Title

Computer-Aided Diagnosis of Hyperacute Stroke with Thrombolysis Decision Support Using a Contralateral Comparative Method of CT Image Analysis

Journal

Journal of Digital Imaging, 27(3)

ISSN

0897-1889

Authors

Shieh, Yao
Chang, Chien-Hung
Shieh, Mengkai
et al.

Publication Date

2014-06-01

DOI

10.1007/s10278-013-9672-x

Peer reviewed

J Digit Imaging (2014) 27:392–406
DOI 10.1007/s10278-013-9672-x

Computer-Aided Diagnosis of Hyperacute Stroke with Thrombolysis Decision Support Using a Contralateral Comparative Method of CT Image Analysis

Yao Shieh · Chien-Hung Chang · Mengkai Shieh · Tsong-Hai Lee · Yeu Jhy Chang · Ho-Fai Wong · Shy Chyi Chin · Scott Goodwin

Published online: 25 January 2014

Abstract New and improved techniques have been continuously introduced into CT and MR imaging modalities for the diagnosis and therapy planning of acute stroke. Nevertheless, non-contrast CT (NCCT) is almost always used by every institution as the front line diagnostic imaging modality due to its high affordability and availability. Consequently, the potential reward of extracting as much clinical information as possible from NCCT images can be very great. Intravenous tissue plasminogen activator (tPA) has become the gold standard for treating acute ischemic stroke because it is the only acute stroke intervention approved by the FDA. ASPECTS scoring based on NCCT images has been shown to be a reliable scoring method that helps physicians to make sound decisions regarding tPA administration. In order to further reduce inter-observer variation, we have developed the first end-to-end automatic ASPECTS scoring system using a novel method of contralateral comparison. Due to the self-adaptive nature of the method, our system is robust and has good generalizability. ROC analysis based on evaluation of 103 subjects who presented to the stroke center of Chang Gung Memorial Hospital with symptoms of acute stroke has shown that our system's dichromatic classification of patients into thrombolysis indicated or thrombolysis contraindicated groups has achieved a high accuracy rate with AUC equal to 90.2 %. The average processing time for a single case is 170 s. In conclusion, our system has the potential of enhancing quality of care and providing clinical support in the setting of a busy stroke or emergency center.

Keywords ASPECTS · Biomedical image analysis · Brain imaging · Computed tomography · Computer-aided analysis · Decision support · Hyperacute stroke · tPA

Y. Shieh (✉)
Department of Electrical Engineering, Chang Gung University, 259 Wen-Hwa 1st Road, Kwei-Shan, Tao-Yuan, Taiwan 333
e-mail: yshieh@mail.cgu.edu.tw

Y. Shieh
e-mail: yshieh@uci.edu

Y. Shieh · S. Goodwin
Department of Radiological Sciences, School of Medicine, University of California Irvine, 101 The City Drive South, Orange, CA 92868, USA

S. Goodwin
e-mail: sgoodwin@uci.edu

C.-H. Chang · T.-H. Lee · Y. J. Chang
Stroke Center and Department of Neurology, Chang Gung Memorial Hospital, Linkou Medical Center and College of Medicine, Chang Gung University, Kwei-Shan, Tao-Yuan, Taiwan 333

C.-H. Chang
e-mail: cva9514@gmail.com

T.-H. Lee
e-mail: thlee@adm.cgmh.org.tw

Y. J. Chang
e-mail: yjc0601@adm.cgmh.org.tw

M. Shieh
Department of Surgery, Russian Medical Academy of Postgraduate Education, ul. Barrikadnaya, d. 2/1, Moscow, Russia
e-mail: mks@alumni.stanford.edu

H.-F. Wong
Department of Neuroradiology, Chang Gung Memorial Hospital, Linkou Medical Center and College of Medicine, Chang Gung University, Kwei-Shan, Tao-Yuan, Taiwan 333
e-mail: hfwong@adm.cgmh.org.tw

S. C. Chin
Department of Diagnostic Radiology, Chang Gung Memorial Hospital, Linkou Medical Center and College of Medicine, Chang Gung University, Kwei-Shan, Tao-Yuan, Taiwan 333
e-mail: b25chin@gmail.com

Introduction

CT and MR are valuable imaging modalities for diagnosis and therapy planning of acute stroke. Non-contrast CT (NCCT) is often used initially to help differentiate between hemorrhagic and ischemic stroke. Conventional MR imaging is also very useful in detecting acute ischemic stroke within the first few hours. Both CT angiography and MR angiography can be used to characterize intravascular thrombi. CT perfusion is a functional imaging technique that can be used to identify salvageable brain tissue [1, 2]. Diffusion-weighted MR imaging can be used for identifying infarcted brain tissue that is irreversibly damaged, whereas perfusion-weighted MR imaging can be used for identifying reversibly injured ischemic brain tissue [3]. Although CT and MR are versatile imaging modalities that are capable of depicting various aspects of acute stroke-associated brain changes, their actual usage at a particular institution depends to a large extent on the availability of equipment and the experience and expertise level of that institution. NCCT, however, is almost always used without exception by every institution as a first-line imaging modality for acute stroke care, due to its high affordability and availability. This fact puts NCCT in a unique position, such that extracting as much clinical information as possible from NCCT images can have exceptionally rewarding results and implications [4], both clinically and economically.

Because it is the only acute stroke intervention approved by the FDA, intravenous tissue plasminogen activator (tPA) has become the gold standard for treating acute ischemic stroke patients if it is able to be administered within 3 h of symptom onset [5]. ASPECTS scoring has been proven to be a quantitatively reliable, standardized method for robust assessment of early ischemic changes on NCCT imaging for the ultimate purpose of helping physicians to make a decision regarding tPA administration [6–8]. Specifically, tPA is recommended if the patient's ASPECTS score is greater than 7, otherwise it is discouraged due to increased intracranial bleeding risk.

The purpose of this study is not to propose an alternative scoring method, different from ASPECTS, but to introduce the first automatic end-to-end ASPECTS scoring system, powered by computer-aided image analysis techniques, which offer the additional benefit of reduced inter-observer variability [9–11] and a potential for systematic improvement of ASPECTS scoring reliability.

In the literature, two groups have published studies which directly address the potential utility and impact of computer-based analysis of NCCT images in the setting of acute stroke care. Maldjian et al. presented an automated method to identify potential acute ischemia in the lentiform nucleus and insula as a demonstration of their method [12]. Lee, Takahashi and colleagues [13–16] introduced an adaptive partial median filter to enhance CT signals and a Z score mapping method that calculates each pixel's distance with respect to its "normalized reference" counterpart as a helpful aid in detecting anomaly. The Z score is calculated on a voxel-by-voxel basis by referencing a standard CT brain atlas using SPM software. There are limitations, however, to using the SPM-based standard atlas. First of all, all patients must be scanned on the same type of scanner with the same imaging parameters to eliminate image-affecting factors not attributable to the subjects themselves [17]. An institution that uses more than one type of scanner may have difficulty in using this approach. Secondly, difficulty in normalizing atypical brains is a known issue [17]. Finally, due to the low contrast of brain tissue in CT images, the SPM-derived standard CT brain atlas may not be of sufficient quality to enable reliable voxel-by-voxel registration. The Z score mapping method seems to show promising early results, but it may need further verification in studies with increased numbers of subjects and scanners.

In this paper, we have developed an automated ASPECTS scoring system, which outputs either a positive or negative recommendation regarding thrombolytic therapy. Our novel strategy of contralateral comparison eliminates the need for an external image pattern standard against which to determine the existence of acute ischemic changes. Each of the ten ROIs defined for ASPECTS scoring consists of a pair of bilaterally symmetric regions which reside on opposite sides of the brain. For each ROI, we compare left and right paired regions based on the two features that most effectively characterize the ischemic condition for that ROI to determine if the difference is large enough to declare an anomaly. Such a comparison method is possible because of the rarity of having a simultaneous, acute bilateral MCA stroke. As a prerequisite of the comparison, both left and right ROI pairs are segmented using an improved version of our previously developed two-step procedure [18] of approximate registration globally followed by fine adjustments locally. Our method is more robust and reliable than the generic SPM normalization. Instead of comparing against an external reference, our determination of anomaly is based on comparison between contralateral ROIs, and therefore, variations arising from different scanners and different imaging parameters such as KVP, window center, and window level, affect both sides equally and tend to cancel out. Consequently, our algorithm is inherently adaptive and robust.

Our automated scoring system has demonstrated a high degree of accuracy based on evaluation of 103 subjects. The average time for image processing, score computation, and anomaly visualization for a single case is 170 s, running on MatLab code. The processing time can be improved by replacing the interpretation language-based code with compiled language code such as C++. Therefore, the system and method have the potential to serve as useful clinical support tools in the busy setting of a stroke or emergency center.

Now that we have demonstrated that our automated scoring system can serve as a stand-alone scoring system with performance comparable with human experts, our future research is to validate the hypothesis that the accuracy of a physician's ASPECTS scoring can be improved when consulting with our automated scoring system as a second opinion in a controlled clinical setting.

Materials and Methods

Subjects

All patients who presented to the stroke center of Chang Gung Memorial Hospital with symptoms of acute stroke between January 2011 and June 2011 were retrospectively considered for this study. IRB approval by Chang Gung Memorial Hospital was obtained before beginning any data collection and associated research involving human subjects in this study. Patients were selected if they (1) presented within 3 h of onset of MCA stroke symptoms, (2) had baseline NCCT imaging, and (3) were later confirmed to have suffered an ischemic stroke. Among these selected patients were patients in which the baseline NCCT appeared normal, but follow-up diffusion-weighted MR imaging confirmed lacunar (15 mm in maximal dimension) infarct within 7 days of ictus. A very small number of patients were excluded due to excessive motion artifacts.

A total of 103 cases were used in this study. NCCT images were acquired using Siemens Sensation-16, Toshiba Aquillon and GE Bright Speed scanners. The images had different imaging parameters depending on their acquisition scanners. All images had 512 rows and 512 columns, with a slice thickness of 5 mm. The Siemens scanners had settings of 120 KVP, 320-mA exposure, dual window centers of 35/350 and corresponding window widths of 80/1,500. The Toshiba scanners had settings of 120 KVP, 230-mA exposure, window center of 40 and window width of 120. The GE scanners had settings of 120 KVP, 61-mA exposure, window center of 45 and window width of 125. Approval for this retrospective study and review of all images and charts were obtained from the Institutional Review Board of Chang Gung Memorial Hospital, Linkou Medical Center.

Image Processing Algorithm

Our method consists of five stages: (1) image preprocessing–extracting the brain image, rotating the brain image if necessary, noise-reduction filtering, and selection of the two slices needed for ASPECTS scoring; (2) unsupervised tissue classification; (3) template-based image segmentation; (4) identification and exclusion of old infarcts; (5) ASPECTS scoring for each ROI using contralateral comparative method; and (6) computation of global ASPECTS score and classification into thrombolysis recommended or thrombolysis contraindicated classes. A block diagram of our method is shown in Fig. 1.

Preprocessing of CT Images

Preprocessing of the CT images comprises the following tasks: (1) stripping away background and skull so that only the brain portion remains; (2) making any necessary rotational corrections to position the image in the vertically-aligned orientation (3) selection of the two pertinent CT image slices, $TARGET_{LS}$, and $TARGET_{US}$; and (4) reducing image noise by a filtering operation. Figure 2 shows an original head image and the brain image after image preprocessing.

The skull is characterized by very high pixel intensity with respect to the rest of the brain image. Consequently, the skull can be located and removed by simple thresholding and characterization as approximately an elliptical-shaped ring. Once the skull is located, the background that resides outside of the skull can be removed as well.

The next step is to determine the bisecting line that separates the left and right hemispheres of the brain. Subsequently, the brain image is rotated in such a way that this bisecting line becomes vertically oriented.

The ASPECTS scoring method requires that two standard axial CT slices, one at the supraganglionic level (upper slice) and the other at the ganglionic level (lower slice), be examined and scored. However, ASPECTS does not prohibit image

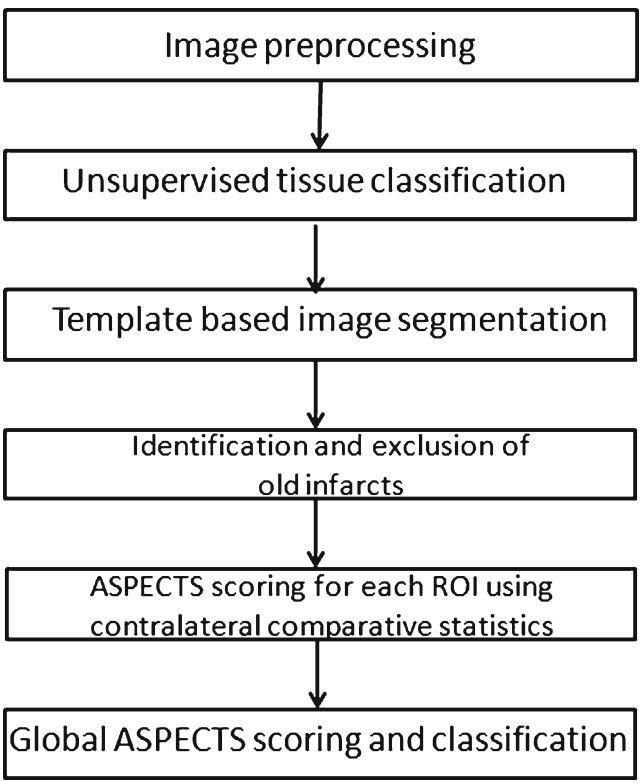

Fig. 1 Block diagram of the automated ASPECTS scoring method

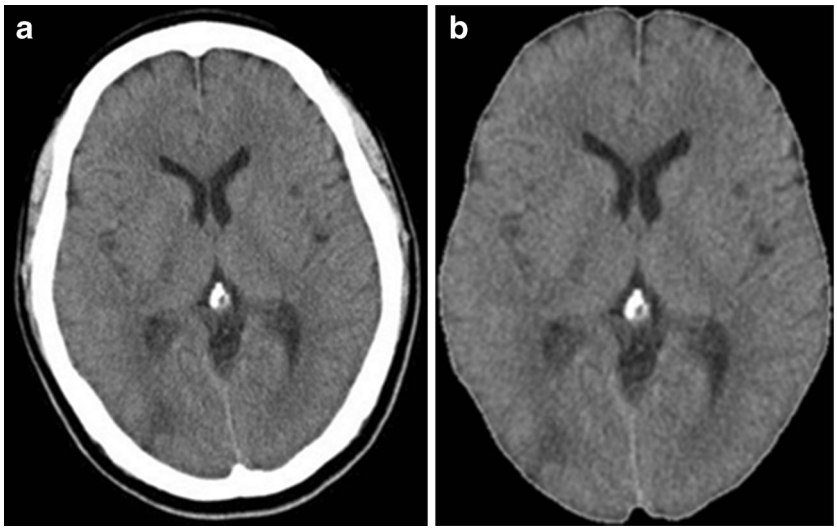

Fig. 2 **a** Original CT head image. **b** Brain image after image preprocessing

interpreters from examining the complete set of CT images. Our algorithm examines two slices only. However, we also consider the slice immediately below the lower slice in determination of focal hypoattenuation for caudate head, internal capsule, and lentiform nucleus regions.

The proper selection of the lower slice is crucial because it affects the concomitant segmentation of the lower slice image into various ROIs, such as caudate head, internal capsule, and lentiform nucleus. A standardized brain image of the lower slice is shown in Fig. 3. It was obtained by averaging the brain images of patients who presented to our medical center for workup of headache and who were subsequently found to have no structural abnormalities on NCCT. A reference template, $TEMP_{LS}$, which is derived from the standardized brain image for the purpose of singling out the lower slice, is shown in Fig. 4. The one slice from the complete stack of images that bears maximum similarity to the reference template is selected as the lower slice for ASPECTS scoring. The reference template is first resized to be the same size as the target image and then overlaid on top of it. The similarity measure is defined by Eq. (1).

$$\text{Similarity} = \sum_{\text{pi} \in \text{Target image}} \text{weight(pi)} \times \text{intensity(pi)} \quad (1)$$

where weight(pi)=1 if pi is white in $TEMP_{LS}$

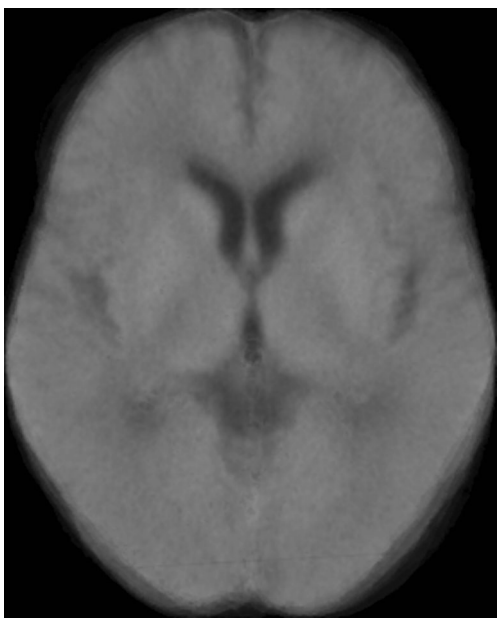

Fig. 3 The normalized brain image of the lower slice

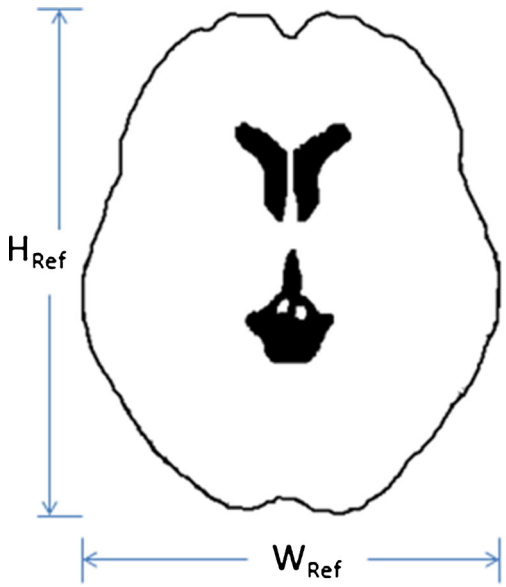

Fig. 4 The reference template, $TEMP_{LS}$, that is used to single out the lower slice from the stack of slices

weight (pi)=0 if pi is black in $TEMP_{LS}$

As an example, Fig. 5 contains a number of slice images, from which Fig. 5(d) can be determined to have maximum similarity with respect to the reference template, $TEMP_{LS}$.

Once the lower slice, $TARGET_{LS}$, is determined, the upper slice, $TARGET_{US}$, can be determined based on the thickness of CT slices because the average distance between upper and lower slices is known. Furthermore, the selection of the upper slice is

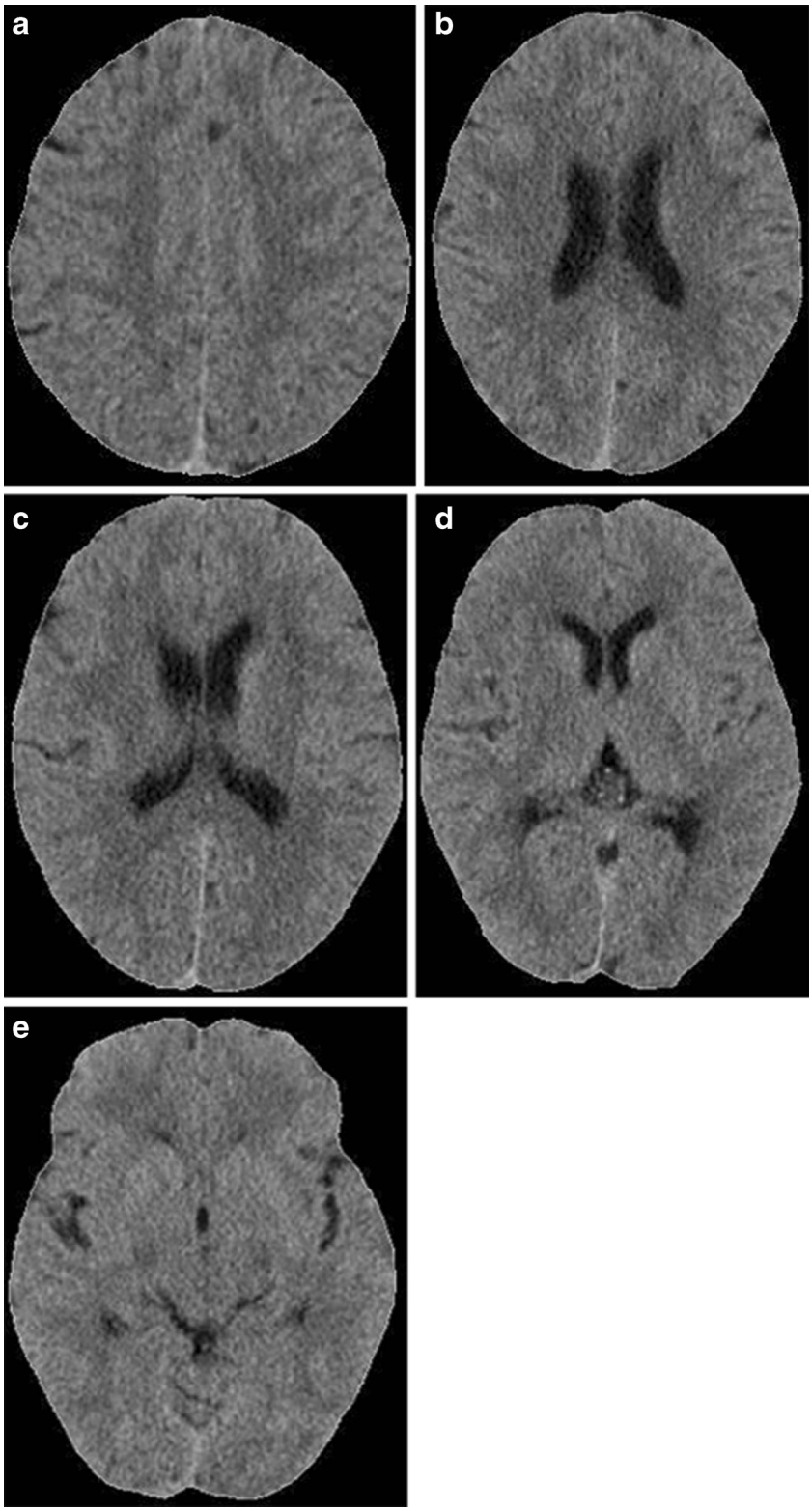

Fig. 5 An example of slice images of a CT study. **d** is determined to have maximum similarity with the reference template, $TEMP_{LS}$

not critical because its segmentation into ROIs does not rely on external templates but on inherent self-contained landmarks.

We have used the 3 by 3 median filter which is effective at retaining sharp edges while reducing image noise [19, 20].

Unsupervised Tissue Classification

Tissue classification of the lower slice brain image, $TARGET_{LS}$, serves two distinct purposes: (1) extraction of the conglomerate cluster, CR_{LV_QC}, comprising lateral ventricles, third ventricle and the quadrigeminal cistern, for the purpose of template-based image registration and (2) identification and exclusion of old infarcts.

Unsupervised tissue classification using the traditional *k*-means clustering technique provides the optimal classification in that it minimizes the total sum of squares of differences between individual intensities and their corresponding intensity cluster means.

Regarding segmentation of the lower slice brain image into various ROIs, the most salient landmark is the low intensity conglomerate region, CR_{LV_QC}, comprising lateral ventricles, third ventricle, and the quadrigeminal cistern. A rectangle, REG_{UTC_L}, of 155 pixel width and 123 pixel length is chosen to perform the *k*-means clustering, with $k=3$. REG_{UTC_L} is the larger green rectangle shown in Fig. 6. The tissue intensity histogram consists of three broad clusters—the low-intensity cluster of CR_{LV_QC}, the high-intensity cluster comprising C, L, and thalami, and the intermediate-intensity cluster comprising IC and other white matter. The real purpose of this *k*-means clustering is to extract the conglomerate region, CR_{LV_QC}. The high contrast between CR_{LV_QC} and its surroundings gives us excellent clues for attaining the optimum image registration, as illustrated in the next section.

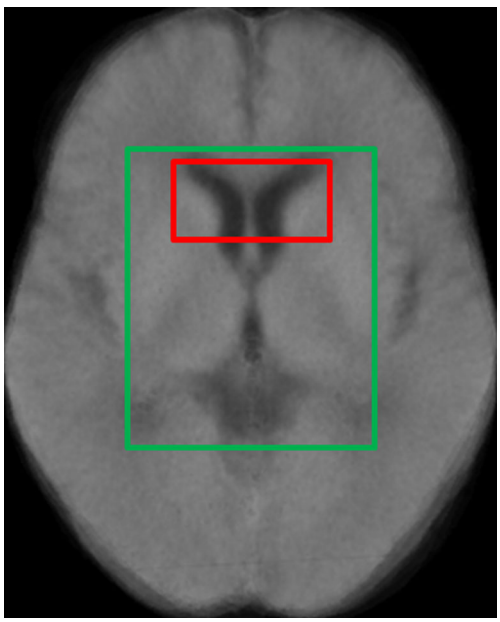

Fig. 6 The two rectangles within which unsupervised *k*-means clustering is performed

Old infarcts manifest a very low image intensity, comparable to that of the lateral ventricles. A rectangle, REG_{UTC_S}, of 50 pixel width and 110 pixel length, which contains the anterior horn of the lateral ventricles, is chosen to perform the *k*-means clustering, with $k=2$. REG_{UTC_S} is the smaller rectangle in red shown in Fig. 6. Based on experimentation, a REG_{UTC_S} that encloses the anterior horns of the lateral ventricles of the lower slice gives consistent and robust clustering results.

Figure 7 illustrates the process with an example. The smaller rectangle, REG_{UTC_S}, is first resized properly to become REG_{UTC_S}' and then superimposed on the target image as shown in Fig. 7a. A *k*-means clustering with $k=2$ is performed within the cropped rectangle of Fig. 7b and the final clustering result is shown in Fig. 7c. Based on the clustering results, the mean μ_{AHLV} and standard deviation σ_{AHLV} of the anterior horn of the lateral ventricles can be determined. Pixels whose intensities are lower than $\mu_{AHLV}+\sigma_{AHLV}$ are identified and grouped into clusters. A cluster of size smaller than a preset threshold is regarded as noise arising from random

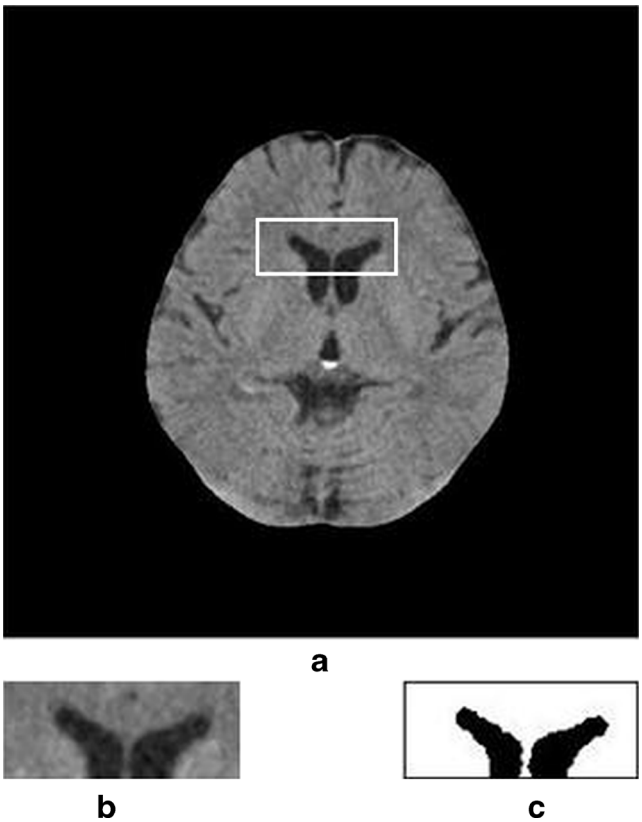

Fig. 7 **a** The rectangle that encloses the anterior horns of the lateral ventricles of the same CT image as in Fig. 6. It is cropped in (**b**) and used for the estimation of mean and standard deviation of lateral ventricles by means of *k*-means ($k=2$) clustering. **c** The *k*-means clustering result

fluctuation. Only clusters that meet the minimum size requirement are retained for further analysis.

Due to the similarity between old infarcts and the lateral ventricles in terms of image intensity, the mean image intensity, μ_{OI}, and standard deviation, σ_{OI}, of old infarcts are set to be equal to μ_{AHLV} and σ_{AHLV}, respectively. Candidate clusters comprising pixels whose intensities are lower than $\mu_{OI}+\sigma_{OI}$ can be either old infarcts or CSF spaces, such as subarachnoid space, ventricles and sulci. Further verification is needed before a cluster can be determined to represent old infarcts. Such verification is related to the ROI to which a pixel in question belongs. Consequently such verification is deferred until segmentation of the lower slice brain image is completed.

Template-Based Image Segmentation

The upper and lower slice images require segmentation into ROIs for further analysis. Segmentation of the upper slice into M4, M5, and M6 regions is accomplished by drawing horizontal lines through the upper and lower poles of the ventricles (Fig. 8). Since the upper and lower poles of the ventricles are inherent salient landmarks derivable from the target image itself, $TARGET_{US}$, no templates are needed.

Instead of using 3D spatial normalization [21], we use a much less computationally-demanding 2D template method for segmentation of the lower slice image. This 2D method is possible since ASPECTS only requires evaluation of two slices, which consequently renders the 3D spatial normalization process unnecessary. Moreover, we have avoided the risk of image quality degradation inherent in re-slicing of 3D image datasets which comprise anisotropic voxels, a case that is common for NCCT images.

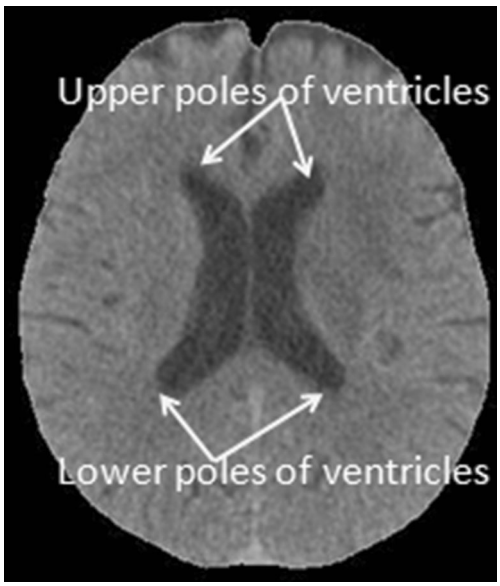

Fig. 8 Segmentation of M4, M5 and M6 regions

Our reference template, $TEMP_{SEG}$, consists of five pairs of components—caudate head (C), internal capsule (IC), lentiform nucleus (L), thalamus, and insular ribbon (I). $TEMP_{SEG}$ of Fig. 9 was derived by manually delineating the boundaries of the five pairs of components on the Reference Image shown in Fig. 3. The insular ribbon (I) region is an area that borders with regions C, IC, and L and extends radially in such a way that these regions collectively form a fan-shaped area. The remaining ROIs—M1, M2, and M3—are segmented by horizontal lines passing through the upper poles of lateral ventricles and the quadrigeminal cistern.

Image registration for the lower slice is a two-step process—initial global registration followed by local refinement. Initially, the Reference Image, along with its corresponding template of ROIs, is scaled by a factor of $W_{Target}/W_{Reference}$ and $H_{Target}/H_{Reference}$ in X and Y directions, respectively, so that it has the same size as the target image, an example of which is shown in Fig. 10. Next, the properly resized reference template is allowed to move within a small two-dimensional range of displacement—$\pm n_x$ pixels in the horizontal direction and $\pm n_y$ pixels in the vertical direction—in search of the best match between the target brain image and the reference template.

Our algorithm calculates the goodness of match based on the four red boundaries shown in Fig. 9. The UR_reg_boundary depicts the high contrast boundary between C and the lateral ventricle on the right hemisphere, whereas the LR_reg_boundary depicts the high contrast boundary between the thalamus and the conglomerate region comprising the third ventricle and the quadrigeminal cistern on the right hemisphere. Likewise, UL_reg_boundary and LL_reg_boundary are the counterparts of UR_reg_boundary and LR_reg_boundary on the opposite hemisphere.

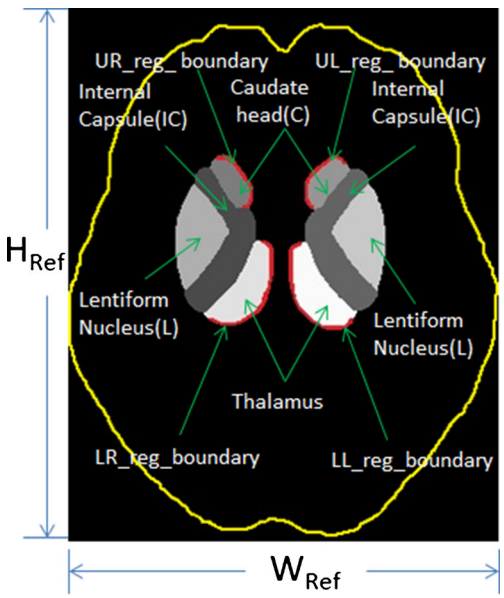

Fig. 9 The reference template, $TEMP_{SEG}$, that is used in image segmentation of the lower slice image

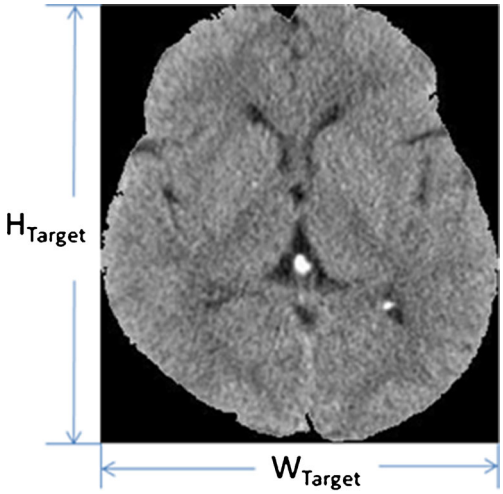

Fig. 10 An example of a lower slice target image to be segmented into various ROIs

At each trial displacement, $(\Delta x, \Delta y)$, the goodness of fit for the four red boundaries is evaluated independently. In other words, the trial displacement, $(\Delta x, \Delta y)_{UR}$, which gives the best match to the UR_reg_boundary determines the segmentation of C in the right hemisphere; whereas the trial displacement, $(\Delta x, \Delta y)_{LR}$, which gives the best match to the LR_reg_boundary determines the segmentation of the thalamus in the right hemisphere. However, $(\Delta x, \Delta y)_{UR}$ and $(\Delta x, \Delta y)_{LR}$ need not be the same, although they do tend to track each other in general. Analogous remarks apply to $(\Delta x, \Delta y)_{UL}$ and $(\Delta x, \Delta y)_{LL}$ on the left hemisphere. Finally, IC, L, and I can be segmented using the reference template and linear interpolation, after C and thalami are segmented. Precise calculation of the goodness-of-match metric is described in the Appendix.

The procedure of our template-based image segmentation is illustrated in Fig. 11. The larger rectangle for unsupervised tissue classification, $REG_{UTC\_L,}$ is properly resized and then superimposed on the target brain image in Fig. 11a, and then a *k*-means clustering (k=3) is performed. It results in three clusters, as shown in Fig. 11b, which is displayed in three colors—a red cluster depicting the conglomerate region, $CR_{LV\_QC,}$ comprising lateral ventricles, the third ventricle and the quadrigeminal cistern, a blue cluster depicting C, L, and thalami, and a green cluster depicting IC and other white matter. Once the conglomerate region, $CR_{LV\_QC,}$ has been segmented, the appropriately resized reference template, $TEMP_{SEG}$', is superimposed over the target image in search of the best registration of C and thalami based on the four

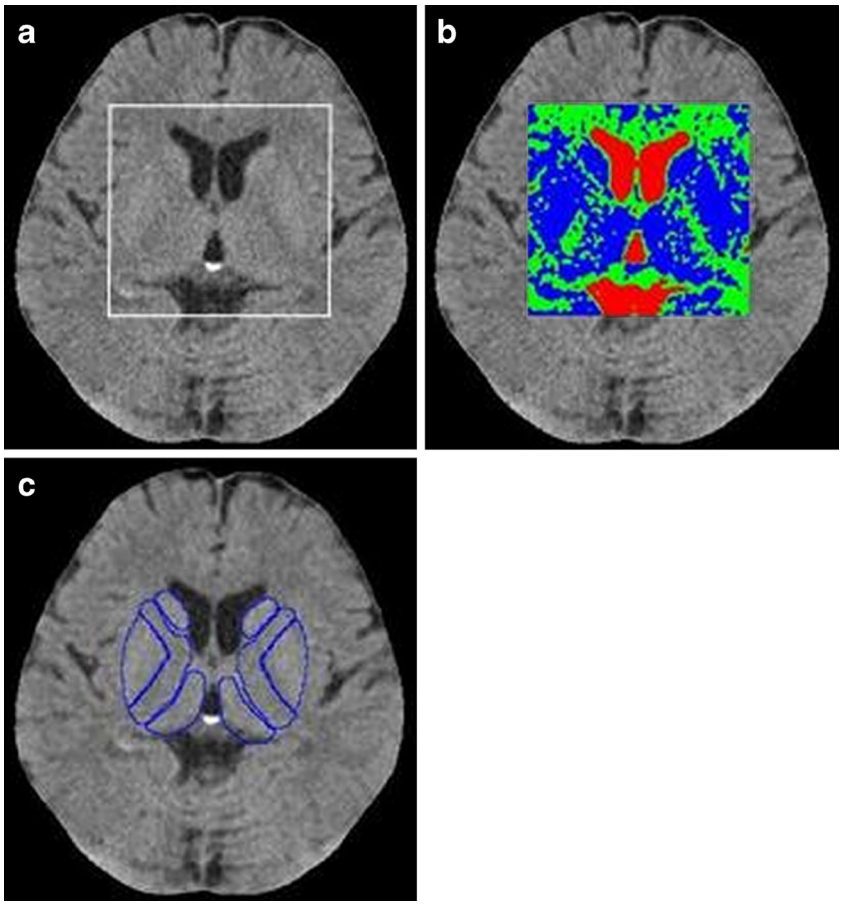

Fig. 11 **a** The larger rectangle for unsupervised tissue classification, $REG_{UTC\_L,}$ is superimposed on the target brain image. **b** The result of *k*-means clustering (k=3) applied to the *rectangle* in (**a**). The *red cluster* depicts the conglomerate region, $CR_{LV\_QC,}$ comprising lateral ventricles, the third ventricle and the quadrigeminal cistern, the *blue cluster* depicts C, L and thalami, and the *green cluster* depicts IC and other white matter. **c** The final result of image segmentation

boundaries, as described above. Once C and thalami have been registered onto the target image, regions IC, L, and I can be segmented according to the reference template. Figure 11c shows the final segmentation result.

Identification of old Infarcts and CSF Spaces

Old infarct regions need to be identified and differentiated from regions of early ischemic hypoattenuation. As illustrated in section D, candidate clusters comprising pixels whose intensities are lower than $\mu_{OI}+\sigma_{OI}$ can be either old infarcts or CSF spaces, such as subarachnoid space, ventricles, and sulci.

Old infarcts that occur in the territories of C, IC, and L are relatively easy to identify because no CSF spaces exist in these regions under normal conditions. Therefore, pixels whose intensities are lower than $\mu_{OI}+\sigma_{OI}$ are determined to be old infarcts and are therefore excluded.

Old infarcts that occur in the territory of I are identified by differentiating them from subarachnoid space pixels based on shape properties. If a candidate cluster's ratio of its major-axis-length to its minor-axis-length is larger than a preset threshold, it is determined to be a CSF space instead of old infarct.

A candidate cluster that occurs in the territories of M1 through M6, is classified as CSF space if it satisfies both of the following conditions: (1) the cluster's ratio of its major-axis-length to its minor-axis-length is larger than a preset threshold and (2) the cluster is connected to the outer boundary of the brain image. Otherwise, the candidate cluster is determined to be an old infarct. Clusters that fail to satisfy the above two conditions are determined to be old infarcts.

Figure 12a shows an old infarct in the territories of C, IC, and L, whereas Fig. 12b shows an old infarct in the territory of M3.

Principle of Contralateral Comparison

The fundamental strategy we have employed in detecting early ischemic anomaly is to compare corresponding ROIs in contralateral hemispheres, a strategy that makes use of the bilateral symmetry present in the human brain. Such a strategy is more robust than that of comparing the target image against a standard, normalized image intensity pattern that is derived from averaging a large number of images of normal subjects. The reason is because such a normalized image depends to a certain extent on the demographic profile of the subjects collected in the database and the scanner parameters that were used in the image acquisition. By contrast, our method is highly adaptive because the reference is derived from within the target image itself. Anatomic peculiarity of a particular individual tends to cancel out when comparing sub-images of opposite hemispheres. Likewise, equipment artifacts arising from different scanner hardware/software and associated scanner parameters also tend to diminish in significance when both sides are affected equally. The fact that a strictly controlled standard - in terms of image scanners, acquisition procedures and resultant output images - is not necessary presents a very favorable situation at an institution like ours, in which a number of different CT scanners made by different manufacturers are in use. We also have cases of referred patients who come from regional hospitals where CT images have already been acquired on different scanners.

The feasibility of using a contralateral comparative method that relies on a self-generated standard derived from the target itself stems from the fact that it is very rare to have an acute bilateral MCA stroke. Consequently, the image pattern of the

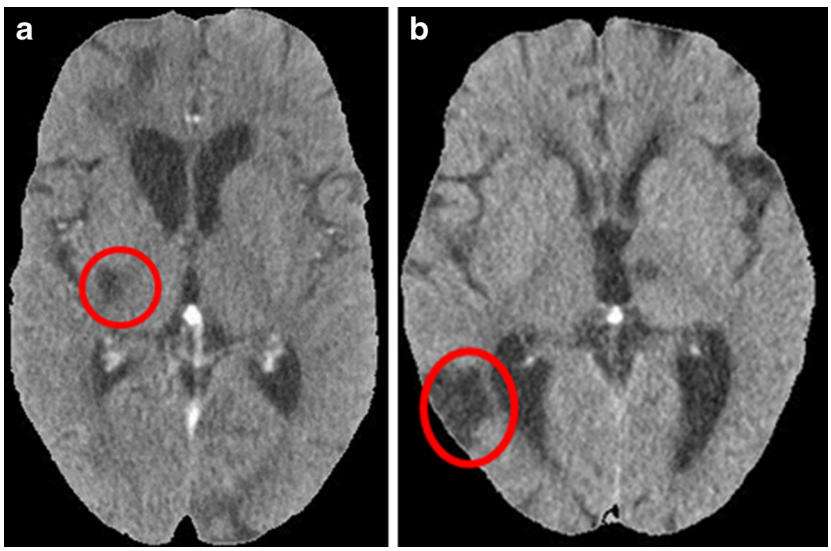

Fig. 12 **a** An old infarct encircled in red that occurs in the territory of C-IC-L. **b** An old infarct that occurs in the M3 region

unaffected hemisphere serves as the gold standard against which its counterpart is to be compared.

Due to imperfect bilateral symmetry, however, pixel by pixel comparison between corresponding ROIs of contralateral hemispheres is not reliable. A pragmatic approach is to compare characteristic properties (or features in the terminology of pattern recognition and classification) of corresponding ROIs of contralateral hemispheres. Generally speaking, the selection of a suitable feature vector involves a tradeoff between complexity and differentiating power. Our goal is to select the feature vector of smallest dimensionality that has sufficient power to detect early ischemic changes. After intensive investigation, we have selected a two-tuple feature vector, <diffuse hypoattenuation, focal hypoattenuation> for the analysis of C, IC, and L, and <diffuse hypoattenuation, GM-WM contrast degradation> for the analysis of I and M1 through M6, where GM and WM stand for gray matter and white matter respectively.

Detection of Early Ischemic Changes for Regions C, IC, and L

To detect diffuse hypoattenuation of a particular region, the two-sample Kolmogorov–Smirnov statistic is computed for the bilateral ROI pair corresponding to that region, after any old infarcts have been removed. The two-sample Kolmogorov–Smirnov statistic is a non-parametric test that can be used on two sets of pixels whose intensities have an arbitrary probability distribution. The null hypothesis is that the two samples are drawn from two populations that share an identical probability distribution. If the computed p value is less than a preset threshold, TH_{dif_hypo}, diffuse hypoattenuation is determined to exist.

To detect focal hypoattenuation pixels in an ROI, we again first exclude old infarcts, as described above. Next, we estimate the mean intensity, μ_{uas}, and standard deviation, σ_{uas}, , based on the ROI on the unaffected side. For a normal subject, the image intensity for each type of region C, IC or L can be modeled to have a Gaussian distribution. We can use the ratio of $\frac{\text{Intensity}(\text{Pixel}(x,y))-\mu}{\sigma}$ as an indicator of a pixel intensity's deviation from the average intensity of the ROI, μ, in units of standard deviation, σ, and call it the "normalized deviation from average", Dev_{Norm}. In a sense, Dev_{Norm} measures the likelihood that pixel(x,y) has focal hypoattenuation. The smaller, or more negative, the value of Dev_{Norm}, the higher the likelihood that it is a focal hypoattenuation pixel.

Next, we compute the normalized deviation from average, Dev_{Norm}, for each pixel of the affected side ROI. A pixel whose Dev_{Norm} is smaller (i.e., more negative) than a preset threshold is considered to be a candidate of focal hypoattenuation and is used in the subsequent attempt to form focal hypoattenuation clusters. In comparison with the Z score method [14–16] which derived a standardized mean, μ, and standard deviation, σ, from a cohort of normal subjects, our approach has the benefit of being more adaptive to the variations of scanner parameters, such as KVP, milliampere exposure, window width and center settings, as well as variations between individual subjects.

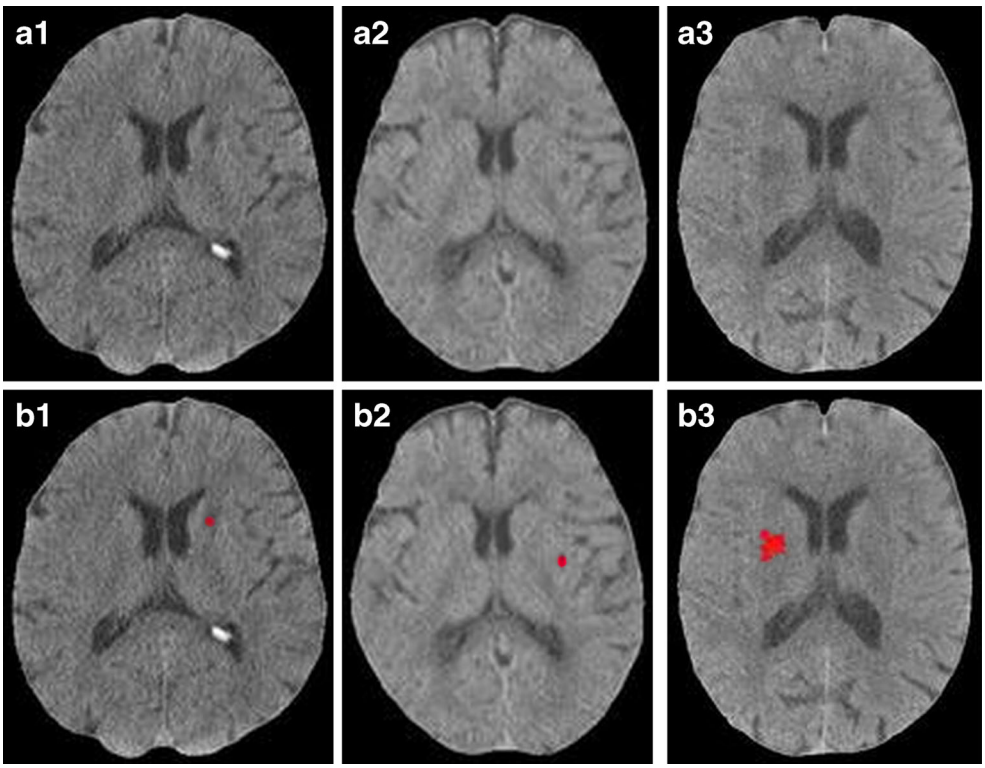

Fig. 13 **a1**, **a2**, **a3** Three CT images. **b1**, **b2**, **b3** Detected focal hypoattenuation changes as shown in *red*

To enhance the reliability of focal hypoattenuation identification, we have also taken the slice immediately beneath the lower slice at the ganglionic level into consideration. Initially, pixels with image intensity lower than $\mu_{\text{uas}} - N \times \sigma_{\text{uas}}$ (where N is a predetermined constant) are identified as potential focal hypoattenuation candidates: pixels_(FH on slice L). Next, potential focal hypoattenuation pixels on the slice immediately beneath the lower slice, pixels_(FH on slice L-1), are also identified in a similar manner. Then pixels_(FH on slice L) and pixels_(FH on slice L-1) are grouped into connected clusters, and the largest cluster is determined. Under this cross-slice clustering scheme, a pixel $P(x,y,z)$ on slice L has nine connectable neighbors $\{P(x-1, y-1, z-1), P(x-1, y, z-1), P(x-1, y+1, z-1), P(x, y-1, z-1), P(x, y, z-1), P(x, y+1, z-1), P(x+1, y-1, z-1), P(x+1, y, z-1), P(x+1, y+1, z-1)\}$ on the slice L-1 in addition to the eight connectable neighbors on the same slice. Focal hypoattenuation is determined to exist for the ROI under investigation if the number of pixels contained in the largest cluster is larger than a preset threshold, $\text{TH}_{\text{focal\_hypo}}$. Figure 13 shows several cases of detected focal hypoattenuation.

Detection of Early Ischemic Changes for Regions I and M1–M6

The same method as used above for regions C, IC, and L is also used for the detection of diffuse hypoattenuation in regions I and M1–M6.

The feature used for the detection of contrast degradation is the measure of dissimilarity of the GM-WM contrasts between an ROI and its bilateral counterpart, or in other words, the loss of contrast between gray matter and white matter in a particular region, using the corresponding contralateral ROI on the unaffected side as a reference. To evaluate this feature, the contrast, CONTRAST(P(x,y)), between each pixel, P(x,y), and its eight neighbors is calculated for all pixels within the ROI in question, according to the SOBEL operator used in edge detection, as shown in Eqs. (2)–(4).

$$\begin{aligned}\text{CONTRAST}_x(P(x,y)) &= \text{int}(P(x+1,y-1)) + 2\times\text{int}(P(x+1,y)) + \text{int}(P(x+1,y+1)) \\ &\quad - \text{int}(P(x-1,y-1)) - 2\times\text{int}(P(x-1,y)) - \text{int}(P(x-1,y+1))\end{aligned} \tag{2}$$

$$\begin{aligned}\text{CONTRAST}_y(P(x,y)) &= \text{int}(P(x-1,y+1)) + 2\times\text{int}(P(x,y+1)) + \text{int}(P(x+1,y+1)) \\ &\quad - \text{int}(P(x-1,y-1)) - 2\times\text{int}(P(x,y-1)) - \text{int}(P(x+1,y-1))\end{aligned} \tag{3}$$

$$\text{CONTRAST}(P(x,y)) = \sqrt{\text{CONTRAST}_x^2(P(x,y)) + \text{CONTRAST}_y^2(P(x,y))} \tag{4}$$

By such a computation, we have obtained a "contrast map" with each pixel's intensity being CONTRAST($P(x,y)$) shown above rather than the original image's intensity. Since we are interested in the contrast between GM and WM, we would like to minimize the confounding effects caused by (1) the high contrast between brain and CSF spaces and (2) the low GM–GM and WM–WM contrasts. In order to eliminate the high contrast between the brain and CSF spaces, we have identified and excluded all pixels whose intensities are less than $\mu_{\text{AHLV}} + \sigma_{\text{AHLV}}$ (as defined in section D), which constitute ventricles, subarachnoid space pixels and sulci pixels, as well as old infarcts, from the computation of contrast. The GM–GM and WM–WM contrasts are smaller than the GM–WM contrasts, and therefore, only pixels that reside within the upper 50th percentile of the contrast histogram are retained for further analysis. Next, the two-sample Kolmogorov–Smirnov (KS) statistic is computed by comparing the upper 50 % of contrast intensities of the affected side against the upper 50 % of contrast intensities of the unaffected side. If the computed p value is less than a preset threshold, $\text{TH}_{\text{contrast}}$, a contrast degradation between GM and WM is determined to exist.

Results

A total of 103 subjects consisting of 61 (59 %) males and 42 (41 %) females were included in our study after excluding nine cases, in which consensus in image interpretation could not be reached. Age ranged from 40 to 95 years, with mean and standard deviation being 67.9 and 12, respectively. The mean and standard deviation of National Institutes of Health Stroke Scale (NIHSS) for our study population were 11.33 and 9.34, respectively. The minimum and maximum NIHSS scores were 0 and 38, respectively. The demographic profile of our study population is listed in Table 1. The gold standard for performance evaluation of our automatic ASPECTS scoring system was the consensus reading of the initial NCCT images acquired within 3 h of stroke together with follow-up CT and MR images by experienced neuroradiologists and neurologists. All 103 subjects had follow-up CT images taken within 1 week. The mean and standard deviation of the gold-standard ASPECTS scores were 7.66 and 2.34, respectively.

All 103 cases were automatically registered with high accuracy, demonstrating that the registration algorithm is very

Table 1 Demographic profile of the study population

Gender	Male=59 %; female=41 %
Age	Min=40, max=95, mean=67.9, S.D.=12
NIHSS	Min=0, max=38, mean=11.33, S.D.=9.34
ASPECTS (gold standard)	Min=0, max=10, mean=7.66, S.D.=2.3

robust, irrespective of variations in scanner types and system parameters such as KVP, milliampere exposures and window center and window level. Figure 14 shows registration results of increasing complexity. Figure 14a is an example of an image that has a high degree of bilateral symmetry, whereas Fig. 14b is an example of an image that has moderate bilateral symmetry. In both cases, independent registration is carried out on both left and right hemispheres. Figure 14c is an example in which excessive deformation of structures on the affected side defies accurate registration. Therefore, the calculated registration on the affected side is discarded and replaced by mirroring the calculated registration of the unaffected side along the hemisphere-dividing line.

Our algorithm has proven to be very effective in identifying old infarcts, which manifest themselves as areas whose pixel intensities are comparable to that of ventricles.

The primary goal of this study has been to develop an automated scoring system that implements ASPECTS scoring for NCCT images to identify early ischemic changes and to facilitate well-founded decisions regarding the suitability of tPA thrombolytic treatment. The scoring performance for ROIs in three broad categories - regions C, IC and L; region I; and regions M1 through M6—are listed in Fig. 15. The performance results for these three categories, in terms of area under the ROC curve (AUC), are 89.1, 90.6, and 86.3 %, respectively.

The issue of decision making regarding thrombolysis can be modeled as a dichromatic classification problem in which the rate of misclassification is to be minimized. The original clinical group that developed ASPECTS observed that the outcome of tPA treatment for patients with an ASPECTS score higher than 7 is significantly better than for patients with an ASPECTS score of 7 or lower. Consequently, the above-mentioned classification into thrombolysis indicated and thrombolysis contraindicated classes is equated to classification into groups of ASPECTS score greater than 7 versus 7 or less. Our system's accuracy rate of dichromatic classification of patients into thrombolysis-indicated or thrombolysis-contraindicated groups is evaluated under the framework of ROC analysis. Sensitivity and specificity are defined as follows:

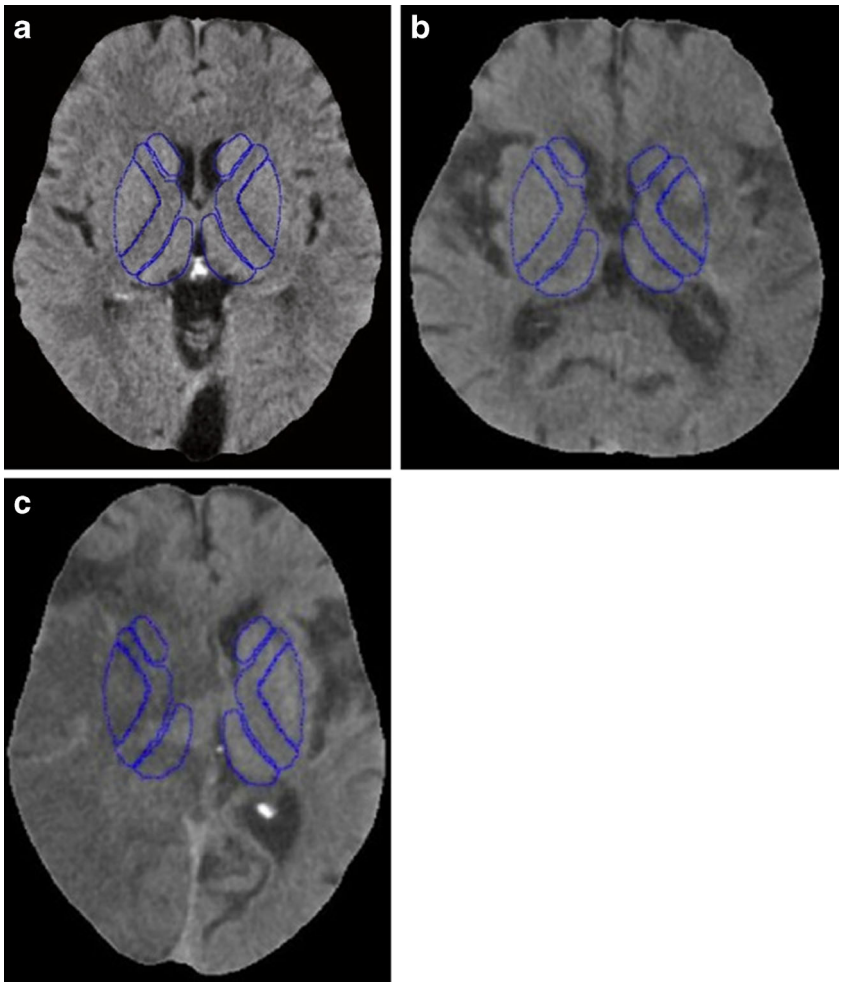

Fig. 14 **a** and **b** are examples in which registration is carried out independently on each side; **c** registration of the unaffected side is mirrored over to the affected side when there exists excessive deformation of structures on the affected side

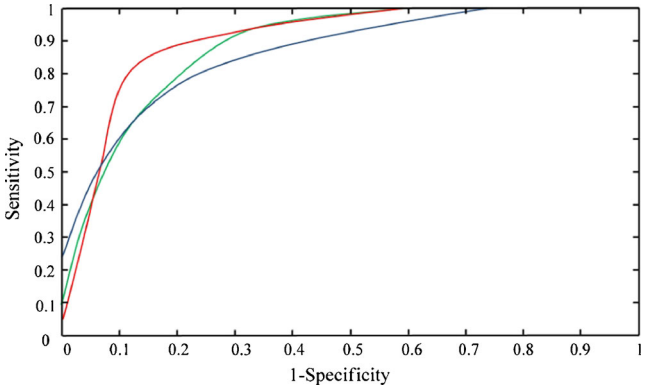

Fig. 15 ROC curves for three broad categories - regions C, IC and L (*Green*); region I (*red*); and regions M1 through M6 (*blue*). Their AUCs are 89.1, 90.6, and 86.3 %, respectively

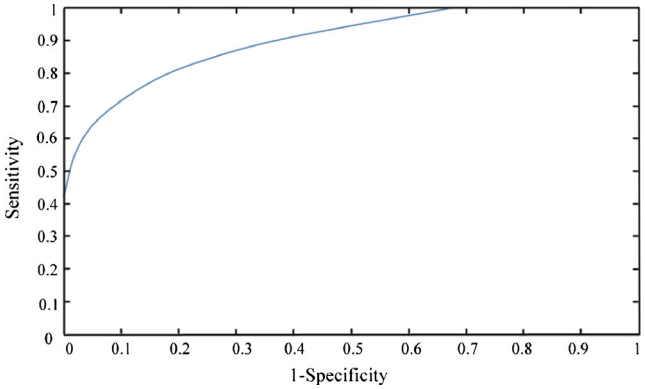

Fig. 16 ROC curve for the global ASPECTS score. AUC is equal to 90.2 %

$$\text{Sensitivity} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}}$$

$$\text{Specificity} = \frac{\text{True negative}}{\text{True negative} + \text{False positive}}$$

where positive and negative means the computed ASPECTS score is between 0 and 7 inclusive, or greater than 7, respectively. The result is shown in Fig. 16. The area under the ROC curve (AUC) is 90.2 %.

The average execution time for end-to-end image processing and displaying of detected early ischemic changes was 170 s, with algorithms implemented as MatLab code and execution on a desktop computer equipped with an Intel i3-2120 CPU running at 3.3 GHz and 4-GB RAM. We avoided computationally demanding 3D re-slicing and also the associated potential risk of image degradation, as described in the "Materials and Methods" section.

Conclusion

In this study, we have reported an automated ASPECTS scoring system which provides an objective assessment of early ischemic changes based on NCCT images. Algorithms that can enhance image quality and improve image segmentation and detection of anomalies have been reported in the literature. Our system, however, is the first stand-alone system that is able to perform end-to-end NCCT image analysis and produce an all-inclusive score that can be used by clinicians to facilitate decision making regarding thrombolytic therapy. Its performance is evaluated based on a retrospective study of 103 consecutive cases that presented to the stroke center of Chang Gung Memorial Hospital, Linkou Medical Center, with symptoms of acute stroke between January 2011 and July 2011. Images involved in these cases were acquired from different scanners with varying parameter settings. Therefore, the robustness and generalizability of our method has been demonstrated. The performance of our system is quite satisfactory, both when evaluated based on individual ROIs and also based on global ASPECTS score. Its potential role of serving as a standalone system that offers recommendations for or against thrombolytic treatment has been demonstrated by its classification accuracy based on ROC analysis.

The system's comparatively quick processing time of 170 s per case on a typical desktop computer suits the clinical setting of a busy stroke or emergency center, and this performance can be significantly improved if needed by rewriting the program in a compiled programming language.

The next step of our research is to evaluate the potential beneficial effects of using our system when tested in a controlled clinical practice environment and verification setting to evaluate its ability to improve the quality of acute stroke care. To be specific, changes in classification accuracy, interpreter confidence, as well as comparisons of response times in connection with utilizing our system as a second opinion consultant will be measured and evaluated quantitatively.

Acknowledgements This work was supported in part by the National Science Council of Taiwan under grant NSC 100-2218-E-182-002 and grant NSC101-2221-E-182-058. The authors are grateful to T.H. Won, C.L. Hsieh, C.F. Hsiao, J.S. Lu, and W.J. Chen for their assistance in computer programming and image data management.

Appendix

Calculation of goodness-of-match metric between target image and reference template

The calculation of goodness of match between the target image and the reference template based on the four red boundaries in Fig. 9 can be expressed mathematically as follows. At each trial displacement, $(\Delta x, \Delta y)$, the goodness

of match for UR_reg_boundary, $G_{UR}(\Delta x, \Delta y)$, is defined by Eq. (5):

$$G_{UR}(\Delta x, \Delta y) = \sum_{pi \in UR\_reg\_boundary} (B_{pi})_{\Delta x, \Delta y} + (More\_than\_10\%\_of\_UR\_reg\_boundary\_in\_lateral\_ventricles) \times (-1,000) \quad (5)$$

where B_{pi}=1 if pixel($x_{pi}+\Delta x$, $y_{pi}+\Delta y$) has at least one but not all of its eight neighbors overlapping with the lateral ventricles of the target image

B_{pi}=0 otherwise

(More_than_10%_of_UR_reg_boundary_in_lateral_ventricles)=1

if more than 10 % of UR_reg_boundary pixels fall within the lateral ventricles,

(More_than_10%_of_ UR_reg_boundary_in_lateral_ ventricles)=0 otherwise;

At each trial displacement, (Δx, Δy), a pixel, pi, may find itself in one of three possible locations with respect to the lateral ventricle in the right hemisphere: (1) the pixel is away from the lateral ventricle if none of the eight neighbors of ($x_{pi}+\Delta x$, $y_{pi}+\Delta y$) belong to the lateral ventricle, (2) the pixel is on the boundary of the lateral ventricle if at least one but not all of its eight neighbors of ($x_{pi}+\Delta x$, $y_{pi}+\Delta y$) belong to the lateral ventricle, and (3) the pixel is inside the lateral ventricle if all of its eight neighbors of ($x_{pi}+\Delta x$, $y_{pi}+\Delta y$) belong to the lateral ventricle. The first term counts the number of pixels of the UR_reg_boundary of the reference template that coincide with the right lateral ventricle boundary on the target image. Each such coincidence is given a point. On the other hand, the number of pixels of UR_reg_boundary of the reference template that fall within the lateral ventricle on the target image is also counted. A very large penalty of −1,000 points is imposed if the number of such pixels exceeds 10 % of UR_reg_boundary. The trial displacement, $(\Delta x, \Delta y)_{UR}$, that has the largest value of $G_{UR}(\Delta x, \Delta y)$ is the best match.

Likewise, at each trial displacement, (Δx, Δy), the goodness of match for LR_reg_boundary, $G_{LR}(\Delta x, \Delta y)$, is defined by Eq. (6) below:

$$G_{LR}(\Delta x, \Delta y) = \sum_{pi \in LR\_reg\_boundary} (B_{pi})_{\Delta x, \Delta y} + (More\_than\_10\%\_of\_LR\_reg\_boundary\_in\_3V\_or\_QC) \times (-1,000) \quad (6)$$

where B_{pi}=1 if pixel($x_{pi}+\Delta x$, $y_{pi}+\Delta y$) has at least one but not all of its eight neighbors overlapping with the conglomerate region comprising the third ventricle and the quadrigeminal cistern

B_{pi}=0 otherwise;

(More_than_10%_of_LR_reg_boundary_in_3V_or_Q-C)=1 if more than 10 % of LR_reg_boundary pixels fall within the conglomerate region comprising the third ventricle and the quadrigeminal cistern,

(More_than_10%_of_upper_boundary_ in_3V_or_QC)= 0 otherwise;

Goodness-of-match metrics for the two boundaries on the left hemisphere, $G_{UL}(\Delta x, \Delta y)$ and $G_{LL}(\Delta x, \Delta y)$, can be calculated in a similar manner.

References

1. Lin K, Rapalino O, Law M, Babb JS, Siller KA, Pramanik BK: Accuracy of the Alberta stroke program early CT score during the first 3 hours of middle cerebral artery stroke: comparison of noncontrast CT CT angiography source images, and CT perfusion. AJNR Am J Neuroradiol 29:931–936, 2008
2. Aviv RI, Mandelcorn J, Chakraborty S, Gladstone D, Malham S, Tomlinson G, Fox AJ, Symons S: Alberta stroke program early CT scoring of CT perfusion in early stroke visualization and assessment. AJNR Am J Neuroradiol 28:1975–1980, 2007
3. Srinivasan A, Goyal M, Al Azri F, Lum C: State-of-the-art imaging of acute stroke. RadioGraphics 26:S75–S95, 2006
4. Symons SP, Cullen SP, Buonanno F, González RG, Lev MH: Noncontrast conventional computed tomography in the evaluation of acute stroke. Semin Roentgenol 37:185–191, 2002
5. Smith MC. Reperfusion Therapy for Acute Ischemic Stroke. 2012 Summer.http://www.neurologyreport.com/stroke/pdf/Smith.pdf. Accessed 6 Feb 2013
6. Demaerschalk BM, Silver B, Wong E, Merino JG, Tamayo A, Hachinski V: ASPECT scoring to estimate 1/3 middle cerebral artery territory infarction. J Neurol Sci 33:200–204, 2006
7. Pexman JH, Barber PA, Hill MD, Sevick RJ, Demchuk AM, Hudon ME, Hu WY, Buchan AM: Use of the Alberta stroke program early CT Score (ASPECTS) for assessing CT scans in patients with acute stroke. Am J Neuroradiol 22:1534–1542, 2001
8. Barber PA, Demchuk AM, Zhang J, Buchan AM: Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. Lancet 355(9216):1670–1674, 2000
9. Grotta JC, Chiu D, Lu M, Patel S, Levine SR, Tilley BC, Brott TG, Haley Jr, EC, Lyden PD, Kothari R, Frankel M, Lewandowski CA, Libman R, Kwiatkowski T, Broderick JP, Marler JR, Corrigan J, Huff S, Mitsias P, Talati S, Tanne D: Agreement and variability in the interpretation of early CT changes in stroke patients qualifying for intravenous rtPA therapy. Stroke 30:1528–1533, 1999
10. Coutts SB, Hill MD, Demchuk AM, Barber PA, Pexman JH, Buchan AM: ASPECTS reading requires training and experience. Stroke 10: 179, 2003
11. von Kummer R, Holle R, Gizyska U, Hofmann E, Jansen O, Petersen D, Schumacher M, Sartor K: Interobserver agreement in assessing early CT signs of middle cerebral artery infarction. AJNR Am J Neuroradio 9:1743–1748, 1996
12. Maldjian JA, Chalela J, Kasner SE, Liebeskind D, Detre JA: Automated CT segmentation and analysis for acute middle cerebral artery stroke. Am J Neuroradiol 22:1050–1055, 2001
13. Lee Y, Takahashi N, Tsai DY, Ishii K: Adaptive partial median filter for early CT signs of acute cerebral infarction. Int J Cars 2:105–115, 2007

14. Takahashi N, Tsai DY, Lee Y, Kinoshita T, Ishii K: Z-score mapping method for extracting hypoattenuation areas of hyperacute stroke in unenhanced CT. Acad Radiol 17(1):84–92, 2010
15. Takahashi N, Tsai DY, Lee Y, Kinoshita T, Ishii K, Tamura H, Takahashi S: Usefulness of z-score mapping for quantification of extent of hypoattenuation regions of hyperacute stroke in unenhanced computed tomography: analysis of radiologist's performance. J Comupter Assited Tomogr 34(5):751–756, 2010
16. Lee Y, Takahashi N, Tsai DY. Computer-Aided Diagnosis for Acute Stroke in CT Images. Computed Tomography—Clinical Applications. 2012
17. Mechelli A, Price CJ, Friston KJ, Ashburner J: Voxel-based morphometry of the human brain: methods and applications. Curr Med Imaging Rev 1(1):9, 2005
18. Shieh Y, Chang CH: Automated ASPECTS Scoring System as a Clinical Support System for Acute Stroke Care. IEEE-EMBS International Conference on Biomedical and Health Informatics, 2012, pp. 691–694
19. Arias-Castro E, Donoho DL: Does median filtering truly preserve edges better than linear filtering? Ann Stat 37:1172–1206, 2009
20. Dougherty G: Digital image processing for medical applications. Cambridge University Press, Cambridge, 2009
21. Bankman IN: Handbook of Medical Image Processing and Analysis, 2nd edition. Academic Press, MA, 2009