Published in final edited form as:

Magn Reson Imaging. 2012 July ; 30(6): 824–836. doi:10.1016/j.mri.2012.02.024.

An approach for computer-aided detection of brain metastases in post-Gd T1-W MRI☆

Reza Farjam^{a,b}, Hemant A. Parmar^c, Douglas C. Noll^a, Christina I. Tsien^b, and Yue Cao^{b,c,*}

^aDepartment of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109-2099, USA

^bDepartment of Radiation Oncology, University of Michigan, Ann Arbor, MI 48103-4943, USA

^cDepartment of Radiology, University of Michigan, 1500 E. Medical Center Drive, Ann Arbor, Michigan 48109-5842, USA

Abstract

Purpose—To develop an approach for computer-aided detection (CAD) of small brain metastases in post-Gd T1-weighted magnetic resonance imaging (MRI).

Method—A set of unevenly spaced 3D spherical shell templates was optimized to localize brain metastatic lesions by cross-correlation analysis with MRI. Theoretical and simulation analyses of effects of lesion size and shape heterogeneity were performed to optimize the number and size of the templates and the cross-correlation thresholds. Also, effects of image factors of noise and intensity variation on the performance of the CAD system were investigated. A nodule enhancement strategy to improve sensitivity of the system and a set of criteria based upon the size, shape and brightness of lesions were used to reduce false positives. An optimal set of parameters from the FROC curves was selected from a training dataset, and then the system was evaluated on a testing dataset including 186 lesions from 2753 MRI slices. Reading results from two radiologists are also included.

Results—Overall, a 93.5% sensitivity with 0.024 of intra-cranial false positive rate (IC-FPR) was achieved in the testing dataset. Our investigation indicated that nodule enhancement was very effective in improving both sensitivity and specificity. The size and shape criteria reduced the IC-FPR from 0.075 to 0.021, and the brightness criterion decreases the extra-cranial FPR from 0.477 to 0.083 in the training dataset. Readings from the two radiologists had sensitivities of 60% and 67% in the training dataset and 70% and 80% in the testing dataset for the metastatic lesions <5 mm in diameter.

Conclusion—Our proposed CAD system has high sensitivity and fairly low FPR for detection of the small brain metastatic lesions in MRI compared to the previous work and readings of neuroradiologists. The potential of this method for assisting clinical decision- making warrants further evaluation and improvements.

Keywords

Brain Metastasis; Computer-Aided Detection; Template Matching; MRI

☆This work is in part presented in annual meeting of the Radiological Society of North America (RSNA 2010).

© 2012 Elsevier Inc. All rights reserved.

*Corresponding author. Department of Radiation Oncology, University of Michigan, Ann Arbor, MI 48103, USA. Tel.: +1 734 647 2914; fax: +1 734 936 7859. yuecao@umich.edu (Y. Cao).

1. Introduction

Brain metastasis is the most prevalent form of cancer in the central nervous system, [1] 60% of which are contributed from lung and breast cancers as primary sites [2]. The number and the size of brain metastases are crucial to determine the suitable treatment for patients [3–5]. Whole brain radiotherapy (WBRT) [3] and stereotactic radio-surgery (SRS) [4] are routine treatments for brain metastases and prolong patient survival. Selection of WBRT versus SRS highly depends on the number of lesions and their size; usually less than 3 lesions with the sizes smaller than 3 cm in diameter may lead to selection of SRS [5]. Failure to accurately diagnosis of the presence of small lesions can result in inappropriate selection of treatment. At present, radiologists diagnose brain metastases using contrastenhanced MRI [6] on which the lesions are highlighted due to disrupted blood-tumor barriers. However, manually screening MR images for diagnosis of brain metastases is prone to potential errors [7] especially for small lesions, given the limitation of 2D screening of a large number of brain MRI slices. Hence, computer-aided automatic screening of MR images is of interest.

Several computer-aided detection (CAD) systems using template matching [8] have been proposed for the detection of lesions, especially in the lung [9–14] and breast [15,16]. In template matching, a set of templates with different resolutions is usually cross-correlated with regions of interest to localize objects of an unknown size. The most crucial task in template design is to optimally determine the shape, size, number of templates and a cross-correlation coefficient threshold in order to achieve high sensitivity and specificity of lesion detection as well as computation speed. Currently, most CAD systems using template matching have been developed for computed tomography (CT) or radiography, in which nodules are usually surrounded by a fairly homogeneous background. One study adopted a CAD system originally developed for the detection of lung nodules on CT to detect brain metastases from contrastenhanced MRI [17]. This work demonstrated the suitability of using spherical templates to detect brain metastatic lesions, which are often observed as bright spherical-like solid masses on contrast-enhanced MRI. However, as a lesion grows, a necrotic region may develop in the center of the mass. This previous work did not take into account the pattern of necrotic regions in their template design, which affected the performance of the CAD system. Most importantly, this previous work did not optimize the template set in design with respect to the size and shape of the lesions, and noise and intensity variation in contrastenhanced MRI, which resulted in a large number of false positives. Also, one important clinical issue for the diagnosis of brain metastases is that small lesions have much higher frequencies of being missed compared to large ones, for which a CAD system could potentially assist a radiologist in diagnosis. All these require a more comprehensive study in the design of CAD systems for brain metastatic lesions.

In this study, we aimed to develop a CAD system for the detection of small brain metastases from contrast-enhanced MRI. We designed spherical shell templates to simulate a tumor growth pattern. We performed theoretical and simulation analyses of effects of size and shape heterogeneity of lesions to optimize the number and size of the templates and the cross-correlation thresholds. Furthermore, we investigated how image factors, e.g., nodule-contrast to intensity variations in the nodule and background; affect the performance of the CAD system. We proposed a nodule enhancement strategy and a set of rule-based criteria to reduce false positives. To develop and evaluate the proposed system, we used nine (including 60 lesions) and 20 patients (186 lesions) in the training and testing steps, respectively. We also include screening results from two neuroradiologists in our study. In this study, we describe the scheme of the proposed CAD system in Section 2, and present the experimental results in Section 3. The conclusion and discussions are finally given in Section 4.

2. Materials and methods

2.1. Template design and selection

2.1.1. Template design—Brain metastases, when small, are seen as sphere-like solid masses on contrast-enhanced MRI. As a lesion grows, a necrotic region starts developing at the center of the lesion. Hence, we designed our templates to mimic a brain tumor growth pattern simulated in a model [18] in which the diameter of the necrotic region, r_n , is non-linearly related to the diameter of the lesion as: $r_n = R - 0.53R^{\frac{2}{3}}$ (Fig. 1A). We create a 3D spherical shell zero-mean mask as following (Fig. 1B):

$$M_r = \begin{cases} 1 & r_n^2 \leq x^2 + y^2 + z^2 \leq R^2 \\ 0 & x^2 + y^2 + z^2 < r_n^2 \text{ or } |x|, |y|, |z| < |a| \text{ and } a = 2R \\ -\beta & x^2 + y^2 + z^2 > R^2 \end{cases} \quad (1)$$

where M_r denotes a mask at resolution r , and β normalizes the mask to have a zero mean. Since the necrotic region is not seen in all lesions, we set intensities in the necrotic zone to be zero, and thereby this area is excluded from analysis. For a given M_r , the mask volume, V_{Mr} is defined by the number of non-negative voxels (including zeros).

2.1.2. Selection of template size & cross-correlation threshold—To localize a lesion by M_r , a cross-correlation coefficient (CCC) map between M_r and signal intensities of MR volume (V_{MR}) is calculated:

$$Corr_{Mr}(V_{MR}, i, j, k) = \frac{\sum_{x=-a}^a \sum_{y=-a}^a \sum_{z=-a}^a [M_r(x, y, z)] \cdot [w_{ijk}(x, y, z) - \bar{w}_{ijk}]}{\sigma_{Mr} \cdot \sigma_w} \quad (2)$$

$$w_{ijk}(x, y, z) = \begin{cases} -a+i < x < a+i \\ -a+j < y < a+j \\ -a+k < z < a+k \end{cases} V_{MR}(x, y, z)$$

where (i, j, k) represent voxel indices, $w_{ijk}(\cdot)$ denotes intensity in a cubic region of V_{MR} with a size of M_r centered at (i, j, k) , $\bar{w}_{ijk}$ is the mean of w_{ijk} , $Corr_{Mr}(V_{MR}, i, j, k)$ is the CCC between M_r and $w_{ijk}(\cdot)$, and σ_{Mr} and σ_w denote standard deviations of M_r and w , respectively. To localize a lesion, the cross correlation map is thresholded, all connected voxels above the threshold are grouped, and each group of voxels represents one candidate lesion.

To localize all lesions, however, a set of multi-resolution masks had to be chosen. To select an optimal set of masks, $\{M_{ri}\}$, and a cross-correlation coefficient threshold, three main factors need to be considered: 1) tumor size variation, and 2) lesion shape variation, and 3) intensity variations within a tumors and its surrounding background. The effects of these factors on selecting an appropriate set of masks are quantitatively investigated and described in the following sub-sections. We designed $\{M_{ri}\}$ and τ in a continuous space first, and then extended the concepts to a digital case.

a. Design of $\{M_{ri}\}$ and τ in a Continuous Space: To design a set of masks $\{M_{ri}\}$ and a CCC threshold, τ , we first considered lesion size variations, but disregarded tumor shape and intensity variations for the time being. Let us consider a set of spherical tumors, $\{T_s, [i, i+1]\}$, with their volumes between two consecutive masks of M_{ri} and M_{ri+1} (Fig. 2) as:

$$V_{M_{r_i}} \leq V_{T_1} < \dots < V_{T_j} < \dots < V_{T_n} \leq V_{M_{r_{i+1}}} \quad (3)$$

where V_{T_j} and $V_{M_{r_i}}$ are the volumes of the j th tumor and the i th mask, respectively. The CCC between T_j and M_{r_i} or $M_{r_{i+1}}$ is determined by the number of mismatched voxels between the tumor and mask (we don't show the proof here). Fig. 3 shows that the CCCs between $\{T_s\}$ and M_{r_i} (or $M_{r_{i+1}}$) form a right-side (or left-side) curve from the i th (or $i+1$ th) mask. Also, the CCCs have the following orders:

$$Corr_{M_{r_i}, Right}(T_j) > Corr_{M_{r_i}, Right}(T_{j+1}) \quad (4)$$

$$Corr_{M_{r_{i+1}}, Left}(T_j) < Corr_{M_{r_{i+1}}, Left}(T_{j+1}) \quad (5)$$

If a threshold, τ , of the CCC is selected at the intercept of the two CCC curves as:

$$\tau = Corr_{M_{r_i}, Right}(T_k) = Corr_{M_{r_{i+1}}, Left}(T_k) \quad (6)$$

All tumors $\{T_s, [i+1]\}$ that have the volumes between M_{r_i} and $M_{r_{i+1}}$ are detected by either M_{r_i} or $M_{r_{i+1}}$, depending upon the volume of the tumor smaller or greater than T_k 's (Fig. 3). Using the same strategy, the threshold for the $i+2$ th mask should be selected as:

$$Corr_{M_{r_{i+1}}, Left}(T_k) = Corr_{M_{r_{i+1}}, Right}(T'_n) = Corr_{M_{r_{i+2}}, Left}(T'_n) = \dots \quad (7)$$

where $M_{r_{i+1}}$ and $M_{r_{i+2}}$ localize the next set of tumors $\{T'_s, [i+1, i+2]\}$ (Fig. 3). Note that a lower threshold value of τ can reduce the number of required masks by increasing the radius increment ($d_{i,i+1} = r_{i+1} - r_i$) between the two adjacent masks of M_{r_i} and $M_{r_{i+1}}$.

So far, we have not considered shape heterogeneity in our analysis. In fact, the CCC is only sensitive to the number of mismatched voxels between the mask and the lesion. This means that the shape heterogeneity of a tumor with respect to a sphere can be quantified by the number of mismatched voxels, and therefore, can be modeled similar to size heterogeneity. A tumor with its shape deviating from a sphere can be localized by the masks designed by Eqs. (6) and (7) if the CCC between the tumor and the mask is still above the threshold. For example, in a simple case, two tumors have the same volume and are located between a pair of adjacent masks, but one is a perfect sphere and another one is not. Since the two tumors have the same CCC, they will have the same probability for localization. This puts a limit on shape deviation of a tumor from a sphere. Since the spherical tumor T_k (at the intercept of the two CCC curves in Fig. 3) has the smallest CCC with either M_{r_i} or $M_{r_{i+1}}$, the number of mismatched voxels of a non-spherical tumor with one of the masks has to be equal to or smaller than the T_k 's in order to be detected by the mask. Thus, $d_{i,i+1}$, the radius difference between the adjacent spherical shell masks limits tolerance of shape variation of a tumor from a sphere. If $d_{i,i+1}$ is set large enough, a set of masks can detect tumors with large shape heterogeneity and result in a high sensitivity for tumor localization. We designed our templates to be able to detect a non-spherical tumor that has the number of mismatched voxels with the mask which is not greater than one of a sphere with the volume twice of the smallest mask. This is also the reason why we set $a=2R$ in Eq. (1) (Fig. 1B). Note that shape deviation corresponding to the inter-shell of the mask can also be considered in a similar manner.

In summary, if $\{T_s\}$ is a set of tumors with various sizes and shapes, to localize $\{T_s\}$, the first mask is chosen as small as the smallest tumor. The second mask is determined by having the volume of the spherical tumor (T_k) at the intercept of the CCC curves of the two

masks, twice as the first mask. The intercept of the two CCC curves also determines the threshold, τ . The remaining masks will be selected based on Eq. (7).

b. Design of $\{M_{ri}\}$ and τ in a Digital Space: In order to find appropriate $\{M_{ri}\}$ and τ in the discrete form, a digital phantom comprised of different spheres with radii ranging from 1 to 6.2mm and a step size of 0.1 is implemented (It will be discussed later in this paper that we are only interested in localizing lesions with a diameter of <5 mm, and hence large spheres are not included in the phantom). We then follow the procedure described previously to find out which digital masks meet the conditions given in Eqs. (3–7). The smallest feasible digital mask M_1 is $r_1=1$. The radius of T_k is 1.26. To identify M_2 , we first compute the CCCs of the digital phantom with M_1 and with all candidates (digitally resolvable) of M_{r2} ($\forall r_2>1$). Next, we approximate the CCCs ($Corr_{r1, Right}(\cdot)$ and $Corr_{r2, Left}(\cdot)$) by an exponential function to find r_2 by having the radius of T_k close to 1.26 (Fig. 4). As a result, we identify $r_2=1.5$ with $\tau=0.8$ and $r_k=1.25$ (radius of T_k). This procedure is repeated to find the remaining masks as $r=\{1, 1.5, 2.3, 3.2, 4.2, \text{and} \dots\}$. Due to digitization, the CCC thresholds for all masks vary from 0.8, but they are chosen slightly below 0.8 to retain maximum sensitivity, e.g., 0.74 for $M_{1.5}$, $M_{2.3}$ and $M_{3.2}$.

c. Design of $\{M_{ri}\}$ and τ with Intensity Variation: To consider effects of intensity variation and noise within a tumor and its surrounding background on the crosscorrelation coefficients, we define a contrast to intensity variation ratio (CIVR) as following:

$$CIVR = \frac{|\mu_t - \mu_b|}{\sqrt{\frac{1}{(N_B + N_T)} [N_T \sigma_t^2 + N_B \cdot \sigma_b^2]}} \quad (8)$$

where N is the number of voxels, μ and σ are the mean and the standard deviation of signal intensities, respectively, and subscripts t and b denote tumor and background, respectively. For a zero-mean mask given in Eq. (1), we have:

$$\frac{N_T - \beta \cdot N_B}{N_T + N_B} = 0 \Rightarrow \beta = \frac{N_T}{N_B} \bar{w} = \frac{N_T \mu_t + N_B \mu_b}{N_T + N_B} \quad (9)$$

Using Eqs. (8) and (9), we rewrite Eq. (2) as:

$$Corr_{M_r}(V_{MR}) = \frac{Corr_{M_r, V_{MR}}(WOIV)}{\sqrt{1 + \frac{(N_B + N_T)^2}{N_B N_T} \times \frac{1}{CIVR^2}}} Corr_{M_r, V_{MR}}(WOIV) = \frac{1}{\sigma_M} \sqrt{\frac{N_T}{N_B}} \quad (10)$$

where σ_m is the standard deviation of the mask, and $Corr_{M_r, V_{MR}}(WOIV)$ denotes the CCC between M_r and V_{MR} without intensity variations within the tumor and in the background ($CIVR=\infty$). Eq. (10) shows that the CCC decreases with CIVR from the value without any intensity variation. We further investigated the effect of intensity variation on choosing the threshold of CCC by simulation, in which we add Gaussian noise with different standard deviations (δ) to the digital phantom described in Section 2.1.2.b, and recalculate $Corr_{ri, Right}(\cdot)$ and $Corr_{ri, Left}(\cdot)$ and the thresholds for the masks. As expected, the CCC thresholds depend upon δ (or CIVR), which will be used to train the system.

2.1.3. Nodule localization—To localize tumors, after computing the CCCs of MRI volumes with a mask M_{ri} at resolution r_i , we identify all voxels above a threshold. Then, we group all connected (26-adjacency) voxels by growing regions [19], and pick the voxel with a maximum CCC from each connected group as the center of the candidate lesion at the r_i th resolution, $C_{ri,k}$.

To refine the definition of the candidate lesion, the k -Means method [19] is applied to classify the voxels within a cubic volume with a size of $2a$ (same size of the mask) and a center at $C_{ri,k}$ (the center of the candidate lesion) of MRI to bright and dark voxels. The candidate lesion is defined by all bright voxels connected to the center. The bright voxels disconnected from the candidate lesion center are discarded.

2.2. False positives reduction

As shown in Section 2.1.2.c, noise and intensity variations in both tumor and background affect the CCC values, and thereby the performance of a CAD system (causing a large number of false positives). In the following sections, we developed strategies to improve the CCCs and reduce the number of false positives.

2.2.1. Nodule enhancement—The intensity variations of a nodule and its surroundings are comprised of three main components: (1) image noise, (2) intra-tissue inhomogeneity and (3) intensity variation of different tissues in the background. For a small lesion, image noise and intra-tissue inhomogeneity are the dominant components. However, as the mask size increases, tissue heterogeneity in the background surrounding the tumor becomes a major factor. In this work, we focused our attention particularly on small lesions and, hence, propose a method to enhance the CIVR of the small lesions in order to employ a higher CCC threshold and reduce false positives.

To increase the contrast of a lesion, histogram equalization [20] Eq. (11), is used to transform bright voxels of MR images to the upper extreme of a new gray scale as I_{he} :

$$I_{he}(k) = \sum_{i=0}^k P(i), k=0, \frac{L}{M}, \frac{2L}{M}, \dots, L \quad (11)$$

where $P(i)$ is the occurrence frequency of the i th intensity scale in the original image, L is the number of gray scales in the original volume, and M is a quantization factor and determines the number of gray levels in I_{he} . Then, I_{he} is normalized to have a value of I'_{he} between 0 and Max (I_{he}):

$$I'_{he} = \frac{I_{he} - \text{Min}(I_{he})}{\text{Max}(I_{he}) - \text{Min}(I_{he})} \cdot \text{Max}(I_{he}) \quad (12)$$

Assuming tumors are in the upper portion of I'_{he} , power transformation is employed to further enhance the nodule and suppress the dark background:

$$I_{he,pt} = I'^{\lambda}_{he} \quad (13)$$

where $I_{he,pt}$ represents the gray scale after applying histogram equalization followed by power transformation, and λ is an exponent. Fig. 5 illustrates that a large λ enhances the contrast of a nodule in the upper scale of I'_{he} by increasing the intensities in the higher gray scales and decreasing those in the lower ones. In addition, histogram quantization [Eq. (11)] decreases the intensity inhomogeneity within a nodule. Combining the two steps, the ratio of the contrast to intensity variation of the nodule is enhanced. Fig. 5 shows that the intensities in the lower half of the gray scale are suppressed almost entirely by $\lambda < 6$. However, a large λ also increases intensity variations within a nodule, and thereby decreases the CIVR. Hence, λ will be optimized in the training step to balance these two effects.

2.2.2. Size criterion—While we develop strategies to increase the sensitivity of the CAD system, we also need to implement methods to reduce the number of false positives. Brain

metastatic tumors are usually isolated lesions within the brain, while other detected nodules are typically part of an extended structure, e.g., contrast-enhanced blood vessels detected by low resolution masks of $r_i=1$ and 1.5, where the size of the cross-section of the small blood vessel is comparable to the mask. Hence, we further grow the initially detected segment by connecting similar voxels in the volume of $(25*2r_i)^3$ beyond the initial VOI by using descriptive statistics (maximum and minimum) of the initial connected voxels. The final size of the segment can differentiate an isolated lesion from an extended and non-tumor like structure. Hence we reject any nodule with a volume eight times ($R_n=2R$) greater than the volume of the applied mask.

2.2.3. Sphere index—To further reduce the number of false positives, e.g., a segment of blood vessel, we compute a sphere index of the detected nodule as the second false positive rejection criterion, which is based upon the fact that metastatic lesions are usually spherical-like instead of cylindrical-like. To do so, we compute eigenvalues (e_1 , e_2 , and e_3) of a covariance matrix of a matrix whose columns are comprised of x , y and z coordinates of the nodule voxels. For each nodule, the eigenvalues were normalized to $\max(e_1, e_2, e_3)=1$. The sphere index, SI , is defined as the geometrical mean of the eigenvalues as following:

$$SI = \sqrt[3]{e_1 e_2 e_3} \quad (14)$$

For a spherical object, $e_1=e_2=e_3=1$ and, hence, $SI=1$. As an object deviates from a sphere, SI decreases from 1. To find an appropriate threshold for SI , we use the descriptive statistics of the lesions in the training step, and define the threshold as: $\tau_{SI}=\mu_{SI}-2.5\sigma_{SI}$ where μ_{SI} and σ_{SI} are the mean and the standard deviation of all the metastatic lesions, respectively.

2.2.4. Brightness criterion—As mentioned in Section 2.2.1, tumors are transferred to the upper extreme of I'_{he} by the histogram equalization. This enables us to further reject the false positives using a brightness index. We exclude any detected nodule with a brightness index $\left(\frac{\mu_k}{\max(I'_{he})}\right)$ less than 0.5, where μ_k denotes the mean intensity of the nodule in I'_{he} .

2.3. Methods of training and evaluation

2.3.1. Human subjects—Thirty one patients diagnosed with brain metastases were enrolled in a prospective MRI study. For each patient, 3D post-Gd-DTPA volumetric T1-W MRI with a voxel size of $0.9375 \times 0.9375 \times 1$ (mm^3) was acquired prior to radiation therapy using a Philips 3T MR scanner. Of the 31 patients, two were excluded from this study since the MR images were corrupted by motion artifact. From the remaining 29 patients, we defined the first 9 patients as a training set and the remaining 20 patients in the testing set. There were four patients, two in the training dataset and two in the testing data set, who had more than 30 lesions each. As a total, there were 1279 and 2753 slices in the training and testing data sets, respectively.

2.3.2. Training and optimization of the CAD system—To train and optimize our CAD system, we identified all the lesions in the initial training dataset. To do so, we put together the results of the initial neuroradiologist as well as a separate retrospective reading by a second highly trained neuroradiologist (H.P.). We also ran the proposed CAD system in an over-sensitive mode on the training dataset by using a fairly low cross correlation coefficient threshold (≈ 0.5) at each resolution in order to detect as many lesions as possible. The second neuroradiologist had two analyses: First, the pre-RT MR images were analyzed without referring to the additional CAD results or further clinical follow ups and imaging studies; and the second analysis of the MRI images with all available clinical information as well as the CAD results. The first analysis was used to evaluate his performance, but the

second analysis was aimed to detect all the lesions in order to train and further optimize the CAD system. In order to document what size and type of lesions that are likely missed by highly trained neuroradiologists and our CAD system, we divided the lesions into three categories: solid mass without necrosis (regular), lesion with necrotic core (hole), and one located near the edge or the surface of the brain (edge). Table 1 summarizes the number of the lesions found in each of the categories in the training data set.

To obtain the statistics of the lesions likely missed by the neuroradiologists, we put together the screening results of the initial diagnosis by the first neuroradiologist as well as the first independent reading of the second neuroradiologist. Since the precise location of each of lesions for the two patients with 30 lesions or more were not available for the initial diagnosis, we excluded the two patients from this analysis. In the initial diagnosis, when a patient has a large number of lesions, the locations of the lesions become irrelevant for treatment, and are usually not reported. Therefore, we cannot use the initial diagnosis of the two patients to find the sizes of the lesions likely missed in clinical screening, and excluded the two patients from the analysis.

To optimize the proposed CAD system, we ran the system on the training dataset at resolutions of $r=1, 1.5, 2.3$, and 3.2 with combinations of all thresholds/parameters (including varying λ from 6 to 9 and τ from 0.5 to 0.75). Then, we generated free-response receiver operating characteristic (FROC) [21] curves to determine the optimal set of parameters for the CAD system, in which a false positive rate was calculated as a ratio of the total of false positives to the number of slices.

To study the effect of each of the proposed false positive rejection criteria, we compared the FROC curves resulted without applying any false positive rejection criteria with the ones by sequentially adding the criteria: (1) the nodule enhancement strategy, (2) the size criterion, (3) the sphere criterion, and (4) the brightness criterion. In this study, we primarily focus on rejecting intra-cranial false positives, since the extra-cranial false positives can be rejected by other methods, e.g., a skull stripping technique [22], which can be implemented, but is currently beyond the scope of this study.

2.3.3. Testing of the CAD system—To evaluate the proposed system, we ran the CAD system on the testing dataset by using the optimal set of the parameters determined from the training dataset. Similar to what was done in the training step, the second neuroradiologist read the MR images twice: once without and then with all available reference materials. Finally, we put together the results of the initial diagnosis and the first independent reading of the second radiologist to obtain the statistics of the lesions likely missed by the neuroradiologists. The results of the second reading of the second neuroradiologist were used as a reference to evaluate the performance of the CAD system. Table 2 summarizes the categories of the lesions in the testing data set.

3. Experimental results

3.1. Training and optimization of the CAD system

Size of lesions likely missed by Neuroradiologists—Table 3 shows the screening results of the initial diagnosis and the first reading of the second radiologist on the training data. Neither of the two neuroradiologists missed any lesions if the diameter of the lesion was > 4 mm. However, the lesion detection rates of the initial diagnosis decreased from 93% to 81% and to 73% for lesion diameters 4 mm, 3 mm, and to 2 mm, respectively. The results of the first independent reading of the second radiologist were similar to the initial diagnosis. If only considering the lesions with diameters less than 5 mm (27 lesions), the two radiologists detected 60% and 66% of the lesions. Based upon these observations, we

trained our proposed CAD system to detect lesions smaller than 5 mm by using the masks with resolutions $r=1, 1.5, 2.3$ and 3.2 .

Optimal Parameters of the CAD system—To optimize the proposed CAD system, the FROC curves of the detected lesions from the training data set were generated by varying the exponent of the power transformation of $\lambda=6, 6.5, 7, 7.5, 8, 8.5$ and 9 and the threshold of CCC from 0.5 to 0.75 (Fig. 6). As seen, for a sensitivity of 96.7% (or 90%), the lowest intra-cranial false positive rate (IC-FPR) is 0.0211 ($27/1279$), or 0.006 ($8/1279$), which is obtained by using $\lambda=8.5$, the CCC threshold corresponding to $CIVR=5.12$ (or 5.85), and the sphere index (SI) 0.24 . Other values of λ produced worse false positive rates when the sensitivity was set greater than 90% . The areas under the FROC curves (AUC) for the FPRs ranged from 0.000 to 0.029 (the maximum shown in Fig. 6) are $0.818, 0.827, 0.852, 0.850, 0.848, 0.884$ and 0.877 for $\lambda=6, 6.5, 7, 7.5, 8, 8.5$ and 9 , respectively. The AUC also indicates that $\lambda=8.5$ performs the best.

Effect of Nodule Enhancement—Next, we compared the FROC curves obtained from the training dataset with and without the nodule enhancement strategy and without applying any false positive rejection criteria. Fig. 7 shows that our nodule enhancement strategy (with the optimal $\lambda=8.5$) improves both sensitivity and specificity significantly. For the IC-FPR of 0.0758 , the sensitivity improved from 81% to 98.3% . However, without using nodule enhancement, the FROC curve reaches a plateau in sensitivity, indicating that sensitivity cannot be improved more than 0.82 even when we increase the false positive rate. For the FPRs ranged from 0.000 to 0.029 , the AUCs are $0.681, 0.784$ and 0.884 for the lesion detections without module enhancement and any false positive rejections, with module enhancement but without false positive rejections, and with all false positive rejection criteria, respectively, in Fig. 7. Fig. 8 shows examples of intensity histograms of a slice of T1-weighted MRI, and intensity variations within and around a lesion with and without applying the nodule enhancement strategy. Note that the histogram equalization transfers gray scales of the nodule to the upper extreme of the image histogram (Figs. 8B and F), and thereby the nodule is bright as well as the vessel and skin are (Figs. 8A and E). The histogram equalization also reduces the noise (standard deviation) within the nodule due to quantization (Figs. 8C and G), and in the surrounding background. The power transformation further increases the contrast of the nodule with respect to the background, and suppresses the intensity variation in the background (Figs. 8L and H). As a result, the $CIVR$ of the nodule is improved from 5.26 (in the original image) to 13.2 (in the $I_{he,pt}$ with $\lambda=6.5$), which enables the selection of a higher threshold of CCC and thereby to have a lower false positive rate. However, a further increase in λ increases false positives and deteriorates the performance of the system. The optimal λ ($=8.5$) has to be determined from the FROC curves as previously shown in Fig. 6.

False Positive Rejection—Table 4 shows the effects of false positive rejection criteria on improvement in specificity of the proposed system. For the 60 lesions with a diameter ≤ 5 mm on 1279 slices in the training dataset, using the nodule enhancement only, intracranial false positives were 97, decreased to 48 after applying the size criterion, to 28 after applying the size and sphere criteria, and to 27 when applying the size, shape and brightness criteria. The size criterion and the sphere index are effective for rejecting both the intra- and extracranial false positives, while brightness criterion is useful for rejecting the extra-cranial false positives. Note that one lesion was rejected mistakenly by the size criterion due to the fact that the lesion was connected to a blood vessel.

3.2. Testing of the CAD system

Size of lesions likely missed by Neuro-Radiologists—The screening results of the testing dataset from the initial diagnosis and the first independent reading of the second radiologist was similar to what was observed in the training dataset: the lesions with a diameter <5 mm had a higher frequency of being missed (Table 5). Again, we excluded the two patients with 30 lesions or more from this analysis due to lack of precise location of each of these lesions. Again if only considering the lesions with diameter less than 5 mm (30 lesions), the two neuroradiologists detected 70% and 80% of the lesions.

Testing of the CAD System—Using the optimal set of parameters obtained in the training step, a sensitivity of 93.5% (174/186) was achieved with an IC-FPR of 0.024 (66/2753) in the testing data (186 lesions with a diameter <5 mm from 20 patients). Fig. 9 shows examples of the lesions (diameters of 2.2, 3.4 and 4.5) missed by the radiologists but detected by the proposed CAD system.

Sources of False Negatives—The lesions missed by our proposed system can be divided into three main categories: (1) lesions with low contrast to intensity variation, (2) lesions attached to a structure and rejected after applying the size criterion, and (3) lesions near the edge of the brain surface. Fig. 10 shows examples from each category. Table 6 shows the number of lesions missed in each category. Of the 186 tested lesions, six lesions having a low contrast to intensity variation ratio (3.8%), 5 lesions attached to a blood vessel (3.1%), and 1 near the edge of the brain surface (3.5%) was missed by our CAD system. Also, all of the five missed lesions in the second category were at the midline of the brain (see the middle panel of Fig. 10), where blood vessels have a higher spatial density. Note that no lesion with necrotic core was missed by our proposed approach.

4. Discussion and conclusion

In this study, a computer-aided detection system has been developed and evaluated for automatic detection and discrimination of small brain metastases in post-Gd T1-weighted MRI, which have a higher frequency of being missed in routine radiological diagnosis. In the development of the system, first we designed a set of multi-resolutions of 3D spherical shell masks, which mimic a bright tumor with or without necrotic core and on a dark background, to localize metastatic lesions by cross-correlation with MR volumes. The analytic and simulation investigations were performed to study the effects of size and shape heterogeneity as well as the intensity variation within a tumor and the surrounding background in order to optimize the template design. Secondly, we developed strategies to improve sensitivity and specificity of the CAD system by using the nodule enhancement strategy to increase the contrast to intensity variation ratio of the lesion and the background, and by using a set of rule-based false positive rejection criteria based upon the size, shape and brightness of the metastatic lesions. Overall, we achieved 93.4% sensitivity and 0.024 intra-cranial FPR for detecting the brain metastases with a size less than 5 mm, which is 10 times less than the intracranial FPR of a previous study [17]. This CAD system has the potential to improve lesion detection especially in lesions less than 5 mm and therefore may aid in the physician's decision-making for selecting appropriate treatment strategy.

Selecting an optimal set of mask sizes and the crosscorrelation threshold depend on three major factors: (1) nodule size, (2) nodule shape, and (3) the intensity variation within the tumor and the background. We showed that two consecutive masks can localize all spherical tumors with volumes between the two masks if the cross-correlation coefficient threshold is selected at the intercept of the two cross-correlation curves. We also showed that the shape heterogeneity of the tumors can be modeled similar to size heterogeneity. Hence, a nodule

can be localized if its shape deviation from a sphere does not exceed the maximum tolerable size variation, which is determined by the radius difference between the two consecutive masks. However, if the radius difference is too large, the masks could be oversensitive to the nodules with a large range of shape variations, resulting in an increase in false positives. Furthermore, our analysis (Fig. 4) shows that it is not necessary to evenly space the masks as stated in a previous study [12]. In fact, evenly spaced masks may not increase sensitivity significantly but definitely increase computation time and possibly false positives. In addition, the data in Tables 1 and 2 indicate that the majority of the small lesions have no necrotic core, and thus, a solid spherical template seems to be a choice for detecting the small lesions. While a spherical shell template, excluding the voxels in the center from analysis, can detect both solid and necrotic lesions, a solid template can result in some non-matching voxels in the center of a lesion with necrotic core. When using a solid template for detecting both solid and necrotic lesions, we need to low the threshold to reach the maximum sensitivity, which results in more false positives than using a spherical shell template (data not shown).

In our CAD system, we implemented several strategies to reduce false positives. Brain metastases are often seen as isolated masses, while other detected nodules are typically be part of the other structures, e.g., blood vessels. Segments of small blood vessels have a high probability of being detected by the masks. In the previous study [17], of the 16 lesions having a diameter b5 mm, 14 lesions were detected, but the false-positive rate is as high as 0.22, in which the majority of them are blood vessels. The detected vessel segments usually belong to a large vascular tree. Therefore, we grew the detected segment beyond the initial VOI by connecting more similar voxels. Hypothetically, the extended structure is unlikely to be a tumor, and can be rejected based on size criterion. Using the size criterion and the sphere index, we are able to reject 71% of the intracranial false positives and 72% of the extracranial false positives in the training data set. In the training processing, we learned that there are some lesions, which have a low contrast with respect to its surrounding background and a large intensity variation, and thus have a greater probability of being missed by a CAD system. Our proposed nodule enhancement strategy improves the low nodule contrast and reduces the intensity variation, and enables us to gain both sensitivity and specificity (Fig. 7). Overall, we are able to reduce intracranial FPR to 0.024, which is a major improvement compared to the previous work. Further reduction of the extracranial false positives can be better achieved by using a skull stripping technique [22], which is not the aim of this paper but can be incorporated into a final CAD package.

In this study, we analyzed what types of lesions are missed by our CAD system (Table 6). The main reasons are typically due to the attachment to another structure, especially in the midline of the brain, or have a low contrast to intensity variation ratio (CIVR). In the former case, segmenting vessels from tissues in the midline could be helpful, e.g., using publicly available software of SPM [23]. One easy solution for the latter problem is to decrease the cross-correlation threshold, but this may increase the false positive rate substantially. For example, the lesion shown in Fig. 10A can be detected by decreasing the cross-correlation threshold at the cost of increasing the false positive rate to 0.1 (that is 5 times greater than we have now). Although our proposed nodule enhancement strategy improves both sensitivity and specificity of the proposed CAD system, developing a technique to locally adapt the CIVR could further improve the performance of the CAD system. A strategy to deal with a lesion near a vessel could also improve the CAD system.

Finally, the CAD system described in this paper is especially designed to localize small brain metastatic lesions, given that the small lesions have a higher probability of being missed by the clinicians (Tables 3 and 5). However, this work can be further extended to localize larger lesions as well. The technical issues as well as clinical needs for detection

and segmentation of large lesions, in a certain extent, differ from the small ones, which require different emphases in the design of the CAD system. For large lesions, segmentation is more clinically interesting than detection. Finally, a clinical study evaluating the impact of our CAD system on overall diagnosis is ongoing.

Acknowledgments

This work is in part supported by the NIH grants of RO1 NS064973 and 3PO1 CA59827, and R21 CA113699.

References

1. Lassman AB, DeAngelis LM. Brain metastases. *Neurol Clin N Am*. 2003; 21:1–23.
2. Nguyen T, DeAngelis LM. Treatment of brain metastases. *J Support Oncol*. 2004; 2:405–416. [PubMed: 15524068]
3. Chao JH, Phillips R, Nickson JJ. Roentgen-ray therapy of cerebral metastases. *Cancer*. 1954; 7:682–689. [PubMed: 13172684]
4. Leksell L. The stereotaxic method and radiosurgery of the brain. *Acta Chir Scand*. 1951; 102:316–329. [PubMed: 14914373]
5. Shaw E, Scott C, Souhami L, Dinapoli R, Kline R, Loeffler J, et al. Single dose radio-surgical treatment of recurrent previously irradiated primary brain tumors and brain metastases: final report of RTOG protocol 90–05. *Int J Radiat Oncol Biol Phys*. 2000; 47:291–298. [PubMed: 10802351]
6. Sze G, Milano E, Johnson C, Heier L. Detection of brain metastases: comparison of contrast enhanced MR with unenhanced MR and enhanced CT. *AJNR Am J Neuroradiol*. 1990; 11:785–791. [PubMed: 2114769]
7. Prastawa M, Bullitt E, Moon N, Leemput KV, Gerig G. Automatic brain tumor segmentation by subject specific modification of atlas priors. *Acad Radiol*. 2003; 10:1341–1348. [PubMed: 14697002]
8. Brunelli, R. Template matching techniques in computer vision: theory and practice. John Wiley and Sons; 2009.
9. Lee Y, Hara T, Fujita H, Itoh S, Ishigaki T. Automated detection of pulmonary nodules in helical CT images based on an improved template matching technique. *IEEE Trans Med Imaging*. 2001; 20(7):595–604. [PubMed: 11465466]
10. Li Q, Kutsuragawa S, Doi K. Computer-aided diagnostic scheme for lung nodule detection in digital chest radiographs by use of a multiple template matching technique. *Med Phys*. 2001; 28(10):2070–2076. [PubMed: 11695768]
11. Bae KT, Kim JS, Na YH, Kim KG, Kim JH. Pulmonary nodule: automated detection on CT images with morphologic matching algorithm—preliminary results. *Radiology*. 2005; 236:286–294. [PubMed: 15955862]
12. Wang P, DeNunzio A, Okunieff P, O'Dell WG. Lung metastases detection in CT images using 3D template matching. *Med Phys*. 2007; 34:915–922. [PubMed: 17441237]
13. Osman XO, Ozekes S, Ucan ON. Lung nodule diagnosis using 3D template matching. *Comput Biol Med*. 2007; 37:1167–1172. [PubMed: 17182027]
14. AMoltz JH, Schwier M, Peitgen HO. A general framework for automatic detection of matching lesions in follow-up CT. *ISBI*. 2009:843–846.
15. Tourassi GD, Vargas-Voracek R, Catarious DM, Floyd CE. Computer-assisted detection of mammographic masses: a template matching scheme based on mutual information. *Med Phys*. 2003; 30(8):2123–2130. [PubMed: 12945977]
16. Lochanambal KP, Karnan M, Sivakumar R. Identifying masses in mammograms using template matching. *International Conference on Communication Software and Networks*. 2010:339–342.
17. Ambrosini RD, Wang P, O'Dell WG. Computer-aided detection of metastatic brain tumors using automated three-dimensional template matching. *J Magn Reson Imaging*. 2010; 31:85–93. [PubMed: 20027576]

18. Kansal AR, Torquato S, Harsh GR, Chioca EA, Deisboeck TS. Simulated brain tumor growth dynamics using a three-dimensional cellular automaton. *J Theor Biol.* 2000; 203:367–382. [PubMed: 10736214]
19. Chen, CH.; Pau, LF.; Wang, PSP. *The handbook of pattern recognition and computer vision.* 2nd ed.. Singapore: World Scientific Publishing; 1998.
20. Gonzalez, RC.; Woods, RE. *Digital image processing.* 2nd edition. Prentice-Hall Inc; 2002.
21. Chakraborty DP, Winter LHL. Free-response methodology: alternate analysis and a new observer-performance experiment. *Radiology.* 1990; 174:873–881. [PubMed: 2305073]
22. Alirezaie J, Jernigan ME, Nahmias C. Automatic segmentation of cerebral MR images using Artificial Neural Networks. *IEEE Trans Nucl Sci.* 1998; 45:2174–2182.
23. <http://www.fil.ion.ucl.ac.uk/spm/>.

Fig. 1.

(A) A simulated brain tumor growth pattern with necrotic core (black), non-proliferative zone (crosshatched) and proliferative zone (white). (B) A spherical shell 3D template designed based upon the simulated brain tumor growth pattern. Here, R , r_n , δ_n , and δ_p represent the radius of the lesion, the radius of the necrotic core, and the lengths of the non-proliferative and proliferative zones [18], respectively.

Fig. 2.

Cross sectional view of a set of spherical tumors ($T_1 < T_2 < \dots < T_n$) with radii (crosshatched) between two consecutive masks M_{ri} (light gray sphere) and M_{ri+1} (dark gray sphere). The two consecutive masks, M_{ri} and M_{ri+1} localize all tumors using a cross-correlation threshold determined by Eq. (7).

Fig. 3.

A scheme on the design of the masks, $\{M_{ri}\}$, and the selection of the cross-correlation coefficient (CCC) threshold, τ , based on Eqs. (3–7). Dashed curves represent CCCs between the masks and the spherical tumors with different sizes. A threshold chosen at the intercept of the CCC curves of the 1st and 2nd masks (at τ) is warranted to detect all tumors with the volumes between the two masks.

Fig. 4.

The optimal set of digital masks, $\{M_{r_i}\}$ with $r=\{1, 1.5, 2.3, 3.2, 4.2, \text{ and } 5.4 \dots\}$, and the cross-correlation coefficient thresholds, τ . In the digital space, the thresholds are approximately 0.74. Solid line: regression fitting line; Solid symbols: data points of cross correlation.

Fig. 5.

The application of power transform to I'_{he} with different values of λ . A large λ increases the contrast between high and low intensities. If a nodule is in the upper portion of I'_{he} , the application of power transform enhances the nodule contrast and suppresses its dark surrounding. For $\lambda < 6$, all intensities lower than a half of the maximum are almost entirely suppressed in $I_{he,pt}$. Here, I_{he} has 1024 gray scales.

Fig. 6. FROC Curves of the detected lesions and false positives from the training dataset by our proposed CAD system. Each of the curves was generated with different values of λ and CCC thresholds. Using $\lambda=8.5$ yields a sensitivity of 90% and a false positive rate of 0.006 (8/1279), or a sensitivity of 96.7% and a false positive rate of 0.021 (27 / 1279). The CCC thresholds were set corresponding to the CIVR=5.12, and sphere index (SI) was >0.24 .

Fig. 7.
The effect of nodule enhancement on the FROC curves.

Fig. 8. The effects of nodule enhancement on post-Gd T1-W MRI (left), intensity histogram (second left), intensity variation within the lesion (second right) and intensity variation in the surrounding background (right) of a typical metastatic lesion. Top row: the original data; Middle row: the data after applying the histogram equalization; Bottom row: the data after applying the power transformation of $\lambda=6.5$.

Fig. 9. Three different lesions with different diameters missed by the radiologists but detected by our proposed CAD system. The lesions diameters are 2.15, 3.37 and 4.51 from left to right. Note that the blood vessels are rejected as a false positive in these images. Top row: the original images; Bottom row: the segmentation map.

Fig. 10.

Examples of three types of lesions missed by our proposed system: Left: a lesion with a low contrast to intensity variation ratio; Middle: a lesion attached to a small blood vessel in the midline of the brain; Left: a lesion near the edge and attaching in some parts to the brain surface.

Table 1

Numbers and characteristics of metastatic lesions in the training data

Dataset	<5 mm			<8 mm			Total		
	Total	R	H	E	Total	R	H	E	Total
Total	60	50	3	9	78	57	7	17	106

R, H, and E denote regular, hole and edge lesions.

Table 2

Numbers and Characteristics of metastatic lesions in the testing data

Dataset	<5 mm		<8 mm			Total	
	Total	R	H	E	Total	R	E
Total	186	157	2	28	250	199	44

R, H, and E denote regular, hole and edge lesions.

Table 3

Numbers of the lesions in the training data detected by radiologists

Method	2 mm	3 mm	4 mm	5 mm	8 mm
1st Radiologist	30/41 (73%)	22/27 (81%)	15/16 (93%)	14/14 (100%)	10/10 (100%)
2nd Radiologist	32/41 (78%)	24/27 (88%)	15/16 (93%)	14/14 (100%)	10/10 (100%)

No lesions with diameter greater than 5 mm were missed by the radiologists. Considering only the lesions smaller than 5 mm in diameter, 60% (16/27) and 66% (18/27) are detected by the two radiologists. Seven patients were included.

Table 4

Effects of false positive rejection criteria in the training data

NE		NE and size		NE, size and SI		NE, Size, SI and B		NE and B	
Miss	TP	FP	Miss	TP	FP	Miss	TP	Miss	TP
1	59	97 (610)	2	58	48 (211)	2	58	27 (106)	59
								1	85 (320)

NE, nodule enhancement; Size, size criterion; SI, sphere index; B, brightness criterion; TP, true positive; FP, false positive. The intra- and extra-cranial FPs are given outside and inside the parentheses, respectively.

Table 5

Numbers of the lesions in the testing data detected by radiologists

Method	2 mm	3 mm	4 mm	5 mm	8 mm
1st Radiologist	66/75 (88%)	56/60 (93%)	48/49 (98%)	45/45 (100%)	29/29 (100%)
2nd Radiologist	69/75 (92%)	58/60 (96%)	49/49 (100%)	45/45 (100%)	29/29 (100%)

No lesions with diameter greater than 5 mm were missed by the radiologists. Considering only the lesions smaller than 5 mm in diameter, 70% (21/30) and 80% (24/30) of the lesions are detected by the two radiologists. Eighteen patients were included.

Table 6

Lesions missed by the proposed system

Dataset	Regular	Hole		Edge	Total
	Low CIVR	Attached to an Object			
Training	0/50=0%	1/50=2%	0%	1/9=11%	2/60=3.3%
Testing	6/157=3.8%	5/157=3.1%	0%	1/28=3.5%	12/186=6.4%
Total	6/207=2.9%	6/207=2.9%	0%	2/37=5.4%	14/246=5.6%