Multi-Stage Osteolytic Spinal Bone Lesion Detection from CT Data with Internal Sensitivity Control

M. Wels[a], B. M. Kelm[a], A. Tsymbal[a], M. Hammon[b], G. Soza[c], M. Suehling[a], A. Cavallaro[b] and D. Comaniciu[d]

[a]Siemens AG, Corporate Technology, Erlangen, Germany
[b]University Hospital Erlangen, Department of Radiology, Germany
[c]Siemens AG, Healthcare, Forchheim, Germany
[d]Siemens Corporate Research, Princeton, NJ, USA

ABSTRACT

Spinal bone lesion detection is a challenging and important task in cancer diagnosis and treatment monitoring. In this paper we present a method for fully-automatic osteolytic spinal bone lesion detection from 3D CT data. It is a multi-stage approach subsequently applying multiple discriminative models, i.e., multiple random forests, for lesion candidate detection and rejection to an input volume. For each detection stage an internal control mechanism ensures maintaining sensitivity on unseen true positive lesion candidates during training. This way a pre-defined target sensitivity score of the overall system can be taken into account at the time of model generation. For a lesion not only the center is detected but also, during post-processing, its spatial extension along the three spatial axes defined by the surrounding vertebral body's local coordinate system. Our method achieves a cross-validated sensitivity score of 75% and a mean false positive rate of 3.0 per volume on a data collection consisting of 34 patients with 105 osteolytic spinal bone lesions. The median sensitivity score is 86% at 2.0 false positives per volume.

Keywords: Computer-aided detection, osteolytic bone lesions, metastases, computed tomography

1. PURPOSE

Metastasization is one of the most critical issues when suffering from cancer. Metastases occur when cancerous cells from a primary tumor spread to initially unaffected body parts. Within the vertebral column they can cause debilitating pain, pathologic fractures, and spinal cord compressions with severe neurological impairments. This makes the detection of such lesions an important task in cancer treatment. It helps assessing the risk of certain courses of the disease and allows to accurately quantify disease progression or response to therapy over time. However, reading and manually identifying, i.e., annotating, spinal bone lesions from 3D CT data is a challenging and labor-intensive task even for experienced radiologists. Also, there may be a significant inter- and intra-user variability among manual annotations. Computer-aided detection (CAD) systems supporting the radiologist's final decision making in a reproducible manner can therefore be regarded as an extremely important tool to overcome these problems.

One way to address object detection tasks of this kind is the use of multiple classifiers each trying to classify object candidates into candidates worth further consideration and candidates to be rejected in turn.[1] Doing so may bear the following risk: consider a multi-stage detector consisting of three classifiers where every subsequent classifier is designed or trained to particularly handle true positive and false positive detection candidates of the previous stage. If we assume that the first classifier achieves 85% sensitivity on previously unseen samples and that the second and third classifier also achieve 85% sensitivity on previously unseen true positive candidates. Apparently, one will end up with an overall sensitivity score of $85\% \cdot 85\% \cdot 85\% = 61.4\%$ for the overall system, which is often too low to allow a practical application. Furthermore, the classifiers' sensitivity on previously unseen true positive candidates is not necessarily assessed during modeling at all, which may have an unpredictable impact on the sensitivity of the overall system.

In order to address this issue we propose a system for fully-automatic osteolytic spinal bone lesion detection in 3D CT images with an internal control mechanism for maintaining sensitivity on unseen data. In a pre-processing step individual vertebral bodies (see Fig. 1) are detected fully-automatically to define regions of
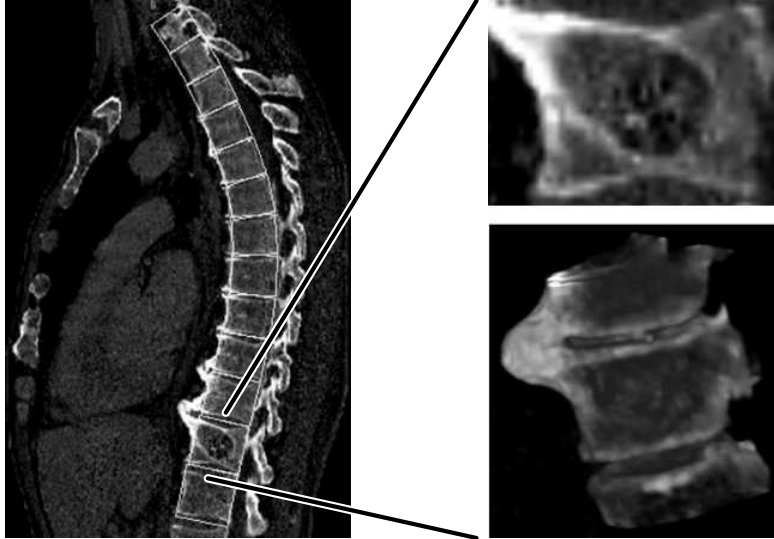
Figure 1. The result of the vertebral body detection and normalization step in sagittal non-normalized and normalized views and in a rendering view.

interest for subsequent lesion detection. The detected vertebral bodies are spatially normalized to their local coordinate systems and input to a multi-stage detector consisting of three random forest-based discriminative models each working on a variety of features associated with individual lesion center candidates. During training each classifier is adjusted to maintain sensitivity on unseen true positive lesion candidates. Nevertheless, all the three classifiers progressively narrow the number of likely osteolytic bone lesion candidates. The remaining candidates are eventually refined by scale estimates through an additional scale detector, grouped together through hierarchical agglomerative clustering, and transformed back to the original coordinate system. The overall system block diagram is depicted in Fig. 2.

2. STATE-OF-THE-ART

Yao et al.[2–4] propose a system for the computer-aided detection of (osteo-)lytic spinal bone metastases. They first automatically extract the spine using thresholding, morphological operations, and region growing. Then, they make use of a watershed transformation to generate lesion candidates, which are finally classified into lesions and non-lesions by a machine learning-based component, which is a Support Vector Machine (SVM). Our multi-stage lesion detection system in contrast is, apart from post-processing, entirely machine learning-based and can therefore conceptually be trained to detect any suspicious spinal bone abnormalities, provided 1) there is sufficient evidence within the data and the derived features to characterize these abnormalities, and 2) the ambiguities between true positive and true negative abnormality candidates are not too large. Furthermore, the system of Yao et al.[2–4] has been trained and evaluated on a data collection containing 50 scans with 28 osteolytic lesions in total. The chosen testing fraction contained 21 scans with 16 osteolytic lesions. We train and evaluate our system for multi-stage osteolytic spinal bone lesion detection on a collection of data sets consisting of 34 scans with 105 lesions. Also, we have chosen cross-validation as the evaluation methodology and do not use predefined training and testing fractions.

Approaches conceptually similar to our method have been successfully applied to detect lymph nodes in CT data[1,5] and ovarian follicles in 3D ultrasound data.[6] However, the difficult problem of osteolytic spinal bone lesion detection so far has not been addressed by a comparable method.
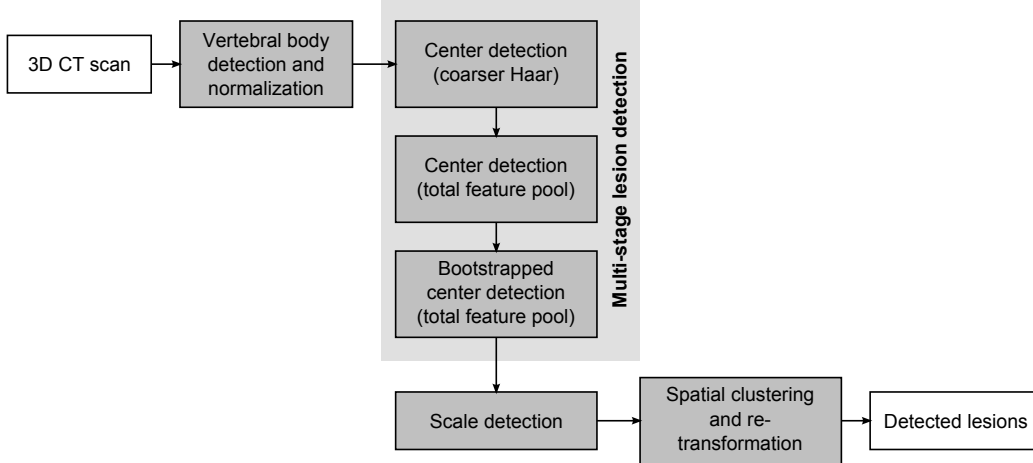
Figure 2. The overall system block diagram.

3. METHODS

3.1 Multi-Stage Detection with Internal Sensitivity Control

Let $n \in \{1, \ldots, N\}$ be the index to a stage of a multi-stage detector consisting of $N \in \mathbb{N}$ discriminative models. Further, let $\mathcal{Z} = \{z\}$ be the set of feature vectors representing training samples. For model generation we have two sets of training data: $\mathcal{C}_{+1}^{(n)} = \{(\pi^{(n)}(z), +1)|z \in \mathcal{Z}\}$, the set of positive training samples, and $\mathcal{C}_{-1}^{(n)} = \{(\pi^{(n)}(z), -1)|z \in \mathcal{Z}\}$, the set of negative training samples. Vectors $\pi^{(n)}(z)$ denote the vectors of components of the overall feature vectors $z \in \mathcal{Z}$ associated with stage n. All the components considered in stage n state the feature pool for model generation. However, not necessarily all components within the feature pool may be considered for classifying a sample in the application stage. For $n \in \{2, \ldots, N\}$ the sets $\mathcal{C}_{+1}^{(n)}$ and $\mathcal{C}_{-1}^{(n)}$ are sampled from the detection results obtained when applying all previous stages to the training data in turn.

For maintaining sensitivity we partition $\mathcal{C}_{+1}^{(n)}$ in disjoint subsets $^a\mathcal{C}_{+1}^{(n)}$ and $^b\mathcal{C}_{+1}^{(n)}$. One for initial model generation and the other for sensitivity control. After having trained a discriminative model $\hat{p}_n(+1|z) \in [0; 1]$ on $^a\mathcal{C}_{+1}^{(n)}$ an inherent threshold $t_n \in [0; 1]$ is adapted such that

$$\frac{|\{(\pi^{(n)}(z), +1) \in {}^b\mathcal{C}_{+1}^{(n)}|\hat{p}_n(+1|z) > t_n\}|}{|{}^b\mathcal{C}_{+1}^{(n)}|} \approx s_n \tag{1}$$

where $s_n \in [0; 1]$ is the pre-defined target sensitivity score of the n-th classifier on unseen data. As $^b\mathcal{C}_{+1}^{(n)}$ is only used for sensitivity control and not for model generation the overall sensitivity on unseen test data of the multi-stage detector can now be estimated as $\hat{s} \approx \prod_n s_n$. Let $s \in [0; 1]$ be the target sensitivity score of a single detection stage. If $s_n = s$ for all n, equation (1) yields $\hat{s} \approx s^n$. This means $s_n = s$ can be chosen based on a given overall target sensitivity score s^* by means of $s = \exp(\ln(s^*)/n)$.

3.2 Pre-Processing: Vertebral Body Detection and Spatial Normalization

In order to effectively search for osteolytic bone lesions we first detect the vertebral bodies in terms of aligned 3D bounding boxes in a pre-processing step. These bounding boxes can be represented by their center, orientation, and scale. Their usage is two-fold: first, they naturally constrain the spatial search range considered for lesion detection. Second, and more importantly, they are used to spatially normalize individual vertebral bodies to a standard coordinate system. The latter is exemplified in Fig. 1. Due to the heterogeneous appearance of pathologic tissue in general, spatial normalization is necessary to take as much advantage as possible of the homogeneity of the non-pathologic background. Thus, varying poses and sizes of the vertebral bodies within the images to be processed are avoided and do not further complicate the detection task. While vertebral body

detection can be performed automatically, e.g., using the approaches described in references[7] and,[8] we rely on manual vertebral body annotations for validation purposes in our present work. That is, we assume faultless pre-processing for unbiased evaluation of our osteolytic spinal bone lesion detection system.

3.3 Discriminative Modeling and Feature Pools

In our system we use a multi-stage detector with three discriminative models. They are all Random Forest classifiers[9] built from a subset of components of the feature pool at hand that most correlate with the training sets' class labels. The online feature filtering technique used in our system is based on Pearson's correlation coefficient. The Random Forest implementation is an extension to the original Random Forest algorithm supporting online learning.

For increased performance we separated the total feature pool for the first stage and only make use of less computationally demanding 3D Haar-like features[1,7,10] at a coarser scale. For the two later stages we additionally use objectness measure-based[11] features, self-aligning features,[1] convergence index-based[12] features, and symmetry-encoding and spatial features taking into account the normalized local vertebral body bounding boxes. Most of the used features have already proven useful in similar object detection tasks. The remaining ones have been designed to exploit the additional information provided by spatial normalization. Each discriminative model within the multi-stage detector faces a challenging classification problem that suffers from class label noise. This is due to the fact that natural degenerative processes often share a similar appearance with osteolytic bone lesions. Random forests are known to be less sensitive to ambiguous labels and are therefore the discriminative modeling technique of choice in our method.

3.4 Online Feature Filtering

In our set-up as well as in many other medical imaging scenarios redundant feature pools of weakly relevant or even irrelevant feature components are generated to represent the objects in question. This includes the mentioned 3D Haar-like features that are often used for object position detection[13] and also simple raw voxel intensities that are another example of a highly redundant representation with many weakly relevant and irrelevant feature components. Even if the learning algorithm used with such representations is relatively robust to the presence of irrelevant feature components, processing these always requires extra computational resources and often causes a decrease in the resulting predictive performance. Besides that, high data dimensionality generally implies several other undesirable effects—a phenomenon known as the "curse of dimensionality". With a fixed number of training instances the predictive power even tends to reduce as the dimensionality increases, which is known as the Hughes phenomenon.

Importance of a feature component $z = z_m$, $\boldsymbol{z} = (z_m)_{m=1,\ldots,M} \in \mathcal{Z}$, may be quantified with its Pearson's correlation with the class feature $y \in \{-1, +1\}$. Pearson's correlation coefficient r_{zy} between z and y can be calculated online with respect to $\mathcal{Z} = \{\boldsymbol{z}_l | l = 1, \ldots, L\}$ as follows:

$$r_{zy} = \frac{L \sum z_l (1 + y_l)/2 - \sum z_l \sum (1 + y_l)/2}{\sqrt{L \sum z_l^2 - (\sum z_l)^2} \sqrt{L \sum ((1 + y_l)/2)^2 - (\sum (1 + y_l)/2)^2}}. \tag{2}$$

The feature importance estimates r_{zy} may be used for feature filtering and need only one pass through the data. Other possibilities to measure feature importance online include the F-Score,[14] and more sophisticated histogram-based estimates for the difference between probability density distributions of the two classes, which can be measured using the Kullback-Leibler divergence or the Hellinger distance. However, the simple linear correlation measure from equation (2) is usually robust enough such that the detectors trained afterwards are at least not worse than detectors trained after using the F-Score measure, the Kullback-Leibler divergence or the Hellinger distance.

It is important to mention that feature selection of that kind allows training detectors for data and hardware set-ups, for which training is not feasible otherwise: detector training becomes scalable with respect to high-dimensional feature pools. Time required for feature selection is neglectable compared to the time needed for subsequent model training. For example, the selection of 20% of the most correlated 3D Haar-like features from the initial set of 100,000 features in the first bone lesion center detector takes only about 5 minutes on a standard
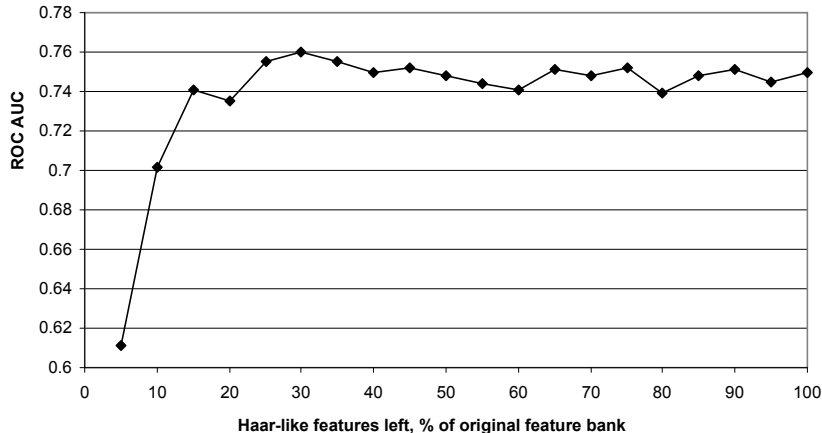
Figure 3. ROC AUC values for bone lesion detection with Pearsons correlation-based feature filtering.

PC. Random Forest model training requires 2 hours. The size of the initial feature vector is only limited by the available memory, which imposes practically no significant restrictions in all possible classification tasks. All the mentioned feature selection strategies are online, with a constant memory requirement kM, where M is the initial size of the feature pool, and k is the size of the data structures required to track the importance of a single feature component.

The online nature of the feature selection approaches is particularly important in medical imaging, allowing to avoid 3D volume caching and thus considerably reducing ultimate memory requirements. Most existing feature selection methods are not online thus requiring all training samples to be present in memory.

Figure 3 demonstrates change in predictive performance when varying the number of features left after feature selection is done. Receiver Operating Characteristic Area Under the Curve (ROC AUC) values are shown for the first detector of our multi-stage detection approach. Pearson's correlation-based feature selection is used in order to select a subset of the most correlated feature components from the initial pool of ca. 100,000 3D Haar-like features. It is interesting to note that the predictive performance drops explicitly only after less than 15% of the features are left. Before this point, even a certain increase in predictive performance in comparison with using the full feature set is observed.

This trend is not specific to this particular classification problem. Similar dependencies are also observed with other feature types (in particular, steerable features in orientation and scale detection[13]) and other subject domains (liver lesion detection and retrieval,[15] fetal brain detection,[16] etc.). Commonly it is safe to reduce the feature pool at least to 30% of the original pool, for features of such nature.

3.5 Online Random Forest

In order to speed up model generation for our learning-based multi-stage osteolytic spinal bone lesion detection framework we make use of an incrementalization of the Random Forest algorithm similar to Saffari.[17] While successful lossless incrementalizations exist for many learning algorithms, most strong techniques applied in real applications are still difficult to incrementalize, and among them perhaps the most prominent example is the Random Forest algorithm itself.[9] A few algorithms for online ensembles of randomized trees have been recently proposed, and their application to vision tasks (in particular tracking) has been considered, despite the fact that they are not lossless and often require considerably more training samples than the corresponding batch techniques in order to converge. In particular, the online Random Forest algorithm of Saffari et al.[17] trains trees of fixed depth and has a fixed structure. We address this issue in our online Random Forest algorithm with the use of primed batch learning to speed up convergence to a reasonable accuracy, different sources of randomness, memory management to avoid exceeding a specified memory limit for the model, and restructuring of the trees according to observed changes in the data distribution. We exploit the memory management scheme proposed for online Hoeffding trees that dynamically activates most promising nodes, for feature distribution observation

and a possible split, and deactivates and removes the less promising ones.[18] Similar to Saffari et al.[17] and unlike in Hoeffding trees, a split is simply generated after observing a certain specified number of instances. For each feature, a Gaussian distribution is assumed and tracked online, and a split threshold value, which maximizes the Information Gain Ratio is selected.

The use of online Random Forests with online correlation-based feature filtering makes our system scalable to the high number of positive and negative training samples, and the size of our feature pools, while achieving competitive predictive performance. A similar online Random Forest implementation has been recently used in a content-based image retrieval system for liver lesion retrieval and characterization.[15]

3.6 Post-Processing

After having detected the centers of likely osteolytic spinal bone lesions a further machine learning-based detector is used to detect the spatial extension of the lesion candidates from a multitude of scale hypotheses, which were generated based on the training data.[10] As lesion candidates located close to each other are likely to represent the same underlying osteolytic bone lesion, these redundant detection results are removed by agglomerative clustering. Also, based on the results of the initial vertebral body detector and a graph cut-based segmentation[19] of the vertebral bodies themselves a rough patient-specific estimation of the spongiosa's intensity distribution within all segmented vertebral bodies is used to additionally reject osteolytic lesion candidates whose center voxels are not sufficiently darker than the immediately surrounding spongious bone tissue. Eventually, the remaining candidates are transformed back to the original image's coordinate system.

4. DATA AND RESULTS

Experiments (9-fold patient-wise cross-validation) have been carried out on a CT data collection consisting of 34 patient-data sets where each patient was suffering from one or more osteolytic lesions within the vertebral bodies. Similar to O'Connor et al.[4] we have restricted our considerations to lesions of substantial size whose ellipsoidal volume is larger than $0.5cm^3$, which roughly corresponds to a sphere with radius 0.5cm. In total there are 105 such lesions in the data collection.

Most of the data has been reconstructed in sagittal slices with slice distances ranging from 1.00mm to 3.09mm with a few axial and one coronal exception (2.00mm slice thickness). Various reconstruction kernels have been used including soft and hard kernel reconstruction. The nearly isotropic in-plane resolution lies between 0.24mm and 1.65mm.

(a)   (b)   (c)   (d)   (e)   (f)
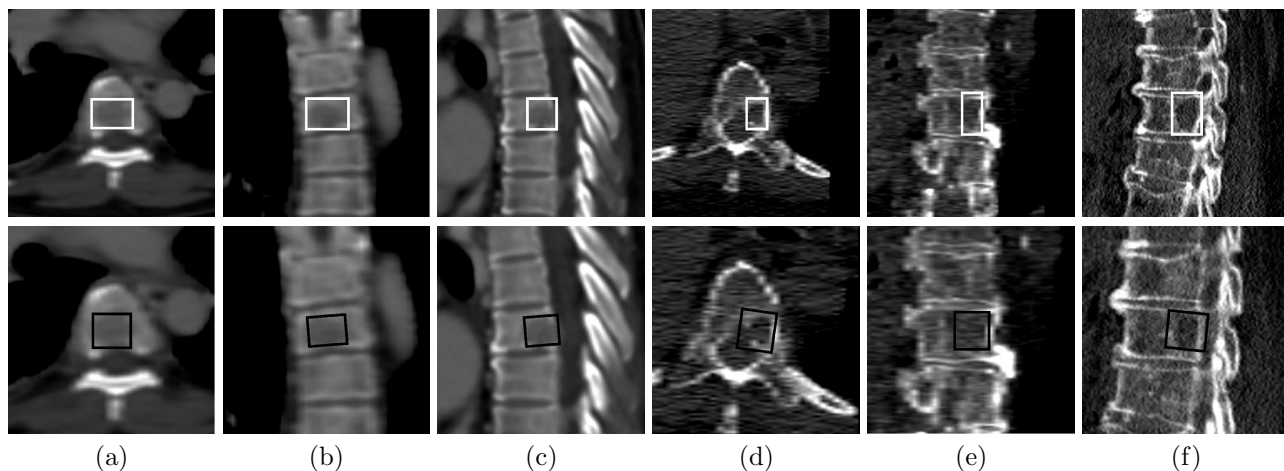
Figure 4. Two exemplary detection results in axial (a, d), coronal (b, e), and sagittal views (c, f). The first row shows the ground-truth annotations in white. The second row shows the detection results in black.

For evaluation ground-truth lesions are considered detected as soon as a detection result's center lies within the axes-aligned bounding box of the ground-truth annotation (see first row of Fig. 4). As osteolytic bone lesions

and other lesions and degenerative processes often look similar, detections that lie within either a previously expert-annotated mixed osteolytic/-blastic lesion, Schmorl's nodule, or hemangioma are not regarded as false positive detections. Further, detections within vertebral bodies that are completely affected by either osteolytic or osteoblastic bone lesions are not counted as false positives. This also holds for detections of osteolytic lesions whose ellipsoidal volume is smaller than 0.5cm^3 or whose limited visibility made the annotators assign a confidence value smaller than 0.5 from the range $[0, 1]$.

Our system achieves 75% sensitivity at 3.0 false positives per volume on average. The median sensitivity score is 86% at 2.0 false positives per volume. The difference between the two sensitivity values is due to the varying contrast and noise characteristics within the used data collection (see Fig. 4) contributing to the degree of difficulty of this detection task. The chosen target sensitivities for each classification stage are 90%, 90%, and 90%. Due to clustering the overall sensitivity of the complete system, which is computed on the lesion level, may differ from the overall target sensitivity on the voxel level. Depending on the number of vertebral bodies within the depicted body portion it takes about 1 to 3 minutes to process one patient data set on standard hardware.

5. CONTRIBUTIONS

We presented a random forest-based multi-stage detector paying respect to label noise in a difficult object detection scenario, i.e., osteolytic spinal bone lesion detection from CT data. The multi-stage detector is modeled with an internal strategy to maintain sensitivity on unseen test data. Each individual detector is an online Random Forest where online correlation-based feature filtering was used before model generation. During post-processing remaining lesion candidates are refined by scale estimates and grouped together through clustering. Additionally, presumed false positive osteolytic bone lesion candidates are rejected based on a patient-specific spongiosa intensity distribution model. Our method has been solidly evaluated on a large data collection consisting of 34 patient data sets with 105 osteolytic bone lesions. The classifiers in each stage are designed to work with an extended feature pool capturing different aspects of the object detection scenario.

6. LIMITATIONS AND CONCLUSIONS

Since osteolytic bone lesions are difficult to identify even for experienced radiologists and the distinction between non-pathologic or other degenerations, e.g., in the case of osteoporosis, may sometimes be unclear, the presented system should be regarded as a tool suggesting suspicious areas for further consideration by an expert. For this purpose the point of operation on a Free-response Receiver Operating Characteristic (FROC) curve of the overall system should be selected such that a sufficiently high sensitivity value is achieved. Under these circumstances a higher FPR is still tolerable as false positive lesion candidates can easily be rejected during manual user interaction. Keeping this in mind we conclude that our system with its current accuracy is close to be clinically applicable for screening examinations.

Furthermore, the discriminative power of our system could potentially be increased by applying our method to combined PET-CT data. Positron emisson tomography (PET) data usually give better visual evidence of osteolytic bone lesions and associated features could be added to the feature pools in a straightforward way.

REFERENCES

[1] Barbu, A., Suehling, M., Xu, X., Liu, D., Zhou, S., and Comaniciu, D., "Automatic detection and segmentation of lymph nodes from CT data," *IEEE Trans. Med. Imag.* **PP**(99), 1 (2011).

[2] Yao, J., O'Connor, S. D., and Summers, R. M., "Automated spinal column extraction and partitioning," in [*IEEE Int. Symp. Biomed. Imag.: Nano To Macro, Arlington, VA, USA*], 390–393 (Apr. 2006).

[3] Yao, J., O'Connor, S. D., and Summers, R., "Computer aided lytic bone metastasis detection using regular CT images," in [*SPIE Med. Imag.: Image Process., Orlando, FL, USA*], **6144**, 614459–1–9 (Mar. 2006).

[4] O'Connor, S. D., Yao, J., and Summers, R. M., "Lytic metastases in thoracolumbar spine: Computer-aided detection at CT—preliminary study," *Radiology* **242**, 811–816 (Mar. 2006).

[5] Feulner, J., Zhou, S. K., Huber, M., Hornegger, J., Comaniciu, D., and Cavallaro, A., "Lymph node detection in 3-D chest CT using a spatial prior probability," in [*IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn., San Francisco, CA, USA*], 2926–2932 (June 2010).

[6] Chen, T., Zhang, W., Good, S., Zhou, K. S., and Comaniciu, D., "Automatic follicle quantification from 3D ultrasound data using global/local context with database guided segmentation," in [*IEEE Int. Conf. Comp. Vis., Kyoto, Japan*], 795–802 (Sept. 2009).

[7] Kelm, B. M., Zhou, S. K., Suehling, M., Zheng, Y., Wels, M., and Comaniciu, D., "Detection of 3D spinal geometry using iterated marginal space learning," in [*Medical Computer Vision (Int. Conf. Med. Image Comput. Comput.-Assist. Interv. 2010 Workshop Proceedings), Beijing, China*], 96–105 (Sept. 2010).

[8] Yao, J., O'Connor, S. D., and Summers, R. M., "Automated spinal column extraction and partitioning," in [*IEEE Int. Symp. Biomed. Imag.: Nano To Macro, Arlington, VA, USA*], 390–393 (Apr. 2006).

[9] Breiman, L., "Random forests," *Mach. Learn.* **45**, 5–32 (Oct. 2001).

[10] Wels, M., Zheng, Y., Carneiro, G., Huber, M., Hornegger, J., and Comaniciu, D., "Fast and robust 3-D MRI brain structure segmentation," in [*Int. Conf. Med. Image Comput. Comput.-Assist. Interv., London, UK*], 575–583 (Sept. 2009).

[11] Antiga, L., "Generalizing vesselness with respect to dimensionality and shape," *The Insight Journal* (July– Dec. 2007).

[12] Kobatake, H. and Hashimoto, S., "Convergence index filter for vector fields," *IEEE Trans. Image Processing* **8**, 1029–1038 (Aug. 1999).

[13] Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., and Comaniciu, D., "Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features," in [*IEEE Int. Conf. Comput. Vis., Rio de Janeiro, Brazil*], (Oct. 2007).

[14] Chen, Y. W. and Lin, C. J., "Combining SVMs with various feature selection strategies," in [*Stud. Fuzz.*], **207**, 315–324, Springer Berlin Heidelberg (July 2006).

[15] Costa, M. J., Tsymbal, A., Hammon, M., Cavallaro, A., Suehling, M., Seifert, S., and Comaniciu, D., "A discriminative distance learning-based CBIR framework for characterization of indeterminate liver lesions," in [*Medical Content-based Retrieval for Clinical Decision Support (Int. Conf. Med. Image Comput. Comput.- Assist. Interv. 2011 Workshop Proceedings)*], (Sept. 2011).

[16] Feng, S., Zhou, S. K., Good, S., and Comaniciu, D., "Automatic fetal face detection from ultrasound volumes via learning 3D and 2D information," in [*IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn., Miami, FL, USA*], 2488–2495 (June 2009).

[17] Saffari, A., Leistner, C., Santner, J., Godec, M., and Bischof, H., "On-line random forests," in [*IEEE Int. Conf. Comput. Vis. Workshops, Kyoto, Japan*], 1393–1400 (Sept. 2009).

[18] Pfahringer, B., Holmes, G., and Kirkby, R., "New options for Hoeffding trees," in [*Aust. Joint Conf. Artif. Intell., Gold Coast, Australia*], 90–99 (Dec. 2007).

[19] Boykov, Y. and Funka-Lea, G., "Graph cuts and efficient N-D image segmentation," *Int. J. Comput. Vis.* **70**, 109–131 (Nov. 2006).