assisted Diagnosis (CAD) system, aiming to improve the performance and reduce the variety of different observations from medical personnel who were assigned to interpret the bone metastases. Meanwhile, Horikoshi *et al.* [16] prove the initial assumption that a database of bone scan image training built from European patients could not be used to detect metastases on Japanese patients. The differences in the bone characteristics of the

European population with the Japanese population [16, 26] are proven to be the cause. Zengin *et al.* [27] also stated that there is a difference in the bone geometry between different ethnicities in the experiment of investigating the bone geometry of White, Black and South Asian men in the UK. White men have more extended diaphyseal sites and hip axis length. At the radius and tibia diaphysis, Black men had more cortical bone within a slightly larger periosteal envelope. For the same size and body weight, South Asian men had similar sized bones compared to White and Black men at the diaphysis but had thinner cortices. Based on the fact that the differences in the characteristics of the bones for several populations affect the quality of segmentation result, in our experiment we use Indonesian patients' dataset, as the type of race and ethnicity in Indonesia is more diverse than in Sweden and Japan [28].

3. CONSTRAINED LOCAL MODEL (CLM)

In this section, we briefly describe the CLM-based method, which consists of two steps, namely, model building and model fitting.

3.1. Model building

CLM is a technique to build a shape and appearance model of a flexible object (such as an organ) in the image. The model is built based on the statistics obtained from labeled training images. Each model consists of a flexible shape template D , which describes how several essential points of the object can vary, and a statistical model of the gray level expected in the region around the model $R(p)$ (response maps) [17].

$$\min_p R(p) + \sum_{i=1}^N D_i(x_i; I)$$

A model is trained from a set of images annotated manually by humans. By analyzing the shape variations and appearance on the training image, a model is built so that it can replicate these variations. To interpret a new image, we must find the right parameters to match the model instances to the image. After the model matches the image, the settings or position of the model can be used to classify or make measurements, or as input for subsequent processing. To make a statistical model of appearance, we took a set of annotated images from several examples. Previously, it must be determined in advance a set of landmarks that can precisely describe the shape of the target and which landmarks can be found on the training image. High curvature or T-junction is the right choice as a landmark. Intermediate points can be added between dominant landmarks. Connectivity of each landmark is also recorded to show how landmarks are connected. For 2D images, landmark points $\{(x_i, y_i)\}$ for a single example can be represented as $2n$ element vectors, x .

$$x = (x_1, \dots, x_n, y_1, \dots, y_n)^T$$

For a training image, a vector of s is produced. Statistical analysis of the vectors must consider that the shape is represented in the same coordinate frame. The shape of an object is deemed to be independent of the position, orientation, and scale of the object. Normalization process such as translation, rotation, and scaling must be applied to each shape so that the total distance D of each shape to the mean shape is minimal.

$$D = \sum |x_i - \underline{x}|^2$$

Each image is annotated by using n landmark points. Each shape model has m parameters, and to obtain variations of the shape model, we have to change the parameter value. Further, principal component analysis (PCA) is used to reduce the model dimension. The position of the model point in image X is given by:

$$X = T_{X_t, Y_t, s, \theta}(\underline{x} + Pb)$$

$T_{X_t, Y_t, s, \theta}$ rotates the images at θ angles, scaling at s , and translation on (X_t, Y_t) . The value of b , the rotating frame, are the shape model parameters in the form of a vector with t dimension.

$$b = P^T(x - \underline{x})$$

The shape variations can be built by changing the values of b . The variance of i and b_i values on the training images were given by λ_i values. The adjustment is limited by $\pm 3\sqrt{\lambda_i}$, to ensure that the generated shape variations would be as similar as possible with the training images. The model variations concerning the i^{th} parameter (b_i) is called i_{th} -mode from the model. P is eigenvector; the rotated coordinate frame is adjusted with the original set of shape vector. P contained several t eigenvectors from covariance matrix P where $P = (p_1|p_2| \dots |p_t)$. For example, if we applied the formula to point (x, y) , we will obtain:

$$T_{X_t, Y_t, s, \theta}(x, y) = (X_t, Y_t) + (s \cos \cos \theta \quad -s \sin \sin \theta \quad s \sin \sin \theta \quad s \cos \cos \theta)(x, y)$$

Supposed we want to obtain the best pose (translation, rotation, scale) and the shape parameter to fit the instance model X to several points on new image Y , we have to minimize the sum of square distance between the model and the targeted points on the images as seen as:

$$|Y - T_{X_t, Y_t, s, \theta}(x + Pb)|^2$$

Meanwhile, to compute a response map on the local region around the corresponding landmark point, a local expert function will be used, such as the Mahalanobis distance, a classifier such as linear SVM [29], or a regressor [22]. The region that supports a local expert could be either one-dimensional (i.e., line) or two-dimensional (such as a rectangular region).

3.2. Model fitting

The fitting process in CLM-based methods consists of two main steps: (1) predict the local displacements of the shape model points, and (2) constrain the configuration of all point to adhere to the shape model. These two steps are iterated until they satisfy a convergence criterion. In this article, we apply the regularized landmark mean-shift (RLMS) as proposed by Saragih *et al.* [17], as its simplicity and efficiency.

4. PROPOSED SEGMENTATION SYSTEM

Our purpose in this research is to derive a model from representing the shapes of particular regions in the whole-body bone scan image. Some samples of whole-body bone scan images are shown in Figure 1. Different examples of bone scan images have different shapes so that a rigid shape model would not be appropriate. We aim to build a model that describes both typical shape and typical variability using the examples such as in Figure 1 as the training set, which further will be used in the whole-body bone scan segmentation. In section 4.1, we briefly explained the bone scan images dan their annotation process. Our proposed segmentation system is described in Section 4.2.

4.1. Bone scan image annotation

Bone is the most common and clinically significant site of metastasis, especially in osteotropic tumors such as prostate cancer and breast cancer [2]. Whole-body scintigraphy with $^{99\text{m}}\text{Tc}$ -methylene diphosphonate ($^{99\text{m}}\text{Tc}$ -MDP) is the most available and cheapest technique for detecting cancerous lesions in bone. The bone scan is the first-line imaging for screening and follow-up of the bone metastases in prostate cancer patients. The diagnosis of a bone scan is generally enforced qualitatively through visual reading.

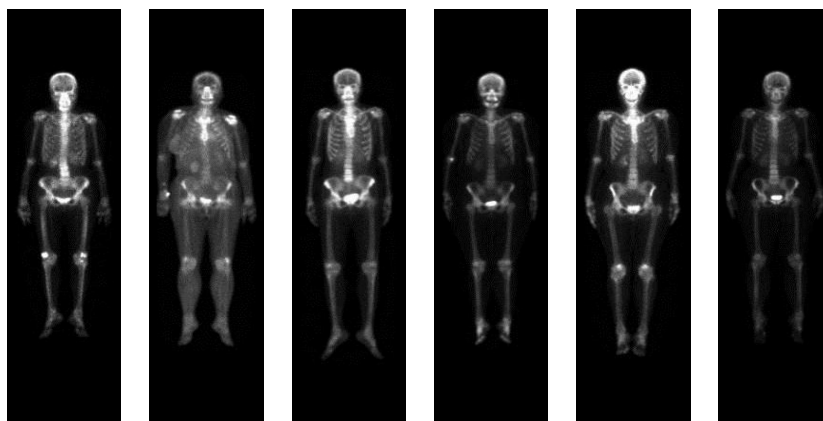


Figure 1. Some examples of whole-body bone scan images

Bone scan test is a nuclear-based imaging test to help diagnose and track the number of bone diseases. The bone scan test is usually recommended by physicians for patients experiencing bone skeletal pain, bone infection, or bone injury that cannot be seen using standard X-rays. Besides, a bone scan test can also be used as an important tool for detecting cancer that has spread (into metastases) to the bone part of the initial location of the tumor, as occurs in breast or prostate cancer. To detect the metastases, firstly, we have to ensure that a mass appeared in a particular region is metastases. Therefore, in this article, we proposed an automatic segmentation system that divides the bone scan image into some particular regions, as shown in Figure 2, which is based on the regions specified by Sadik *et al.* [3]. Sadik *et al.* [3] divided the anterior view of the whole-body bone scan image into four regions, namely head and spine region, chest region, proximal arms and clavicle region, and pelvis and proximal legs region.



Figure 2. Predefined regions of whole-body bone scan image [3]

To model those regions, we represent each region by a collection of points. For the bone scan images, we have chosen to place the points around the boundary of predefined regions, as shown in Figure 3. Hence, each labeled point represents a particular part of the region boundary. The method works by modeling how different labeled points tend to move together as the shape varies. If the labeling is incorrect, with a particular point placed at different sites on each training shape, the method will fail to capture shape variability reliably. We call these labeled points as landmark points. The purpose of the annotation process is to provide important landmark points to be used in the segmentation process. This process consists of three main steps: (1) determine the regions for segmentation, (2) determine the landmark points for each region, and (3) represent the region by determining the connectivity between landmark points already defined in step (2). For this purpose, we divide the bone scan image into 9 (nine) predefined separated segments, as can be seen in Figure 3, namely (i) head_ant segment, (ii) R.chest_ant segment, (iii) L.chest_ant segment, (iv) R.arm_ant segment, (v) L.arm_ant segment, (vi) hips_ant segment, (vii) C.hipsHole_ant segment, (viii) R.hipsHole_ant segment, and (ix) L.hipsHole_ant segment. Those nine segments represent the four regions, as defined by Sadik *et al.* [3] in Figure 2. Further, we define the landmarks points for each segment, along with their connectivity as follows:

- a. "Head_ant" segment: head and spine segment; the segment is represented by the green line; the number of landmark points is 26, with the index number starts from 0 to 25.
- b. "R.chest_ant" segment: right chest segment; the segment is represented by the red line; the number of landmark points is 11, with the index number starts from 26 to 36.
- c. "L.chest_ant" segment: left chest segment; segment with the purple line; the number of landmark points is 11, with the index number starts from 37 to 47.
- d. "R.arm_ant" segment: right proximal arms and clavicles segment; segment with the navy line; the number of landmark points is 10, with the index number starts from 48 to 57.
- e. "L.arm_ant" segment: left proximal arms and clavicles segment; segment with the cyan line; the number of landmark points is 10, with the index number starts from 58 to 67.
- f. "hips_ant" segment: pelvis and proximal legs segment; segment with the orange line; the number of landmark points is 27, with the index number starts from 68 to 94.
- g. "C.hipsHole_ant" segment: center hole of pelvis segment; segment with the yellow line; the number of landmark points is 8, with the index number starts from 95 to 102.
- h. "R.hipsHole_ant" segment: right hole of pelvis segment; segment with the brown line; the number of landmark points is 4, with the index number starts from 103 to 106.
- i. "L.hipsHole_ant" segment: left hole of pelvis segment; segment with the light purple line; the number of landmark points is 4, with the index number starts from 107 to 110.

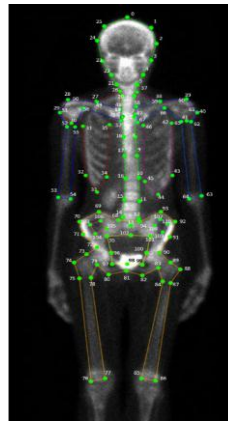


Figure 3. Landmark points for each segment

4.2. Segmentation system

Our proposed segmentation system consists of two phases: (1) the model-building phase, a model is learned from the appearance variations to the shape variations, and (2) the fitting phase, the learned model is applied to a test image to localize landmark points. An overview of the segmentation system is shown in Figure 4. Firstly, in the model building phase, several images are annotated by determining several landmark points to indicate the segments we want to identify. Further, the annotated image will be trained using the CLM method to produce shape and patch model. The shape model is constructed by applying principal component analysis (PCA) to the point distribution model (PDM) from the training images. The texture models were created by extracting several mean-shift vectors from each patch around each landmark point. The next phase, namely the fitting phase, is the phase in which the system fits the right shape model to the test image. In the shape fitting process, we use the regularized landmark mean-shift (RLMS) method [17]. This process produces an estimate of the location of landmark points in the test image. Furthermore, segmentation is done by combining the results of the estimation of landmark locations with metadata files that contain the definitions for each segment.

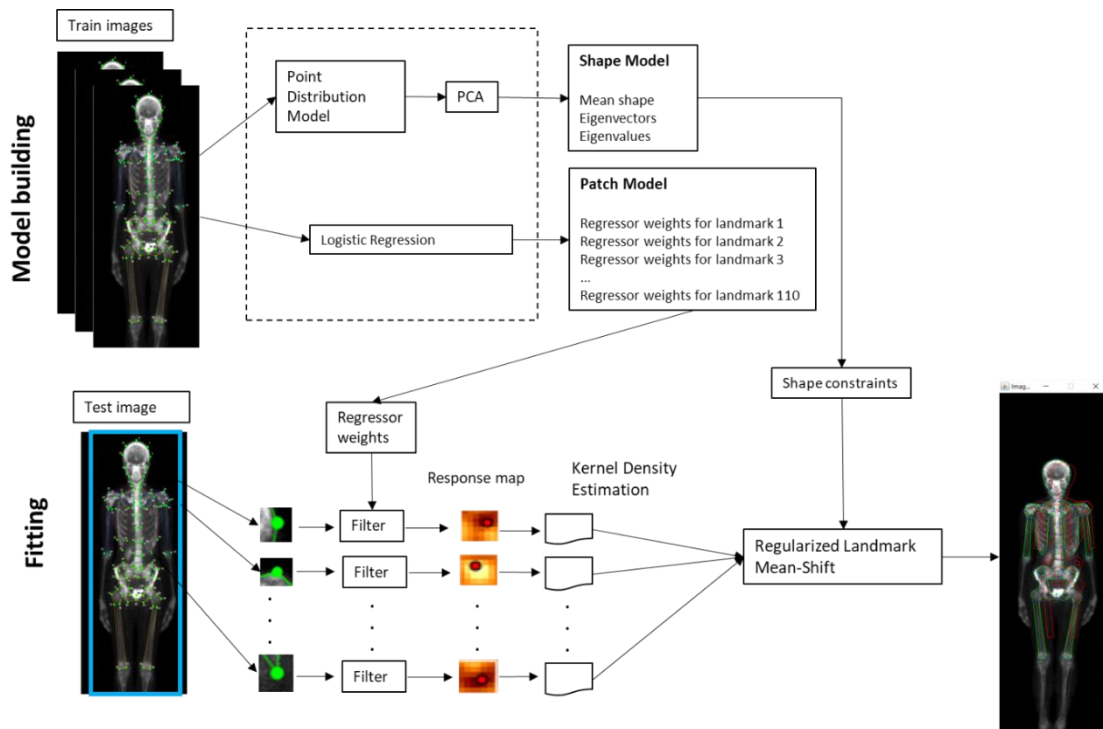


Figure 4. Proposed segmentation system

5. RESULTS AND DISCUSSION

In this section, we explain the experimental result of segmenting the whole-body bone scan image. We use menpo [30] for the implementation of CLM. We explained the preprocessing step in section 5.1, while the result and analysis in section 5.2.

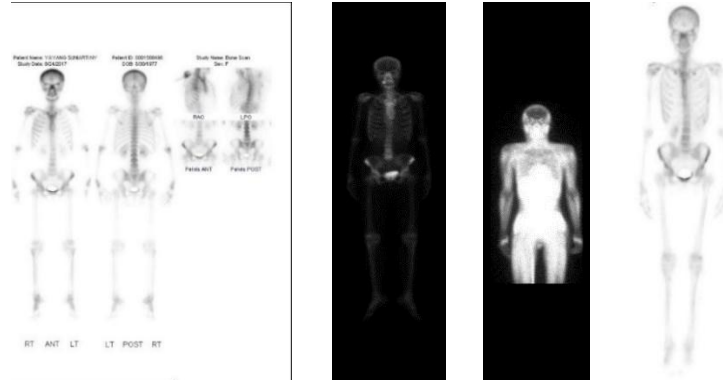


Figure 5. Examples of some variations of the raw bone scan image

5.1. Data preprocessing

For the experiment purpose, the bone scan images were obtained from the image database of the Faculty of Nuclear Medicine, Universitas Padjajaran, in DICOM format. These images were bone scan images from 20 patients. The bone scan images obtained are quite diverse, as can be seen in Figure 5, which is described as follows:

- Whole-body bone scan image anterior view
- Whole-body bone scan image posterior view
- Bone scan image of the pelvis
- Bone scan image of the chest
- Bone scan image of the head

Aside from the diversity of raw bone scan images, for this research, we only process a whole-body bone scan image with a black background. Therefore, we crop the bone scan image from the raw image to obtain the bone scan image having a single object in it, with a resolution of 256x1024 pixels. An example of cropping results can be seen in Figure 6 (a). Meanwhile, as can be seen in Figure 5, some bone scan images were suffered from poor quality. We increase the contrast of the raw image to enhance the bone parts in the bone scan image because the annotation process requires a precise location of bone parts. Examples of the result of changing the contrast can be seen in Figure 6 (b).

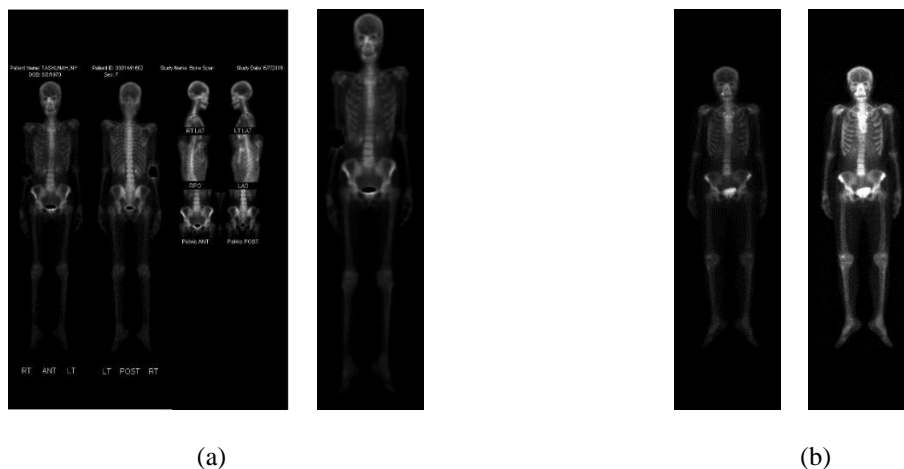


Figure 6. Preprocessing result, (a) Cropping result, (b) Changing contrast

5.2. Result and analysis

For the experiment, we use parameters as follows: the size of the patch shape is 10 x 10 pixels; diagonal is 200; the number of iterations for the fitting process is 20; the size of the context shape is 34x34 pixels. We use image gradient orientation [31] as features which are extracted from each patch and orthogonal PDM [32] in the construction of the shape model. The input of the model building phase is 20 annotated bone scan images, with some examples that can be seen in Figure 7. As we can see in Figure 7, there are some shape variations for chest, proximal arm, clavicles, and pelvis.

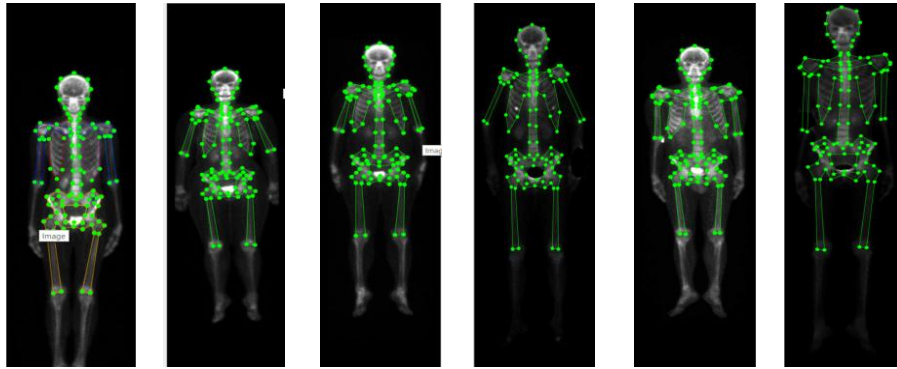


Figure 7. Shape variations of training images

The purpose of this phase is to get the shape model and patch model. The shape model consists of a mean shape and several possible shape variations. The shape model produced at this phase becomes the segment representation, which will then be used in the segmentation process. In Figure 8, we provide the mean shape for each fold, which will be used as an initial reference shape in the fitting process. Some examples of shape model variations produced from the model building phase can be seen in Figure 9.

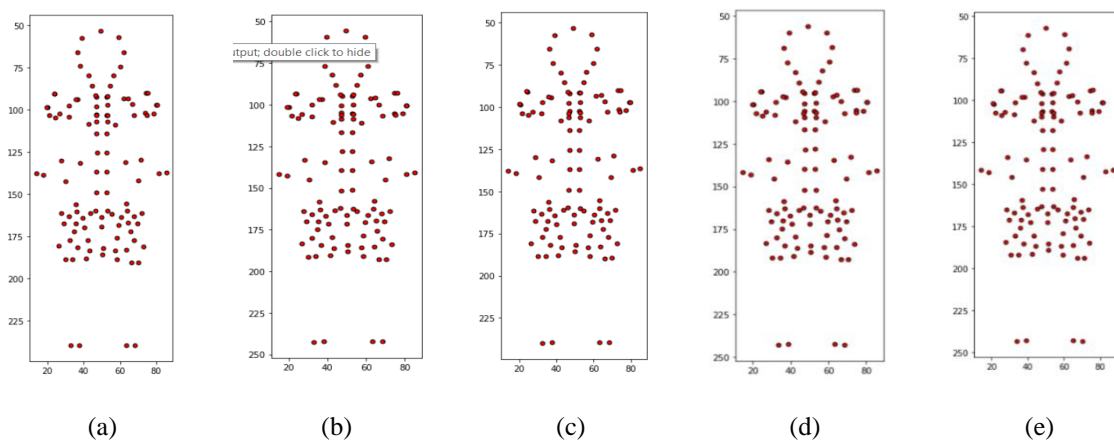


Figure 8. Mean shape for each fold, (a) Fold-1, (b) Fold-2, (c) Fold-3, (d), Fold-4, (e) Fold-5

The segmentation process was defined as fitting the shape models onto the test image. The fitting mechanism used 5-fold cross-validation with cumulative error distribution (CED) as the error measurement. In each fold, there were four testing images and 16 training images. For each fold, we define initial image reference as depicted in Figure 8. This image reference was used as the mean shape model, which is constructed by applying PDM from training images in each fold. The error for five iterations fitting phase can be seen in Table 1, with an average error is 0.023. The segmentation process is conducted by connecting the 111 landmark points resulted from the fitting process. Figure 10 shows some examples of segmentation results, the red line represents the initial shape before the fitting process, and the green line represents the final shape after the fitting process.

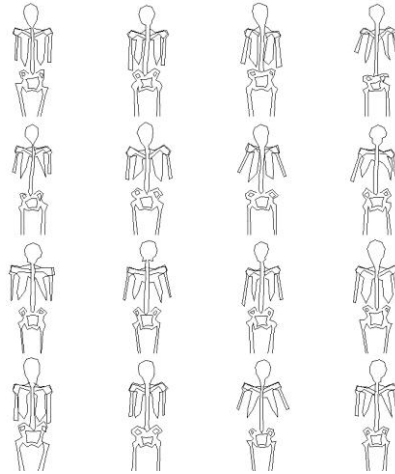


Figure 9. Some shape models resulted from the model building phase

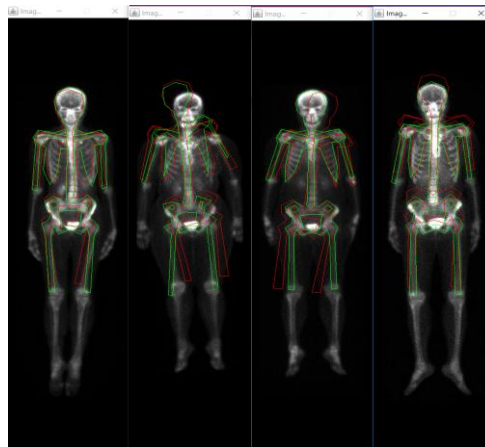


Figure 10. Samples of the segmentation result

Table 1. Cumulative error distribution (CED) score on 5-fold cross validation

Fold	Test image	iter-1		iter-2		iter-3		iter-4		iter-5		average	
		initial error	final error	initial error	final error	initial error	final error	initial error	final error	initial error	final error	initial error	final error
fold-1	Img_1	0.043	0.022	0.065	0.032	0.046	0.018	0.080	0.062	0.077	0.071	0.062	0.041
	Img_2	0.050	0.019	0.062	0.020	0.058	0.021	0.044	0.019	0.058	0.019	0.055	0.019
	Img_3	0.080	0.018	0.046	0.016	0.039	0.015	0.060	0.017	0.047	0.016	0.054	0.016
fold-2	Img_4	0.050	0.018	0.091	0.019	0.039	0.018	0.075	0.034	0.075	0.022	0.066	0.022
	Img_5	0.066	0.015	0.068	0.017	0.054	0.016	0.067	0.016	0.060	0.024	0.063	0.017
	Img_6	0.047	0.039	0.045	0.017	0.059	0.036	0.065	0.023	0.047	0.015	0.053	0.026
fold-3	Img_7	0.027	0.020	0.060	0.037	0.059	0.022	0.035	0.017	0.035	0.017	0.043	0.022
	Img_8	0.068	0.017	0.039	0.016	0.047	0.016	0.030	0.016	0.045	0.017	0.046	0.016
	Img_9	0.080	0.020	0.033	0.023	0.067	0.025	0.051	0.020	0.052	0.021	0.056	0.022
fold-4	Img_10	0.068	0.045	0.048	0.023	0.040	0.023	0.083	0.019	0.048	0.022	0.057	0.026
	Img_11	0.057	0.023	0.053	0.026	0.041	0.022	0.058	0.025	0.023	0.023	0.046	0.024
	Img_12	0.088	0.045	0.066	0.020	0.042	0.020	0.053	0.020	0.040	0.021	0.058	0.025
fold-5	Img_13	0.049	0.016	0.052	0.017	0.039	0.016	0.019	0.016	0.039	0.017	0.040	0.016
	Img_14	0.023	0.020	0.075	0.050	0.073	0.018	0.036	0.021	0.096	0.059	0.061	0.034
	Img_15	0.055	0.022	0.062	0.052	0.050	0.015	0.070	0.027	0.074	0.013	0.062	0.026
average	Img_16	0.060	0.013	0.049	0.013	0.074	0.024	0.047	0.013	0.036	0.014	0.053	0.015
	Img_17	0.071	0.014	0.064	0.021	0.075	0.014	0.048	0.018	0.081	0.030	0.068	0.019
	Img_18	0.046	0.018	0.031	0.019	0.022	0.019	0.042	0.020	0.034	0.020	0.035	0.019
	Img_19	0.063	0.016	0.071	0.021	0.028	0.016	0.063	0.016	0.052	0.018	0.055	0.017
	Img_20	0.047	0.024	0.059	0.020	0.032	0.021	0.056	0.053	0.042	0.020	0.047	0.028
	average	0.057	0.022	0.057	0.024	0.049	0.020	0.054	0.023	0.053	0.024	0.054	0.023

As we mentioned before, the initial shape is the mean shape of the training images, which is used as an initial reference shape. Further, by using the landmark points definition for each region (see Figure 3) of the bone scan image, we can define four regions, namely head and spine region, chest region, proximal arms and clavicle region, and pelvis and proximal legs region. As this experiment applies the RLMS, we compare our experiment result with ASM [21], as used in Sadik *et al.* [15]. Both ASM and RLMS were categorized into CLM based fitting method. With the result, which is depicted in Table 2, we can see that our proposed segmentation system using RLMS appeared to have better performance with error=0.0283.

Table 2. CED comparison of RLMS and ASM

Fold	ASM		RLMS	
	initial error	final error	initial error	final error
fold-1	0.1572	0.1501	0.0485	0.0176
	0.1583	0.1681	0.0456	0.0196
	0.1156	0.1414	0.0471	0.0167
	0.0594	0.0133	0.0507	0.0138
fold-2	0.1005	0.1092	0.0576	0.0199
	0.053	0.0214	0.053	0.0441
	0.0662	0.0245	0.0515	0.0369
fold-3	0.0965	0.112	0.0411	0.0167
	0.0829	0.0654	0.0585	0.0207
	0.0331	0.0246	0.0654	0.0851
fold-4	0.0608	0.0191	0.0837	0.0633
	0.0629	0.0176	0.0626	0.0192
	0.1274	0.1155	0.0465	0.0175
	0.0668	0.0184	0.0499	0.0383
fold-5	0.0523	0.0129	0.0378	0.0133
	0.0228	0.0124	0.0327	0.0137
	0.0644	0.0164	0.0755	0.0145
	0.0434	0.0183	0.0907	0.0217
average	0.1424	0.1423	0.0247	0.0164
	0.1065	0.1022	0.0605	0.0563
	0.08362	0.065255	0.05418	0.028265

6. CONCLUSION

In this article, we propose the whole-body bone scan image segmentation using the CLM-based approach. In that system, we apply the optimization strategy, namely RLMS that was shown to be efficient and straightforward. From the experimental result, we show that our proposed system achieves better performance than the commonly used CLM-based approach, which is the active shape model. With this promising segmentation result, we managed to prove that we succeed in obtaining a specific predefined anatomical region, which will be very useful in the detection of cancer lesions on the bone, which frequently appears in the form of metastases.

ACKNOWLEDGEMENTS

We would like to express our sincere thanks to Telkom University for supporting a Fundamental and Applied Research Grant "Penelitian Dasar dan Terapan" with a contract number KWR4.111/PNLT3/PPM/2018. Also, we would like to thank the Department of Nuclear Medicine and Molecular Theranostic, Dr. Hasan Sadikin General Hospital, Faculty of Medicine, Universitas Padjadjaran, Indonesia, for assistance with the collection of data.

REFERENCES

- [1] J. Ferlay *et al.*, "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," *Int. J. Cancer*, vol. 136, no. 5, pp. 359-386, 2015.
- [2] R. E. Coleman, "Clinical Features of Metastatic Bone Disease and Risk of Skeletal Morbidity," *Clin. Cancer Res.*, vol. 12, no. 20, pp. 6243s-6249s, 2006.
- [3] M. Sadik *et al.*, "Computer-assisted interpretation of planar whole-body bone scans," *J. Nucl. Med.*, vol. 49, no. 12, pp. 1958-1965, 2008.
- [4] D. Ulmert *et al.*, "A novel automated platform for quantifying the extent of skeletal tumour involvement in prostate cancer patients using the bone scan index," *Eur. Urol.*, vol. 62, no. 1, pp. 78-84, 2012.
- [5] A. Kikuchi and T. Kawakami, "Future of Artificial Intelligence and Nuclear Cardiology," *Ann. Nucl. Cardiol.*, vol. 4, no. 1, pp. 79-82 2018.
- [6] I. Hamadeh, P. Nordblom, and Karl Sjöstrand, "System for Detecting Bone Cancer Metastases," RE47609, 2019.
- [7] F.-G. Elfarra, M. A. Calin, and S. V. Parasca, "Computer-aided detection of bone metastasis in bone scintigraphy images using parallelepiped classification method," *Ann. Nucl. Med.*, vol. 33, no. 11, pp. 866-874, 2019.

- [8] H. Wakabayashi *et al.*, "Bone scintigraphy as a new imaging biomarker: the relationship between bone scan index and bone metabolic markers in prostate cancer patients with bone metastases," *Ann. Nucl. Med.*, vol. 27, no. 9, pp. 802-807, 2013.
- [9] K. Nakajima, L. Edenbrandt, and A. Mizokami, "Bone scan index: A new biomarker of bone metastasis in patients with prostate cancer," *Int. J. Urol.*, vol. 24, no. 9, pp. 668-673, 2017.
- [10] S. M. Larson, "EXINI Quantitative Bone Scan Index: Expanded Utility for the Planar Radionuclide Bone Scan," *J. Nucl. Med.*, vol. 57, no. 1, pp. 5-6, 2016.
- [11] A. Inaki, K. Nakajima, H. Wakabayashi, T. Mochizuki, and S. Kinuya, "Fully automated analysis for bone scintigraphy with artificial neural network: usefulness of bone scan index (BSI) in breast cancer," *Ann. Nucl. Med.*, vol. 33, no. 10, pp. 755-765, 2019.
- [12] C. Xu, D. Pham, and J. Prince, "Image Segmentation Using Deformable Models," *Handb. Med. Imaging*, vol. 2. *Med. Image Process. Anal.*, pp. 129-174, 2010.
- [13] Z. Guo, "Deep learning meets graph: novel hybrid methods for improved medical image analysis," Ph.D. dissertation, The University of Iowa, United State, 2020.
- [14] M. R. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI," *Med. Image Anal.*, vol. 30, pp. 108-119, 2016.
- [15] M. Sadik, M. Suurkula, P. Hoglund, A. Jarund, and L. Edenbrandt, "Improved Classifications of Planar Whole-Body Bone Scans Using a Computer-Assisted Diagnosis System: A Multicenter, Multiple-Reader, Multiple-Case Study," *J. Nucl. Med.*, vol. 50, no. 3, pp. 368-375, 2009.
- [16] H. Horikoshi, A. Kikuchi, M. Onoguchi, K. Sjöstrand, and L. Edenbrandt, "Computer-aided diagnosis system for bone scintigrams from Japanese patients: Importance of training database," *Ann. Nucl. Med.*, vol. 26, no. 8, pp. 622-626, 2012.
- [17] J. M. Saragih, S. Lucey, and J. F. Cohn, "Deformable model fitting by regularized landmark mean-shift," *Int. J. Comput. Vis.*, vol. 91, no. 2, pp. 200-215, 2011.
- [18] D. Cristinacce and T. Cootes, "Feature detection and tracking with constrained local models," in *BMVC 2006 - Proceedings of the British Machine Vision Conference*, pp. 929-938, 2006.
- [19] A. Asthana, S. Zafeiriou, S. Cheng and M. Pantic, "Robust Discriminative Response Map Fitting with Constrained Local Models," *2013 IEEE Conference on Computer Vision and Pattern Recognition*, Portland, OR, pp. 3444-3451, 2013.
- [20] T. Baltrušaitis, P. Robinson and L. Morency, "OpenFace: An open source facial behavior analysis toolkit," *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Placid, NY, pp. 1-10, 2016.
- [21] T. F. Cootes, a Hill, C. J. Taylor, and J. Haslam, "The Use of Active Shape Models for Locating Structures in Medical Images," *Image Vis. Comput.*, vol. 12, no. 6, pp. 355-366, 1994.
- [22] D. Cristinacce and T. F. Cootes, "Boosted Regression Active Shape Models," in *Proceedings of the British Machine Vision Conference 2007*, pp. 79.1-79.10, 2007.
- [23] E. Ebhotemhen, "Medical Image Segmentation using an Extended Active Shape Model," *Int. J. Comput. Appl.*, vol. 69, no. 19, pp. 24-29, 2013.
- [24] B. van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny and M. A. Viergever, "Active shape model segmentation with optimal features," in *IEEE Transactions on Medical Imaging*, vol. 21, no. 8, pp. 924-933, 2002.
- [25] B. J. Friedman, T. Hastie, and R. Tibshirani, "Additive Logistic Regression: a Statistical View of Boosting," *Ann. Stat.*, vol. 28, no. 2, pp. 337-407, 2000.
- [26] K. Nakajima *et al.*, "Enhanced diagnostic accuracy for quantitative bone scan using an artificial neural network system: A Japanese multi-center database project," *EJNMMI Res.*, vol. 3, no. 1, pp. 1-9, 2013.
- [27] A. Zengin *et al.*, "Ethnic differences in bone geometry between White, Black and South Asian men in the UK," *Bone*, vol. 91, pp. 180-185, 2016.
- [28] L. L. Cavalli-Sforza, A. Piazza, P. Menozzi, and J. Mountain, "Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data," in *Proceedings of the National Academy of Sciences*, vol. 85, no. 16, pp. 6002-6006, 1988.
- [29] Yang Wang, S. Lucey and J. F. Cohn, "Enforcing convexity for improved alignment with constrained local models," *2008 IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, AK, pp. 1-8, 2008.
- [30] J. Alabort-i-Medina, E. Antonakos, J. Booth, P. Snape, and S. Zafeiriou, "Menpo: A Comprehensive Platform for Parametric Image Alignment and Visual Deformable Models," *Proceedings of the 22nd ACM international conference on Multimedia - MM '14*, Orlando, pp. 679-682, 2014.
- [31] G. Tzimiropoulos, S. Zafeiriou and M. Pantic, "Subspace Learning from Image Gradient Orientations," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 12, pp. 2454-2466, 2012.
- [32] D. A. Ross, J. Lim, Rwei-Sung Lin and Ming-Hsuan Yang, "Incremental Learning for Robust Visual Tracking," *International Journal of Computer Vision*, vol. 77, no. 1-3, pp. 125-141, 2008.

BIOGRAPHIES OF AUTHORS



Ema Rachmawati received the B.Sc. on Informatics Engineering from Institut Teknologi Bandung (ITB), Indonesia in 2004, the M.Sc on Informatics Engineering from Institut Teknologi Bandung (ITB), Indonesia in 2008, and the Ph.D. on Electrical Engineering and Informatics from Institut Teknologi Bandung (ITB), Indonesia in 2018. Since 2010, she joined Telkom University as a lecturer in the School of Computing. Her research interests include machine learning, object recognition, and image understanding.



Jondri is a senior lecturer in the School of Computing at Telkom University, Bandung Indonesia. He received a Master's degree in Mathematics from Institut Teknologi Bandung, Bandung, Indonesia, in February 1999. His research is on machine learning for biomedical signal classification.



Kurniawan Nur R received the B.Sc on Informatics Engineering from Telkom University, Indonesia, in 2008 and M.Sc on Informatics Engineering from Institut Teknologi Bandung (ITB), Indonesia in 2013. Since 2014, he became a lecturer in the School of Computing, Telkom University. His research interests revolve around machine learning, specifically in computer vision.



Achmad Hussein S. Kartamihardja graduated as a medical doctor (MD) from the Faculty of Medicine, Universitas Padjadjaran, in 1984, and since 1985 is continuously participating in various IAEA-based training courses all over the world, including the IAEA Expert Mission in Sudan, 2001. He joined as a lecturer at the Faculty of Medicine Universitas Padjadjaran from 2004. As one of the pioneers of nuclear medicine field in Indonesia, he was appointed as the Head of Department of Nuclear Medicine, Hasan Sadikin Hospital (1995-2003), the president of Indonesian Society of Nuclear Medicine (2004-2012), the National Project Coordinator IAEA Project on PET from 2008, the Country Principle of Asian School of Nuclear Medicine from 2009, and contributed as an editorial board member of *World Journal of Nuclear Medicine* (2002-2010). He obtained his Master in Medical Law (MH.Kes) from Unika Soegijapranata in 2007 and his doctorate at Faculty of Medicine Universitas Padjadjaran in 2014. His long contribution to scientific research and education in nuclear medicine earned him a full professorship in 2018. He is currently the Head of Department of Nuclear Medicine, Hasan Sadikin Hospital (2014-), and also a Vice Dean of Asian School of Nuclear Medicine (2014-) and an editorial board member of *Asia and Oceania Journal of Nuclear Medicine* (2013-) and *Jurnal Sains dan Teknologi Nuklir Indonesia* (2014-).



Arifudin Achmad graduated as a medical doctor (MD) from the Faculty of Medicine, Universitas Gadjah Mada in 2007, and obtained his doctorate at Gunma University Graduate School of Medicine, Department of Diagnostic Radiology and Nuclear Medicine, Japan, in 2013, focusing in nuclear medicine and molecular imaging. After two years of a post-doctoral fellowship in molecular imaging research (2013-2015) and work as an assistant professor at the same department (2015-2017), he returned to Indonesia to join as a lecturer at Faculty of Medicine Universitas Padjadjaran and at the same time pursue his medical residency in Department of Nuclear Medicine and Molecular Theranostics (2018-present).



Rini Shintawati obtained her Diploma on Radiation Technologist from Ministry of Health's Academy of Radiography – Jakarta in 1992, B.Sc on Physics from Universitas Diponegoro in 2009, and Master of Biomedical Science (MBS) from Gunma University Graduate School of Medicine, Japan in 2015. She joined the Department of Nuclear Medicine, Hasan Sadikin Hospital as a radiographer in 1996 and later became a medical physicist from 2009. She also participated in several IAEA-based training courses and contributing to nuclear medicine research.