The Hong Kong Polytechnic University

Department of Health Technology and Informatics

Computer Aided Diagnosis of Acute Intracranial Hemorrhage on Brain CT

by

CHAN TAO

A thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

February 2007

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____ (Signed)

Chan Tao (Name of student)

Dedicated to my beloved wife, Joanna,

my delightful children, Bernard and Chelsea

Abstract

Acute intracranial hemorrhage (AIH) is a major cause of neurological disturbance or complication of head injury. Its presence dictates different management strategy. In modern medicine, detection of AIH relies on the use of brain computed tomography (CT). But diagnosis of AIH can become difficult when the lesion is inconspicuous or the reader is inexperienced.

The objective of the current project is to develop a computer aided diagnosis (CAD) system that improves the diagnostic performance of AIH on CT by clinicians.

A total of 186 cases, including all 62 continuous cases that showed AIH not more than 1cm in size obtained during a 6 month period, and 124 randomly selected controls that were obtained during the same period, were retrospectively collected from the CT archive of Princess Margaret Hospital. The imaging diagnoses were established by consensus of two experienced radiologists.

A CAD was designed and implemented. It reads and processes standard DICOM image files. Intracranial contents are segmented from the CT images, which are then subjected to denoising and adjustment for CT cupping artifacts. AIH candidates are extracted from the intracranial contents based on top-hat transformation and subtraction between two sides of the image about the mid-sagittal plane. AIH candidates are registered against a normalized coordinate system such that the candidates are rendered anatomical information. True AIH is differentiated from mimicking normal variants or artifacts by a knowledge based classification system incorporating rules that make use of quantified imaging features and anatomical information.

The CAD algorithm was manually trained using 40 positive and 80 control cases. In the validation test using the remaining 22 positive and 44 control cases, the system achieved sensitivity of 83% for small (< 1cm) AIH lesions on a per lesion basis or 100% on a per case basis, and false positive rate 0.020 per image or 0.29 per case.

In the observer performance study, 7 emergency physicians, 7 radiology residents and 6 radiology specialists were recruited as readers of 60 sets of brain CT selected from the same 186 case collection, including 30 positive cases and 30 controls. Each reader read the same 60 cases twice, first without, then with the prompts produced by the CAD system. The clinicians rated their confidence in diagnosing each case of showing AIH in both reading modes. The results were analyzed using the multiple-reader, multiple-case receiver operating characteristic (MRMC ROC) paradigm, which showed significantly improved performance for emergency physicians, average area under the ROC curve (Az) significantly increased from 0.83 to 0.95 without and with the support of CAD. Az for radiology residents and specialists also improved, from 0.94 to 0.98 and from 0.97 to 0.98 respectively..

In summary, a CAD system which boasted high sensitivity and low false positive rate has been developed. MRMC ROC study confirmed that it can improve diagnostic performance of clinicians, especially emergency physicians. It is anticipated that such a system can reduce diagnostic errors and improve patient care when it is integrated in the clinical environment for daily operation.

Publications arising from the thesis

Peer reviewed papers

1. Chan T., *Computer Aided Diagnosis of Small Intracranial Hemorrhage*, to be published in the Special Issue on CAD, Computerized Medical Imaging and Graphics, 31(4/5) 2007.

2. Chan T., Huang H.K., *Effect of a Computer-aided Diagnosis System on Clinicians' Performance in Detection of Small Acute Intracranial Hemorrhage on Computed Tomography* (submitted).

Conference Presentations

1. Chan T., Huang H.K., *Computer Aided Diagnosis of Acute Intracranial Hemorrhage*, RSNA 2005

2. Chan T., Huang H.K., *An Illustrative Application of CAD in Neuroradiology: Intelligent Automated Detection System of Intracranial Hemorrhage*, RSNA 2005

3. Chan T., Huang H.K., *Computer-aided Diagnosis of Small Acute Intracranial Hemorrhage*, RSNA 2006

4. Chan T., Huang H.K., *ROC Analysis on the Effect of CAD on Clinicians' Performance in Detecting Acute Intracranial Hemorrhage*, RSNA 2006

Acknowledgements

I must express my deepest gratitude to my supervisor, Professor Bernie H.K. Huang. Professor Huang has not only been instrumental in my study, but inspirational to my outlook in life. He has enlightened me in the area of research in imaging informatics and guided me tirelessly throughout my period of study. I am most impressed by Bernie's enthusiasm towards the development of imaging informatics technology. He has always been unbelievably energetic. I am most proud to have been his student and hope to follow his footsteps in pursuing my aspirations.

My colleagues and friends in the Health Technology and Informatics Department, Dr. Maria Law, Dr. F.H. Tang, Dr. Phoebe Chan, and Dr. Lawrence Chan have all been extremely kind to me. My zeal could not last long without their support and guidance. They have provided me with practical tips in various technical problems at different stages of my project.

Thanks to Bernie, I had the opportunity of visiting the Image Processing and Informatics Laboratory of University of Southern California. Dr. Brent Liu has taught me a lot regarding PACS, both in classroom and visits to clinical facilities. Dr. Arkadiusz Gertych has advised me on image processing and programming skills. I have also enjoyed the hospitality and company of Michael, Lucy, and Bing.

Finally, I would like to thank my previous colleagues at Princess Margaret Hospital, Hong Kong for their kind assistance in collecting clinical materials and participating in the observer performance tests, especially to Dr. Lily Chiu, Dr. M.C. Au Yeung, and Dr. C.M. Chan.

Contents

List of Figures

List of Tables

List of Abbreviations

AC	anterior commisure
AIH	acute intracranial hemorrhage
AUC	area under ROC curve
Az	area under ROC curve
CAD	computed aided diagnosis/detection
CSF	cerebral spinal fluid
CT	computed tomography
DICOM	Digital Imaging and Communication in Medicine
EDH	extradural hemorrhage
GUI	graphical user interface
HU	Hounsfield unit
ICH	intracerebral hemorrhage
IVH	intraventricular hemorrhage
IPH	intraparenchymal hemorrhage
kV	kilovolt
mA	milliampere
MRI	magnetic resonance imaging
MSP	mid sagittal plane
MRMC	multiple readers multiple cases
OML	orbital-meatal line
PC	posterior commisure
ROC	receiver operating characteristic
SAH	subarachnoid hemorrhage
SDH	subdural hemorrhage

Chapter 1 INTRODUCTION

1.1 CLINICAL BACKGROUND

Definition of Acute Intracranial Hemorrhage (AIH)

Acute intracranial hemorrhage (AIH) literally means recent bleeding inside the confine of the skull. This includes bleeding inside or outside the brain, which are termed intraaxial and extraaxial hemorrhage respectively. Intraaxial hemorrhage include bleeding into the brain substance, or intraparenchymal hemorrhage (IPH) and bleeding into the ventricular system, or intraventricular hemorrhage (IVH). Intraparenchymal hemorrhage can be further specified according to the exact anatomical location of bleeding, e.g. thalamic hemorrhage and brainstem hemorrhage. Extraaxial hemorrhage is classified according to the anatomical layer of meninges where bleeding occurs, namely extradural hemorrhage (EDH) – between the periosteal leaf and meningeal leaf of dura mater, subdural hemorrhage (SDH) – between dura mater and arachnoid matter, and subarachnoid hemorrhage (SAH) – between arachnoid mater and pia mater. The descriptive terms for the timing of intracranial hemorrhage is defined by the time elapsed since the onset: hyperacute hemorrhage refers to bleeding between 0-4 hours, acute: 5-72 hours, subacute: 4-21 days, and chronic: > 22 days(Cohen and Wayman 1992). In the following sections,

both hyperacute and acute hemorrhage are considered as AIH.

Causes of AIH

The most common cause of AIH is trauma. Both blunt and penetrating head injury can cause all kinds of AIH, depending on the mechanism and extent of injury. EDH occurs when an artery supplying the dura is torn and is frequrently associated with skull fracture. SDH is usually result of torn veins that bridge the brain and venous sinuses. SAH can result from trauma to vessels in the leptomeninges or surface vessels or the brain surface itself. Intracerebral hematoma may occur in shear injury in both the acute and delayed stage. IVH results mostly from rotationally induced tearing of subependymal veins (Taveras 1996; Zimmerman, Gibby et al. 2000).

In addition, major causes of AIH include hemorrhagic infarction, hypertensive hemorrhage, aneurysms, vascular malformations, intratumoral hemorrhage, vasculitis, dural sinus thrombosis, mycotic aneurysm, amyloid angiopathy, bleeding dyscrasias or anticoagulation therapy(Taveras 1996; Zimmerman, Gibby et al. 2000; van Gijn and Rinkel 2001; Panagos, Jauch et al. 2002).

AIH can produce variable neurological symptoms, dependent on the different functions served by the affected regions. In general, it is not possible to differentiate between hemorrhage from other causes of neurolgoical disturbances, e.g. ischemia,

central nervous system infection, or neoplasm, based on clinical findings (Taveras 1996; Mader and Mandel 1998; Zimmerman, Gibby et al. 2000; Panagos, Jauch et al. 2002; Perry, Stiell et al. 2005; Mullins 2006).

Clinical Significance of AIH

Identification of AIH is of crucial clinical significance, because its presence necessitates very different management strategies. For example, patients suffering from ischemic type of stroke can benefit from thrombolytic therapy, anticoagulant and/or antiplatelet therapy. On the contrary, such treatments are contraindicated in patients suffering from hemorrhagic type of stroke (Beauchamp, Barker et al. 1999; Adams, Adams et al. 2005), whilst patients suffering from hemorrhagic stroke may rapidly deteriorate due to increased intracranial pressure and may require neurosurgical intervention (Panagos, Jauch et al. 2002). Also, victims of head injury and intracranial hemorrhage require close monitoring and may benefit from surgical evacuation of blood clots; whilst people suffering from minor head injury without intracranial hemorrhage evident on CT can be discharged after shorter period of neurological observation (Jagoda, Cantrill et al. 2002). Similarly, patient presenting with headache may suffer from SAH that requires hospitalization and further treatment, whereas patient with normal imaging findings can be followed up (Mark and Pines 2006). Therefore it is obvious that a reliable method that demonstrates AIH

is essential for evaluation and proper management of patients presenting with neurological symptoms or head injury.

In addition, it is also important to predict the prognosis even for patients with known problems, which is also often related to the presence or absence of AIH. For example, presence of AIH could significantly alter the outlook of stroke patients, because the hemorrhagic type fare much worse that the ischemic type (Broderick, Adams et al. 1999). Also hemorrhage complicating central nervous system infection or neoplasm can significantly worsen the prognosis.

1.2 AIH ON COMPUTER TOMOGRAPHY (CT)

Clinical Issues

CT has been the imaging method of choice for evaluation of patients suffering from suspected stroke, head injury, sudden headache, and altered level of consciousness, with strong evidence from randomized trials (Broderick, Adams et al. 1999). This is because CT is widely available in most hospitals and many imaging centers, quick to perform, lacks contraindications, adaptable to patients requiring life support equipment, and sensitive to AIH (Diehl 1978; Taveras 1996; Bagley 1999; Zimmerman, Gibby et al. 2000).

With virtually ubiquitous availability, the trend has been to liberally scan patients for

the investigation of suspected neurological emergencies or head injury. For certain clinical scenarios, e.g. minor head trauma and seizure, there are guidelines for judicious of CT based on some clinical indicators (Jagoda, Cantrill et al. 2002). However, it is generally accepted that low threshold for for CT is necessary to avoid missing clinically silent but potentially lethal pathology. It has been shown that nearly a third of acute abnormalities demonstrable on CT would be missed if emergency physicians were to select patients based on their clinical impressions. The same study indicated that the type of pathology as predicted from clinical information matched the actual abnormality in only two-thirds of the time. (Reinus and Zwemer 1994).

It was recognized early in the inception of CT that it is extremely valuable in differentiating between hemorrhagic and ischemic disease (Ambrose 1973). The clinical utility of CT for evaluation of acute hemorrhage has been remarkable even for the earliest scanners. EDH mortality dropped by more than a half in one center before and after CT was introduced (Cordobes, Lobato et al. 1981). With increasing recognition that CT can safely rule out AIH and other acute abnormalities, use of CT is also attributed for the declining hospitalization rate of patients suffering from head injury (Thurman and Guerrero 1999).

Other imaging modalities are supplementary for initial evaluation of patients suffering

from neurological emergencies. Skull radiograph has little role to play even for patients suffering from head injury, because it does not correlate well with presence of intracranial lesion (Jagoda, Cantrill et al. 2002). Angiography is considered for patients without clear cause of hemorrhage who are surgical candidates. MRI and MRA are helpful and may obviate the need for angiography in selected patients (Broderick, Adams et al. 1999). For patients with subacute or chronic symptoms, the choice between CT and MRI is less clear. This is because MRI is better suited for detection of underlying pathology of hemorrhage, whilst CT becomes less sensitive as blood clot resolves with time (Shah and Kelly 1999).

Physical Principle

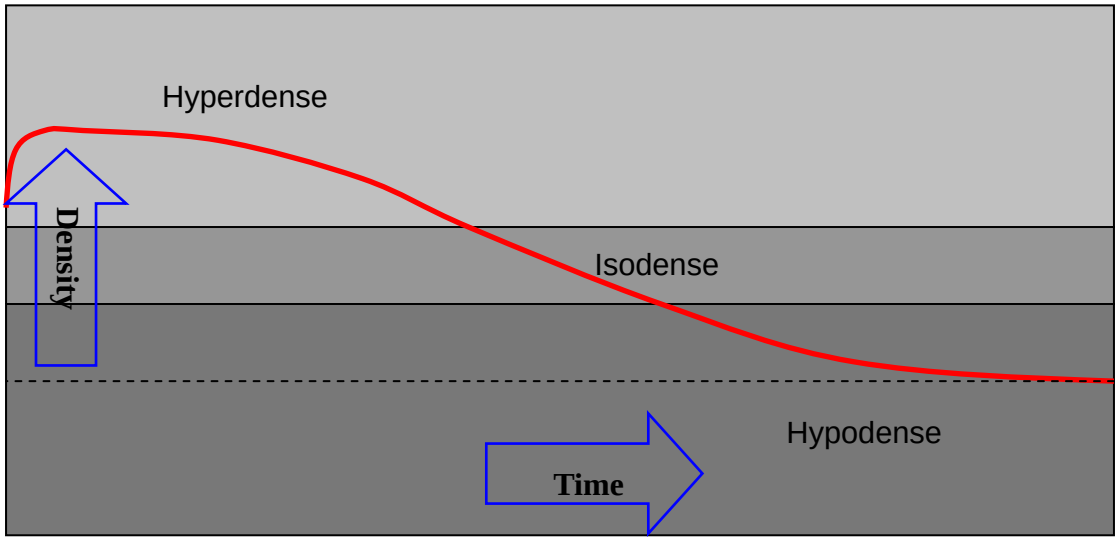
In CT, thin beams of x-ray pass through the subject from multiple different projections. The intensity of x-ray beams is attenuated to different degrees, due to absorption by both photoelectric interaction and Compton scattering interactions, which in turn depend on the electron density of the material being irradiated. The intensity of the exiting x-ray is measured, the data integrated, and a digital image is produced.

The digital image contains pixels which intensity is proportional the degree of attenuation of x-ray in the corresponding volume of tissues. The pixel values are

normalized against the attenuation of water, which is assigned the number of zero, and magnified by a constant. When the constant is 1000, the unit of pixel value or CT number becomes the Hounsfield unit (HU), which is definied as $1000 \times (\mu_x - \mu_w)/\mu_w$, where μ_x = pixel value of tissue and μ_w = pixel value of water.

By this convention, water is 0 HU, air is -1000 HU, and dense cortical bone or calcicum can be above +2000 HU. AIH, is usually shown as regions that show higher attenuation relative to the lipid rich brain parenchyma, as illustrated in figure 1.

Figure 1 - Change in density of intracranial hemorrhage with time
Acute intracranial hemorrhage is typically hyperdense relative to brain parenchyma, due to concentration of hemoglobin in blood clot. Hyperacute hemorrhage (< 4 hours) may be isodense to brain occasionally. With time, density of blood clot gradually decreases and becomes the same as that of brain at around 1 week. Blood clot lysis further progresses, hence its density further decreases, until it is similar to that of CSF months to years afterward the initial event.



Visualization of an acute clot on CT depends on its intrinsic physical properties including the density, volume, location, and relationship to surrounding structures, and technical factors including scanning angle, slice thickness, and windowing (Cohen and Wayman 1992). The attenuation value of acutely clotted blood depends on its density, which is primarily determined by concentration of hemoglobin (New and Aronow 1976; Norman, Price et al. 1977). The average attenuation value of AIH has been reported by different authors to be in the range of between 50-90 (Ambrose 1974; Scott, New et al. 1974; Cohen and Wayman 1992). These values need to be considered with reference to those of the brain, which forms the background on which AIH is shown. The brain comprises white matter and gray matter, which have been reported to show average attenuation of between 29-30 HU and 35-38 HU respectively (Phelps, Hoffman et al. 1975; Weinstein, Duchesneau et al. 1977). It is noteworthy that these values are much lower in infants, when the brain are richer in lipids (Boris, Bundgaard et al. 1987). The difference in average attenuation between AIH and brain parenchyma is hence resolvable by all current scanners that readily demonstrate differences in contrast of approximately 0.4% or 4 HU ((Zimmerman, Gibby et al. 2000). This is why CT is sensitive for showing AIH which may be of barely higher attenuation than that of adjacent parenchyma. In fact, it was suggested that a difference of more than 4 HU to the surrounding tissue should be enough for

detection of hematoma that might necessitate surgery (Bergstrom, Ericson et al. 1977).

It is intuitive that AIH can become difficult to identify when it is small. Yet there is no accepted convention for classification of the size of intracranial hematoma. The long axis diameter and thickness in the transverse plane represent the conventionally measured dimensions of intraaxial hematomas and extraaxial hematomas respectively. Following this practice conveniently, the current study defines a lesion as small if it is (a) intraaxial hemorrhage having a long axis diameter 1cm or below, or (b) extraaxial hemorrhage having a thickness equal or less than 1cm. All the sizes quoted in the following discussions refer to either the long axis diameter for intraaxial hematomas or the thickness of extraaxial hematomas measured in the transverse plane.

Difficulties in Diagnosis of AIH on CT

It is obvious that diagnosis of AIH requires correct interpretation of the demonstrable AIH on CT. The aforementioned figures of attenuation values of AIH and brain may suggest that AIH is present whenever regions that show attenuation of 40 HU or above are found. But the actual range of attenuation values for AIH can be much wider, closer to the range of 40-110 HU as written by Taveras in his standard reference in neuroradiology (Taveras 1996). This is confirmed by our own experience

of measuring a large number of acute blood clots during the course of this study, the attenuation varies from 35 to 105 HU. Likewise, attenuation values of gray and white matter also vary widely. Although no such data is available from literature, we have found that the range (instead of the mean) of CT no. for gray and white matter range from 20-40 HU and 30-55 HU. Therefore CT numbers of AIH and normal brain can overlap significantly.

In addition, there are many occasions when regions that show high attenuation actually represent calcifications or brain substance that show higher than expected attenuation due to artifacts like CT cupping artifact, volume averaging with adjacent skull, and beam hardening artifact. Therefore, the knowledge of different patterns of normal variants or artifacts that may mimic AIH becomes essential when the judgment between genuine AIH and its mimics is required.

For these reasons, it is expected that considerable training is required for accurate interpretation of brain CT with respect to detection of AIH. Yet the experts in image interpretation, namely radiologists, are in short supply. In most parts of the world, acute care physicians, including emergency physicians, internists, or neural surgeons, are the only ones to read the CT images, especially during odd hours, when radiologists' expertise may not be immediately available (Sarkarati and Reisdorff

2002). However, the skill of acute care physicians regarding interpretation of brain CT has been shown to be imperfect. In one study, emergency physicians detected hemorrhage only 73% of the time (Schriger, Kalafut et al. 1998). Even in the better staffed institutes, radiology residents rather than specialists assume the responsibility of radiological diagnosis during nighttime. Another study has shown that even radiology residents could, albeit infrequently, overlook hemorrhage on brain CT (Wysoki, Nassar et al. 1998).

1.3 CAD FOR AIH

Concepts of Computer Aided Diagnosis

It has long been recognized that errors in image interpretation, including erroneous perception or analysis, are inevitable even for the best human observers. Such errors may be exacerbated as results of fatigue, inexperience, or environmental factors (Erickson and Bartholmai 2002). CAD has been developed based on the notion that computers making objective or quantitative use of image-derived and clinical information can improve the accuracy in detection of abnormalities or to provide a list of probable differential diagnoses.

Use of computer for quantitative analysis of medical images has been reported in the 1960's, when the ambition was to replace radiologists in detecting abnormalities.

Such attempts, known as automated computer diagnosis, were unsuccessful and now considered unrealistic. Starting in the 1980's, a different approach that aimed to produce computer output that could be useful for, but not replace, radiologists have come into being. This approach became known as computer aided diagnosis or CAD (Doi 2004).

CAD schemes are generally composed of several basic components, which include 1. image processing for enhancement and extraction of lesions, 2. quantification of image features e.g. size, contrast, shape, etc., and 3. data processing for distinction between normal and abnormal patterns, based on image features obtained in previous step. (Doi, MacMahon et al. 1999).

CAD is best suitable for tasks that are tedious, e.g. looking for small lung nodule in chest CT, and tasks that require assimilation of multiple imaging features, e.g. breast mass on mammography. CAD is also useful when trained observers are lacking (Summers 2003).

Applications of CAD

CAD for lesions on mammogram has been the most prominent development in the realm of CAD. It was well known that radiologists miss a significant proportion of cancers on screening mammography. The current accepted paradigm for CAD in

mammography is to use it as second reader. Various abnormalities demonstrable on mammogram have been tapped in previous and continuing studies. There have been efforts focusing on detection of abnormal microcalcifications, mass, and architectural distortion. It was the first CAD device to receive the Food and Drug Administration (FDA) approval in 1998. With availability of commercial products, the use of CAD has moved from laboratory to clinical practices around the world. Since then, clinical studies that assess effect of CAD on radiologists' performance and clinical outcomes have accumulated (Freer and Ulissey 2001; Brem, Baum et al. 2003; Sacks 2003; Gur, Sumkin et al. 2004; Klym, King et al. 2004). Cancer detection rate has been increased by 8-19% in moderate sized clinical trials.

Another area that has seen FDA approved device is in detection of lung nodule, using chest radiograph or computer tomography. The products aim to find nodules that may represent early cancers that are amenable to treatment. Approaches based on measurement of change over time and evaluation of features from a single examination have been investigated. In one clinical study, CAD has detected significant lesions missed by routine clinical interpretation of chest CT in 1/3 of patients (Peldschus, Herzog et al. 2005).

Polyp detection in virtual colonoscopy is another area where CAD development has

been active. Features based on surface distortion, attenuation, and wall thickness have been found to be useful. A number of early clinical trials for this application have shown that sensitivities range from 70%-100%, associated with false positive findings in 2-8 per patient (Summers, Johnson et al. 2001; Yoshida, Nappi et al. 2002).

Assessment of bone age by hand radiography based on conventional atlas comparison method has been tedious and inexact task that radiologists have to live with for decades. Early success have been reported for automatic identification of bone age based on digital hand atlas (Pietka, Gertych et al. 2001).

Application of CAD in Neuroradiology

Applications in neuroradiology are mainly for quantitation of disease rather than for detection or diagnosis (Erickson and Bartholmai 2002). There was one published CAD system by Maldjian that aimed to detect acute middle cerebral artery (ischemic) stroke, making use of registration against the MNI coordinate system and comparison of attenuation over same anatomical region on both sides of the brain (Maldjian, Chalela et al. 2001).

Recent Attempts in CAD for AIH

It is envisaged that CAD may help to improve the accuracy in detection of AIH and hence decrease the risk of misdiagnosis and mismanagement. To date, there has been

no published work in the CAD of AIH to the best of our knowledge. Some related presentations on this topic do exist in recent conferences. Hodgson introduced a system that detects intracerebral hemorrhage based on median filtering and thresholding and reported sensitivity of 97% in more than 100 clinical cases showing hematomas of variable sizes, but their scheme essentially ignored extra-axial hemorrhage by discarding peripheral most portion of intracranial contents (Hodgson, Wilson et al. 2004). Goto et al. developed a CAD system which aimed to detect extra-axial hematoma by subtraction between two sides of the brain after morphing using thin plate spline. Their system also attempt to differentiate EDH from SDH based on morphology of the hematoma; however they have not reported on the actual sensitivity and specificity of the system (Goto, Aizawa et al. 2005). Yang et al. developed a CAD for AIH which took into account of clinical information provided by clinician, including age and blood pressure, etc., but was only capable of showing some large hematomas in 18 clinical cases (Yang, Lim et al. 2005). It is noteworthy that none of these systems have reported success in detecting small AIH, which are the lesions that could be difficult to diagnose by both human observers and computer systems. It is believed that impact in clinical practice is limited if such systems could only detect sizable lesions, because sizable lesions are unlikely to be missed by clinicians in the first place.

1.4 OBSERVER PERFORMANCE STUDY

Current CAD schemes aim to assist readers in making diagnoses by providing quantitative analysis of radiological images. It does not have to be higher or even be on par with the performance level of radiologists. Instead, the potential gain is due to the synergistic effect obtained by the effort of both human and computer. The ultimate test of a CAD system is its additive value. In other words, a CAD system can only prove to be useful should the human observers using the system achieve better diagnostic accuracy than the observers would achieve on their own. It is possible that an apparently accurate CAD system may lead to wrong diagnosis rather than correct ones, e.g. when the reader takes wrong clues from the CAD and ignores correct CAD outputs, thus reducing the diagnostic accuracy of the human-computer diagnostic system. For that reason, investigations of possible human-computer interaction such as the receiver operating characteristic (ROC) studies are necessary (Doi 2005). The multiple-reader multiple-case receiver operating characteristic MRMC ROC paradigm has been commonly used in the evaluation of CAD systems, which is not only efficient in terms of resource requirement (i.e. fewer readers and cases are required for a specified precision), but also produces results that can generalize to the populations of readers and cases from which the samples were drawn (Dorfman, Berbaum et al. 1992; Wagner, Beiden et al. 2002).

1.5 OBJECTIVES

The primary objective of this research project is to develop a CAD system that accurately identifies small AIH to help in the management of patients suffering from head injury or acute neurological disturbance in an emergency setting.

The secondary objective is to perform an observer performance study on the CAD system using MRMC ROC paradigm, in order to establish the usefulness of the system to potential users.

Chapter 2 MATERIALS AND METHODS

2.1 CASE COLLECTION

A total of 186 brain CT studies, including 62 cases showing small AIH and 124 cases showing no AIH, were retrieved from the CT archive of the Princess Margaret Hospital in Hong Kong. All were cases performed on an emergency setting for evaluation of head injury or acute neurological disturbance. Prior or follow-up studies of the same patients were also retrieved when available. But these were used only for establishment of diagnoses, not as separate entries into the dataset. Follow-up studies that show AIH which has been imaged in the same episode of hospitalization were excluded. In other words, only the first set of images showing AIH was included for each patient. This is to avoid having multiple representations of the same patient. The studies were subsequently anonymized, apart from the sex and age of the patients.

All studies were acquired with a single detector CT scanner (HiSpeed CT, GE Medical Systems, Milwaukee, WI, USA). All images were axial images obtained parallel to the orbito-meatal line (OML) or cantho-meatal line, which is the line joining lateral canthus of the eye to the external acoustic meatus, the conventional plane used for acquisition of clinical brain CT images (Silverman and Brink 1998). They were obtained using 120kV and 80-200mA. Most (159) of the examinations

comprised 10mm thick sections throughout the brain, the standard protocol used in the local institutes for emergency studies. Nineteen of the examinations comprised 5mm sections through posterior fossa and 10mm sections through the rest of the head, performed for patients with suspected lesions in the posterior fossa. Eight were 5-7 mm sections obtained for small children.

All emergent brain CT studies performed within a 6 month period from September 2004 to March 2005 that showed small acute intracranial hemorrhage have been retrospectively collected. It is noteworthy that studies containing AIH larger than 1cm in addition to smaller lesions were excluded from this dataset . In total, 62 positive cases were collected. Of these, 40 cases were collected at the first phase of data collection, which constituted the training dataset of the CAD system. The remaining 22 cases were collected at the second phase and used for validation of the system.

The 22 cases were not included initially because they were not available in the temporary archive in the CT suite during the scheduled time of collection, probably being deleted from the workstations as routine housekeeping procedures between the intervals of the scheduled visits. They were subsequently retrieved from the permanent archive. Although the training and validation cases were not randomly allotted, we believe there were no plausible systemic biases involved in the process.

124 emergency CT cases, twice the number of positive cases, were selected randomly and used as controls. Final radiological diagnoses in this group included normal (93), chronic ischemia (18), acute ischemic stroke (10), and tumor (3). The composition of cases is listed in table 1.

Table 1 – Radiological diagnosis of collected cases

Radiological Diagnosis	**No. of Cases**	**Training**	**Validation**
Acute intracranial hemorrhage	62	40	22
Control			
Normal	93	65	28
Chronic ischemia	18	8	10
Acute ischemic stroke	10	6	4
Tumor	3	1	2

The separation of data into a training and a validation dataset provides unbiased estimates of the estimates of the ability of the algorithm to classify ‘previously unseen’ cases. It is recognized that more sophisticated cross-validation methods with resampling including the leave one out method have the advantage of producing estimates with lower standard error and hence results of more statistical robustness. But such cross-validation methods are difficult to perform because the classifiers used in the current scheme are not based on algorithms like artificial neural networks (ANNs) or discriminant functions (DFs) that can be trained and then erased automatically.

Before the aforementioned dataset was collected, 46 brain CT studies, including 25 cases that show AIH and 21 cases showing no AIH have been arbitrarily chosen and retrieved from the CT archive of the Princess Margaret Hospital in Hong Kong. These cases include hematomas that show variable sizes and types, including SAH of 1mm thick and IPH of 5.6cm long. These cases were used only during the early algorithm establishment but were not included in the subsequent training or validation processes, and should hence be ignored in the subsequent sections.

Approval has been obtained from the institute review board , Kowloon West Cluster Research Ethics Committee of Hospital Authority, Hong Kong, for retrospective collection of clinical and imaging data and their use in this project.

Establishment of Radiological Diagnoses

The radiological diagnoses in all the cases were established by consensus of two radiologists, who had 7 years and 11 years of experience in reading brain CT respectively. These radiologists did not participate in the subsequent observer performance study. In majority of the cases, the diagnoses were considered unequivocal by the radiologists. In a minority of cases, consensus was only established after reviewing prior or follow-up CT or MR images of the brain. These included 4 of the subsequently proven positive cases and 3 controls. In addition,

dimensions, locations, and type of individual AIH volumes were measured and recorded by radiologists for each study. One blood clot that spanned across several axial sections was counted as one volume in this study, rather than considered as several separate lesions. Altogether there were 123 contiguous volumes of small AIH, 77 in the training cases and 46 in the validation dataset, with well represented samples of each different type of AIH and different sizes.

2.2 CAD SYSTEM

All the anonymized DICOM CT images were transferred from the CT archive to a Pentium based PC running the Windows XP operating system. The CAD has been developed using MATLAB (The MathWorks, Inc., Natick, MA, USA). The system was programmed to read and process DICOM images in their indigenous format and file structure. DICOM is the acronym for Digital Imaging and COmmunication in Medicine. It is the de facto industry standard for file structure and communication protocol in various medical imaging equipment (*http://medical.nema.org*). The user selects the folder containing the series of images in question, all DICOM images in that particular folder are automatically loaded into the system. The images are subsequently sorted according to the table position where individual image is obtained. They are also scaled, and adjusted to the desirable contrast according to information

residing in DICOM headers of individual images.

Overview of the CAD Scheme

The flow chart of the CAD algorithm is illustrated in figure 2. The image processing and analysis methods used in the scheme are listed in table 2. To begin, intracranial contents are segmented by global thresholding and morphological operations, followed by image analysis methods. Noise reduction using median filtering and adjustment for CT cupping artifacts are performed. The intracranial contents are realigned into the conventional orientation after automatic localization of 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 left-right subtraction. Image features of the candidates are quantified. The candidate AIH are given anatomical context by registration against a normalized coordinate system purposely developed for this project. The features and coordinates are then used in the rule based classification system to reduce false positives due to normal variants and artifacts.

Figure 2 – Flow chart of the CAD scheme

Individual components and intermediary outputs after successive steps of the algorithm are illustrated. Basic components of a usual CAD, including image preprocessing, image segmentation, image analysis, and classification are all utilized. 1. intracranial contents segmented using thresholding and morphological operations; 2. preprocessing steps that reduce noise and CT cupping artifacts; 3. intracranial contents aligned by locating mid-sagittal plane and boundaries of the brain; 4. AIH candidates extracted using combined method of tophat transform and left-right comparison; 5. AIH candidates rendered anatomical meaning by registration against a purposely developed coordinate system; 6. genuine AIH distinguished from mimicking variants or artifacts by the rule based classification system, using both image features and and anatomical information. The ICH in right basal ganglia is correctly identified as genuine AIH and outlined in red, whilst the mimics are outlined in blue.

Details of individual steps are outlined in Table 2.

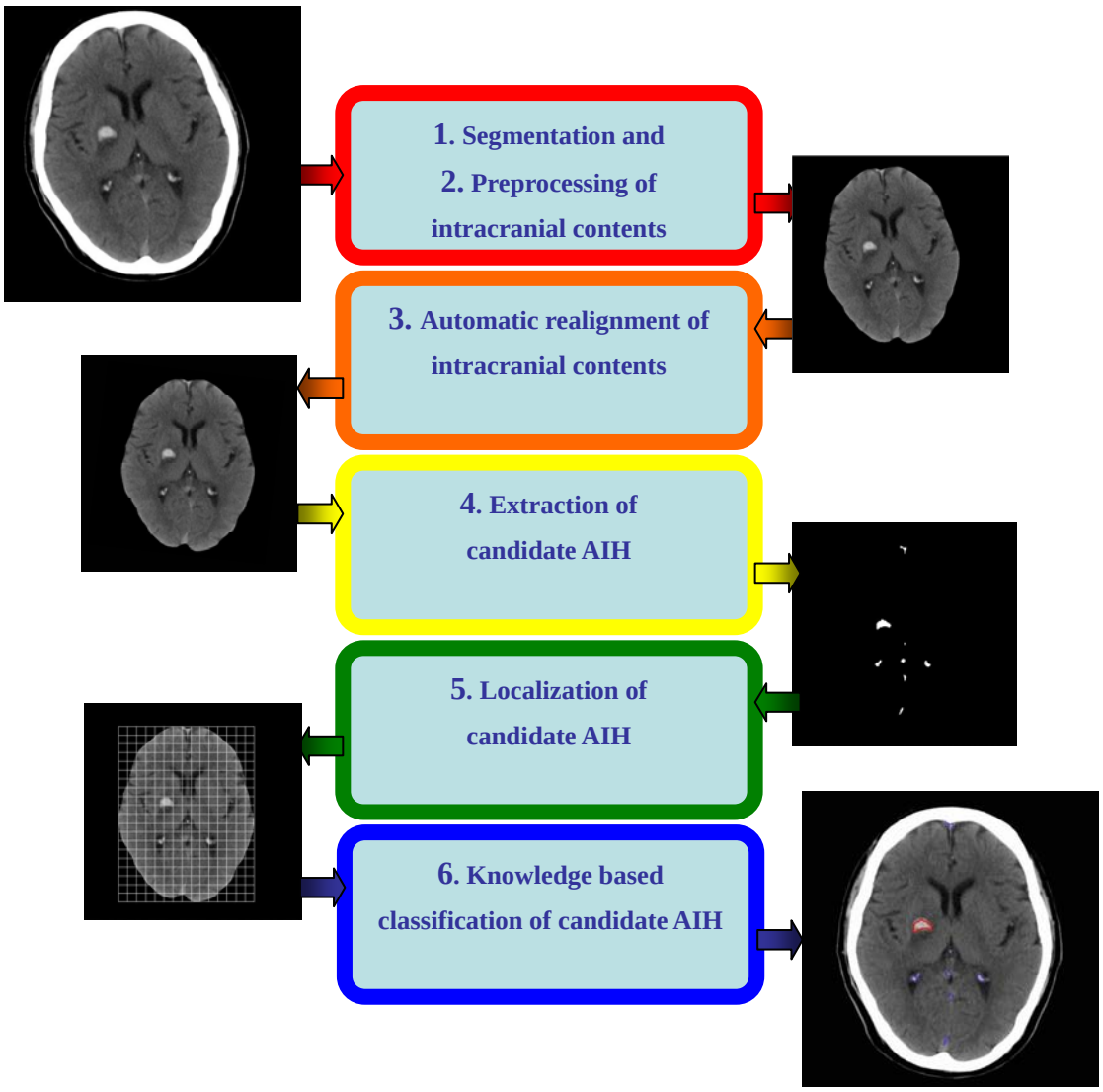


Table 2 – Details of image processing and analysis steps used in the CAD

Details of individual image processing and analysis steps of in the CAD algorithm as outlined in Figure 2. The numerals corresponds to those shown in figure 2.

Steps	Methods	Purposes
1. 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 tissues
2. Preprocessing of intracranial contents	Median filtering Adjustment of intensity according to distance from the skull	Denoising Correction for CT cupping artifacts
3. Automatic realignment of images	Automatic localization of limits of brain, ventricles, floor of anterior intracranial fossa, mid-sagittal plane	Align the brain into the normal position
4. Extraction of candidate AIH	Top hat transformation Subtraction between the two sides	Highlight local high density regions Extract asymmetrically high density region
5. Localization of candidate AIH	Registration of the brain in question against a normalized coordinate system	Render the candidate AIH anatomical information
6. 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 noises, artifacts, and normal variants

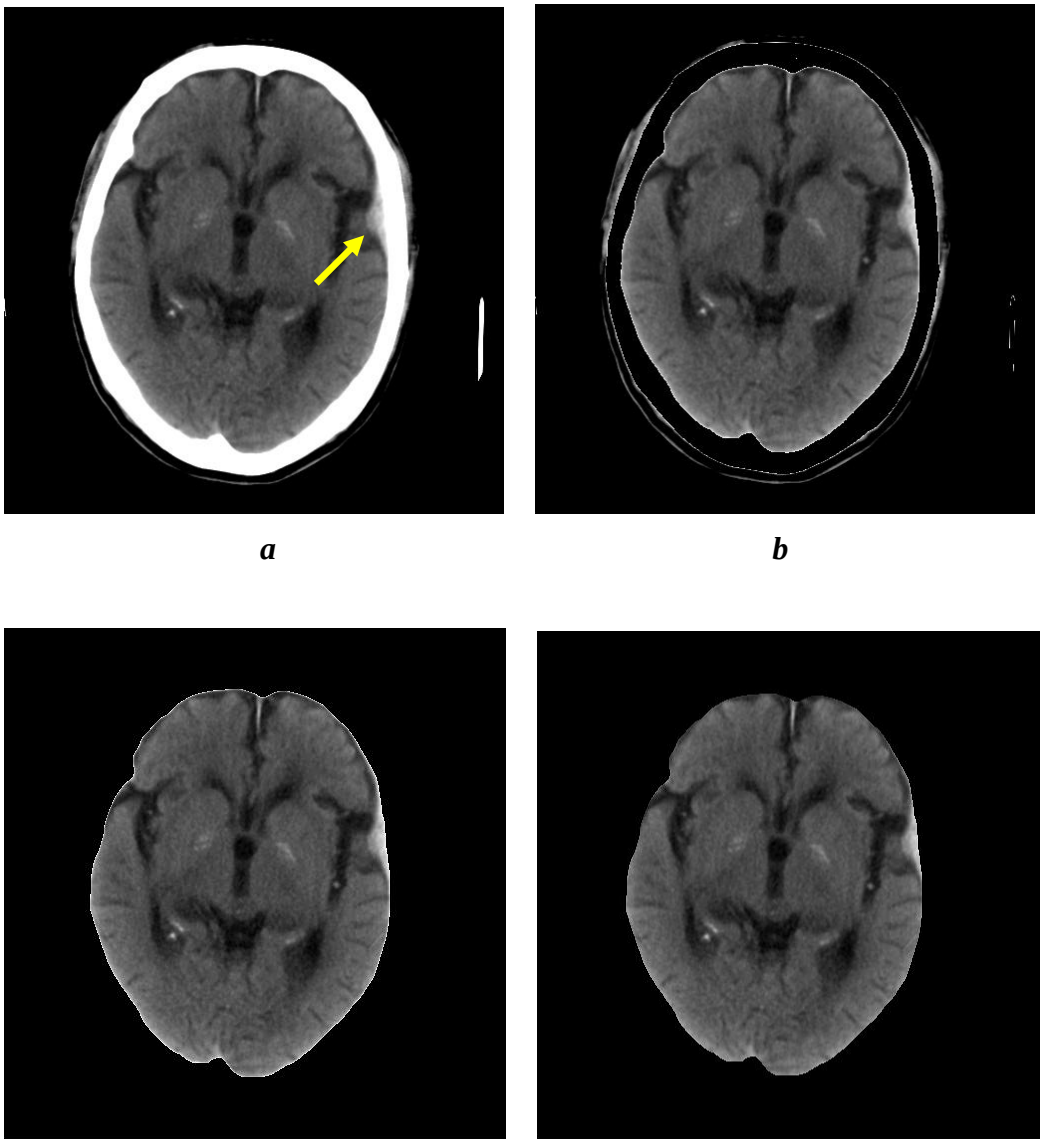
Segmentation of Intracranial Contents

Skull, by virtue of its exceptionally high attenuation values that range from 80 to > 2000 typical of bones, is first removed using global thresholding with a threshold value of 100. Lower thresholds that can more completely remove skull should not be used because this would remove significant proportion of AIH. Morphological opening using a relatively large structural element is performed to remove scalp or other small external structures. Finally, small remaining bits of bone around the perimeter of intracranial content are removed using morphological erosion using a small structural element. The process is illustrated in figure 3.

Note that extracranial structures may not be entirely removed if they are larger than the structural element used in the morphogical opening. These include unusually thick scalp thissues, e.g. resulting from swelling resulting from head injury, structures in the orbits and paranasal sinuses, and other extrinsic structures e.g. head rest. Such large extracranial structures can be distinguished from intracranial content by the fact that they are separated from the centrally located intracranial contents by regions of void that represent removed bones. The intracranial contents, including the brain and cerebral spinal fluid (CSF) containing spaces can subsequently be segmented by selectively removing elements which are not contiguous with the central component.

Figure 3 – Segmentation of intracranial contents

The segmentation of intracranial content is illustrated for the original image as shown in ***a***, which contains a small acute left extradural hematoma marked by yellow arrow. Note the high density band representing the head rest on the left of the image. First, global thresholding was used to remove the majority of the skull as shown in ***b***. Then, the scalp and external structures are removed by morphological opening using a relatively large structural elememt, resulting in ***c***. Finally, the small residual rim of skull around the perimeter of the intracranial content is removed using morphological erosion, producing the segmented intracranial content as shown in ***d***.



a ***b***

c ***d***

In the levels below the level of petrous temporal bone, the temporal lobes may become separated from the central component of intracranial content just like the orbits, but they are different in that the temporal lobes in the middle cranial fossa would be contiguous with the rest of the brain one section above, not so for the orbits or other extracranial structures. The process is illustrated in figure 4.

Realignment of Images

CT studies obtained in an emergency setting as in this study tend not to be perfectly positioned. This may occur when the patients are confused or in distress. But the images need to be properly realigned such that subsequent anatomical labeling process can proceed. To do this, location and orientation of the brain need to be determined.

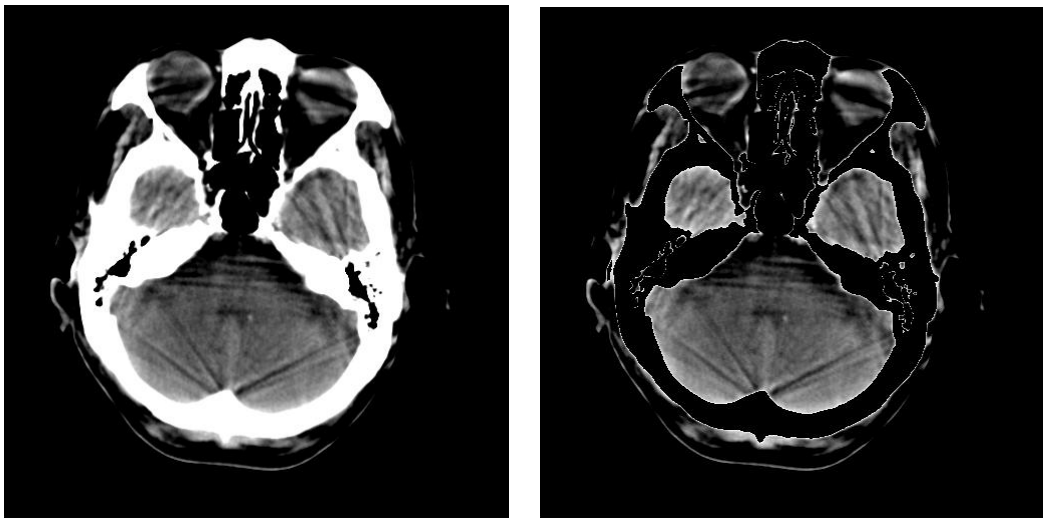
Boundaries of the Brain

The superior, anterior, posterior, and both lateral limits are easily located by finding the margins of the segmented brain.

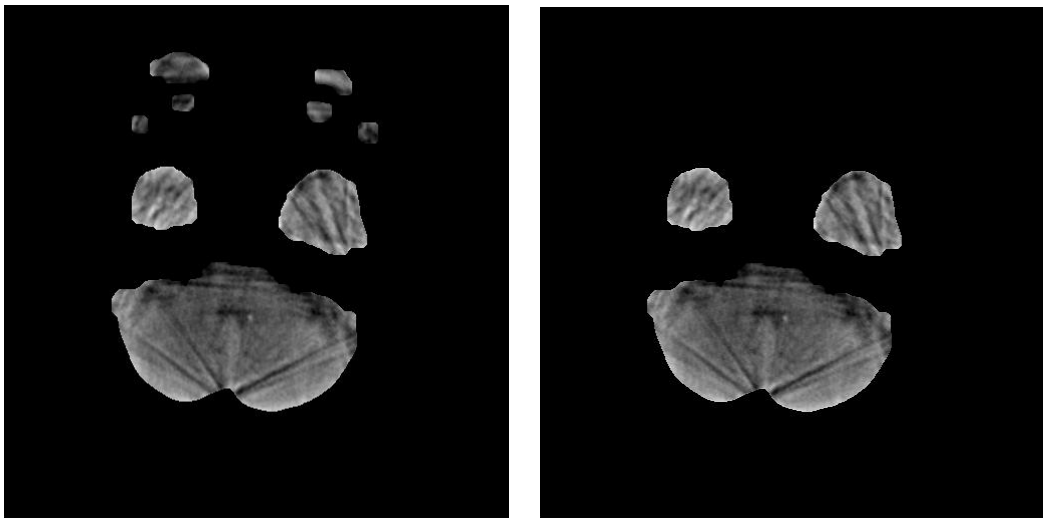
The landmark for the inferior limit of the cerebral hemisphere is at the lowest level of temporal lobes; however, during the development of the system, it was found that the floor of the anterior cranial fossa correlates more consistently with the location of

Figure 4 – Segmentation in lower portion of cranial cavity

The segmentation of intracranial content in sections below the level of petrous temporal bone is slightly more complicated. A such section of the same patient as shown in **figure 3*a*** is shown in ***e***. Bone is removed using thresholding, producing ***f***. The segmentation process as described also segments the orbits and scalp tissues over temporal fossa, as in ***g*** , which are removed based on their lack of continuity with the intracranial content in superior sections, providing the final result ***h***.



e ***f***



g ***h***

most of the internal structures that are relevant for localization of AIH. Because the floor of the anterior cranial fossa approximately parallels the axial plane on brain CT,

the area of the brain decreases significantly below the floor as the frontal lobe disappears from the section. Therefore the section where a significant drop in cross sectional area of the segmented brain, on comparison with the section above, marks the level of the floor of anterior cranial fossa. In addition, at the level of the anterior floor of cranial fossa, the shape of the cross section of the brain becomes irregular and less elliptical, because it is where the irregular bony floor of orbits come into the picture. In other words, the level of the anterior cranial fossa can be accurately identified as the plane where there is significant 1. decrease in cross sectional area of the segmented intracranial contents; and 2. decrease in relative area of the segmented intracranial content to its convex hull, going from superior to inferior axial sections.

Location of Mid Sagittal Plane by Symmetry

There were several reports of midline location based on symmetry of brain parenchyma (Junck, Moen et al. 1990; Minoshima, Berger et al. 1992; Ardekani, Kershaw et al. 1997), but the proposed approaches were not suitable for brains containing pathologies which alter the symmetry such as AIH. Our system locates the mid sagittal plane (MSP) based on the assumption of symmetry of CSF spaces, predominantly the ventricular system. It was observed that the ventricles are not significantly distorted when the AIH is small and there is no sizable IVH.

The image of the brain that contains the main bulk of the lateral ventricles is selected by locating the section that contains the highest ratio of CSF containing space to solid brain substance. This particular plane is chosen because the bulk of the lateral ventricles are large structures usually situated symmetrically about the midline, thus allowing the comparison of both sides for locating the midline. This section is then binarized with a threshold such that CSF containing regions take on the value of zero whilst the brain parenchyma takes on the value of one. The binarized image is then rotated about its centroid to a range of angles. The absolute differences between the rotated image and its mirror image are obtained. The angle of the MSP is the one that produces the least difference between the two halves of the rotated brain thus obtained. This process is illustrated in figure 5.

Manual Readjustment

Automatic localization of the mid sagittal plane may seldom be inaccurate due to exaggerated asymmetry in some normal subjects or distortion of anatomy due to pathological conditions such that the basic assumption of symmetry about the midline used by the algorithm is violated. Therefore, human intervention, although infrequently, may be required such that subsequent analysis can proceed. The system provides for this by incorporating an intermediary step when the observer can check

and adjust the computer deduced MSP.

Figure 5 – Automatic localization of mid-sagittal plane

The angle of the mid-sagittal plane is obtained by finding the line of symmetry of the body of lateral ventricles. The section containing the body of lateral ventricles ***a*** is automatically selected as described in text. The central portion of the image is binarized with threshold at CSF attenuation. It is rotated over a range of angles. The differences between the rotated image ***b*** and its mirror image ***c*** are obtained ***d***. The angle which gives the least difference ***e*** is the angle of the mid sagittal plane ***f***.

a ***b*** ***c***

d ***e*** ***f***

After the midline and level of upper boundary of anterior cranial fossa are obtained, both are displayed for the user to decide if the localization is satisfactory. If not, an interactive tool is available for the user to manually align the images.

We believe identification of basic anatomical structures is a task that emergency physicians can comfortably perform. In fact, during development of the system, 7 emergency physicians had taken part in an informal trial of the manual adjustment tool. All were able to accurately localize the mid-sagittal plane and lowest section of anterior cranial fossa in randomly selected image datasets.

Preprocessing of Intracranial Contents

Median Filtering

To reduce noise, median filtering using a 3-by-3 square kernel was applied. Median filter was chosen because it is less sensitive to extreme values and capable of removing outliers without reducing sharpness of the image.

Adjustment for Cupping Artifacts

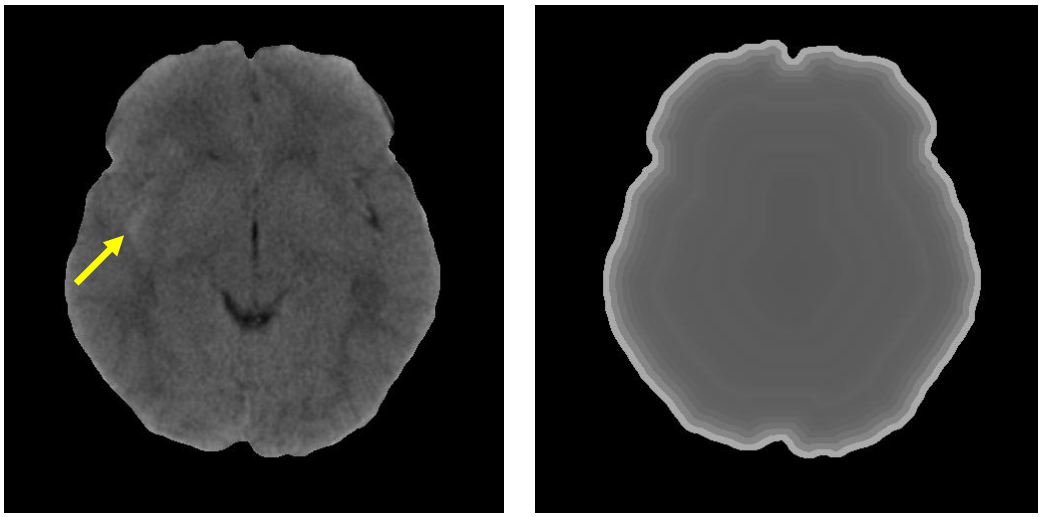
CT cupping artifacts apparently increase attenuation over regions subjacent to skull, which degree is about inversely proportional to distance from the soft tissue-bone interface (Barrett and Keat 2004). This artifact stems from directional dependent

variation in the amount of beam-hardening near the soft tissue-bone interface. Although most scanners attempt to correct for these errors, the correction is typically incomplete. CT cupping artifact has been a recognized problem for intensity based thresholding of intracranial structures (Ruttimann, Joyce et al. 1993).

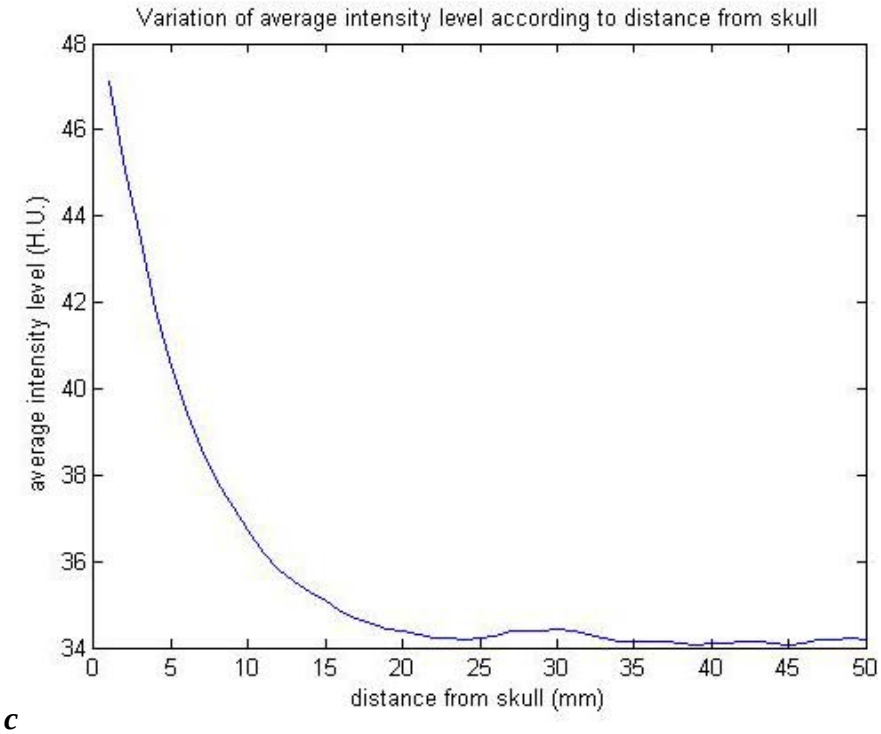
In the current application, these artifacts are particularly troublesome because immediately subjacent to skull are the locations where thin extraaxial hemorrhage are found, hence the peripheral regions cannot simply be discarded. Instead, regions closer to skull are identified as abnormal only if they show intensity significantly higher than those at a similar distance from the skull. To perform this adjustment, the average intensity as a function of distance from skull is obtained as a reference against which abnormal regions are compared. A series of successive morphological erosions is first performed on the binarized image of the intracranial contents. The difference between the image after the n^{th} erosion and the $n+1^{th}$ erosion yields the ring like mask at a distance proportional to n from the skull. Then average intensity of this ring can be obtained by correlation of this mask with the original image and divided by number of pixels in the mask. This essentially builds a map of average density for regions equidistant from the soft tissue-bone interface. The process is illustrated in figure 6. When the original image subtracts the map, the cupping artifacts is reduced. The median filtered image with reduced cupping artifacts **I(x,y)** is then ready for

Figure 6 – CT cupping artifacts and its adjustment

Image ***a*** represents the original image. Yellow arrow marks the subarachnoid hemorrhage. Image ***b*** is the map of concentric rings, which represent regions equidistant from the skull. The brightness of each ring reflects the average intensity level of corresponding regions in image a. The rings are obtained by successive morphological erosion of a mask of the original image. Image ***c*** is the plot of average intensity level as a function of distance from skull calculated for image ***a***.



a ***b***

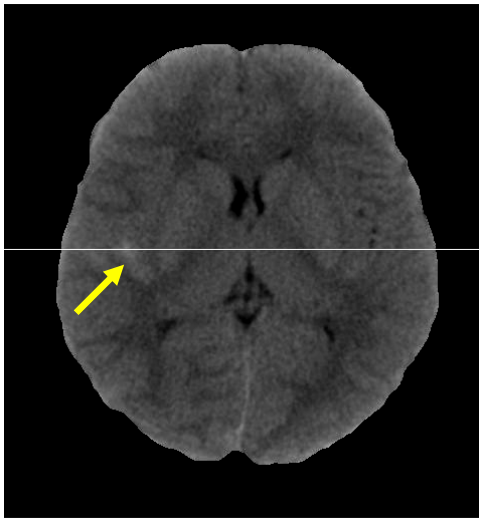


c

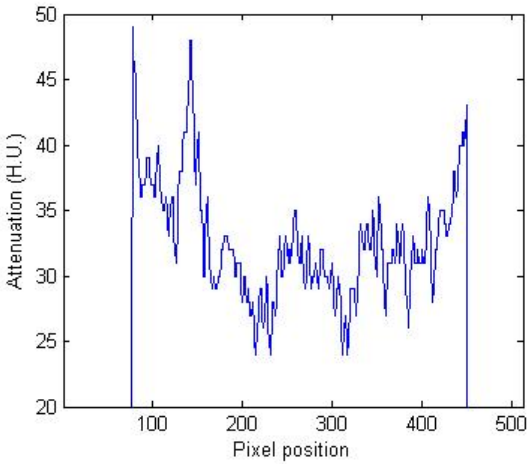
subsequent segmentation processes. The effect of the aforementioned preprocessing steps is illustrated in figure 7.

Figure 7 – Effects on images after preprocessing steps

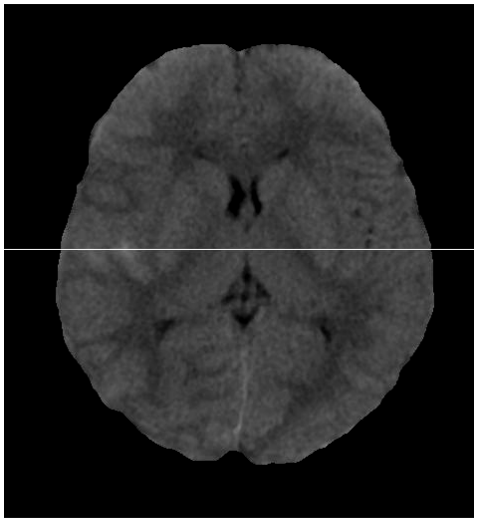
The original image with artifactual increase in signal intensity towards the brain skull interface is shown in ***a***. The horizontal line indicates the position from where the intensity profile ***b*** is obtained. The image after correction of cupping artifacts is shown in ***c***. The intensity profile along the same horizontal line is shown in ***d***. It can be appreciated by comparison of ***b*** and ***d*** that the peak intensity at the pixel position of around 140 due to AIH is more prominent after correction of cupping artifacts.



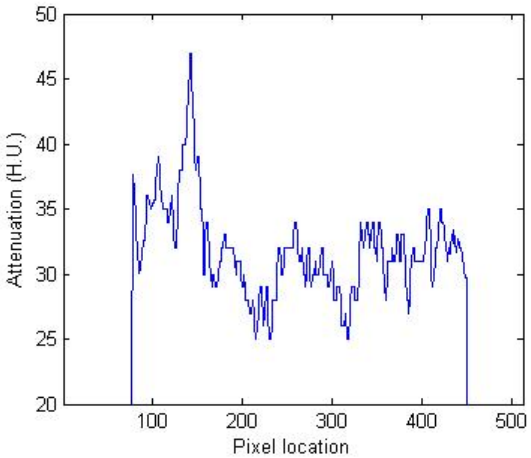
a



b



c



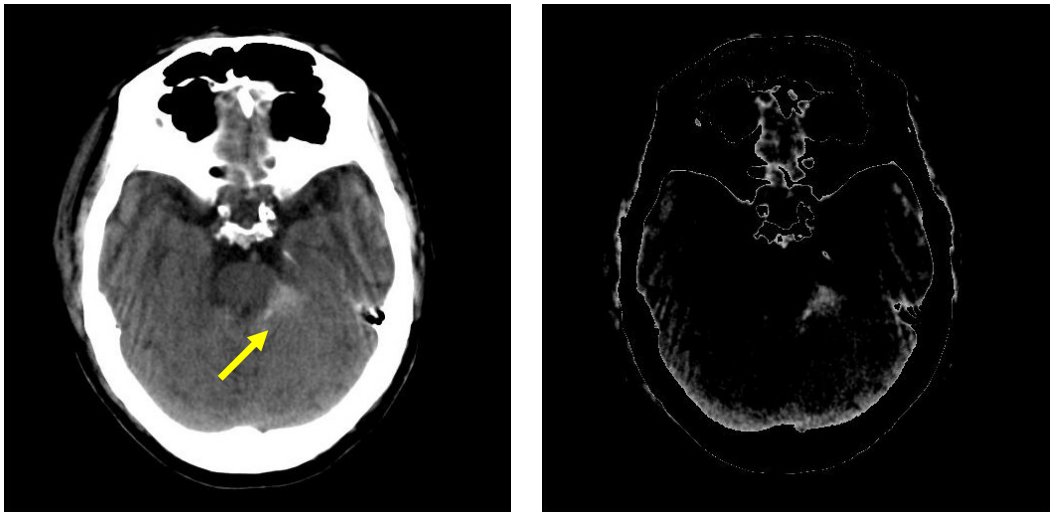
d

Segmentation of Candidate AIH

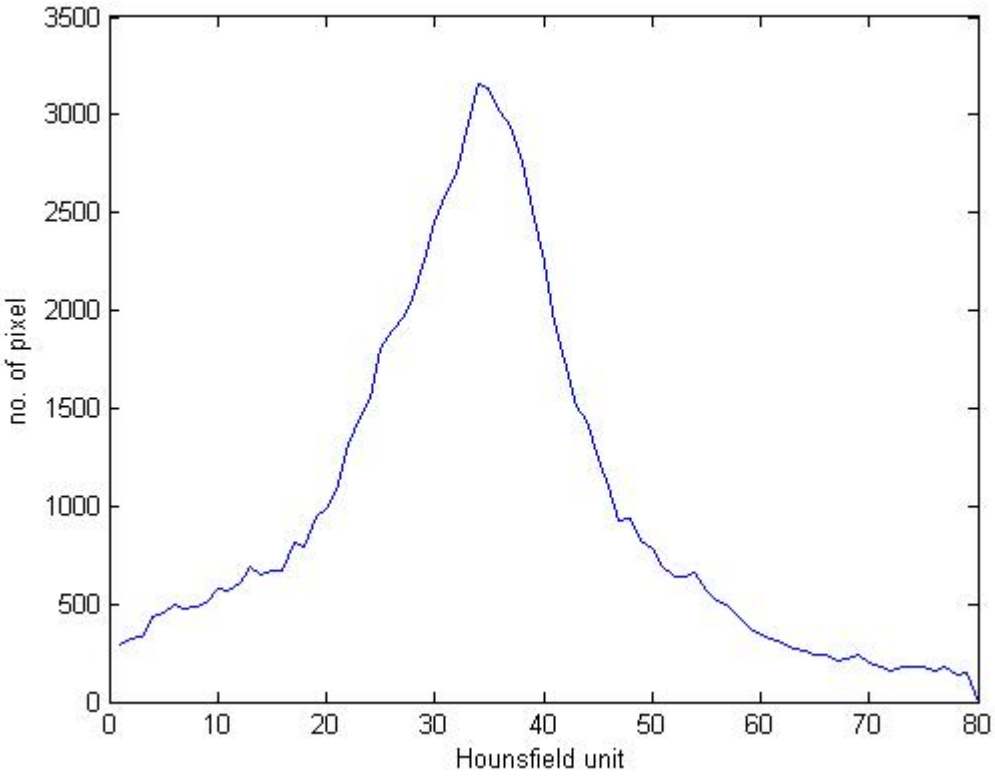
AIH is distinguished by its increased attenuation relative to normal intracranial structures. As mentioned in the introduction, we have found that the range (instead of the mean) of CT no. for AIH, gray and white matter to be 35-105, 30-55 and 20-40 respectively. It was therefore self-evident that simple thresholding of attenuation value would not work because the wide range of attenuation of AIH overlaps those of normal parenchyma. An illustration of the inadequacy of global thresholding for the segmentation of AIH is shown in figure 8. The problem is especially important for small AIH because the smaller lesions are more affected by partial volume averaging, which reduces the contrast and blurs the edge between the lesion and its adjacent parenchyma. Therefore global thresholding is especially difficult for detection of small AIH. In this system, two further considerations in addition to actual intensity value of the preprocessed image **I(x,y)** are obtained: 1. the intensity of a pixel above that of its immediate surroundings, obtained using image top-hat transformation; 2. its intensity difference above that of its contralateral anatomical region, obtained by subtraction of the flipped image from the original image.

Figure 8 – Problems with global thresholding

Image ***a*** shows acute subarachnoid hemorrhage in left ambient cistern as marked by yellow arrow. Image ***b*** is the tresholded image of *a* using lower and upper threshold of 50 and 70 respectively. The hematoma, as well as portions of the brain that show high attenuation due to artifacts are extracted. Image ***c*** is a plot of the number of pixels against the attenuation value shown on each pixel for the intracranial contents of ***a***.



a ***b***



Top-hat Transformation

Top-hat by opening or white top-hat transformation is the difference between the original image and its morphological opening (Gonzalez and Woods 2002; Soille 2003). Morphological opening, (denoted by o) of a image f by a structuring element b is the erosion (denoted by $\ominus$) of f by b, followed by the dilation (denoted by $\oplus$) of the result by b.

$$f \circ b = (f \ominus b) \oplus b$$

The gray-scale dilation of f by b is defined as

$$(f \oplus b)(x,y) = \max\{f(x-x',y-y')+b(x',y') \mid (x',y') \in Db\}$$

where Db is the domiain of b and f(x,y) is assumed to $-\infty$ outside the domain of f. This can be thought of rotating the structuring element about its origin and translating it to all locations in the image. At each translated location, the rotated structuring element values are added to the image pixel values and the maximum is computed. In practice, gray-scale dilation is usually performed using flat structuring elements, meaning that the value of b is 0 at all coordinates over which Db is defined. Hence the gray-scale dilation is simplified as

$$(f \oplus b)(x,y) = \max\{f(x-x',y-y') \mid (x',y') \in Db\}$$

And the gray-scale dilation becomes a local maximum operator, where the maximum is taken over a set of pixel neighbors determined by the shape of Db.

The gray-scale erosion of f by b is defined as

$$(f \ominus b)(x,y) = \min\{f(x+x',y+y')+b(x',y') \mid (x',y') \in Db\}$$

where Db is the domiain of b and f(x,y) is assumed to $+\infty$ outside the domain of f. Similar as discussed above, gray-scale erosion using flat structuring elements can become a local minimum operator and be simplified as

$$(f \ominus b)(x,y) = \min\{f(x+x',y+y') \mid (x',y') \in Db\}$$

Top-hat transformation is used to extract signal peaks on background of variable intensity levels, which are not entirely covered by the structuring element used for morphological opening. For the current application, as visualization of AIH depends primarily on the difference of attenuation between the lesion and its surrounding brain parenchyma that are of variable attenuation values, top-hat transform using a relatively large disk shaped structuring element is performed. A pixel value in the top-hat transformation image represents the difference of a signal peak above that of its surrounding in the original image, **I(x,y)**. Hence the difference in attenuation rather

than the absolute value is highlighted. This generate the first transformation for the segmentation **F(x,y)**.

F(x,y) = **top-hat transformation of I(x,y)**……....Equation 1

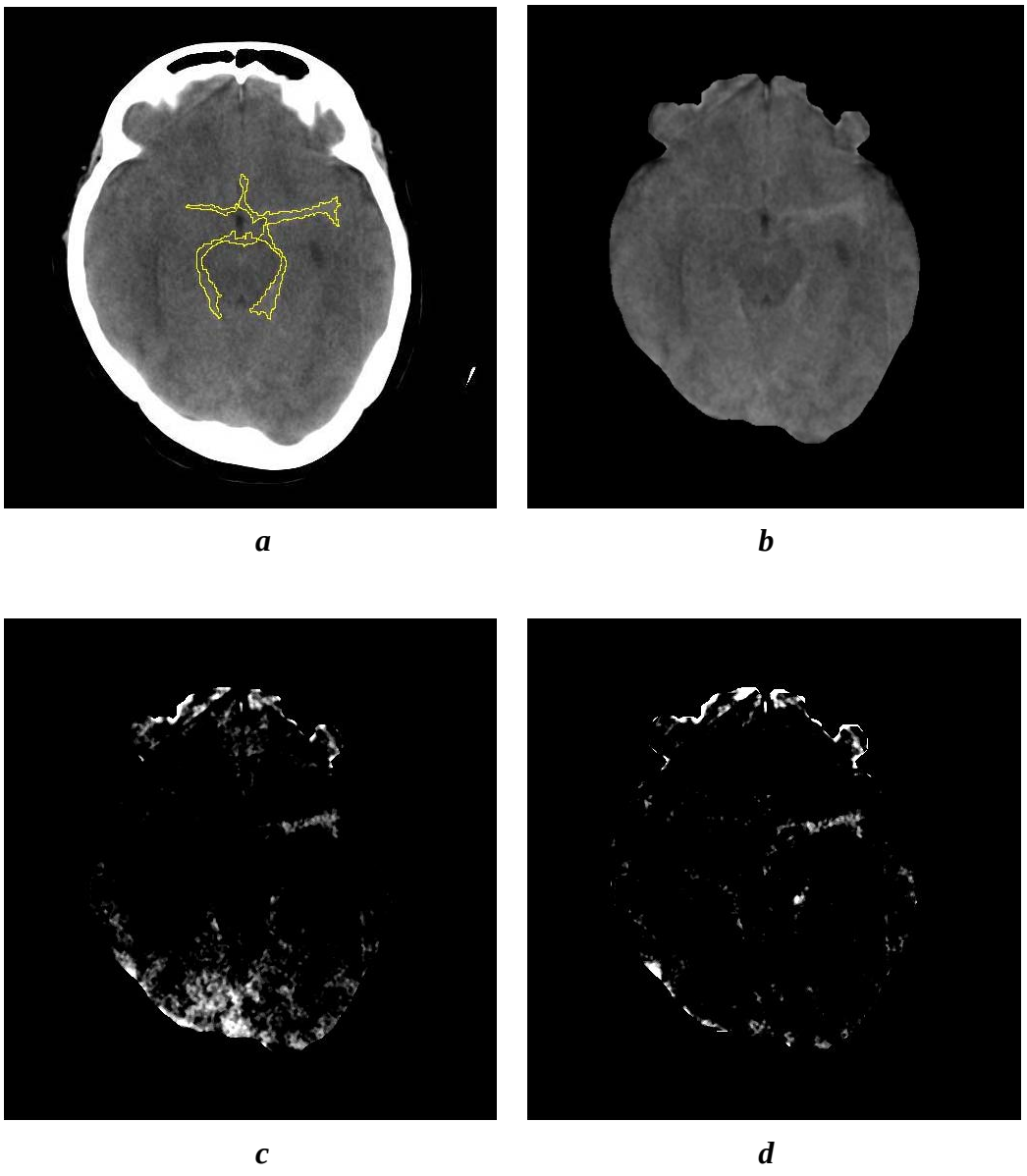
The effect of top-hat transformation is illustrated in figure 9.

Extraction of Asymmetrical High Intensity Region

In addition, regions which are brighter than their presumably normal contralateral anatomical regions are more likely to be abnormal. It is especially important for small extraaxial hematomas, which are only of marginally increased signal intensity than the adjacent gray matter, but are much brighter than the contralateral CSF containing space. Subtraction between the two sides is therefore performed to locate the regions that show higher attenuation than their contralateral counterparts. First, the brain is inverted along the mid sagittal plane. Minor asymmetry between the two sides of the brain is adjusted by elastic transformation of the flipped image. The control points for the transformation are obtained by intersections of 1. diagonals crossing the centroid of the mid sagittal plane and 2. the perimeter of the brain. This produces elastically transformed mirror image **J(x,y)**. Then the difference between the original image **I(x,y)** and the morphological closing transformation of **J(x,y)** is obtained for

Figure 9 – Effect of top-hat transformation

The section in ***a*** shows a case of subarachnoid hemorrhage, in which the AIH is outlined in yellow. The intracranial content after preprocessing is shown in ***b***. The contribution of top-hat transformation can be appreciated from the comparison between ***c***, which is obtained after global thresholding of ***b***, and ***d***, which is obtained after top-hat transformation of ***b***. While both ***c*** and ***d*** show the AIH in left Sylvian fissure, ***d*** , although not exactly like the outlined AIH in *a* is much cleaner than ***c***.



a ***b***

c ***d***

individual pixels to produce the image **G(x,y)**. This step is illustrated in figure 10.

G(x,y) = **I(x,y)** – closing of **J(x,y)**…….…….Equation 2

With these, a pixel, **p(x,y)** of an axial image is segmented as candidate AIH if the sum of the weighted averages of **F(x,y)** and **G(x,y)** of the pixel values in the aforementioned transformation images, exceeds a predefined threshold T (figure 11), T being determined through the use of the training set.

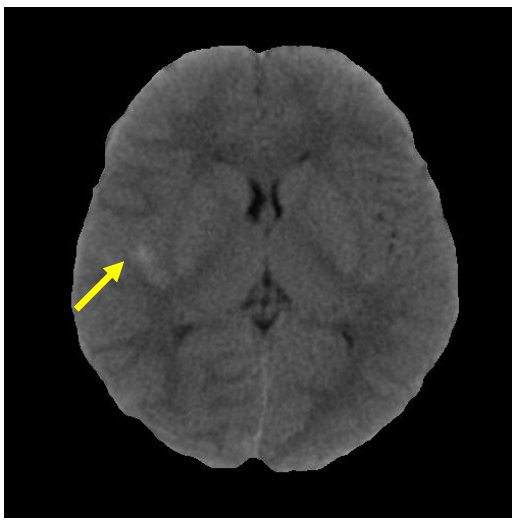
p(x,y) is a candidate AIH if w_1**F(x,y)** + w_2**G(x,y)** > T……..…..Equation 3

where w_1 and w_2 are the weightings of **F** and **G** respectively

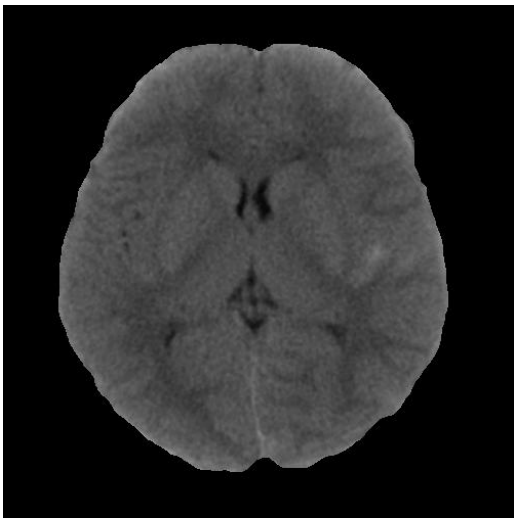
After thresholding with threshold value obtained by training, the AIH regions that are smaller than a predefined size are removed, to reduce artifacts due to noise that tend to be isolated peaks of small size instead of confluent hyperdense regions. The lower limit of size for AIH to be recruited for subsequent processing is variable depending on the level of the section and whether the region is at the periphery or central of the axial image, being higher towards the superior and inferior sectons and higher at the periphery or than the central portion of images. It is to reduce effect of noise and limit the number of AIH candidates for subsequent analysis. Then the margins of candidate

Figure 10 – Subtraction between sides of the image

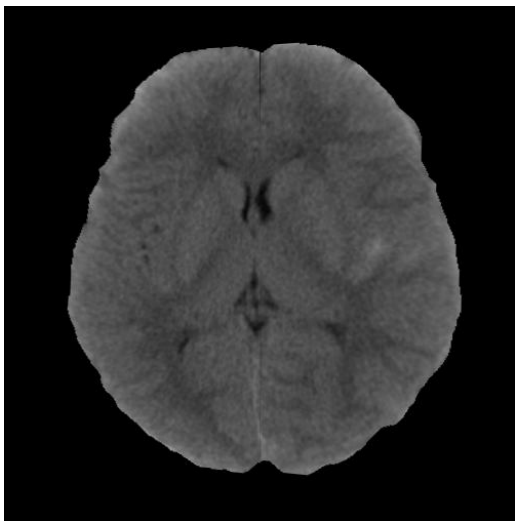
The procedure of highlighting asymmetrical high density region is illustrated by the intermediary outputs of the CAD scheme. The original is shown in ***a***. The flipped image ***b*** is elastically transformed to reduce the normal structural asymmetry between the two sides of the brain ***c***. The difference between ***a*** and the morphological closing transformation of ***c*** highlights AIH within the right Sylvian fissure is of higher signal than CSF in the left contralateral Sylvian fissure ***d***.



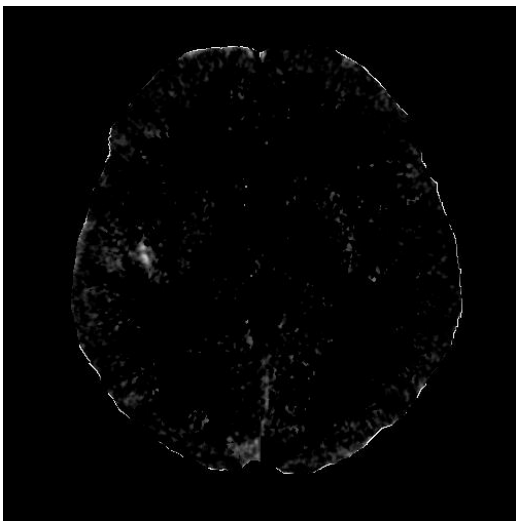
a



b – flipped ***a***

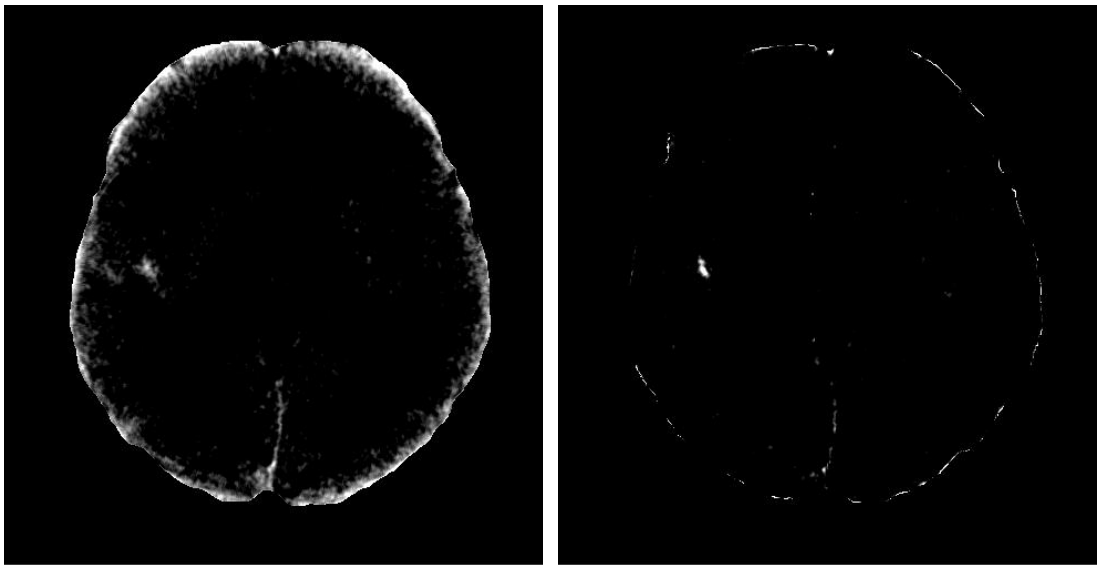


c – elastic transformation of ***b***



d – ***a*** subtracts closing of ***c***

Figure 11 – Effect of combined steps on extraction of AIH
Comparison of global thresholding of original image after removal of skull and extracranial tissues ***a***, and thresholding of the combined processing steps including correction of cupping artifacts, image top hat transformation, and left-right comparison ***b*** as discussed in the text. It is evident that the result of ***b*** is much cleaner and manageable for subsequent processing.



a ***b***

AIH can be better approximated using an iterative intensity based region growing method. At each turn, pixels bordering the perimeter are checked if they show attenuation of less than 4 HU below that of its adjacent pixel on the AIH perimeter. Those pixels that fit the criteria are included as part of the new perimeter pixels and the process is repeated until no more pixel could be added to the AIH according to the criteria. Subsequently, various descriptors, including intensity, size, long axis length, short axis length, eccentricity, orientation of the lesion in terms of the angle between

the long axis length and the mid sagittal plane are extracted from individual candidate AIH region. These parameters, as well as some secondary parameters derived from them, e.g. average intensity (intensity/area), form the basis of subsequent knowledge based analysis.

Localization of Candidate AIH

In order to render the candidate AIH anatomical positional information such that subsequent classification can be performed in light of anatomical context, the brain need to be registered against a coordinate system in which anatomical information of a normalized brain model is embedded. Construction of the coordinate system is outlined in the following paragraph. The control points for registration are sizable structures which can be automatically obtained. After registration by local weighted mean transformation, candidate AIH can be checked against the positions where AIH or mimics frequently appear.

Development of the Coordinate System

Since the current project is designed for images representing relatively thick axial sections as prescribed in clinical protocols, anatomical labeling based on existing coordinate systems, e.g. Talairach atlas (Talairach and Tournoux 1988), is unreliable. This is because the registration against the Talairach atlas requires localization of the

anterior commisure and posterior commisure, which are tiny structures of no more that 5mm across. They can be defined on 3D images. But such small structures are usually not well shown on the thick section images, let alone accurate localization. Therefore a coordinate system that is applicable for clinical brain CT images has been purposely developed.

It was observed from a large number of images that the anatomical structures relevant for the current application can be adequately localized with reference to some anatomical landmarks that can be automatically extracted from thick axial sections, including the mid sagittal plane, the boundaries of the brain, the level of anterior cranial fossa, and the boundaries of the lateral ventricles.

Based on these landmarks, the brain can be represented by a 3D coordinate system composed of a 15 x 20 x number of axial section (left-right x anterior-posterior x superior-inferior) matrix. The size of the matrix in each axial section is chosen such that the size of each coordinate position is fine enough to locate the anatomical structures without confusion with adjacent structures, yet not so small that normalization becomes difficult. The mid-sagittal plane is put into the central column of the coordinate. The cross-correlation between a coordinate position and the actual anatomical label is obtained by normalization using 65 of the training cases which

show normal findings. The normalization process involved registration of the coordinate system to the training cases based on the automatically extracted control points from individual case. Control points include the centroid of the mid sagittal plane, bounding box of the largest axial section, and bounding box of the lateral ventricle. The matrix was displayed as grids overlaid onto the original images, we then recorded the coordinates of the relevant structures after visual inspection of the composite images. After normalization process, relative frequency of occurrence of particular structure at each coordinate was transferred to the coordinate system.

Knowledge Based Classification

The main aim of the classification is to reduce false positives produced by the segmentation program. In fact, this is a common strategy used by different CAD systems to rely on classification system to reduce false positives produced by earlier steps. The approach used in the current classification is novel in that anatomical context in addition to various imaging features is incorporated. There are several common causes of mimickers of AIH that the system needs to deal with: calcifications, normal variants that show high attenuation, and partial volume averaging of bones.

Calcifications mimicking AIH

Calcification is one of the most frequent mimicker of AIH (Reisdorff and Schwartz

2000). Calcium deposits in previous insults, tumors, and vascular malformations. It is shown as regions of high attenuation, usually >120, but may range anywhere from 50 to >1000 HU on axial CT. The actual value depends on the relative concentration of calcium within each voxel and technical factors including volume averaging. When the attenuation is on the high side, e.g. above 120, it can be confidently distinguished from AIH. The differentiation becomes difficult in the range where the attenuation of calcium and AIH overlaps, i.e. in the range of 50-105. It is when the area of the abnormality needs to be considered. For example, if a very small AIH produces a hyperdense region, it tends to produce attenuation much lower than the usual density of 60-70 HU due to partial volume averaging; hence hyperdense regions showing moderately high attenuation, e.g. 65, are more likely due to AIH rather than calcium if its area or volume is larger, and vice versa. Therefore a region more likely represents calcium rather than AIH if its attenuation divided by area is higher than a predetermined threshold.

The probability of being calcification also increases if the AIH candidate is located in structures that are known to produce calcifications as normal variants. The most important examples are basal ganglia, choroids plexus in ventricles, and falx cerebri. This is when the inputs of anatomical context come into play. When an AIH candidate is localized over these anatomical regions, a different set of rules are invoked to

evaluate the candidates. For example, at the basal ganglia, lower attenuation, larger area, and lower attenuation/area ratio are required for the AIH candidate to be classified as calcification (not genuine AIH), as compared against postulated calcifications in other regions of the brain. The criteria are even looser if there appears to be similar candidate on the contralateral side. This rule was laid down in recognition of the tendency of normal calcification to occur symmetrically on both sides of the brain. The effect produced by these rules are illustrated in the examples as shown in figure 12 .

Normally high attenuation regions mimicking AIH

Besides calcifications, regions may show higher than normal attenuation due to stagnant flow, e.g. in the venous sinuses. Most frequently confused anatomical structures are the great vein of Galen, superior sagittal sinus, straight sinus, and transverse sinuses. These anatomical structures can be defined by their relative position in the brain, i.e. anatomy, and their shape, e.g. triangular for superior sagittal sinus, and oblong for straight sinus.

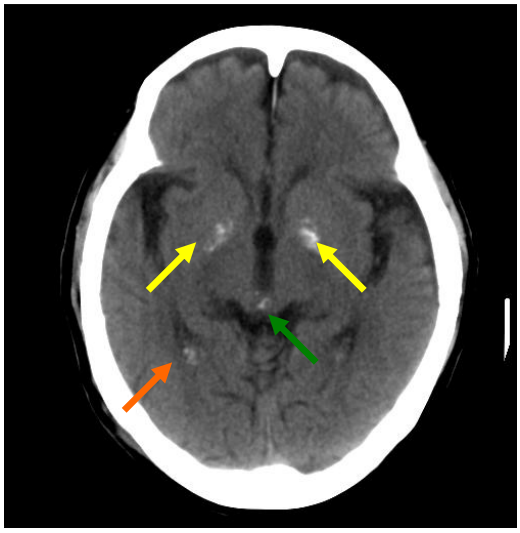
Artifacts mimicking AIH

In addition to genuine structures, artifacts may also produce regions of elevated

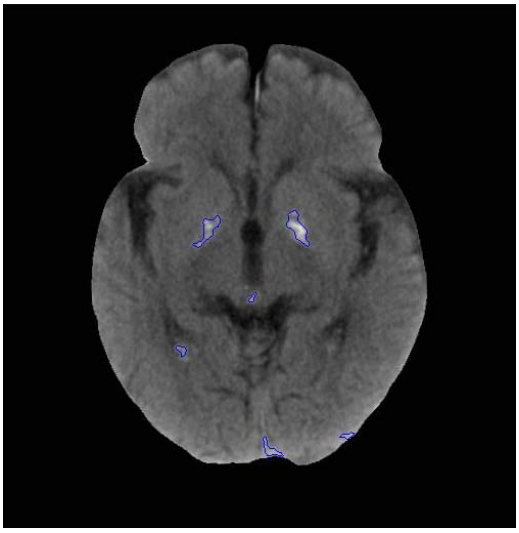
Figure 12 – Calcifications identifiable by the classification system

Images ***a*** and ***b*** show that the calcifications in basal ganglia, choroids plexus, and pineal gland calcifications, marked by yellow, orange, and green arrows respectively, are correctly classified as mimics of AIH. AIH candidates subsequently discared as mimics rather than genuine AIH are outlined in blue by the system. They show relatively high attenuations relative to their sizes (image features) and are located at the areas susceptible to calcification deposition (anatomical information).

Images ***c*** and ***d*** show that the calcifications or ossification of falx cerebri, pink arrow, is also correctly classified, based on their linear configuration, vertical orientation (image features) and central location (anatomical information).

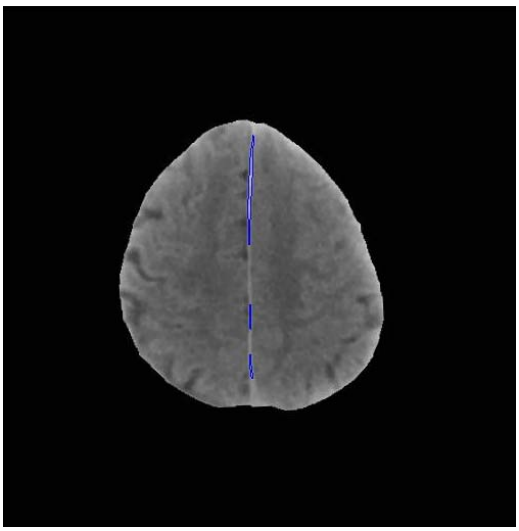


a



b

c



d

attenuation values. Beam hardening artifacts produce alternating bright and dark straight bands. These are most commonly shown in posterior fossa, where the bones are relatively thick and irregular and attenuation across one particular orientation may be significantly higher due to bones along their path.

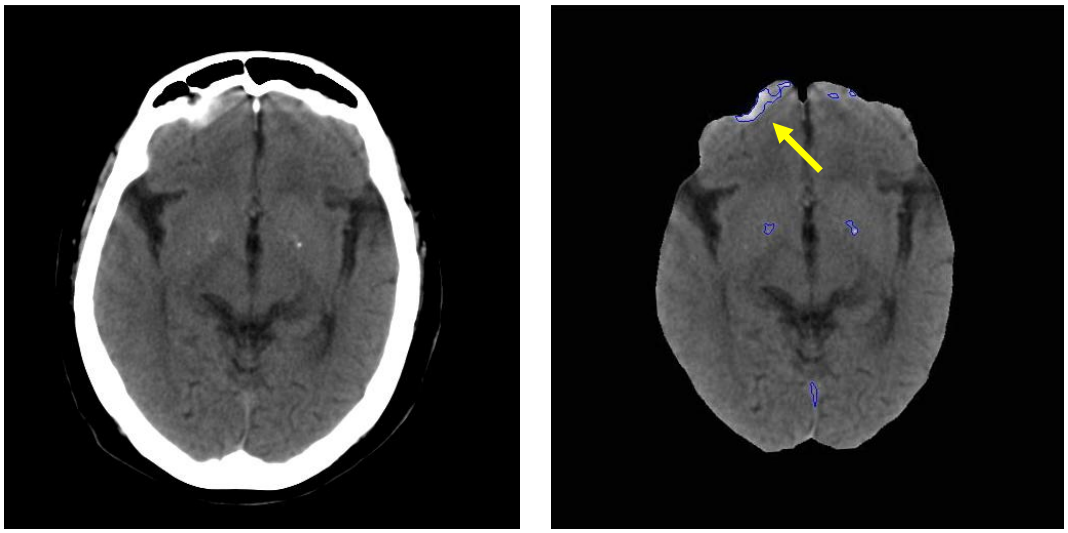
Partial volume averaging may produce an apparent AIH, especially at the floor of the anterior cranial fossa, where the bone is irregular. The voxels just above the floor can include portions of the underlying bone, hence its attenuation is increased by bits of bones. These voxels may produce regions with attenuation in the range of the AIH and hence extracted as AIH candidates. The important features of these partial volume averaging lies not in their imaging features, but their anatomical location in regions known to produce these artifacts, namely immediately above orbital floor and petrous temporal bone. An example is shown in figure 13.

In summary, there are known constellations of imaging features and anatomical positions for different confounding normal variants, e.g. basal ganglia calcifications and venous sinuses, rules that lower probability of calling those suspected AIH as genuine are invoked when the criteria are met. For example, unusually high density regions smaller than a predetermined area and located symmetrically at the expected anatomical positions of bilateral basal ganglia would be rated a low probability for

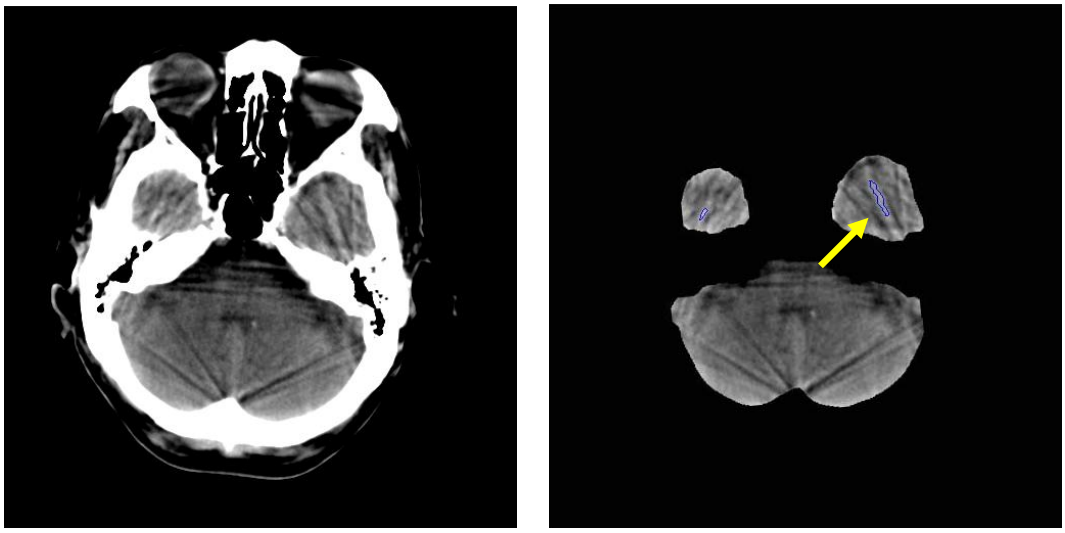
AIH, because they very much satisfy the rule for excluding basal ganglia calcification.

Likewise, there are artifacts with some specific combinations of imaging and

Figure 13 – Partial volume and beam hardening indentifiable by the system
Images ***a*** and ***b*** are the original image and CAD output of an image just superior to the floor of anterior cranial fossa that show partial volume average at the right frontal region. Images ***c*** and ***d*** are the original image and CAD output of an image through temporal lobes and posterior fossa that show beam hardening artifact. The artifacts in both cases, yellow arrows, were extracted as AIH candidates, but subsequently discarded by the knowledge base classification, outlined in blue.



a ***b***



c ***d***

anatomical features, e.g. straight linear beam hardening artifacts between bones in posterior fossa, which are taken into account by another set of rules. In addition, correlation between different image planes have also been built in, e.g. partial volume artifacts are considered when a high density region lies immediate above bones, especially over petrous temporal bone and anterior cranial fossa.

Genuine AIH can also be positively identified from the candidates of suspected AIH, again based on both imaging features and anatomical location. Despite the fact that AIH can take on myriads of different size and shape, there are certain definable patterns that certain type of AIH may follow. For example, subarachnoid hemorrhage produces blood clots that match the configuration of subarachnoid spaces, which are located at the basal cisterns, Sylvian fissures, and sulcal spaces; epidural hemorrhage produce blood clots with elliptical configurations which long axis parallels that of the overlying skull. Therefore if a candidate AIH volume matches these descriptions, e.g. a AIH volume with high eccentricity and vertically oriented (image features) and is located at the coordinate that corresponds to either Sylvian fissure (anatomical context), it would be rated a higher probability of being a genuine lesion by the rule-based classification system.

Generally the rules that incorporate both imaging features and anatomical information

take the following form:

$$P \rightarrow Q \quad \text{Equation 4}$$

$$Q \rightarrow \text{alters probability of AIH} \quad \text{Equation 5}$$

Where P are the set of rules which check for the anatomical locations of the candidate AIH and Q are the corresponding rules which subsequently evaluate the imaging features of the candidate appropriate for the particular anatomical positions. When the imaging features satisfy some defined pattern, probability of AIH is increased or decreased for the candidate. Some sample rules are listed in table 2.

The knowledge based classification provides a combined rating of probability derived from every invoked rule for each AIH candidate, which is then determined to be genuine AIH if the probability is higher than the predetermined threshold or a mimicker if the probability is lower.

Table 3 – Sample rules used in the classification

Sample rules used in the knowledge base classification system. P are the set of rules which check for the anatomical locations and Q are the rules which evaluate the imaging features appropriate for some particular anatomical positions. A candidate AIH is first checked for the anatomical position. If the position is one that satisfies a particular P, the corresponding Q will be invoked to evaluate the image features of the candidate AIH. If an appropriate pair of P and Q is satisfied, the probability of AIH for the candidate is adjusted accordingly. The magnitude of change being determined by the training process.

Anatomy rules (P)	Imaging feature rules (Q)	Interpretation	Probability of AIH
Rules that lower probability of AIH for candidates that conform to calcifications or normal high density structures			
mid sagittal plane, supracranial fossa	vertically aligned, ↑attenuation, ↑eccentricity, ↑ long axis length ↓short axis length	falx calcification	↓
mid sagittal plane, supracranial fossa periphery	intermediate attenuation, intermediate eccentricity, ↓convex hull	superior sagittal sinus	↓
medial portion of basal ganglia	↑attenuation (↑ if area ↑), ↓area (↑ if symmetrical), symmetrical	basal ganglia calcification	↓
central portion of cerebellum	↑attenuation (↑ if area ↑), ↓area (↑ if symmetrical), symmetrical	dentate nuclei calcification	↓
Rules that lower probability of AIH for candidates that conform to artifacts			
posterior cranial fossa	↑ eccentricity, ↑ long axis length ↓ short axis length	beam hardening artifact	↓
above anterior cranial fossa above temporal bone periphery near vertex	↑ attenuation ↓ area beyond adjacent bone in contiguous section	partial volume averaging	↓
Rules that increase probability of AIH for candidates that conform to particular type of AIH			
sylvian fissure	vertically aligned, intermediate attenuation, ↑ eccentricity, intermediate long axis length ↓ short axis length	sylvian fissure subarachnoid hemorrhage	↑
periphery	perpendicular to perimeter of brain ↓ long axis length ↓ short axis length	sulcal space subarachnoid hemorrhage	↑
anterior portion of posterior cranial fossa	horizontally alighed, intermediate attenuation, ↓ area	basal cistern subarachnoid hemorrhage	↑

Display of Output for Perusal

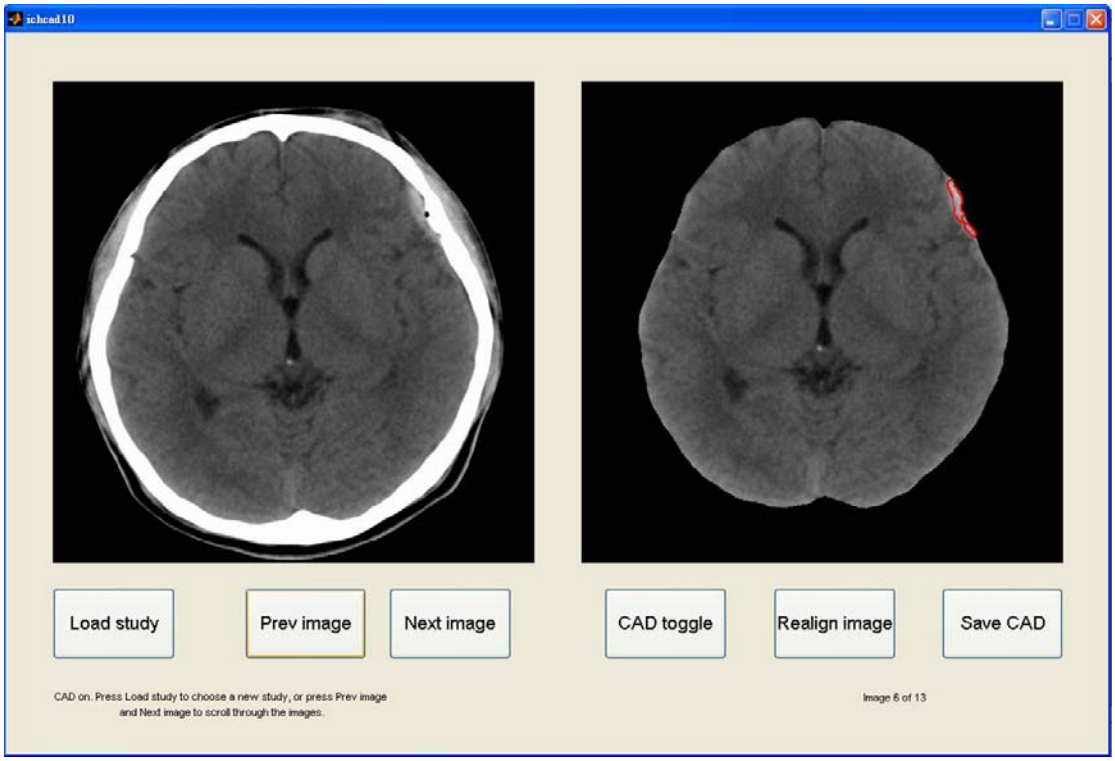
The outlines of the genuine AIH as classified by the CAD are colored in red whilst the mimickers are colored in blue. The outlines are overlaid onto the original images. The original images without and with color coded outlines of AIH are displayed side by side, such that the user can compare the output of the system with the original images to better appreciate the results and make the clinical decision of whether to call the lesion a genuine AIH and act upon it. A screen capture of the graphical interface is shown in figure 14.

Storage of Output

The CAD output of images annotated with the detected AIH, as well as the textual data of locations and image features, e.g. area of AIH can be saved at the click of a button. This allows integration of the CAD outputs into structured reports in DICOM. Such integration allows transfer and display using standard DICOM compliant devices. The storage of textual data also allows comparison of studies, e.g. monitoring size of AIH as prognostic factor of further deterioration/improvement, and future data mining in the picture arching and communiation system (PACS) environment (Huang 2004).

Figure 14 - Screen capture of the CAD system graphical user interface

The original images are displayed on the left window, whilst the output images with overlay of the outlines of AIH are displayed on the right. The original and output images are displayed in stripe mode and linked such that they can be scrolled in synchrony for review of the whole series.



2.3 OBSERVER PERFORMANCE STUDY

Readers

Seven emergency physicians, with 5 to 9 years (average 6.4 years) of experience of working in the ER and the same number of years of experience in emergent brain CT interpretation, 7 radiology residents of experience from 1 to 4 years (average 2.3 years), and 6 board certified radiology specialists of experience from 7 years to 30 years (average 17.8 years), were invited to participate in the evaluation of the CAD system. All the participants regularly interpret CT obtained on emergency basis.

Cases

Sixty sets of images were randomly selected from the 186 case database for the observer performance study. These include 30 cases that showed AIH and 30 cases that were either normal or showed pathology but no hemorrhage. In 26 of the positive cases, presence of AIH was considered unambiguous by the radiologists. In the other 4, the diagnoses were concluded with follow-up CT/MRI. AIH of different types as dicussed in Chapter 1 have been included.

All intraparenchymal hematomas included in this study were smaller than 1cm in long axis diameter, while all extraaxial hematomas were thinner than 1cm. The other 30 cases revealed either normal findings or pathology other than hemorrhage, which

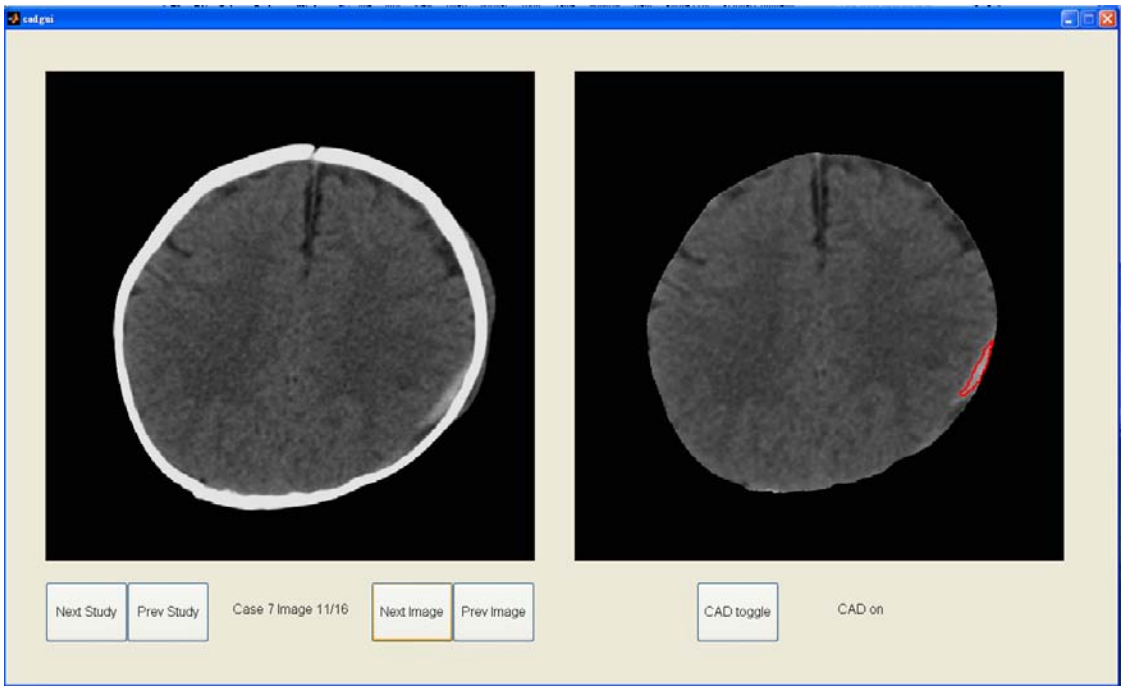
included acute and chronic infarct, ischemia, and tumor.

Receiver Operating Characteristic

The readers were asked to read the original images using a graphical user interface (GUI) specifically implemented for this study, as shown on figure 15.

One axial section was displayed at one time. The readers could scroll through the images of a particular case back and forth. The experiment was conducted in a radiologist's reporting room, where ambient light was low. The window level and width was set at the usual level for viewing brain CT, i.e. widow level of 40 and window width of 100. They could adjust the brightness of the screen to suit their individual needs, but image windowing has not been provided because this on its own could be considered one form of diagnostic aide. In particular, increasing the window level and widening the window width increase conspicuity of extraaxial hematomas (Lev, Farkas et al. 1999). The readers were instructed to record their confidence in detecting AIH on a scale of 1 (absolute absence of AIH) to 10 (absolute presence of AIH). Readers were also instructed to interpret the score of 5 and 6 as indeterminate, with 5 erring on the side of absent AIH and 6 otherwise. Immediately after they have finished rating all sixty cases, they would re-read the images, now with the CAD output images displayed side by side with the corresponding original images.

Figure 15 - Screen capture of the graphical user interface
GUI used in the observer performance study. The original images were displayed in the left window in stack mode. In the second reading, the output images of CAD were displayed in the right window. An output image contained the segmented and realigned intracranial contents, and AIH was outlined. The original and CAD output images were scrolled in synchrony.



Both the original and the CAD output images would scroll together in synchrony.

They again recorded their confidence level in the same way.

The readers were explicitly informed that the CAD had produced sensitivity of

80-85% on a per lesion basis and a false positive rate of less than 1 in 3 cases in

earlier tests, but performance for individual case may depend on size and contrast

difference of lesion(s) it contains. They were also reminded that the actual accuracy of the CAD output in the sample cases that they were going to read might be better or poorer than the quoted figures, reflecting difference in case selection.

The recorded data were subject to MRMC ROC analyses using the public domain software LABMRMC developed by the University of Chicago The program was based on the Dorfman-Berhaum-Metz method (Dorfman, Berbaum et al. 1992) that uses analysis of variance so that results are generalizable to the population of readers and cases. The ROC curve was obtained by maximum likelihood estimation of the binormal distributions that best fit the rating data of the readers.

Conventional Indicators

Since it was believed by many, including most of the participants in the test, that the diagnosis of AIH is an all or none question, it is also desirable to present the results in some conventional indicators that are more familiar to clinicians and are based on a yes/no type of response. The scores were placed into two categories of 1-5 and 6-10, which dichotomize the results in to absence/presence of AIH. The sensitivity/ specificity pair and positive/negative predictive values are calculated accordingly.

Number of Change in Diagnosis

In addition, the frequency when the use of CAD resulted in actual change of diagnosis

during the experiment, as opposed to mere change in confidence of one particular diagnosis or another, was examined. The diagnosis of absence/presence of AIH for individual case was determined based on the aforementioned method of dichotomizing the score ratings. Frequency of the change in diagnosis and the correctness of such changes were recorded. This information can reflect the impact that use of CAD may have in actual clinical practice, with altered diagnostic decisions that affect management options.

Chapter 3 RESULTS

3.1 PERFORMANCE OF THE CAD

Anatomical Localization

This method correctly located the level of the floor of middle cranial fossa in 97.5% (117/120) and 95.3% (61/64) of the training and validation cases respectively. All the other cases were off by one axial section only.

The mid sagittal planes (MSP) were accurately localized in 69.1% (83/120) and 65.6% (42/64) of the training and validation dataset respectively, which are defined as system output that is within 1mm of displacement and 1 degree of rotation from the genuine MSP. In 22.5% (27/120) and 25% (16/64) of the cases, the MSP were satisfactorily localized, off by within 3mm of displacement and 3 degrees of rotation, such that subsequent analysis was not adversely affected.

Altogether the system was successful in automatically putting the brain into the reference frame in 90.8% (109/120) and 85.9% (55/64) of the training and validation cases respectively, such that subsequent anatomical labeling was satisfactory. For the rest of the cases, the brain can be readily realigned by the user through the interactive interface such that subsequent analysis can be performed satisfactorily, which took no

more than a few seconds. Summary of the aforementioned results in anatomical localization is listed in table 4.

Table 4 – Accuracy of anatomical localization
Summary of success rate in identification of middle cranial fossa, mid sagittal plane, and putting the brain into the reference frame for anatomical localization. Manual intervention is required when the system fails to automatically put the brain into the reference frame.

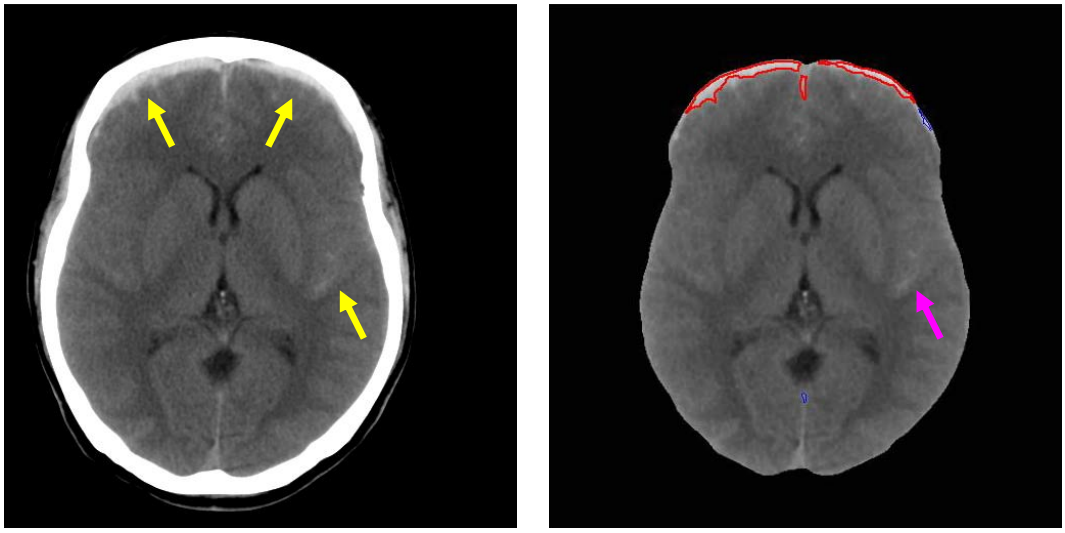
	Training	Validation
Middle cranial fossa	97.5%	95.3%
Mid sagittal plane	91.6%	90.6%
Reference frame	90.8%	85.9%

Performance in Diagnosis of AIH

The performance of the system was described by sensitivity and specificity pairs on both per lesion and per case bases. The per lesion basis 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, e.g. SAH between 2.5mm and 5mm, is of interest. On the other hand, the performance on per patient or per case basis is of more clinical relevance than the performance on a per lesion basis for the diagnosis of AIH. It is because the management options depend on the presence or absence of lesions rather than the quantity of lesions. An example of the difference between sensitivity on per lesion and per case basis is shown in figure 16. It is

Figure 16 – Sesitivity per case vs per lesion basis

This case illustrates the difference between sensitivity on per case/patient basis and per lesion basis. There are two AIH lesions in the original image ***a***, as denoted by the yellow arrows. In the CAD output ***b***, only the frontal region subdural hemorrhage was detected, outlined in red. The subarachnoid hemorrhage in left Sylvian fissure is missed, pink arrow. For this case, it would still be classified as a true positive, but the sensitivity on a per lesion basis on this image is only 1/2.



a ***b***

recognized that the per patient measurement can be affected by the different number of lesion present in each case, e.g. a case showing multiple lesions of different size and type would be more likely to be correctly diagnosed as showing AIH than another case showing one single lesion.

In the following section, a CAD output is counted as true positive if it intersects the genuine blood clot as determined by radiologists. It is counted as false positive if it does not overlap any of the genuine blood clots. The sensitivity and specificity on a per lesion basis are calculated accordingly.

Since classification of the results on a per patient/case basis may not be immediately apparent, its definition as used in the current study are elaborated in the following section. 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 positives. If there are one or more CAD output(s), but none of which is a true positive, then the patient is considered false positive. On the contrary, if there is no CAD output for a particular patient, but there is a genuine blood clot, the case is considered a false negative. If there is no CAD output for a patient in whom no blood clot existed, the case is counted as a true negative.

For the training cases, there were 77 contiguous AIH volumes in the 40 patients. The overall sensitivity was 84.4% (65/77). This increased significantly with increase in size of the lesion. The system correctly identified all AIH lesions larger than 5mm (28/28). Sensitivity dropped to 90% (27/30) for lesions between 2.5mm and 5mm, and 52.6% (10/19) for lesions smaller than 2.5mm (Table 5).

Forty-six contiguous AIH volumes were present in the 22 positive validation cases, averaging more than 2 lesions per case. The sensitivity on a per lesion bases was 82.6%(38/46). This increased significantly with increase in size of the lesion (Table 6). The sensitivity were 57.1% (4/7) for lesions 2.5mm or smaller, 84.2% (16/19) for

Table 5 – CAD result in the training cases

Summary of CAD result in the training cases according to type and size of individual AIH vol on a per lesion basis. IPH – intraparenchymal hemorrhage, IVH – intraventricular hemorrhage SAH – subarachnoid hemorrhage, SDH – subdural hemorrhage, EDH – extradural hemorrhage, AIH vol – number of contiguous AIH volumes, CAD+ - number of AIH vol correctly identified by the CAD.

Training Cases	*IPH*		*IVH*		*SAH*		*SDH*		*EDH*		*All Types*	
	AIH vol	*CAD +*	*AIH vol*	*CAD +*	*AIH vol*	*CAD +*	*AIH vol*	*CAD +*	*AIH vol*	*CAD +*	*AIH vol*	*CAD +*
0-2.5mm	0		1	1	17	9	1	0	0		19	10
>2.5-5.0mm	0		2	0	11	10	13	13	4	4	30	27
>5-7.5mm	1	1	0		6	6	3	3	1	1	11	11
>7.5-10.0mm	4	4	0		5	5	7	7	1	1	17	17
All Sizes	5	5	3	1	39	30	24	23	6	6	77	65

Table 6 – CAD result in the validation cases

Summary of CAD result in the validation cases according to type and size of individual AIH vol on a per lesion basis.

Validation Cases	*IPH*		*IVH*		*SAH*		*SDH*		*EDH*		*All Types*	
	AIH vol	*CAD +*	*AIH vol*	*CAD +*	*AIH vol*	*CAD +*	*AIH vol*	*CAD +*	*AIH vol*	*CAD +*	*AIH vol*	*CAD +*
0-2.5mm	0		0		7	4	0		0		7	4
>2.5-5.0mm	0		3	2	12	10	3	3	1	1	19	16
>5-7.5mm	2	2	1	1	6	5	2	2	1	1	12	11
>7.5-10.0mm	1	0	1	1	2	2	4	4	0	0	8	7
All Sizes	3	2	5	4	27	21	9	9	2	2	46	38

lesions between 2.5mm and 5mm, 91.7% (10/11) for lesions between 5mm and 7.5mm, and 87.5% (7/8) for lesions 10mm or smaller.

There were altogether 23 false positive lesions detected in all the training cases, which were present in both positives and controls. False positive rate was 0.19 (23/120) per case. For the validation cases, the false positive rate was 0.29 (19/66) per case. On a per patient basis, the sensitivity and specificity were 95% (38/40) and 88.8% (71/80) respectively for the training cases (Table 7).

Table 7 - Summary of CAD results on a per patient basis for training cases

	Validation Positive	Validation Negative
AIH present	38	2
AIH absent	9	71

Sensitivity =	95.0%	Specificity =	88.9%
Positive Predictive Value =	80.8%	Negative Predictive Value =	97.3%
Accuracy =	90.8%		

The system achieved sensitivity of 100% (22/22) and specificity of 84.1% (37/44) for the diagnosis of AIH for the validation cases (Table 8).

Table 8 - Summary of CAD results on a per patient basis for validation cases

	Validation Positive	Validation Negative
AIH present	22	0
AIH absent	7	37

Sensitivity =	100.0%	Specificity =	84.1%
Positive Predictive Value =	75.9%	Negative Predictive Value =	100.0%
Accuracy =	89.3%		

The current system, although not optimized in terms of speed, takes an average of approximately 15 seconds per image to produce the output. Actual time varies substantially for each case, depending on number of image and number of candidate AIH produced to be evaluated by the classification system.

3.2 OBSERVER PERFORMANCE STUDY

MRMC ROC

Significantly improved performance is observed in emergency physicians, average area under the ROC curve (Az) increased from 0.8288 to 0.9484 ($p<0.05$) when they make the diagnosis without and with the support of CAD. Az for radiology residents increased from 0.9478 to 0.9831. Az for radiology specialists increased from 0.9657 to 0.9776, but was statistically insignificant. The results are shown in figure 17.

The Az values scored by individuals before and after CAD are presented in figure 18. Only 1 of the 20 subjects scores marginally lower Az after CAD. The other 19 people all attained a variable degree of increment after use of CAD.

It was observed that performance of emergency physicians with support of CAD approached that of the radiology residents without CAD, whilst performance of the radiology residents with CAD approached that of the radiology specialists without CAD. This signifies that the CAD can improve reader performance as well as reduce

variability amongst readers of different expertise levels.

The data from all 20 readers were analyzed together using the public domain software DBM MRMC based on PROPROC area analysis, assuming random readers and cases. All the clinicians as a whole group show statistically significant of increase in Az from 0.9150 to 0.9630 ($p = 0.0015$).

Figure 17 – MRMC ROC Curve
ROC of detection of AIH amongst different groups of clinicians. EP – emergency physicians. RR – radiology residents, RS – radiology specialists, UA – unaided reading mode, CAD – CAD reading mode.

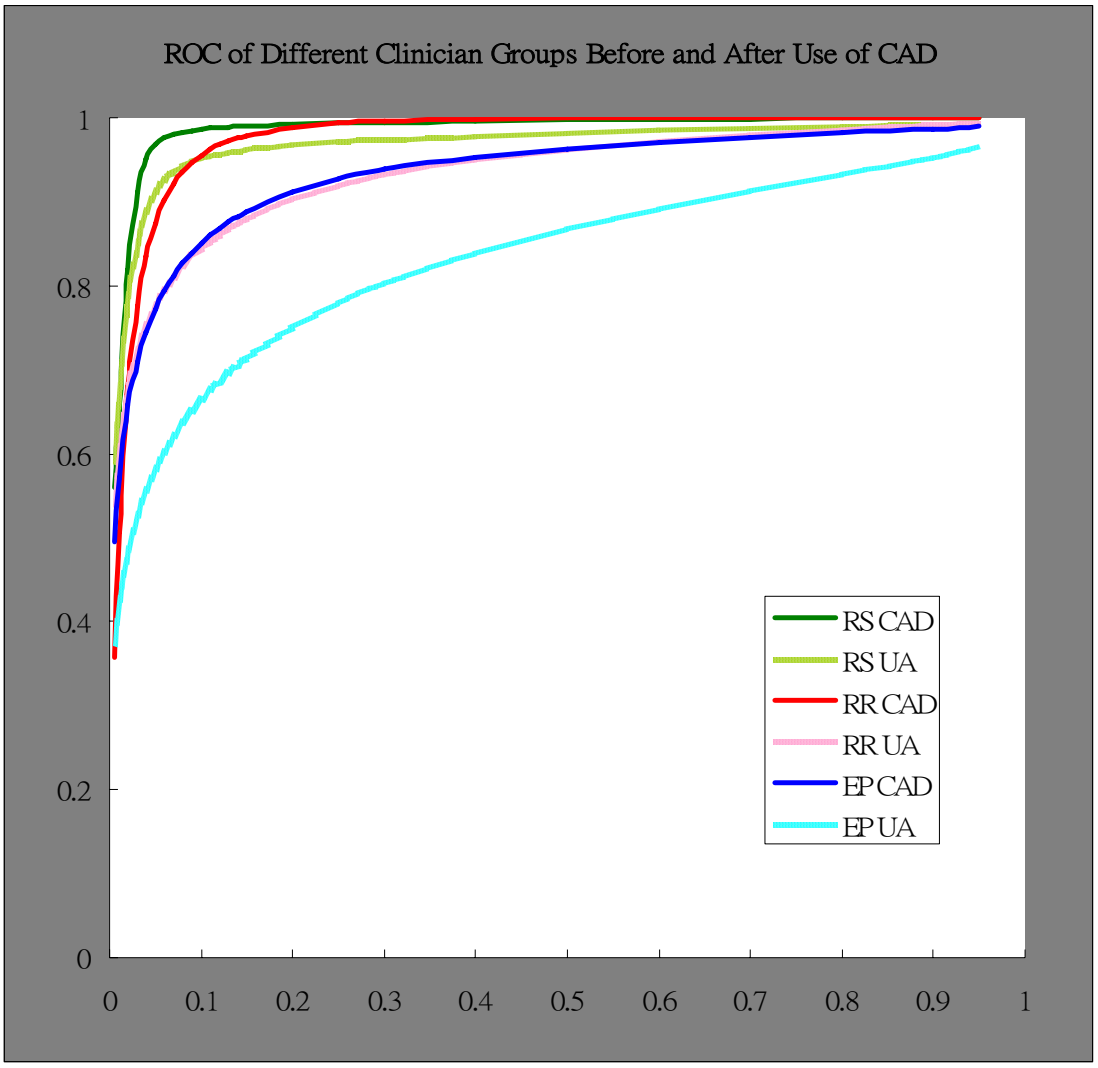
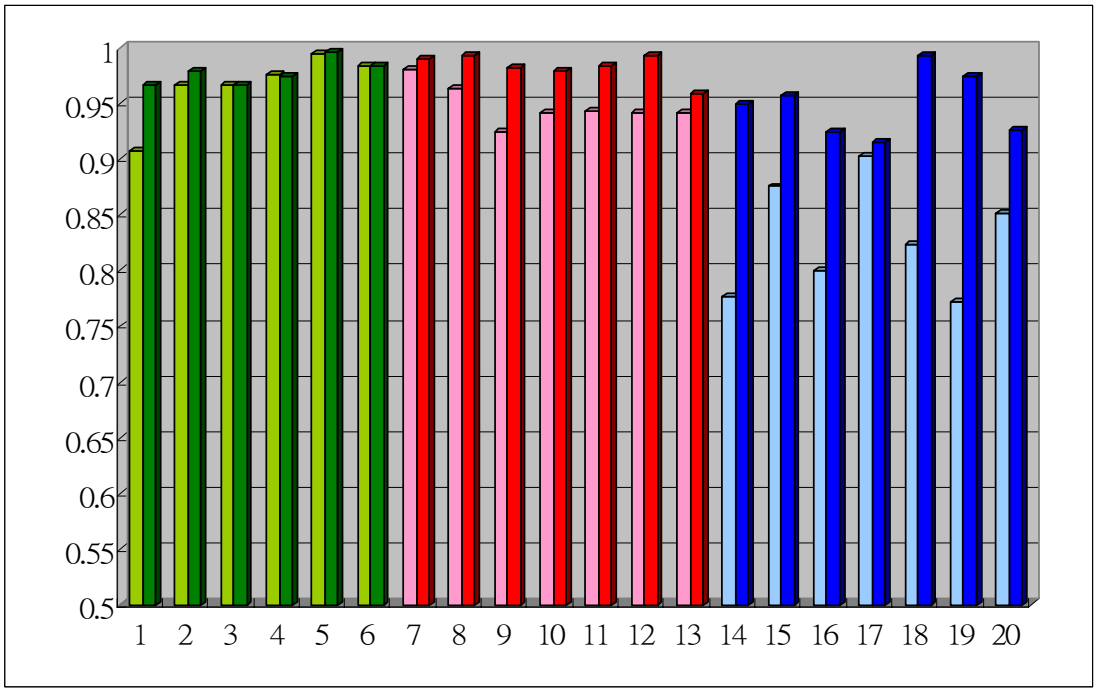


Figure 18 - Bar chart showing the Az before and after use of CAD
The marginal increase shows greatest increase in emergency physicians, less for the radiology residents, and least for the radiology specialists. Subjects 1-6 (green) are radiology specialists, 7-13 (red) are radiology residents, and 14-20 (blue) are emergency physicians.



Conventional Indicators

The sensitivity, specificity, positive and negative predictive values were calculated for each reader both before and after use of CAD (Table 9). It was demonstrated that both sensitivity and specificity improved for each group. The improvement is most remarkable for emergency physicians, in whom the average sensitivity/specificity rose from 73.3%/81.4% to 80.5/90.5%; for the radiology residents: from 86.2%/88.1% to 93.8%/92.9%; and for the radiology specialists: from 92.2%/93.3% to 95%/94.4%.

After use of CAD, the positive predictive values improved from 80.1% to 89.5% for ER, from 88.4% to 93.0% for radiology residents, and from 93.3% to 94.5% for radiology specialists respectively. The negative predictive values also improved from 75.7% to 82.5% for ER, from 86.7% to 93.8% for radiology residents, and from 92.6% to 95.1% for radiology specialists.

Again, it was observed that these conventional indicators of diagnostic accuracy of emergency physicians with support of CAD approached that of the radiology residents without CAD, and the results of the residents with CAD approached those of the specialists without CAD.

Table 9 – Performance indicators of clinicians using CAD
Average performance indicators including sensitivity, specificity, positive predictive value, and negative predictive value for different clinician groups with and without CAD support. All indicators in all clinician groups are improved after use of CAD.

%	Emergency Physician		Radiology Residents		Radiology Specialists	
	Unaided	**CAD**	**Unaided**	**CAD**	**Unaided**	**CAD**
Sensitivity	73.3	80.4	86.2	93.8	92.2	95.0
Specificity	81.4	90.5	88.1	92.9	93.3	94.4
Positive Predictive Value	80.0	89.5	88.4	93.0	93.3	94.5
Negative Predictive Value	75.7	82.5	86.7	93.8	92.6	95.1

Number of Change in Diagnosis

When diagnostic decision for each individual case was considered, it was found that use of CAD corrected the diagnosis far more frequently than misled the readers to a wrong diagnosis. For the emergency physicians, use of CAD led to 46 correct changes (beneficial effect) in diagnosis and 12 wrong changes (detrimental effect) out of the maximum number of possible change of 420 (7 readers x 60 cases). For the radiology residents, the figures were 29 versus 3 out of 420. For the radiology specialists, the figures were 7 and 0 out of 360. Thus use of CAD is associated with change in diagnosis in decreasing order of relative frequency from emergency physicians (13.8%), to radiology residents (7.0%) to radiology specialists (1.9%). On the other hand, the relative frequency of correct change vs. incorrect change show increasing trend from 79.3%:20.7% for emergency physicians, to 90.6%:19.4% for radiology residents, to 100%:0% for radiology specialists (Table 10). Hence although emergency physicians tend to benefit more frequently from the use of CAD, they also seem less capable of avoiding detrimental effect from its application.

Table 10 – Frequency of change in diagnosis after CAD

Number of cases in which clinicians change their diagnostic decision after CAD. The proportion of correct change relative to incorrect change increased from emergency physicians to radiology residents to radiology specialists. The total and relative number of change decreased from emergency physicians to radiology residents to radiology specialists.

	Emergency Physicians	**Radiology Residents**	**Radiology Specialists**
Correct change (% of actual no. of change)	46 (79.3%)	29 (90.6%)	7 (100%)
Incorrect change (% of actual no. of change)	12 (20.7%)	3 (9.4%)	0 (0%)
Frequency of change in decision	58	32	7
% Change in decision/total possible change	13.8% (58/420)	7.6% (32/420)	1.9% (7/360)

Chapter 4 DISCUSSION

4.1 DIFFERENCES FROM EXISTING CAD SCHEMES

The target for detection in the current system is small AIH. It differs from existing CAD products, e.g. malignancy detection in mammography and nodule detection in chest radiograph or CT, in that the system is intended to be used by clinicians other than radiologists and that the system rates the authenticity of candidate lesions in different portions of the image dataset differently, depending on their anatomical positions and imaging features.

As noted before, in emergent settings, expert radiologists may not be readily available to provide the often crucial image interpretation. Therefore the duty is shifted to clinicians who may not be best equipped for the task. It is therefore believed that CAD may become useful in these situations, in addition to its proven value for screening examinations. Special considerations need to be made because observers of less expertise may not be confident or knowledgeable enough to judge the correctness of CAD outputs. Therefore CAD systems targeted for non-radiologists need to minimize the false positive rates.

During the development stage of the system, it was found that the myriad

combinations of imaging features of AIH in different parts of the brain could not be adequately described without reference to the anatomical positions where the lesions are found. But when the candidate lesions are divided up based on their anatomical positions, classification between genuine AIH and mimicking variants or artifacts becomes feasible. It is because specific combination of imaging features and anatomical information produce patterns which can be used for classification. This contrasts against target lesions of many CAD systems which are well-described with relatively little variation in their configurations, which are hence less dependent on the anatomical information as opposed to the local imaging features.

4.2 Issues Regarding Choice of Small Lesions

Clinical Relevance

The CAD system itself and the observer performance study have focused on detection of small AIH, because it is conceivable that the small ones are those that cause diagnostic difficulty. In fact, during the developmental stage of the CAD, pilot study comprising 25 AIH cases of a wide range of sizes confirmed the intuition that sensitivity for cases including large (> 1cm) AIH reached 100% for both CAD and emergency physicians. Therefore detection of large AIH by the CAD may not be clinically useful as such lesions pose no diagnostic challenge to clinicians.

In an internal audit conducted at the author's institute, a total of 3341 emergent brain CT were performed for initial evaluation of head injury or neurological disturbance over a 6 month period. Of these, 279 cases were reported to show AIH, making the prevalence of AIH amongst emergent brain CT 8.35% (279/3341). Furthermore, 62 of the 279 cases showed only blood clots defined as small by the criteria mentioned earlier; hence prevalence of small AIH was 1.86% (62/3341). These figures illustrate that negative cases usually predominate in real clinical settings. It can be expected that the relatively infrequent detrimental effect of wrongly label AIH as a result of false positive output of the CAD system may magnify with the much higher proportion of normal cases in clinical practice versus the 50% normal cases in the current MRMC ROC protocol. Hence false positives produced by the CAD may cause more mistaken diagnosis. We have provided for this possible detrimental effect by keeping the false positive rate to a remarkable 0.29 per set of brain CT, which is much lower than that reported for many CAD systems. Also, since both sensitivity and specificity were improved, it is expected that the system can improve the performance even when the proportion of cases showing only small AIH in clinical setting is much lower than that in the current experiment.

It may be assumed that small AIH would pose less immediate danger to the patient. But missing the small AIH at an early stage may do harm to the patient by leading to

wrong management, especially when thrombolytic therapy or antiplatelet medication is prescribed. Also, patient may not be given appropriate level of care and monitoring and suffer when AIH subsequently grow or rebleeding occurs.

Size Measurement

There has been no reported objective method of classifying AIH according to its size. The target size of detecting AIH smaller than 10mm is arbitrarily chosen. The width rather than the length is chosen for extraaxial hemorrhage because it is the dimension which is clinically relevant and customarily reported. Although it does imply that 10mm extraaxial AIH would be larger in area/volume than a 10mm intraaxial AIH as defined in this study, because in the former the measured thickness is usually equivalent to the short axis, whilst in the later the long axis is measured. Nonetheless, it is believed that convention used in the current study better reflect the radiologists' assessment and reporting standard. It is understood that measurement of area/volume may be more robust scientifically speaking, and they can easily be calculated by the computer. But area and volume are simply not the figures that clinicians or radiologists use for describing lesions during their daily practice, for the pragmatic reason that such measurements cannot be easily obtainable.

Small AIH often shows ill-defined boundaries, variable configurations, and little

contrast difference with adjacent structures, mainly as a result of partial volume averaging and relative important contribution by noise and artifacts. These features present challenges to the development of CAD based on low level processing alone. Yet for a system to be clinically useful, it must be able to detect small lesions. The approach adopted in the current system is to be oversensitive in locating candidate AIH volumes using image processing techniques such that the vast majority of abnormal foci are extracted; however, this inevitably generates too many false positives. The average number of false positives generated by this CAD system based on the aforementioned processes alone was 62.1 per set of images or 4.2 per image in the training set. It is especially important for a CAD system intended for use by non-radiologists, who may not be highly competent in telling the difference between genuine lesion and false positives, to be highly specific. It was found that genuine AIH may not be all that different from false positives considering simple image features. But with the input of anatomical information, classification becomes feasible. It is because the candidates with similar image features may mean AIH at one particular location, or artifact if situated at another. With the knowledge based classification system in place, the false positive per case was dramatically reduced to 0.19 in the training dataset.

Effect on Observer Performance Study

The choice of small AIH also facilitates comparison between unaided and CAD assisted reading. By employing more difficult cases, statistical power of the experiment would be increased (Metz 1989). The downside is that such a ‘stress test’ contains non-representative samples with disproportionate number of difficult cases, which affects the generalizability of the results to the general population. Strictly speaking, the results as obtained from the ROC studies may not be readily generalizable to the population of cases containing AIH of various and predominantly larger size lesions.

4.3 USAGE OF ANATOMICAL INFORMATION

Rationale

There are different sets of rules for different parts of the brain because the probability of having a particular type of AIH or certain variant and artifact depends on anatomical position. For example subdural hemorrhage can be present over the convexity of the brain, along the falx cerebri or tentorium, but not within the brain or ventricular system. Also, the configuration of the same type of hematoma may depend on the anatomical location, e.g. SAH fills and takes on the configuration of the CSF space where bleeding occurs. Therefore rules for identifying typical image feature of

SDH can only be applied at the appropriate regions. Likewise, calcifications are usually present in basal ganglia, pineal gland, and dentate nuclei and tend to be symmetrical. Accordingly rules for identifying normal calcifications are applied over the said regions with additional criteria of symmetry. With these rules in place, the false positive rate has been kept at low level.

Coordinate System for Thick Section Images

There are several established methods to map digital images to standard brains. The most well known are the Talairach atlas and the Montreal Neurological Institute (MNI) brains. Software packages are available for mapping to these brains (MEDIC, SPM 5). However, these available systems cannot be effectively utilized in this project because the thick (5-10mm) sections used in clinical protocols preclude the accurate localization of landmark structures, especially the anterior commisure (AC) and posterior commisure (PC) required to define the AC-PC plane, which is the prerequisite for registration. In addition, the images are acquired along the orbital canthal – meatal plane (OM) conventionally used for clinical brain CT imaging, which lies at a variable angle of around 9 degrees from the AC-PC plane (Weiss, Pan et al. 2003). The thick axial sections thus obtained cannot be consistently converted to the AC-PC plane.

The current system maps the individual images to a coordinate system specifically developed based on axial sections that parallel the OM plane, which is readily applicable to images obtained using standard clinical protocols for brain CT imaging. Although it lacks definition of the AC and PC, such that anterior-posterior relationship of some internal structures may be less accurate, the system does contain additional control points that modify the registration process according to the configuration of the lateral ventricles. It was observed that this significantly reduces the variability of coordinate positions for some relevant internal structures, especially the choroids plexus and basal ganglia. It is postulated that inclusion of control points based on the lateral ventricle positional information more accurately reflects the change in relative position of internal structure resulting from age-related brain atrophy. On the other hand, a more accurate age corrected coordinate system may be developed using the aforementioned scheme, with collection of a large sample containing enough number of sample for each age group. It is also hoped that inclusion of more anatomical labels in addition to the ones which have been obtained in this development may make the coordinate system valuable for other applications.

It is recognized that the accuracy of the anatomical labeling process in the current project is limited by the relatively large size represented by each coordinate position, which is about $1cm^3$ for a normal sized adult brain. In addition, tilt and yaw in the

coronal and sagittal plane cannot be effectively corrected because the axial sections are thick. However, it must be understood that pinpoint accuracy in brain mapping is a daunting task on its own, and is impractical for images obtained using present day clinical protocols. Adding to the problem is that the current CAD deals with abnormal brain, which makes accurate mapping to a normalized atlas even more difficult. It is believed that the simple coordinate system generating relatively rough estimate of anatomical position is more efficient and reliable for the current application.

4.4 UNEXPECTED RESULTS

The results in the training cases show lower sensitivity than those of the validation cases on per patient basis, 95% versus 100%. This was probably a chance occurrence because all the missed lesions in the validation cases are present in cases where at least one other lesion was picked up by the system. The very similar overall sensitivity on per lesion basis of 84.4% and 82.6% for the two groups might be affected by the different distribution of size of lesions in individual patients.

It is admitted that the allocation of training and validation cases were not randomized in the strict sense. But we believe the collection process was not systemically biased, as the allocation into each group was essentially decided by the independent schedule between data collection by the investigator and housekeeping clearance of the

temporary CT archive. However, there does seem to be difference in the composition of the lesions in the two datasets regarding the size of lesions, the training cases contain a lower proportion of larger (>5mm) lesions, 36.4% (28/77), as compared against the validation cases, 43.5% (20/46). This might have inflated the apparent overall sensitivity on per lesion basis for the validation cases. Therefore, we believe the results of sensitivity by size best represent the performance of the system. A similar trend of increase from 50-60% for lesions 2.5mm or smaller to 90-100% for lesions larger than 5mm was observed for both the validation cases and training cases.

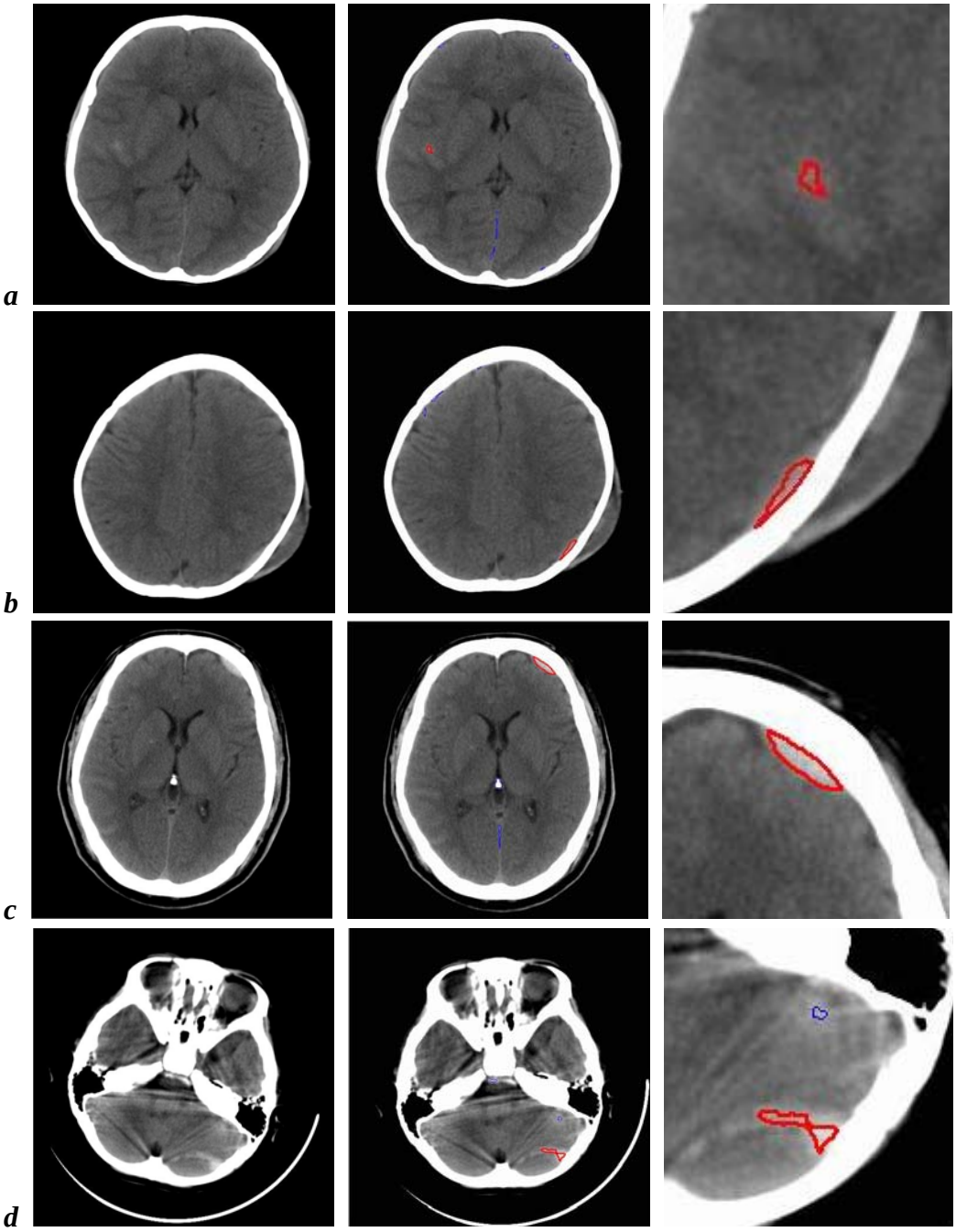
4.5 CHALLENGING CASES

Successful Examples

It was shown that the system can detect lesions that can be difficult for acute care physicians or even radiology residents. Some such lesions that are missed by most of the emergency physicians in the observer performance study are shown in figure 19. Lesions may be missed simply because they are too small as in figure 19-a. Or else they may become inconspicuous because they show low attenuation approaching that of surrounding normal brain, figure 19-b or high attenuation approaching that of adjacent skull, as in figure 19-c. In other cases the lesion may be misinterpreted by some to be some kind of artifact or normal variant mimicking AIH as in figure 19-d.

Figure 19 – Difficult AIH successfully diagnosed with CAD

AIH can be difficult to detect if they are small, of similar attenuation to adjacent structures, or confused with normal variant and artifact. Examples of some difficult cases (the left column) with their CAD results (the middle column) and magnified views (the right column) are shown. The system outlines AIH with red perimeters. High density regions which are segmented but subsequently classified as unlikely to be AIH are outlined with blue perimeters.



Unsuccessful Examples

There are 16 false positive cases and 2 false negative cases in all the 186 cases that include both the training and validation dataset. The false positive cases are most frequently due to calcifications at falx (6) being idenfied as AIH because they are too thick or lying at an angle to the vertical mid sagittal plane. Other causes include hyperdense peripheral gray matter (4), mistaken AIH in choroids plexus (4), partial volume averaging due to bone (2), beam hardening artifact (2), hyperdense transverse sinus (1), and abnormal calcium deposition in old infarct (1). The false negatives occur in two cases, showing very thin subarachnoid hemorrhage in the falx and small intraventricular hemorrhage respectively.

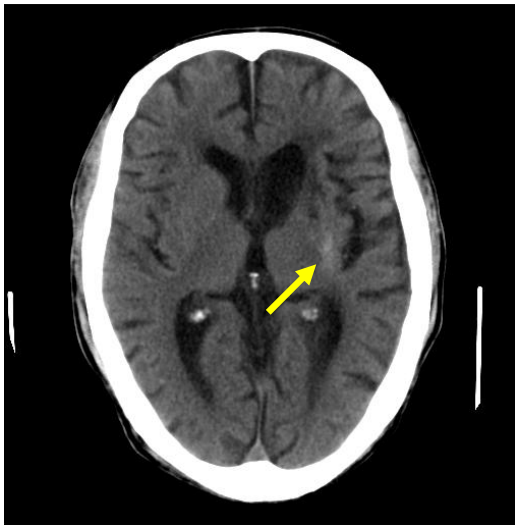
There are cases when the same errors could be committed by many experienced radiologists, like the examples as shown in figure 20. These cases show abnormalities that are difficult to diagnose even for human experts, and are hence considered reasonable, albeit wrong, conclusions. However, whether such errors may contribute to false reassurance is debatable.

Figure 20 – Examples of wrong diagnosis by the CAD

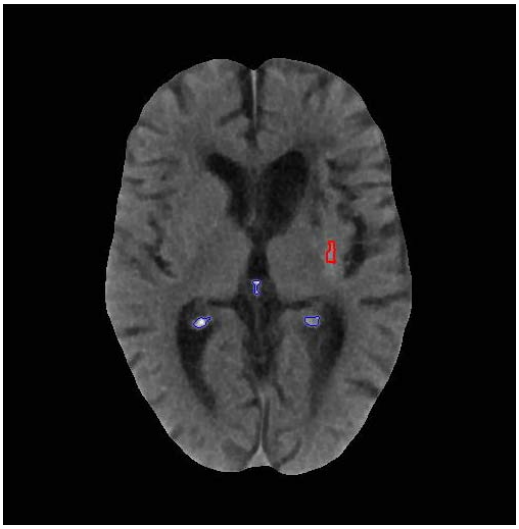
Image ***a*** and ***b*** are the original and CAD outputs of a case that shows calcification rather than AIH. This case is classified as a false positive because the highlighted lesion (in red) is not a genuine AIH.

Image ***c*** and ***d*** are the original and CAD outputs of a case that shows genuine brainstem hemorrhage. This case is classified as a false negative because the genuine lesion (in blue) is not highlighted.

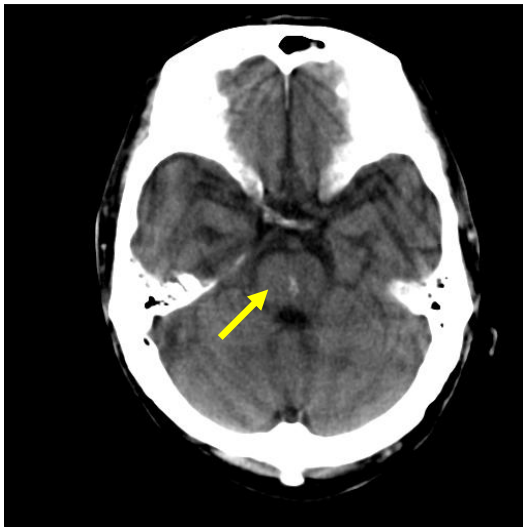
It is noteworthy that both of these cases have their diagnoses confirmed after review of prior and follow-up images respectively, and the majority of radiology specialists participated in this study made wrong diagnosis on one or both of the cases.



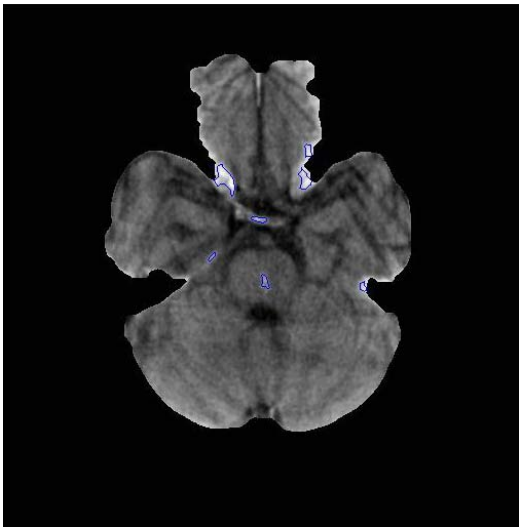
a



b



a



b

4.6 CAD FOR CLINICIANS OTHER THAN RADIOLOGISTS

Many CAD schemes have focused on high volume screening examinations, where detection of abnormality is rare and tedious but crucial. We envisage that CAD can also help in emergency situations, when expert human observers, in our case neuroradiologists, may not be readily available. This is when observers of less expertise need to promptly make the crucial decisions, and hence second opinions in the form of a CAD output may mean the difference between life and death.

When the contexts and users are different, the performance requirement of the CAD can be different and need to be adjusted so as to produce the most accurate human-computer diagnostic system.

It has been pointed out that CAD may achantaoctually degrade human interpretation under certain conditions (Alberdi, Povykalo et al. 2004), which design and assessment should therefore take this into account. The target users of the current CAD system are emergency physicians and other acute care physicians, and probably radiology residents as well. It is different from other existing CAD targeted for expert radiologists. This can potentially raise the stake of a wrong suggestion provided by CAD, as the less skilled reader are probably less capable of judging whether the output is correct or not. Indeed, results from the current study supported that

observers less skillful in image interpretation are more prone to be affected by the wrong outputs of a CAD. This was reflected by difference in the number of correct change in diagnosis versus the number of wrong change for the three groups of clinicians studied. As such, the potential of detrimental effect attributable to CAD may require more careful consideration. Fortunately, the number of mistakes were outweighed by the correct changes attributable to use of CAD.

Effective use of CAD requires training and experience with the system. For the cases used in this ROC study, the CAD system on its own achieved sensitivity of 90%(27/30) and specificity of 90%(27/30) on a per case basis. It was noteworthy that emergency physicians can improve their performance in terms of both sensitivity and specificity with CAD, just like the radiology residents and radiology specialists, but the sensitivity of emergency physicians with CAD is still lower than that of CAD alone. It has also been pointed out that they were more prone to accept some of the wrong outputs. It is therefore inferred that more education about AIH and the CAD system is more important for users of less expertise in image interpretation and is necessary before the CAD system can be effectively used.

4.7 LIMITATIONS OF THE OBSERVER PERFORMANCE STUDY

Independent Mode vs. Sequential Mode

The experimental procedure required that readers read all the cases without CAD support first, before they reread the same cases with CAD support after minimal delay. This design is more closely related to the sequential than the independent reading mode as described by Kobayashi(Kobayashi, Xu et al. 1996). The independent mode is the conventional method for conducting observer studies, when reading of images without and with CAD support are separated by a period so long that readers should have no recollection of the cases. For evaluation of CAD or other forms of adjuncts, the sequential mode, in which readers read each case first without, then with CAD support immediately afterwards, is the favored mode. It is because it mimics the way the CAD is supposed to be used, hence the potential benefit of the first reading to the second reading is a realistic experimental design rather than a bias. In addition, the sequential mode is more efficient in terms of reader time. We have modified the method by having the readers complete the readings in unaided mode first because we suspect that readers may change their level of suspicion or vigilance during the experiment should they become affected by feedback available from the CAD in a similar previous case. In other words, we hoped to avoid the training effect during the course of the study. The bottom line is that reader variance and measured results are

virtually the same for both reading modes (Beiden, Wagner et al. 2002).

Lesion Localization

It is recognized that the current study did not measure the performance in terms of lesion localization. The major flaw is that a ‘true positive’ response may be result of detection of noise or other mimicking artifacts rather than the genuine lesion. To provide this information, a location specific ROC (LROC) or free response ROC (FROC) study is required (Wagner, Beiden et al. 2002). This type of study would be more demanding due to the requirement of establishing the lesion database and the additional investment in readers’ reading time. In many instances the AIH may span several sections and hence the exact definition of correction localization can be elusive. After all, the presence or absence of AIH is more important than the quantity and precise localization of lesion. Therefore we believe a patient level ROC evaluation is sufficient to decide if the CAD can be beneficial for clinical management under the current local environment.

Actual Impact in Clinical Environment

It is widely accepted that MRMC ROC is very efficient for evaluation of diagnostic systems including CAD. However, even when the results are generalizable to the reader and case population under the study condition, whether or not the gain in

performance can realize its beneficial impact in the clinical environment is a matter of debate (Gur 2003).

4.8 FUTURE DEVELOPMENT

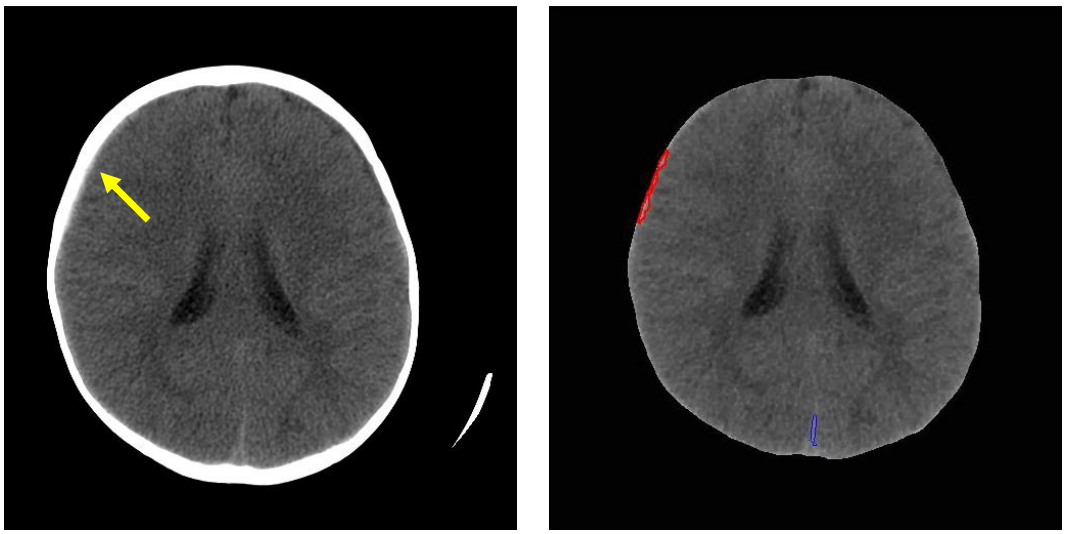
Application on Thin Section Images

The CAD system has been designed to work on conventional axial sections obtained from an old machine. The authors are currently working on adapting the system to images obtained from multidetector row CT machine. Initial results were promising. The impression was that more accurate results could be achieved because the images are less prone to artifacts (Jones, Kaplan et al. 2001) and contrast between AIH and normal parenchyma is higher, mainly because thinner sections make volume averaging less of a problem even for small lesions.

As a result. correction of CT cupping artifact can be less aggressive, so that sensitivity for very thin (< 2mm thick) extra-axial hemorrhage is increased. Also, a lower threshold of attenuation difference can be used such that smaller candidate lesions with lower attenuation can be detected. A successful case is shown in figure 21 .

In addition to the techniques already implemented in the current project, AIH can often be better characterized by 3 dimensional morphology or interrelation. For example, disc like lesions that span different sections would make extra-axial

Figure 21 – Successful conversion for MDCT images
An example of very thin subarachnoid hemorrhage, yellow arrow, in an original image ***a*** obtained from a 64 slice MDCT, being correctly detected by the system in ***b***, outlined in red.



a ***b***

hemorrhage far more likely than linear beam hardening artifacts that tend to be different even in contiguous plane. Such differentiation can be difficult on axial sections alone because both would appear as peripherally located linear hyperdensity.

Application on thin section images may benefit from more accurate classification system due to more exact anatomical localization based on some more developed brain atlases that are applicable to 3D images.

Possible Extension of Techniques to Other Applications

The concept of diagnosis based on anatomical information is well-recognized by

clinicians. The component of anatomical labeling developed in the current project can be used for systems aiming to detect other types of intracranial lesions that show different frequency of occurrence in different anatomical reigon. A very useful application that may also complement the current system would be CAD of acute infarct. Of course, the target lesion would be ill-defined low density regions rather than high density foci, Even tumor diagnosis can utilize the concept, because different kinds of tumors tend to affect different anatomical regions.

Improvement in Localization of Mid Sagittal Plane

The system requires human intervention in localization of MSP in less than 10% of the time. Emergency physicians participating in the ROC have been invited to use the manual alignment tool. It was observed that all of them localize the MSP accurately and efficiently.

Besides locating the MSP by finding the plane about which some parameters, e.g. intensity levels or some other features, are symmetrical, the MSP can be located by finding the interhemispheric fissure direct (Prima, Ourselin et al. 2002). The former method is the most commonly used and is used in the current CAD scheme. Its major limitation is that the brain is not necessarily symmetrical, in both normal and diseased individuals alike. Contrary to detection of symmetry, the latter method is insensitive to normal or abnormal asymmetry. However, the fissure often show some degrees of curvature, which may produce some meaningless planes. Also such methods are not readily adaptable to different imaging modalities. Yet another new method was

presented in RSNA 2004, which locate the MSP by finding the bony attachments of falx, which is essentially a modification of locating the interhemispheric fissure.

Since both approaches are imperfect due to different kinds of limitations, it is anticipated some combinations of both methods may produce more accurate localization of the MSP, which remains to be determined.

Development of More Versatile CAD

It has been noted that CAD research and development have been fragmented, in the sense that researchers concentrate on development of applications focusing on some unrelated tasks of narrow scopes (Summers 2003; Partain, Chan et al. 2005). It is expected that development of more versatile CAD systems or combinations of CAD schemes can provide comprehensive evaluation of the same images. For the current project, possible extension may include integration with CAD of infarct when the obvious alternate cause of stroke is ischemia or infarct, and CAD of aneurysm, as the cause of AIH. Such useful combinations can provide comprehensive evaluation of patients and better answer some important clinical questions of particular relevance to stroke management.

Integration with PACS

Another issue that needs to be considered before such system can be put into clinical practice is integration into the workflow of acute care physician or the emergency

room. It is important for the acceptance of CAD into clinical practice that it involves minimal additional cost of training. Doctors can become apprehensive when they have to relearn each and every type of CAD and even reluctant to use the CAD. In addition, Also, the current system is different from most of the current CAD systems designed for screening or routine reporting purpose, for which speed is of less concern. For applications to be useful for immediate management, immediate availability without significant additional cost of time is of utmost importance.

It is desirable that CAD is packaged with the PACS, which allows users to perform the interpretive task on their usual platform. It has been defined in the DICOM standard two structured report (SR) templates for mammography and chest CAD results. The Integrating the Healthcare Enterprise (IHE) has also published a post-processing workflow profile for integrating CAD applications with clinical PACS workflow (*http://www.ihe.net*). The outputs of the current CAD include pictures of the images together with overlay of the AIH, as well as textual data that indicates position and size of lesions. The pictures and data can be converted to screen capture and structured report data format supported by the DICOM standard and subsequently stored as a new series under the same study of the same patient from where the original images were obtained. These data can be integrated into the PACS using integration software. In fact, the current system has been one of the CAD

schemes being used for the testing of a recently developed CAD-PACS intergration toolkit (Zhou 2006; Zhou, Liu et al. 2007).

Implemetation for Daily Clinical Practice

We envisage that the CAD system can be implemented into daily clinical practice especially in the emergency room. It is believed that the system can be used as a triage tool for patients suffering from minor neurological disturbance or head injury. After CT is performed, clinicians can read the images with help from the system. If AIH has been excluded, the patients may be safely observed for a shorter period of time before discharge. Also, patients may be admitted to neurosurgical units or other units based on the presence of absence of AIH.

Parameters and hence performance of the CAD system may require adjustment according to the way the system is used. For example, supposing the CAD is intended to be used in ER for triaging patients suffering from head injury, who may be admitted if AIH is demonstrated, or undergo a short period of observation otherwise, high sensitivity is more important than low false positive rate. The system can also be adjusted to suit the users' characteristics, e.g. when the CAD is to be used by radiologist as a 'second reader', it can also afford to be more sensitive than specific, as suggested by our results that radiologists tend to be less affected by wrong outputs.

Chapter 5 CONCLUSION

A CAD system capable of identifying small intracranial hemorrhage has been developed. It extracts AIH candidates based on correction of its high intensity relative to its surrounding brain and contralateral anatomical structures. The system then classifies genuine AIH from mimicking normal variants or artifacts based on both image features and anatomical information, which is made possible by construction of a coordinate system that incorporates positional information of normal brain structures.

Results from observer performance study confirmed that clinicians, especially emergency physicians, improved their performance in detection of small AIH when they read brain CT with CAD.

It is expected that this system can benefit patient care especially in emergency situations when timely management decision need to be made by acute care physicians.

REFERENCES

DICOM. *http://medical.nema.org*.

IHE. *http://www.ihe.net/*.

LABMRMC & DBM MRMC. Available at *http://xray.bsd.uchicago.edu/krl/KRL_ROC/software_index.htm*.

MEDIC. Available at http://medic.rad.jhu.edu/download/public/.

SPM 5. Available at www.fil.ion.ucl.ac.uk/spm/.

Adams, H., R. Adams, et al. (2005). *Guidelines for the early management of patients with ischemic stroke: 2005 guidelines update a scientific statement from the Stroke Council of the American Heart Association/American Stroke Association*. Stroke **36**(4): 916-23.

Alberdi, E., A. Povykalo, et al. (2004). *Effects of incorrect computer-aided detection (CAD) output on human decision-making in mammography*. Acad Radiol **11**(8): 909-18.

Ambrose, J. (1973). *Computerized transverse axial scanning (tomography). 2. Clinical application*. Br J Radiol **46**(552): 1023-47.

Ambrose, J. (1974). *Computerized x-ray scanning of the brain*. J Neurosurg **40**(6): 679-95.

Ardekani, B. A., J. Kershaw, et al. (1997). *Automatic detection of the mid-sagittal plane in 3-D brain images*. IEEE Trans Med Imaging **16**(6): 947-52.

Bagley, L. J. (1999). *Imaging of neurological emergencies: trauma, hemorrhage, and infarction*. Semin Roentgenol **34**(2): 144-59.

Barrett, J. F. and N. Keat (2004). *Artifacts in CT: recognition and avoidance*. Radiographics **24**(6): 1679-91.

Beauchamp, N. J., Jr., P. B. Barker, et al. (1999). *Imaging of acute cerebral ischemia*. Radiology **212**(2): 307-24.

Beiden, S. V., R. F. Wagner, et al. (2002). *Independent versus sequential reading in ROC studies of computer-assist modalities: analysis of components of variance*. Acad Radiol **9**(9): 1036-43.

Bergstrom, M., K. Ericson, et al. (1977). *Variation with time of the attenuation values of intracranial hematomas*. J Comput Assist Tomogr **1**(1): 57-63.

Boris, P., F. Bundgaard, et al. (1987). *The CT (Hounsfield unit) number of brain tissue in healthy infants. A new reliable method for detection of possible degenerative disease*. Childs Nerv Syst **3**(3): 175-7.

Brem, R. F., J. Baum, et al. (2003). *Improvement in sensitivity of screening mammography with computer-aided detection: a multiinstitutional trial*. AJR Am J Roentgenol **181**(3): 687-93.

Broderick, J. P., H. P. Adams, Jr., et al. (1999). *Guidelines for the management of spontaneous intracerebral hemorrhage: A statement for healthcare professionals from a special writing group of the Stroke Council, American Heart Association.* Stroke **30**(4): 905-15.

Cohen, W. and L. Wayman (1992). *Computed tomography of intracranial hemorrhage*. Neuroimaging Clinics of North America **2**: 75-87.

Cordobes, F., R. D. Lobato, et al. (1981). *Observations on 82 patients with extradural hematoma. Comparison of results before and after the advent of computerized tomography*. J Neurosurg **54**(2): 179-86.

Diehl, J. T. (1978). *CT scanning in traumatic and emergency patients*. Comput Tomogr **2**(3): 183-7.

Doi, K. (2004). *Overview on research and development of computer-aided diagnostic schemes*. Semin Ultrasound CT MR **25**(5): 404-10.

Doi, K. (2005). *Current status and future potential of computer-aided diagnosis in medical imaging*. Br J Radiol **78 Spec No 1**: S3-S19.

Doi, K., H. MacMahon, et al. (1999). *Computer-aided diagnosis in radiology: potential and pitfalls*. Eur J Radiol **31**(2): 97-109.

Dorfman, D. D., K. S. Berbaum, et al. (1992). *Receiver operating characteristic rating analysis. Generalization to the population of readers and patients with the jackknife method*. Invest Radiol **27**(9): 723-31.

Erickson, B. J. and B. Bartholmai (2002). *Computer-aided detection and diagnosis at the start of the third millennium*. J Digit Imaging **15**(2): 59-68.

Freer, T. W. and M. J. Ulissey (2001). *Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center*. Radiology **220**(3): 781-6.

Gonzalez, R. C. and R. E. Woods (2002). Digital image processing. Upper Saddle River, N.J., Prentice Hall.

Goto, H., H. Aizawa, et al. (2005). *CAD System in the Emergency Medical Care for Abdominal and Head Trauma*. RSNA 2005, Chicago.

Gur, D. (2003). *ROC-type assessments of medical imaging and CAD technologies: a perspective*. Acad Radiol **10**(4): 402-3.

Gur, D., J. H. Sumkin, et al. (2004). *Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system*. J Natl Cancer Inst **96**(3): 185-90.

Hodgson, R., J. P. Wilson, et al. (2004). *CAD System for Detecting Haemorrhage in CT of Stroke*. RSNA 2004.

Huang, H. K. (2004). PACS and imaging informatics : basic principles and applications. Hoboken, N.J., Wiley-Liss.

Jagoda, A. S., S. V. Cantrill, et al. (2002). *Clinical policy: neuroimaging and decisionmaking in adult mild traumatic brain injury in the acute setting*. Ann Emerg Med **40**(2): 231-49.

Jones, T. R., R. T. Kaplan, et al. (2001). *Single- versus multi-detector row CT of the*

brain: quality assessment. Radiology **219**(3): 750-5.

Junck, L., J. G. Moen, et al. (1990). *Correlation methods for the centering, rotation, and alignment of functional brain images*. J Nucl Med **31**(7): 1220-6.

Klym, A. H., J. L. King, et al. (2004). *The effect of routine use of a computer-aided detection system on the practice of breast imagers: a subjective assessment*. Acad Radiol **11**(6): 711-3.

Kobayashi, T., X. W. Xu, et al. (1996). *Effect of a computer-aided diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs*. Radiology **199**(3): 843-8.

Lev, M. H., J. Farkas, et al. (1999). *Acute stroke: improved nonenhanced CT detection--benefits of soft-copy interpretation by using variable window width and center level settings*. Radiology **213**(1): 150-5.

Mader, T. J. and A. Mandel (1998). *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.

Maldjian, J. A., J. Chalela, et al. (2001). *Automated CT segmentation and analysis for acute middle cerebral artery stroke*. AJNR Am J Neuroradiol **22**(6): 1050-5.

Mark, D. G. and J. M. Pines (2006). *The detection of nontraumatic subarachnoid hemorrhage: still a diagnostic challenge*. Am J Emerg Med **24**(7): 859-63.

Metz, C. E. (1989). *Some practical issues of experimental design and data analysis in radiological ROC studies*. Invest Radiol **24**(3): 234-45.

Minoshima, S., K. L. Berger, et al. (1992). *An automated method for rotational correction and centering of three-dimensional functional brain images*. J Nucl Med **33**(8): 1579-85.

Mullins, M. E. (2006). *Modern emergent stroke imaging: pearls, protocols, and pitfalls*. Radiol Clin North Am **44**(1): 41-62, vii-viii.

New, P. F. and S. Aronow (1976). *Attenuation measurements of whole blood and*

blood fractions in computed tomography. Radiology **121**(3 Pt. 1): 635-40.

Norman, D., D. Price, et al. (1977). *Quantitative aspects of computed tomography of the blood and cerebrospinal fluid*. Radiology **123**(2): 335-8.

Panagos, P. D., E. C. Jauch, et al. (2002). *Intracerebral hemorrhage*. Emerg Med Clin North Am **20**(3): 631-55.

Partain, C. L., H. P. Chan, et al. (2005). *Biomedical Imaging Research Opportunities Workshop II: report and recommendations*. Radiology **236**(2): 389-403.

Peldschus, K., P. Herzog, et al. (2005). *Computer-aided diagnosis as a second reader: spectrum of findings in CT studies of the chest interpreted as normal*. Chest **128**(3): 1517-23.

Perry, J. J., I. G. Stiell, et al. (2005). *Attitudes and judgment of emergency physicians in the management of patients with acute headache*. Acad Emerg Med **12**(1): 33-7.

Phelps, M. E., E. J. Hoffman, et al. (1975). *Attenuation coefficients of various body tissues, fluids, and lesions at photon energies of 18 to 136 keV*. Radiology **117**(3 Pt 1): 573-83.

Pietka, E., A. Gertych, et al. (2001). *Computer-assisted bone age assessment: image preprocessing and epiphyseal/metaphyseal ROI extraction*. IEEE Trans Med Imaging **20**(8): 715-29.

Prima, S., S. Ourselin, et al. (2002). *Computation of the mid-sagittal plane in 3-D brain images*. IEEE Trans Med Imaging **21**(2): 122-38.

Reinus, W. R. and F. L. Zwemer, Jr. (1994). *Clinical prediction of emergency cranial computed tomography results*. Ann Emerg Med **23**(6): 1271-8.

Reisdorff, E. J. and D. T. Schwartz (2000). Emergency radiology. New York, McGraw-Hill Health Professions Division.

Ruttimann, U. E., E. M. Joyce, et al. (1993). *Fully automated segmentation of cerebrospinal fluid in computed tomography*. Psychiatry Res **50**(2): 101-19.

Sacks, W. M. (2003). *Estimating the effect of computer-aided detection on the sensitivity of screening mammography*. Radiology **226**(2): 597-8; author reply 598-9.

Sarkarati, D. and E. J. Reisdorff (2002). *Emergent CT evaluation of stroke*. Emerg Med Clin North Am **20**(3): 553-81.

Schriger, D. L., M. Kalafut, et al. (1998). *Cranial computed tomography interpretation in acute stroke: physician accuracy in determining eligibility for thrombolytic therapy*. Jama **279**(16): 1293-7.

Scott, W. R., P. F. New, et al. (1974). *Computerized axial tomography of intracerebral and intraventricular hemorrhage*. Radiology **112**(1): 73-80.

Shah, S. M. and K. M. Kelly (1999). Emergency neurology : principles and practice. Cambridge ; New York, Cambridge University Press.

Silverman, P. M. and J. A. Brink (1998). Helical (spiral) computed tomography : a practical approach to clinical protocols. Philadelphia, Lippincott-Raven.

Soille, P. (2003). Morphological image analysis : principles and applications. Berlin ; New York, Springer.

Summers, R. M. (2003). *Road maps for advancement of radiologic computer-aided detection in the 21st century*. Radiology **229**(1): 11-3.

Summers, R. M., C. D. Johnson, et al. (2001). *Automated polyp detection at CT colonography: feasibility assessment in a human population*. Radiology **219**(1): 51-9.

Talairach, J. and P. Tournoux (1988). Co-planar stereotaxic atlas of the human brain : 3-dimensional proportional system : an approach to cerebral imaging. Stuttgart ; New York, Thieme Medical Publishers.

Taveras, J. M. (1996). Neuroradiology. Baltimore, MD, Williams & Wilkins.

Thurman, D. and J. Guerrero (1999). *Trends in hospitalization associated with traumatic brain injury*. Jama **282**(10): 954-7.

van Gijn, J. and G. J. Rinkel (2001). *Subarachnoid haemorrhage: diagnosis, causes and management*. Brain **124**(Pt 2): 249-78.

Wagner, R. F., S. V. Beiden, et al. (2002). *Assessment of medical imaging and computer-assist systems: lessons from recent experience*. Acad Radiol **9**(11): 1264-77.

Weinstein, M. A., P. M. Duchesneau, et al. (1977). *White and gray matter of the brain differentiated by computed tomography*. Radiology **122**(3): 699-702.

Weiss, K. L., H. Pan, et al. (2003). *Clinical brain MR imaging prescriptions in Talairach space: technologist- and computer-driven methods*. AJNR Am J Neuroradiol **24**(5): 922-9.

Wysoki, M. G., C. J. Nassar, et al. (1998). *Head trauma: CT scan interpretation by radiology residents versus staff radiologists*. Radiology **208**(1): 125-8.

Yang, G. L., C. T. Lim, et al. (2005). *A Practical Computer-aided Diagnosis System for Intracranial Hemorrhage Detection in Acute Stroke*. RSNA 2005, Chicago.

Yoshida, H., J. Nappi, et al. (2002). *Computer-aided diagnosis scheme for detection of polyps at CT colonography*. Radiographics **22**(4): 963-79.

Zhou, Z. (2006). *A PACS-CAD toolkit for integrating an independent CAD workstation to diagnostic workflow based on DICOM SR and IHE profiles*. RSNA 2006, Chicago, USA.

Zhou, Z., B. J. Liu, et al. (2007). *CAD-PACS Integration Tool Kit Based on DICOM Secondary Capture, Structured Report and IHE Workflow Profiles*. Comput Med Imaging Graph **31**(4/5).

Zimmerman, R. A., W. A. Gibby, et al. (2000). Neuroimaging : clinical and physical principles. New York, Springer.