Evaluation of a computer-aided detection algorithm for timely diagnosis of small acute intracranial hemorrhage on computed tomography in a critical care environment

Joon K. Lee[a], Tao Chan[b], Brent J. Liu[a], H.K. Huang[a]
[a]Image Processing and Informatics Laboratory, Department of Radiology
University of Southern California, Doheny Eye Institute Building,
1450 San Pablo Street, Suite 2100, Los Angeles, CA, USA 90033;
[b]Dept. of Health Technology and Informatics, The Hong Kong Polytechnic University,
Hong Kong, SAR, China

ABSTRACT

Detection of acute intracranial hemorrhage (AIH) is a primary task in the interpretation of computed tomography (CT) brain scans of patients suffering from acute neurological disturbances or after head trauma. Interpretation can be difficult especially when the lesion is inconspicuous or the reader is inexperienced. We have previously developed a computer-aided detection (CAD) algorithm to detect small AIH. One hundred and thirty five small AIH CT studies from the Los Angeles County (LAC) + USC Hospital were identified and matched by age and sex with one hundred and thirty five normal studies. These cases were then processed using our AIH CAD system to evaluate the efficacy and constraints of the algorithm.

Keywords: CAD, Brain, Computed Tomography, CT, Intracranial Hemorrhage, Algorithm

1. INTRODUCTION

Intracranial hemorrhage within the brain can be divided into intraaxial and extraaxial categories. Intraaxial hemorrhages include intraparenchymal hemorrhages (IPH) and intraventricular hemorrhages (IVH) while extraaxial hemorrhages encompass epidural hemorrhages (EDH), subdural hemorrhages (SDH), and subarachnoid hemorrhages (SAH).

AIH is often a sequela of head trauma and can result in acute neurological disturbances. Identification of AIH is of paramount clinical importance as its presence and nature dictates distinct management and treatment strategies. However, clinical signs, symptoms, and other parameters are insufficient to accurately differentiate AIH from other etiologies of neurologic disturbance [1,2]. Therefore, CT has long been the primary modality for detection and characterization of AIH due to its wide availability, quick performance, and ready depiction of AIH.

In most parts of the world, emergency department physicians, internists, or neurosurgeons are often the first to read CT studies, particularly when a radiologist's expertise is not immediately available. The skill of such acute care physicians has been shown to be imperfect based on receiver operating characteristic (ROC) studies [3]. Further, radiology residents have been shown to infrequently overlook hemorrhage on brain CT [4]. Among even the best human observers, errors in image interpretation are inevitable [5].

It is envisioned that CAD will help improve the accuracy of detecting AIH and decrease the risk of misdiagnosis and mismanagement. The ultimate goal of this study is to develop a CAD system to integrate within the clinical environment and to evaluate and refine a CAD algorithm [6] developed to detect small AIH. However, this system differs from other CAD products in that it is intended to be used by clinicians other than radiologists and that the likelihood of true AIH in identified candidate lesions is rated based on anatomical positions and imaging features.

Medical Imaging 2009: Advanced PACS-based Imaging Informatics and Therapeutic Applications
edited by Khan M. Siddiqui, Brent J. Liu, Proceedings of SPIE Vol. 7264, 726416 · © 2009 SPIE
CCC code: 1605-7422/09/$18 · doi: 10.1117/12.815390

In 2007, we presented a CAD system for the detection of small AIH. The purpose of this study is to validate the efficacy of that system and use the results to develop a CAD system for clinical implementation and evaluation using a larger data set derived from an acute care setting.

2. METHODOLOGY

2.1 Study Population

Final reports signed by staff radiologists for all non-contrast CT's of the head performed from 2005 to 2008 were extracted from the LAC+USC hospital information systems, parsed, and imported into a custom database for data mining. One hundred thirty eight patients were identified who were reported to have AIH and whose studies were available for export. Each patient was matched by age, sex, and ethnicity with a control patient who was reported to have a normal study. Available studies were exported out of the Department of Radiology PACS as DICOM files and the headers were anonymized. This retrospective study was approved by our institutional review board.

2.2 CT Acquisition

All CT studies included in the present study were acquired employing a Picker PQ 5000 or 6000 single-slice CT scanner. Acquisition parameters include 5 mm collimation, 130 kV, and beam currents of 30 mA.

2.3 CAD System

The CAD system was developed using MATLAB (The MathWorks, Inc., Natick, MA, USA) and modified from an earlier iteration to run in batch mode without any user intervention. The system ran on a 1.8 Ghz Intel Core 2 Duo laptop with 2 megabytes of RAM. A flowchart of the CAD algorithm is illustrated in Figure 1. Image processing and analysis methods used by the system are listed in Table 1.

Intracranial contents are segmented by global thresholding and morphological operations followed by contiguity analysis. Noise reduction using median filter and adjustment for CT cupping artifacts are performed. The intracranial contents are realigned into the conventional orientation after automatic localization of the mid-sagittal plane and boundaries of the series of images. Then, high attenuation components are segmented as candidate AIH from each of the axial sections based on top-hat transformation and subtraction between the two halves of the images. Image features of the candidate regions are then quantified. The candidate AIH are given anatomic context by registration against a normalized coordinate system developed for this project. Then, the features and coordinates are used in a rule-based classification system to reduce false positives due to normal variants and artifacts.

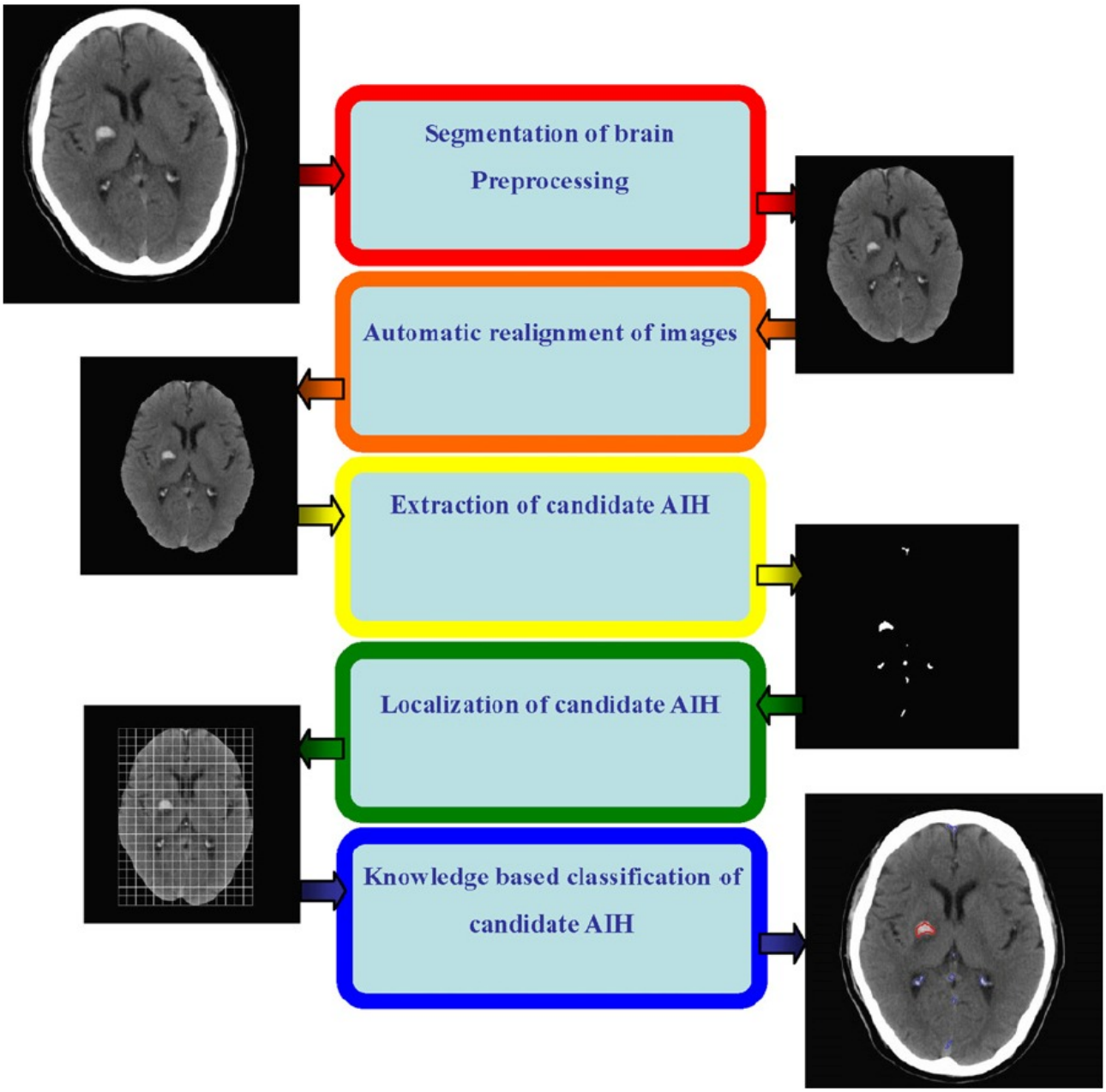

Fig. 1. Schematic diagram of the CAD system. Intermediary outputs of an image showing right basal ganglia hemorrhage illustrate the effect of individual steps.

Table. 1. Details of individual image processing and analysis steps in the CAD system as outlined in Fig. 1.

Steps	Methods	Purposes
Segmentation of intracranial contents	Global thresholding and morphological operations Remove structures not contiguous with the main central bulk of intracranial contents	Remove bones of skull and face Remove scalp, orbits, and other head and neck soft tissue
Preprocessing of intracranial contents	Median filtering Adjustment of intensity according to distance from the skull	Denoising Correction for CT cupping artifacts

Automatic realignment of images	Automatic localization of limits of brain, ventricles, floor of anterior intracranial fossa, mid-sagittal plane	Align the brain into normal position
Extraction of candidate AIH	Top-hat transformation Subtraction between the two sides	Highlight local high density regions Extract asymmetrically high density regions
Localization of candidate AIH	Registration of the brain in question against a normalized coordinate system	Render the candidate AIH anatomical information
Knowledge-based classification of AIH	Rule-based system with inputs of image features and anatomical coordinates of the extracted candidates	Distinguish genuine AIH from false positives resulting from noise, artifacts, and normal variants

2.4 Display of output

The contours of potential and genuine AIH are overlaid in different colors onto a duplicate of the original images. This permits juxtaposition of original input and CAD output images. A custom visualization tool was developed to facilitate inspection and to record analysis. A screenshot of the graphical interface is shown in Figure 2.

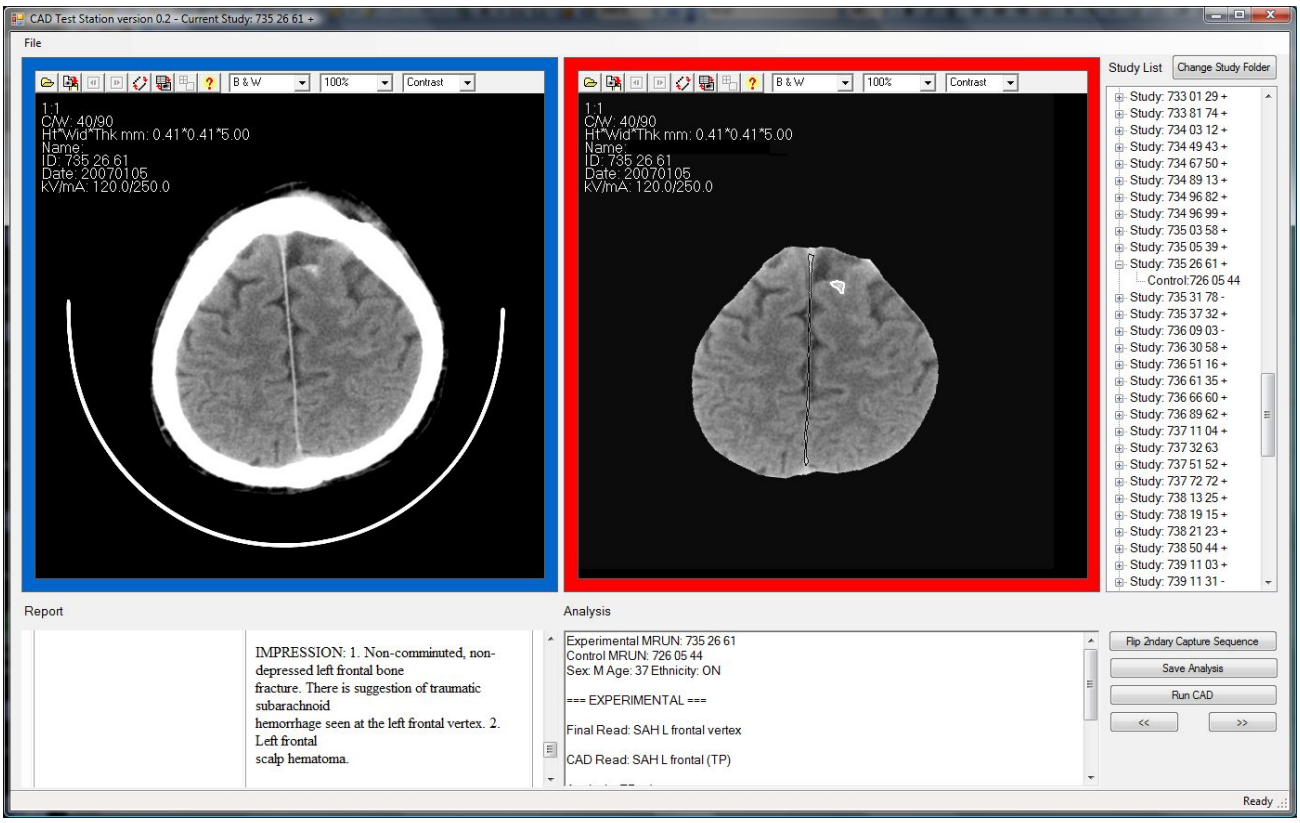

Fig. 2. Screenshot of CAD visualization tool. The original images are displayed in the left window while the output images with overlay of the outlines of AIH are displayed on the right.

3. RESULTS

The performance of the system was described by sensitivity and specificity pairs on both per lesion and per case bases. The per lesion descriptors are more informative when number or size of lesions need to be quantified, or when the performance of detecting some particular type of lesion is of interest. On the other hand, the performance on a per patient or per case basis is of more clinical relevance for the diagnosis of AIH as management options depend on the presence or absence or lesions rather than the quantity of lesions.

In the following section, a CAD output is counted as true positive if it overlaps at least part of the true AIH as determined by radiologists.

The actual presence of a lesion was based on the gold standard of the final report signed by a staff neuroradiologist. For each combination of lesion and location, a unique lesion was noted to exist in this study. Thus, a single study could possibly represent multiple AIH categories at multiple locations. The success or failure of the CAD system was based on whether such lesions derived from the final report were correctly identified. Specifically, all distinct lesions described in the final report were compared against CAD-identified candidate lesions by visual inspection. A successful detection was confirmed if any region of the lesion, regardless of the size, was correctly outlined by the CAD system. A logical next step would be establishment of an ROC study.

On a per patient basis, if any one of the CAD outputs for a particular patient is a true positive, then the case is counted as true positive, disregarding whether the other outputs are true or false positive. If there are one or more CAD output(s), none of which are true positives, the study is considered a false positive. On the contrary, if there is no CAD output for a particular patient, but there is a genuine AIH, the case is considered a false negative. If there is no CAD output for a patient in whom an AIH exists, the case is counted as a true negative.

Represented in the 135 studies were 20 studies with IPH, 14 with IVH, 91 with SAH, 53 with SDH, and 13 with EDH. The overall sensitivity in lesion detection was 69.6% (167/240) (Table 2). However, in some cases, the system correctly identified the location of the AIH but subsequently categorized the candidate lesion as not being a true AIH. With this in consideration, the overall sensitivity in lesion identification was 84.2% (202/240).

Table 2. Summary of CAD results according to type of individual AIH on a per lesion basis. TP: True positive; TP » FN: Correctly identified lesion but incorrectly classified as not being true AIH; FN: False negative.

	IPH	IVH	SAH	SDH	EDH	Total AIH
Studies	20	14	91	53	13	191
TP	20	10	72	53	12	167
TP » FN	5	6	17	5	2	35
FN	3	6	19	10	0	38

On a per patient basis, the overall sensitivity was 77.0% (104/135) (Table 3). Similar to the per lesion basis, when correctly identified but incorrectly classified AIH is taken into consideration, the overall sensitivity in lesion identification was 89.6% (121/135).

Table 3. Summary of CAD results on a per patient basis.

	AIH Cases
Studies	135
TP	104
TP » FN	17
FN	14

Our control cases returned 100% false positive results primarily originating around the falx most likely due to an error in the registration of the brain against the internal, normalized coordinate system used by the CAD system. Incorrect application of the coordinate system would apply inconsistent or erroneous rules-based logic in determining the nature of a candidate lesion. Interestingly, corresponding false positives around the falx were not replicated in the CAD output of true AIH cases. To address this possible source of error, we will have to audit that segment of the algorithm to determine whether the coordinate system is being oriented properly and consistently.

The current system has not been optimized for speed and continues to take an average of approximately 15 s per image to produce the CAD output. Actual times vary substantially for each case and depend on the number of images and candidate AIH regions to be evaluated by the classification system. This can mean that the time required to perform the CAD can be an important factor in the design of an integrated CAD system.

4. CONCLUSIONS

4.1 Detection of AIH

In a departure from our prior assessment, we selected random AIH cases to test against the CAD system from an institution in the United States as opposed to the original study that was performed in Hong Kong. It appears that some component of the CAD system may not be functioning optimally in response to these differences which is a normal outcome in CAD when changing population groups. That the core of the algorithm has not changed suggests that the performance decrement may stem from discrepancies in acquisition protocols between the two sites, in the knowledge-based classification system, or the combination thereof. More investigation is needed and is ongoing.

If we compare the sensitivity of the CAD system on a per lesion basis from our previous study [6] (training cases with 84.4% and validation cases with 82.6%) to the sensitivity of candidate lesion identification in this study (84.2%), we find that the performance of the identification component of the CAD system appears consistent. This may suggest that the source of our underlying performance issue lies in the elements associated with the classification system.

This dataset of random AIH cases included studies in which multiple AIH lesions existed concurrently and throughout the entire brain. The CAD system identifies candidate AIH regions in part by comparing pixel intensity in presumably normal contralateral anatomic regions. This process is particularly useful in identification of lone AIH. However, when multiple AIH exist in bilateral distributions across hemispheres, this procedure can be compromised yielding unexpected results. These conditions may account for some the false negative results.

4.2 Future Development

CAD systems development continues to be a major activity in the field of radiology. Recently, CAD systems have been applied to polyp detection in CT colonoscopy [7], mammography [8], pulmonary nodule detection [9], and traumatic brain injury [10]. Detection of small AIH continues to be an area amenable to the development of computer algorithms and CAD systems accurate enough to assist clinicians in emergent conditions. The combination of the results of this study and our prior study still support the idea that our CAD system can be refined into an integrated practice aid.

In the near term, we will need to isolate the source of the performance decrement in the CAD system to re-attain the levels established in our prior study. The success of AIH lesion identification seems unaltered lending support to the overall approach. A more in depth comparison of the captured images from LAC+USC versus Princess Margaret hospital in Hong Kong [6] and a review of the knowledge-based classification system may yield a quick solution to this current issue and a resolution to our false positive controls. We will likely rerun these acquired cases and controls through a new iteration of the CAD algorithm before continuing with our plan to perform a multisite assessment of the system.

Next, as most institutions upgrade their equipment to multislice CT scanners, the adaptation of this CAD algorithm to thin-slice images generated from these new machines may yield more accurate results. Such images are less prone to artifact and the contrast between AIH and normal parenchyma is higher as thinner sections make volume averaging less of a problem, even for small lesions. Furthermore, AIH may be better characterized in these newer acquisitions by their three dimensional morphology or interrelation. This is planned for the near future as we collect more cases during the clinical evaluation.

The design of a fully integrated CAD system will need to take into consideration the current processing time of the algorithm. The MATLAB code is currently not optimized for speed and takes an average of 15 s per slice. This means that the aggregate time to completion for the CAD algorithm can extend into the minutes. Such duration affects workflow considerations, particular when we consider the intended audience and role of this tool within the clinical milieu. The future incarnation of the CAD system will need to reflect either code migration to a better performing programming language or optimized MATLAB code to achieve gains in speed. If these remedies and/or hardware investment do not decrease the latency, the architecture of the CAD system may require non-real time or parallel processing prior to a study's arrival at a reading station.

We continue to envision a CAD system that can be implemented into daily clinical practice in the emergency department. It is believed that the system can be used as a triage tool for patients suffering from minor neurological disturbances or head trauma. After CT is performed, clinicians can read the images with support from the CAD system and make the appropriate treatment and management decisions.

REFERENCES

[1] Perry JJ, et al., "Attitudes and judgment of emergency physicians in the management of patients with acute headache," Acad Emerg Med 12(1), 33-7 (2005).

[2] Mader TJ, Mandel A., "A new clinical scoring system fails to differentiate hemorrhagic from ischemic stroke when used in the acute care setting," J Emerg Med 16(1), 9-13 (1998).

[3] Schriger DL, et al., "Cranial computed tomography interpretation in acute stroke: physician accuracy in determining eligibility for thrombolytic therapy," JAMA 279(16), 1293-1297 (1998).

[4] Mysoki MG, et al., "Head trauma: CT scan interpretation by radiology residents versus staff radiologists," Radiology 208(1), 125-128 (1998).

[5] Erickson BJ, Bartholmai B., "Computer-aided detection and diagnosis at the start of the third millennium," J Digit Imag 15(2), 59-68 (2002).

[6] Chan T., "Computer aided detection of small acute intracranial hemorrhage on computer tomography of brain," Computerized medical Imaging and Graphics 31, 285-298 (2007).

[7] Baker ME, et al., "Computer-aided detection of colorectal polyps: can it improve sensitivity of less-experienced readers? Preliminary findings," Radiology 245, 140-149 (2007).

[8] Fenton, JJ, et al., "Influence of computer-aided detection on performance of screening mammography," N. Engl. J. Med 356, 1399-1409 (2007).

[9] Beyer F, et al., "Comparison of sensitivity and reading time for the use of computer-aided detection (CAF) of pulmonary nodules at MDCT as concurrent or second reader," Eur. Radiol. 17, 140-149 (2007).

[10] Yuh EL, et al., "Computer-aided assessment of head computed tomography (CT) studies in patients with suspected traumatic brain injury," J. Neurotrauma 25, 1163-1172 (2008).