

# Progressive Data Transmission for Anatomical Landmark Detection in a Cloud

M. Sofka<sup>1</sup>; K. Ralovich<sup>2</sup>; J. Zhang<sup>1</sup>; S. K. Zhou<sup>1</sup>; D. Comaniciu<sup>1</sup>

<sup>1</sup>Siemens Corporate Research, 755 College Road East, Princeton, NJ 08540, USA;

<sup>2</sup>Technical University of Munich, Boltzmannstr. 3, 85748 Garching, Germany

## Keywords

Cloud computing, machine learning, pattern recognition system, computer-assisted image processing, image compression

## Summary

**Background:** In the concept of cloud-computing-based systems, various authorized users have secure access to patient records from a number of care delivery organizations from any location. This creates a growing need for remote visualization, advanced image processing, state-of-the-art image analysis, and computer aided diagnosis.

**Objectives:** This paper proposes a system of algorithms for automatic detection of anatomical landmarks in 3D volumes in the cloud computing environment. The system addresses the inherent problem of limited bandwidth between a (thin) client, data center, and data analysis server.

**Methods:** The problem of limited bandwidth is solved by a hierarchical sequential detection algorithm that obtains data by pro-

gressively transmitting only image regions required for processing. The client sends a request to detect a set of landmarks for region visualization or further analysis. The algorithm running on the data analysis server obtains a coarse level image from the data center and generates landmark location candidates. The candidates are then used to obtain image neighborhood regions at a finer resolution level for further detection. This way, the landmark locations are hierarchically and sequentially detected and refined.

**Results:** Only image regions surrounding landmark location candidates need to be transmitted during detection. Furthermore, the image regions are lossy compressed with JPEG 2000. Together, these properties amount to at least 30 times bandwidth reduction while achieving similar accuracy when compared to an algorithm using the original data.

**Conclusions:** The hierarchical sequential algorithm with progressive data transmission considerably reduces bandwidth requirements in cloud-based detection systems.

tion of anatomical landmarks in 3D volumes in the cloud computing environment. The system, dubbed Detection in a Cloud (DiC), is used by thin-client devices that request a set of anatomical landmarks from patient data for visualization or analysis [6]. The patient data stored in a data center are transmitted to a high performance data analysis server that runs the detection algorithm. (In the medical domain and also in this paper, these servers are referred to as the Picture Archiving and Communication System (PACS) and Computer Aided Detection (CAD) servers, respectively). The image with the anatomy highlighted is returned back to the client for display and analysis to support diagnosis and treatment planning (►Fig. 1) [7].

Inherent difficulties in designing such a system are large image sizes (often hundreds of megabytes) and limited bandwidth among thin client, data analysis server (e.g. CAD server), and data center (e.g. PACS server). Depending on the bandwidth, the transmission of large datasets can take tens of seconds or even minutes. Since it is not possible to process the data on the thin client directly, an obvious solution is to efficiently transmit the data from the PACS server to the CAD server for processing (to support software as a service (SaaS) model [8]). However, transmitting data between PACS and CAD servers is prohibitive already when several detection requests are made simultaneously demanding bandwidths of tens of GBit / second. This complicates the workflow in interactive applications where the results must be available immediately. Finally, limited memory and insufficient CPU power of the client necessitates remote data processing.

We propose a hierarchical sequential detection algorithm to avoid the problem of

## Correspondence to:

Michal Sofka  
Siemens Corporate Research  
755 College Road East  
Princeton, NJ 08540  
USA  
E-mail: michal.sofka@siemens.com

Methods Inf Med 2012; 51: 268–278

doi: 10.3414/ME11-02-0017

received: March 2, 2011

accepted: December 6, 2011

prepublished: April 5, 2012

## 1. Introduction

In the recently emerging concept of cloud-computing-based systems, various authorized users have secure access to patient records from a number of care delivery organizations such as hospitals, urgent care centers, doctors, laboratories, and imaging centers from any location [1, 2]. Remote visualization of data sets in the clinical

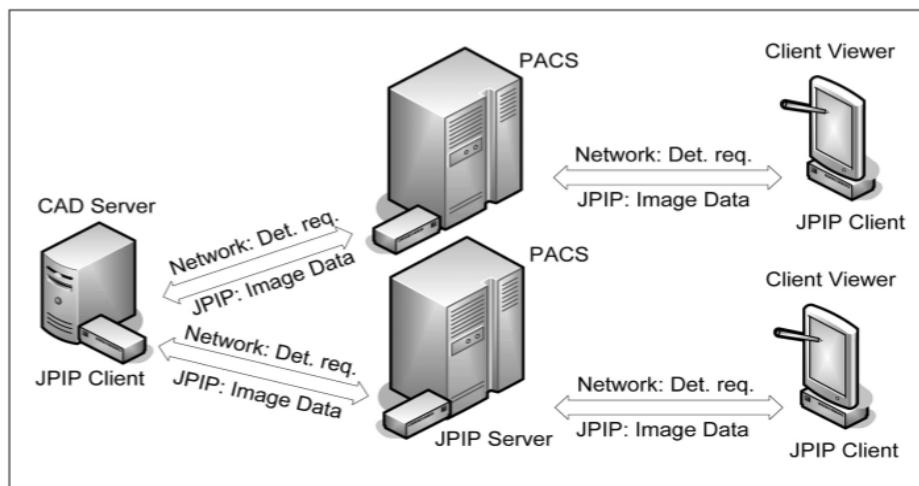
practice is possible through the integration based on the Digital Imaging and Communication in Medicine (DICOM) and Health Level 7 (HL7) standards [3]. This integration allows for advanced image processing, state-of-the-art image analysis, and computer aided diagnosis distributed over several systems [4, 5]. This paper takes advantage of this integration and proposes a system of algorithms for automatic detec-

transmitting large datasets. First, the algorithm only processes candidate regions with high prior probability rather than the entire images. Second, all image regions are compressed with a lossy compression. The lossy compression does not hinder the final detection accuracy. The hierarchical algorithm is robust and accurate, partly due to using the compressed images also in training. We provide an evaluation of training and detection on images compressed with lossy 3D JPEG 2000 and analysis of the data transmission size when detecting multiple landmarks.

## 2. Related Work

Locally-hosted servers are typically used in laboratories for data-intensive tasks, and for sharing data and applications in collaborative research, e.g., in the Biomedical Informatics Research Network (BIRN) and Cancer Biomedical Informatics Grid (caBIG), both funded by the National Institutes of Health (NIH) [9]. The recent years have witnessed large increases in computing power, data storage capacity and network speed, giving birth to medical applications which may handle large data volumes of increased complexity, distributed over the Internet [6]. The paper in [6] describes an infrastructure to handle queries into a large distributed database to aid in mammogram analysis. In [10], large scale medical image registration is run on a distributed system to evaluate accuracy of the final alignment compared to statistically obtained reference standard. In [11], a system for analyzing images of cells from time-lapse microscopy images is run on a cloud computing cluster to visualize living specimen's dynamic processes. The paper in [12] presents a cloud-based system for mobile health care information management. In this paper, we present a system for automatic detection of anatomical landmarks in CT and MRI volumes in the cloud computing environment.

Previous discriminative approaches [13, 14] detect objects by testing entire images exhaustively at all locations. Hierarchical modeling has focused on exploiting multiple feature levels of different resolutions [15–17] and on part-based [18] or



**Fig. 1** Detection in a Cloud (DiC) system. The client sends a request for the detection of an anatomy for a specific patient. The detection algorithm on the CAD server requests data from the PACS server. The detection results are efficiently visualized by the client via the JPEG 2000 JPIP protocol.

region-based [19] architectures. In our multi-resolution hierarchy, the position candidate hypotheses are propagated during training and detection. This results in a more robust detector than when the levels are trained independently [20] and can be extended to multiple landmarks. Multi-object detection in [21] is realized through the Hierarchical Detection Network (HDN) which makes it possible to predict candidate object poses based on previously detected objects. This approach is suitable for progressive data transmission, since only image regions surrounding the candidate locations are needed for the detection of additional landmarks. Using only sparse set of regions is important to keep the transmission data size small.

The JPEG 2000 standard [22] also includes client/server Interactive Protocol (JPIP) for transmitting image regions at desired resolutions using the least bandwidth required. The JPIP protocol is useful for visualizing large Dicom images remotely [22] and has a potential to be used in image analysis applications. The quality of JPEG 2000 images after lossy compression have been previously evaluated for reading radiology datasets [23]. In this paper, we evaluate the robustness of a learning-based algorithm using compressed images in training and detection.

Operating under bounded bandwidth and computational power has been pre-

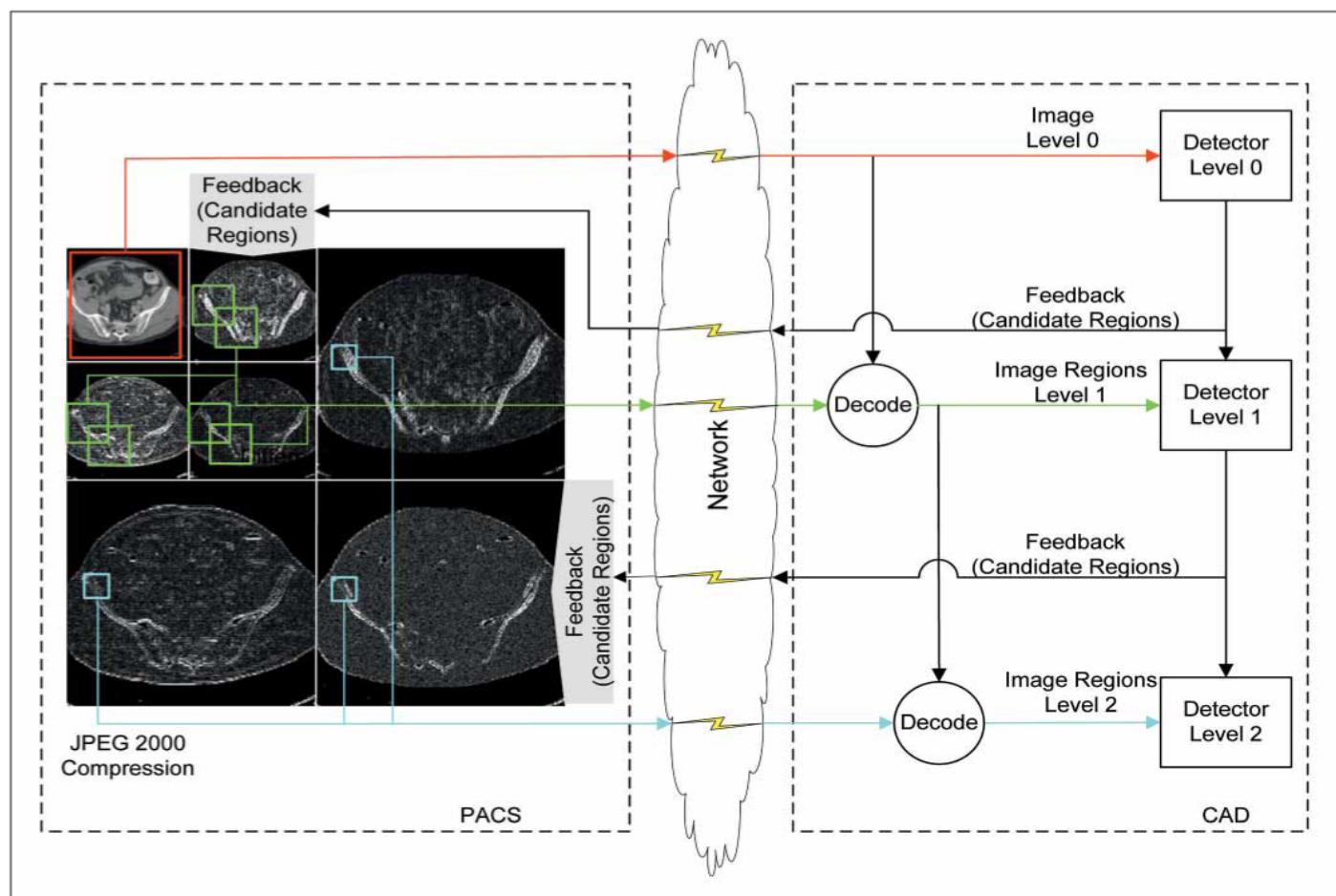
viously addressed in visual surveillance applications [24, 25]. The extracted information (regions [24] and detected objects [25]) has much smaller size than the original images and can be transmitted efficiently over a wide-area network.

## 3. Methods

The core of the DiC system (►Fig. 2) is a hierarchical sequential learning algorithm (Section 3.2) with a sequence of detectors trained for different resolutions and landmarks. The search range of a detector is determined based on the prior probability of the landmark location estimated using the previous detector. The image regions corresponding to the search range are progressively obtained over the network from the PACS server. Since they are encoded with a JPEG 2000 image compression, only high frequency wavelet components need to be transmitted at each subsequent level.

### 3.1 Hierarchical Multiple Landmark Detection

In our multiple landmark detection algorithm, we adopt Hierarchical Detection Network (HDN) that estimates unknown landmark states (e.g., landmark locations) as a sequential decision process [21]. The



**Fig. 2** Overall DiC system diagram. The hierarchical detection algorithm progressively obtains image regions required for detection at each level and for every landmark.

formulation is similar to Markov chain approaches to object tracking, but instead of a temporal motion model with temporal observations, there is a spatial dependence (or prior relationship) between landmarks. For each landmark  $t$ , the input to the algorithm is a set of observations obtained from the image neighborhood  $V_t$ . The neighborhood  $V_t(r, q, R)$  is specified by the coordinates of a bounding box with size  $R$  within a  $d$ -dimensional image  $V, V:R^d \rightarrow [0, 1]$  of resolution  $r$  and quality  $q$  (such as measured by peak signal-to-noise ratio, PSNR). The quality  $q$  is lower for images with artifacts caused by image compression. The PSNR value is determined with respect to the uncompressed image, which has the highest quality  $q$ . Each landmark is represented by its position  $\theta_t = (p_x, p_y, p_z)$ . Let's denote the sequence of multiple land-

mark detections as  $\theta_{0:t} = \{\theta_0, \theta_1, \dots, \theta_t\}$ . The sequence of observations (features) is denoted as  $V_{0:t} = \{V_0, V_1, \dots, V_t\}$ . The multi-landmark detection problem is solved by recursively applying prediction and update steps to obtain the posterior distribution  $f(\theta_{0:t}|V_{0:t})$ . The prediction step computes the probability density of the state of the landmark  $t$  using the states of the previous landmarks, up to  $t-1$ , and previous observations of all landmarks up to  $t-1$ .

The prediction approximates the detection up to landmark  $t$  using the transition probability,  $f(\theta_t|\theta_{0:t-1})$ , and the posterior up to landmark  $t-1$ :

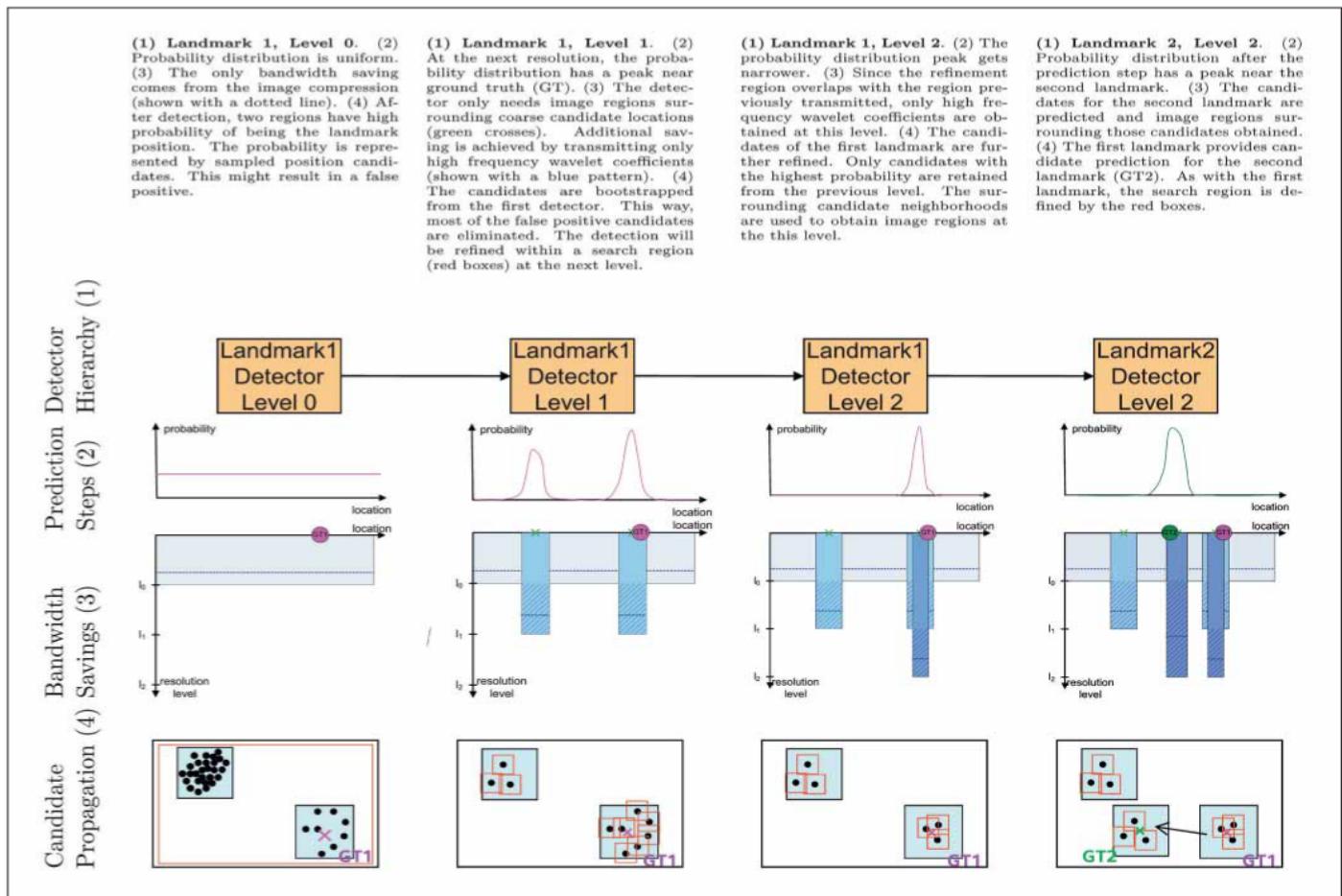
$$\begin{aligned} f(\theta_{0:t}|V_{0:t-1}) &= \\ f(\theta_t|\theta_{0:t-1})f(\theta_{0:t-1}|V_{0:t-1}) & \end{aligned} \quad (1)$$

The update then fuses the results with the new observation region,  $V_t$ :

$$f(\theta_{0:t}|V_{0:t}) = \frac{f(V_t|\theta_t)f(\theta_{0:t-1}|V_{0:t-1})}{f(V_t|V_{0:t-1})} \quad (2)$$

The likelihood (observation model),  $f(V_t|\theta_t)$ , is empirically modeled by training a discriminative model. The transition prior  $f(\theta_t|\theta_{0:t-1})$  approximates the sequential dependence of landmark  $t$ . These terms are further defined in the next section.

During training, the landmark position candidate hypotheses are propagated from the coarser levels to the finer levels using the above steps as follows. At the coarsest resolution  $r_0$ , a classifier  $D(r_0, q_0)$  is trained using the volume region  $V(r_0, q_0, R_0)$ . The



**Fig. 3** Schematic diagram showing the robustness of the hierarchical processing (1). Prediction step of the sequential sampling and the schematic diagram of the probability distribution (2). Bandwidth savings achieved by the hierarchical processing and bootstrapping feedback (3). The search regions in the candidate propagation are shown with red rectangles (4). Only three resolution levels and two landmarks are shown to demonstrate the concept.

size  $R_0$  of the region is the size of the whole image at resolution  $r_0$ . The detector is then used to obtain position candidate hypotheses at this level. The candidates with the highest probability are bootstrapped to train a detector  $D(r_1, q_1)$  at the next level with resolution  $r_1$ . The volume region  $V(r_1, q_1, R_1)$  is composed of the union of neighborhood regions of size  $R_1$  surrounding the position candidates. The bootstrapping procedure continues until all levels  $\{r_j\}$  have been processed. When training detectors for multiple landmarks, the candidates are propagated between landmarks in a similar way. As a result, hierarchical detection and multiple landmark detection is formulated in a unified way.

The hierarchical processing has several advantages. First, the decreasing context re-

gion size helps to avoid local maxima of the probability distribution that would otherwise cause false positives. This results in a more robust and efficient algorithm that operates on datasets of reduced size. Second, the search step depends on the resolution at each level and does not need to be chosen to balance the accuracy of the final result and computational speed.

### 3.2 The Observation and Transition Models for Adaptive Learning

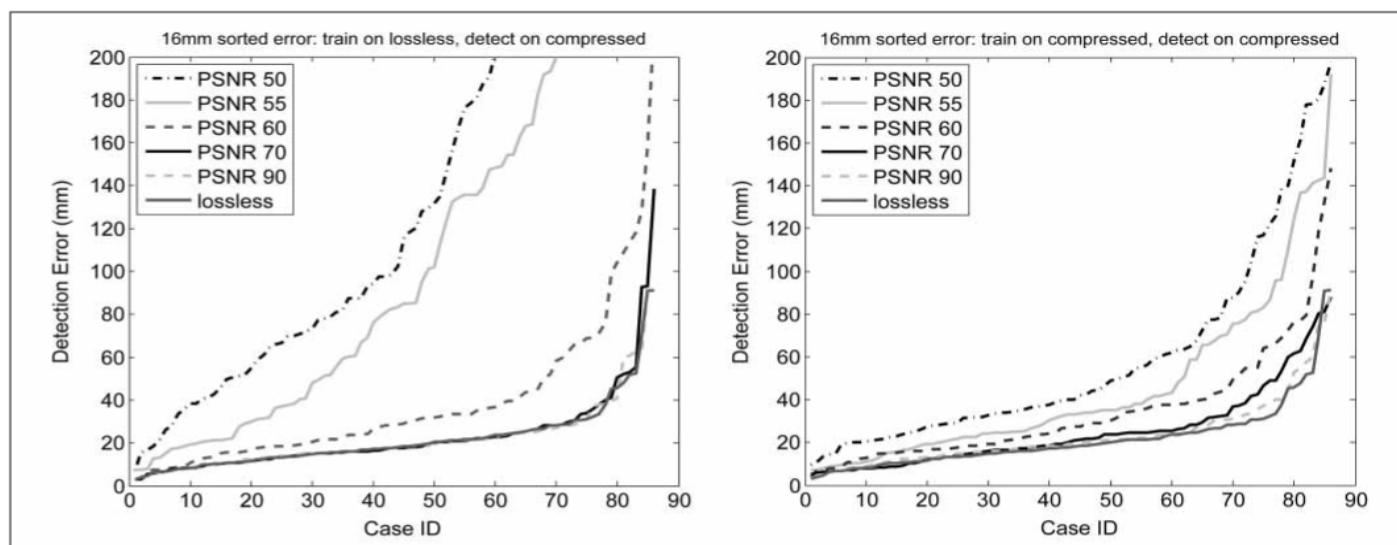
Let us now define a random variable  $y \in \{-1, +1\}$ , where  $y = +1$  indicates the presence and  $y = -1$  absence of the landmark. We train a Probabilistic Boosting Tree classifier (PBT) [13] with nodes composed of

AdaBoost classifiers trained to select features that best discriminate between positive and negative examples of the landmark. We can then formulate the posterior distribution of a landmark presence as  $f(y_t = +1 | \theta_t, V_t)$ . Therefore, the observation model (Section 3.1) has the form

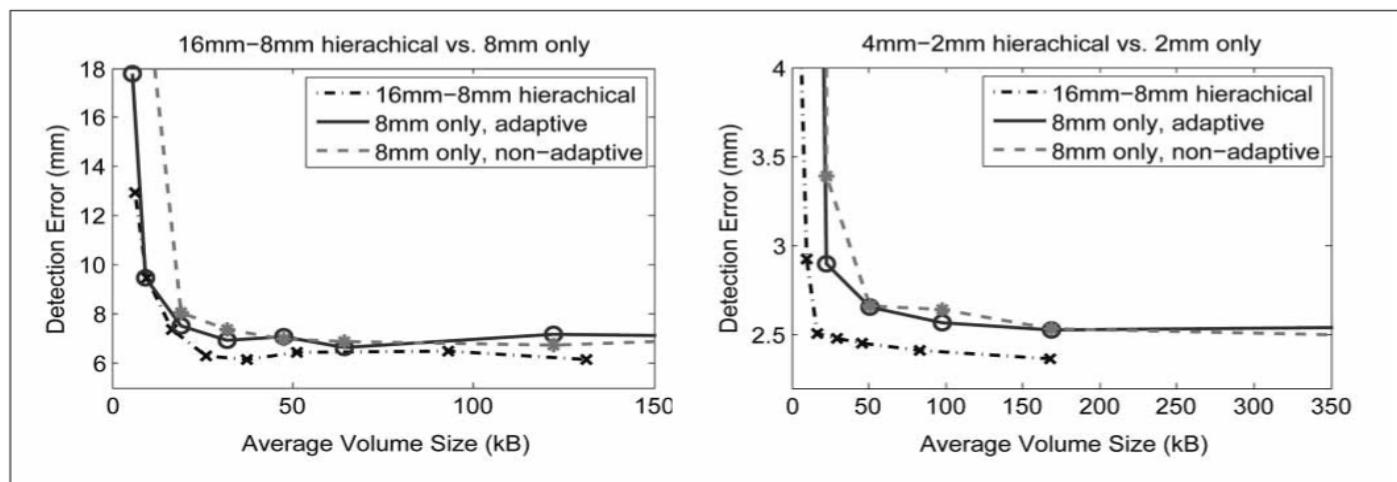
$$f(V_t | \theta_t) = f(y_t = +1 | \theta_t, V_t). \quad (3)$$

The best transition kernel might stem from any previously detected landmark, depending on the anatomical context. In this paper, we use a pairwise dependency

$$f(\theta_t | \theta_{0:t-1}) = f(\theta_t | \theta_j), \quad j \in \{0, 1, \dots, t-1\} \quad (4)$$



**Fig. 4** Sorted detection error for different levels of compression in testing when trained on uncompressed (left) and compressed (right) images. By using compressed images in training, the classifier is more robust to the compression artifacts during detection.



**Fig. 5** Detection error vs. average volume size for hip bone landmark in CT (left) and crista galli landmark in brain MRI (right). The images were compressed in training and testing with the same pSNR level (adaptive) and uncompressed in training and compressed in testing (nonadaptive). The hierarchical processing results in lowest detection error through the focused coarse-to-fine search and training on compressed volumes. The average size of the uncompressed and lossless-compressed volumes is: 404 kB and 189 kB (8 mm CT), 3334 kB and 985 kB (2 mm MRI).

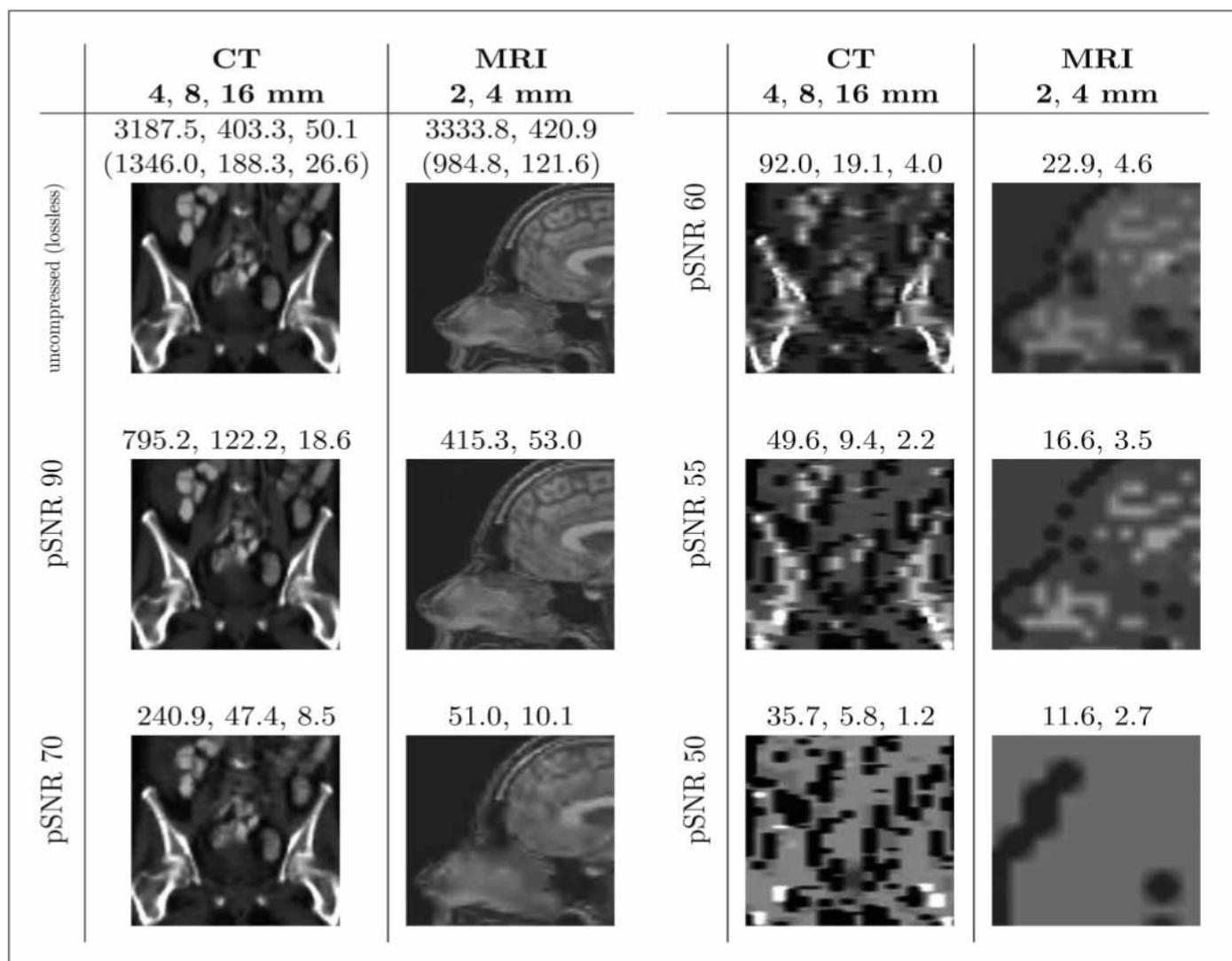
We model  $f(\theta_t | \theta_{0:t-1})$  as a Gaussian distribution estimated from the training data.

When applying a classifier trained using uncompressed data to detect landmarks in a lossy-compressed data, the compression artifacts decrease the detection accuracy. This is especially true when high compression rate of 3D JPEG 2000 introduces strong artifacts. To improve the detection accuracy, we make the detector adapted to the compression artifacts by training the

classifier on the compressed data instead of the uncompressed data. In the learning procedure, AdaBoost automatically discards image features sensitive to compression artifacts and selects robust features that are consistent around anatomical structures overall. In this paper, we do not study the impact of the compression on the diagnostic reading and refer reader to [23]. The performance of the adaptive learning is analyzed in Section 4.

### 3.3 Progressive Data Transmission

The hierarchical detection algorithm allows for enormous bandwidth savings between the CAD and PACS servers. First, since the images are encoded with a lossy 3D JPEG 2000 compression, only high frequency wavelet components are transmitted at the higher resolution levels. The robustness of the system is not sacrificed thanks to adaptive discriminative learning



**Fig. 6** Average compressed volume sizes (in kB) at different compression levels and resolutions. The first row also shows sizes after lossless compression (in parentheses). These statistics are computed over all testing volumes.

which focuses on consistent anatomical structures and ignores compression artifacts (Section 3.2). Second, when we incorporate the bootstrapping of the candidates across levels and multiple landmarks, image at the coarsest resolution  $r_0$  is transmitted in its entirety and only image regions surrounding the candidates with the highest probability  $f(\theta_{0:t}|V_{0:t-1})$ , of the prediction step are used at subsequent levels (Section 3.1). See ▶Figure 3 for an overview of the savings achieved by hierarchical detection and progressive data transmission.

The DiC system must compromise between data bandwidth and the detection accuracy. Denoting  $q_l$  image quality used at the level  $l$  with resolution  $r_l$ , the total size of all image regions for transmission is

$$S_{total} = S(V(r_0, q_0, R_0)) + S(V(r_1, q_1, R_1)) + \dots + S(V(r_n, q_n, R_n)), \quad (5)$$

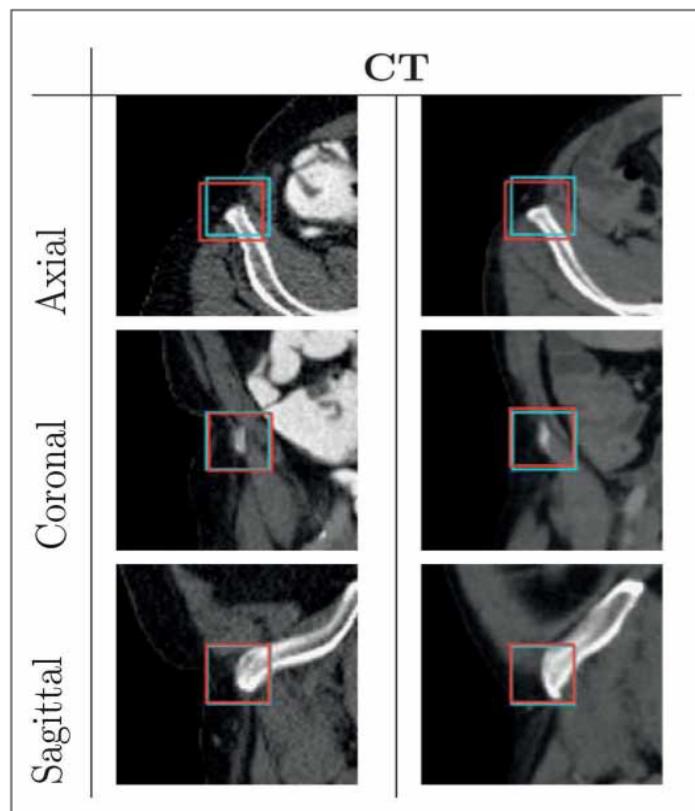
where  $n$  is the total number of levels used. An algorithm only using the finest resolution  $r_n$  without the bootstrap feedback would require size  $S(V(r_n, q_n))$ . The final detection score  $P_{final} \in [0, 1]$  is computed from the scores at each level

$$P_{overall} = P(D(r_0, q_0)) \cdot (P(D(r_1, q_1)) \dots P(D(r_n, q_n))). \quad (6)$$

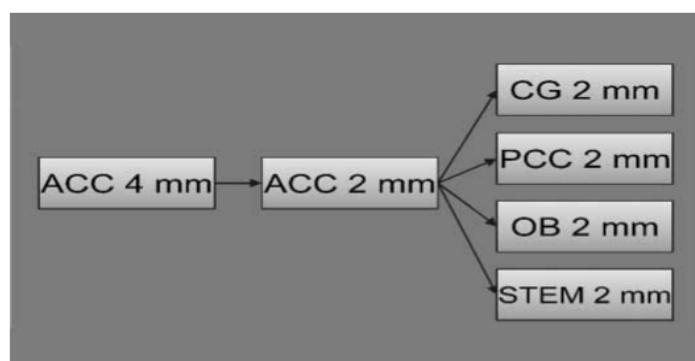
Using these definitions, we can now formulate the following optimization problems to determine the quality used at each resolution level. First, given a fixed-size budget  $S'$ , the goal is to maximize the detection performance by choosing different quality values  $(q_0, \dots, q_n)$  for different resolutions:

$$(q_0, \dots, q_n) = \underset{S \leq S'}{\operatorname{argmax}} P_{overall} \quad (7)$$

Second, given a required detection performance  $P'$ , the goal is to minimize the



**Fig. 7**  
Example images with ground truth locations (red boxes) and detection results (cyan boxes). The CT images correspond to the right hip bone landmark.



**Fig. 8**  
The diagram of the Hierarchical Detection Network for five brain landmarks: crista galli (CG), occipital bone (OB), the anterior of the corpus callosum (ACC), the posterior of the corpus callosum (PCC), and a landmark in the brain stem (STEM). The training and testing datasets are disjoint and the volumes in each were chosen randomly.

total size of all image regions by choosing different quality values ( $q_0, \dots, q_n$ ) for different resolutions:

$$(q_0, \dots, q_n) = \underset{P_{overall} \geq P^*}{\operatorname{argmin}} S_{total} \quad (8)$$

Solutions to these problems are complicated by the fact that the choice of the quality at the level not only directly influences the detection score  $D(r_l, q_l)$ , but also the detection score  $D(r_{l+1}, q_{l+1})$  and the selection of  $q_{l+1}$  at the next level,  $l + 1$ .

In this paper, we adopt a simpler approach by setting the quality to the same pSNR value at each resolution. We study the effect of this selection on the average detection error and on the total size  $S_{total}$  of all image regions for transmission.

## 4. Results

Our experiments start by showing the advantage of training on lossy and losslessly

compressed images. We will then show that hierarchical learning improves the robustness of the algorithm and loosens the bandwidth requirements. After this, we will present experiments comparing the full detection pipeline evaluating the hierarchical learning on compressed and uncompressed images. Finally, we will evaluate the bandwidth requirements in multiple landmark detection.

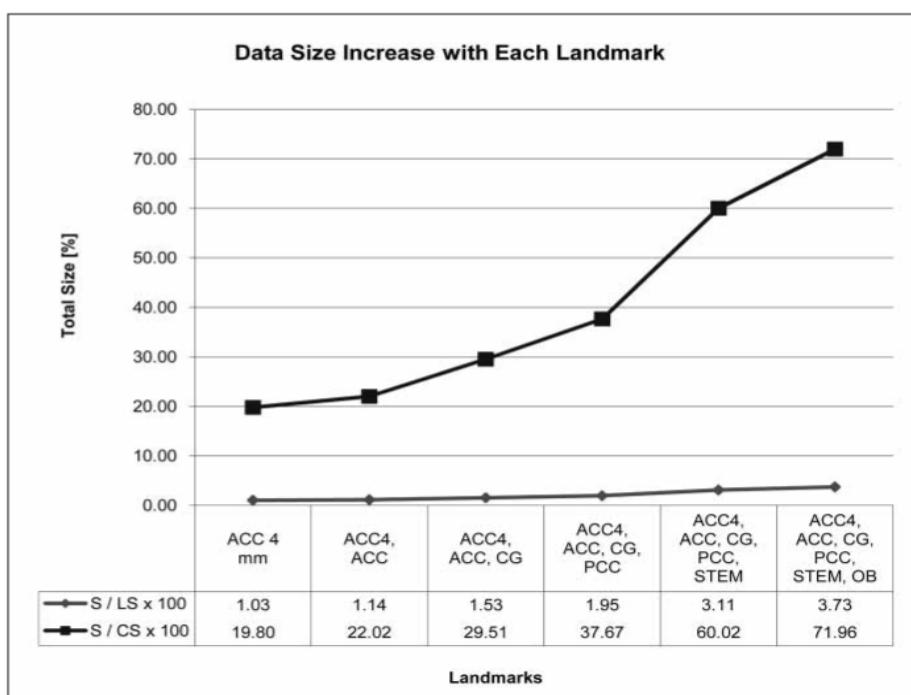
### 4.1 Data Sets

Our experiments are on two datasets. The first dataset has 247 CT volumes (161 for training and 86 for testing) with average size  $97 \times 80 \times 165$  voxels after resampling to 4 mm isotropic resolution. The landmark of interest is the right hip bone landmark (► Fig. 7). The second dataset has 511 MRI volumes (384 for training and 127 for testing) with average size  $130 \times 130 \times 101$  voxels after resampling to 2 mm isotropic resolution. In each volume, we detect the crista galli (CG) landmark of the brain (► Fig. 11). In the multi-landmark detection experiment, we also detect the tip of the occipital bone (OB), the anterior of the corpus callosum (ACC), the posterior of the corpus callosum (PCC), and a landmark in the brain stem (STEM). The training and testing datasets are disjoint and the volumes in each were chosen randomly.

### 4.2 Progressive Data Transmission in Hierarchical Detection

In the first experiment, we test the detection on images compressed at different pSNR levels; see ► Figure 6 for examples. The detection error statistics were computed for images of different pSNR levels with classifiers trained on a) images with the same pSNR level, and b) uncompressed images.

Our second experiment demonstrates the robustness of the hierarchical detection. A single level classifier trained on CT images with 8 mm resolution is compared to the hierarchical classifier (Section 3.1 and ► Fig. 3) trained on images with 16 mm and 8 mm resolution. This experiment is repeated for MR images with 2 mm



**Fig. 9** The average volume size for hierarchical multi-landmark detection relative to average size of whole lossless-compressed volumes (LS) and volumes of pSNR 70 (CS). When adding each landmark the average volume size increases.

resolution and a 4 mm-2 mm hierarchy. In ►Figure 5, the median of the 80% smallest errors<sup>a</sup> are plotted against the average volume size computed for each pSNR level.

The next experiment compares the overall hierarchical detection using uncompressed images and images compressed at pSNR 70. The hierarchical system is also compared to a simpler algorithm operating

on uncompressed images with a single resolution. The results comparison is in ►Table 1.

#### 4.3 Progressive Data Transmission in Multiple Landmark Detection

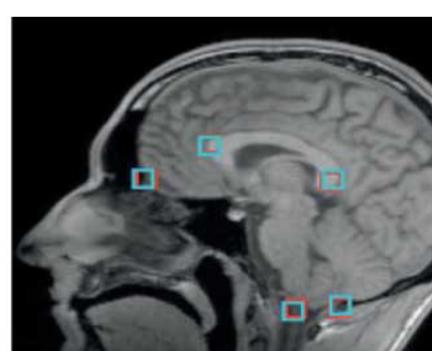
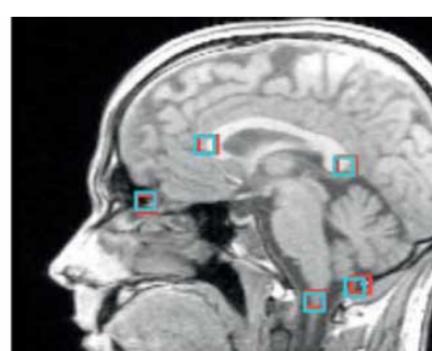
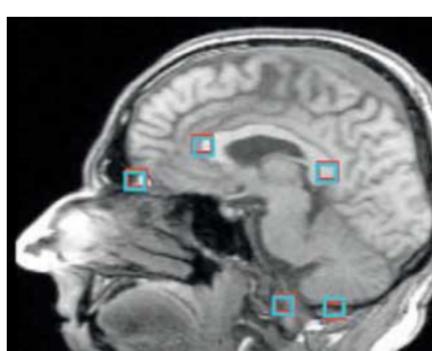
The final experiment analyzes bandwidth requirements when detecting multiple landmarks. Clearly, with more landmarks detected in a single volume the data size

that needs to be transmitted is increased. We use the brain MRI dataset (pSNR 70) and five landmarks described above. The landmarks correspond to distinct brain structures that are useful for automatically estimating the mid-sagittal plane [26]. The diagram of the Hierarchical Detection Network (HDN) [21] is in ►Figure 8.

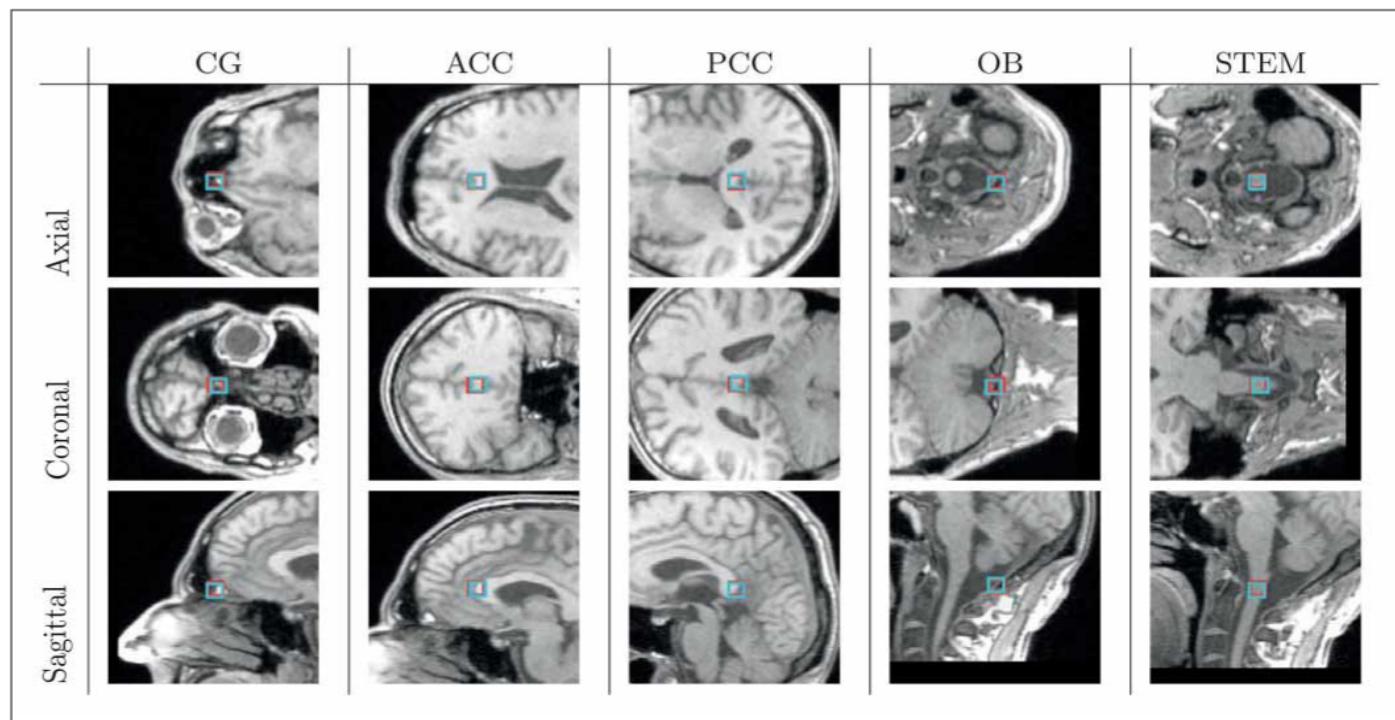
The landmarks are projected onto the mid-sagittal plane for display in ►Figure 10. They are correctly detected despite anatomical and intensity variations. Individual landmarks for the first case from this figure are shown in ►Figure 11.

## 5. Discussion

The proposed progressive data transmission for the hierarchical detection has been analyzed in several experiments as described in the previous section. The plots in ►Figure 4 show that we can obtain better detection performance when training on compressed images thanks to the classifier's ability to adapt to the training data and ignore inconsistencies caused by the compression artifacts (see Section 3). In ►Figure 5, the overall image sizes in the hierarchical detection were compared to a single level algorithm. By training on compressed images we can achieve smaller detection errors for a given average volume size than by training on uncompressed images. The detection errors decrease further when using the hierarchical approach due to the robust search strategy. We found that images at pSNR 70 provide accurate results with significant reduction of overall image size. The results summarized in ►Table 1



**Fig. 10** Example images with ground truth locations (red boxes) and detection results (cyan boxes). Five landmarks are used to estimate the mid-sagittal plane shown for three different cases.



**Fig. 11** Example images with ground truth locations (red boxes) and detection results (cyan boxes) for five brain MRI landmarks of the first case from Figure 10: crista galli (CG), occipital bone (OB), the anterior of the corpus callosum (ACC), the posterior of the corpus callosum (PCC), and the brain stem (STEM). All five landmarks are accurately detected.

**Table 1** The median detection error of the hierarchical detection on images compressed at pSNR 70 (2nd and 5th column), on uncompressed images (3rd and 6th column), and on a single resolution lossless-compressed images (4th and 7th column). The average size of uncompressed volumes is 3188 kB (4 mm CT) and 3334 kB (2 mm MRI). The hierarchical algorithm trained with images of pSNR 70 requires the least amount of data without sacrificing the detection accuracy.

	CT			MRI		
	16–8–4 hier pSNR 70	16–8–4 hier lossless	4 mm lossless	4–2 hier pSNR 70	4–2 hier lossless	2 mm lossless
Error [mm]	3.87	3.54	3.98	2.50	2.37	2.27
Avg. Data Size [kB]	106.22	393.45	1345.96	16.99	168.07	984.76

show that an algorithm trained on images compressed with lossy compression achieves data size reduction by a factor of 3.7 (9.9 for MRI) when compared to a hierarchical training on lossless-compressed images, by a factor of 12.7 (58.0 for MRI) when compared to an algorithm operating on a single resolution, and by a factor of 30.0 (196.2 for MRI) when the original (uncompressed) images are used. The median detection error is comparable for all three cases.

When the number of landmarks increases, the data size that needs to be trans-

mitted also increases. We analyzed the performance of multiple landmark detection in ►Table 2. The table shows that the detection error is low (2.12 mm on average) even though images with pSNR 70 are used in training and detection. The table also shows that the average volume size used in detecting each landmark is 7.34 kB. This is on average 0.75% of the overall size of lossless-compressed volumes (LS) and 14.39% of the overall size of volumes of pSNR 70 (CS). As before, large bandwidth savings can be achieved when detecting individual landmarks. The plot in ►Figure 9 analyzes

the total size when all landmarks are detected together. Relative to CS, the total volume size is 71.96%. Therefore, when detecting many landmarks in the same volume, the benefit of the feedback loop (►Fig. 2) is lower. However, using lossy compression still helps, since the average volume size relative to LS is only 3.73% (reduction of the original size 26.8 times). Relative to the original uncompressed size (3334 kB), this size amounts to 1.10% (reduction 90.8 times).

**Table 2** The detection accuracy of five brain MRI landmarks (2nd column), average volume size (S) in detection (3rd column), S relative to average size LS = 984.76 kB of lossless-compressed volumes (4th column), and S relative to average size CS = 51 kB of volumes of pSNR 70 (5th column). The detection errors are low despite using images with pSNR 70. The average volume size S is small.

Landmark	Err [mm]	Size (S) [kB]	S/LS×100 [%]	S/CS×100 [%]
ACC 4 mm	2.74	10.10	1.03	19.80
ACC	1.83	1.13	0.11	2.22
CG	2.26	3.82	0.39	7.49
PCC	1.81	4.16	0.42	8.16
STEM	2.12	11.40	1.16	22.35
OB	2.57	6.09	0.62	11.94
Average	2.12	7.34	0.75	14.39

## 6. Conclusions

Cloud computing is creating transformational shift in the health care industry. Easy sharing of information within a hospital, clinic, community, or region is becoming possible through the cloud-based Health Information Exchanges (HIEs). This allows multiple health systems to share a single PACS server where all the medical images are stored. Rural areas can benefit by connecting to the HIEs through the advancement of telecommunications-based health solutions and the broadband infrastructure available through telehealth networks [27, 28]. Data transfer bottlenecks remain one of the challenges of data-intensive applications spread across the boundaries of the cloud. In order to minimize cost, it is necessary to consider implications of traffic at every level of the system [29]. This paper addressed the problem of large data transfer sizes in the application of anatomical landmark detection. Such application is useful for visualizing and analyzing image regions surrounding these landmarks.

We presented *Detection in a Cloud (DiC)* system for anatomical landmark detection in the cloud computing environment. At the core of the system is a hierarchical learning algorithm that propagates position candidate hypotheses across a hierarchy of classifiers during training and detection. The algorithm only requires image regions surrounding the candidates which results in less bandwidth for remote data access. Further bandwidth savings (without sacrificing the detection accuracy) are

achieved by compressing the images regions with lossy JPEG 2000. The total bandwidth savings for retrieving remotely stored data amount to 30.0 times (CT data) and 196.2 times (MRI data) reduction when compared to the original data size and 12.7 times (CT) and 58.0 times (MRI) when compared to data size after lossless compression. We showed that during detection of five MRI landmarks total bandwidth amount to 90.8 times reduction compared to the original volume size and 26.8 times reduction compared to data size after lossless compression.

The proposed approach makes it possible to shift the integration, maintenance, and detection software updates from the client to the CAD server. Therefore, when the classifiers are updated, they are immediately available to all clients. In the clinical environment, detected anatomical parts can be reviewed on the client devices remotely. The current system opens many exciting future research directions both on the algorithmic side as well as on the systems side. We are interested the most in building more complicated models with several landmarks of interest trained for different modalities. Such large scale systems will require coordination of multiple CAD servers possibly distributed in a wide-area network.

There are challenges remaining to be addressed. First, quality level in the DiC system has been chosen manually based on experimental evidence. For more complicated systems, it will be necessary to select the quality automatically at all levels and

thus minimize the total bandwidth for desired detection accuracy. In other systems, it will be useful to automatically select the quality levels such that detection accuracy is maximized given a fixed bandwidth. Second, current multi-landmark detection algorithm uses one of previously detected landmarks to predict an approximate location of the next landmark. Robustness and the overall accuracy would be further improved if multiple previously detected landmarks were used for this prediction. Solving these problems is the main focus of our future work.

## References

- Agarwal A, Henehan N, Somashekharappa V, Pandya AS, Kalva H, Furht B. A Cloud Computing Based Patient Centric Medical Information System. In: Furht B, Escalante A, editors. Handbook of Cloud Computing. Springer US; 2010. pp 553–573.
- Haux R. Health information systems – past, present, future. Int J Med Inform 2006; 75 (3–4): 268–281.
- Blobel BGME, Engel K, Pharow P. Semantic Interoperability – HL7 Version 3 Compared to Advanced Architecture Standards. Methods Inf Med 2006; 45 (4): 343–353.
- Faggioni L, Neri E, Castellana C, Caramella D, Bartolozzi C. The future of PACS in healthcare enterprises. European Journal of Radiology 2011; 78 (2): 253–258.
- Ohmann C, Kuchinke W. Future Developments of Medical Informatics from the Viewpoint of Networked Clinical Research. Methods Inf Med 2009; 48 (1): 45–54.
- Estrella F, del Frate C, Odeh THRM, Rogulin D, Amendolia SR, Schottlander D, et al. Resolving Clinicians Queries Across a Grids Infrastructure. Methods Inf Med 2005; 44 (2): 149–153.
- Handels H, Ehrhardt J. Medical Image Computing for Computer-supported Diagnostics and Therapy. Methods Inf Med 2009; 48 (1): 11–17.
- Weitzel M, Smith A, de Deugd S, Yates R. A Web 2.0 Model for Patient-Centered Health Informatics Applications. Computer 2010; 43 (7): 43–50.
- Rosenthal A, Mork P, Li MH, Stanford J, Koester D, Reynolds P. Cloud computing: A new business paradigm for biomedical information sharing. J Biomed Inform 2009; 43: 342–353.
- Glatard T, Pennec X, Montagnat J. Performance evaluation of grid-enabled registration algorithms using bronze-standards. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). Lecture Notes in Computer Science; 2006. pp 152–160.
- Kanade T, Yin Z, Bise R, Huh S, Eom SE, Sandbothe M, et al. Cell Image Analysis: Algorithms, System and Applications. In: IEEE Workshop on Applications of Computer Vision (WACV); 2011.
- Doukas C, Pliakas T, Maglogiannis I. Mobile Healthcare Information Management utilizing

- Cloud Computing and Android OS. In: Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE. Buenos Aires, Argentina; 2010. pp 1037–1040.
13. Tu Z. Probabilistic Boosting-Tree: Learning Discriminative Models for Classification, Recognition, and Clustering. CVPR 2005; 2: 1589–1596.
  14. Viola P, Jones MJ. Rapid object detection using a boosted cascade of simple features. CVPR 2001; 1: 511–518.
  15. Schiele PSMFSRB. Discriminative structure learning of hierarchical representations for object detection. CVPR 2009; 2238–2245.
  16. Samaras WZGZD. Real-time Accurate Object Detection using Multiple Resolutions. ICCV 2007.
  17. Zhu L, Yuille AL. A Hierarchical Compositional System for Rapid Object Detection. NIPS; 2005. pp 1633–1640.
  18. Sudderth EB, Torralba A, Freeman WT, Willsky AS. Describing Visual Scenes Using Transformed Objects and Parts. IJCV 2008; 77 (1–3): 291–330.
  19. Salembier VVFMP. Binary Partition Trees for Object Detection. TIP 2008; 17 (11): 2201–2216.
  20. Butko NJ, Movellan JR. Optimal scanning for faster object detection. CVPR 2009. pp 2751–2758.
  21. Sofka M, Zhang J, Zhou SK, Comaniciu D. Multiple Object Detection by Sequential Monte Carlo and Hierarchical Detection Network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, CA; 2010. .
  22. Schelkens P, Skodras A, Ebrahimi T. JPEG 2000 Suite. John Wiley and Sons; 2009.
  23. Ringl H, Schernthaner R, Sala E, El-Rabadi K, Weber M, Schima W, et al. Lossy 3D JPEG2000 Compression of Abdominal CT Images in Patients with Acute Abdominal Complaints: Effect of Compression Ratio on Diagnostic Confidence and Accuracy. Radiology 2008; 248: 476–484.
  24. Dufaux F, Ebrahimi T. Scrambling for Video Surveillance with Privacy. In: CVPR Workshop; 2006.
  25. Fleck S, Busch F, Biber P, Strasser W. 3D Surveillance A Distributed Network of Smart Cameras for Real-Time. In: CVPR Workshop; 2006.
  26. Schwing A, Zheng Y, Harder M, Comaniciu D. Method and System for Anatomic Landmark Detection Using Constrained Marginal Space Learning and Geometric Inference; 2009. US patent filed. Application number: 12/604,495, Publication number: US 2010/0119137 A1.
  27. Foran DJ, Meer PP, Papathomas T, Marsic I. Compression guidelines for diagnostic telepathology. Information Technology in Biomedicine, IEEE Transactions on. 1997; 1 (1): 55–60.
  28. Thielst CB. At the crossroads: NRTRC white paper examines trends driving the convergence of tele-health, EHRs and HIE. World Hosp Health Serv 2010; 46 (4): 17–23.
  29. Armbrust M, Fox A, Griffith R, Joseph AD, Katz R, Konwinski A, et al. A view of cloud computing. Commun ACM 2010; 53: 50–58.