Automatic Detection and Classification of Brain Hemorrhages

MAHMOUD AL-AYYOUB, DUAA ALAWAD, KHALDUN AL-DARABSAH & INAD ALJARRAH
Jordan University of Science & Technology
Irbid 22110, Jordan
maalshbool@just.edu.jo, {duaa.alawad,khaldoun.mattar}@gmail.com, inad@just.edu.jo

Abstract: Computer-aided diagnosis systems have been the focus of many research endeavors. They are based on the idea of processing and analyzing images of different parts of the human body for a quick and accurate diagnosis. In this paper, the aforementioned approach is followed to detect whether a brain hemorrhage exists or not in a Computed Topography (CT) scans of the brain. Moreover, the type of the hemorrhage is identified. The implemented system consists of several stages that include image preprocessing, image segmentation, feature extraction, and classification. The results of the conducted experiments are very promising. A recognition rate of 100% is attained for detecting whether a brain hemorrhage exists or not. For the hemorrhage type classification, more than 92% accuracy is achieved.

Key–Words: brain hemorrhage, brain ct scans, machine learning, image processing, image segmentation

1 Introduction

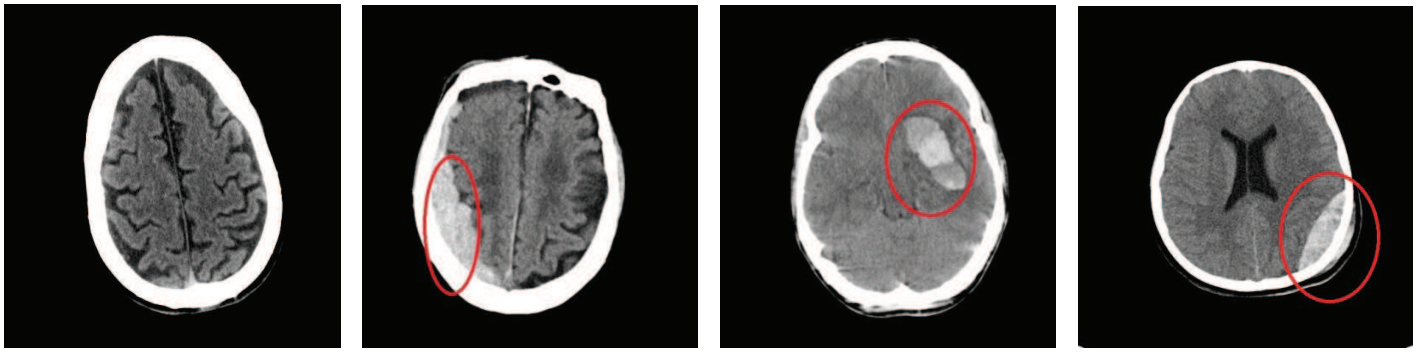
Nowadays, cerebrovascular diseases are the third cause of death in the world after cancer and heart diseases [16]. One of the cerebrovascular diseases is brain hemorrhage which is caused by a brain artery busting causing bleeding. There are many types of brain hemorrhage such as: epidural, subdural, subarachnoid, cerebral, and intraparenchymal hemorrhage. These types differ in many aspects such as the size of the hemorrhage region, its shape, and its location within the skull. In this work, a machine learning technique is proposed to diagnose brain hemorrhages based on medical images.

Brain hemorrhage can be diagnosed by two kinds of images: Clinical Head Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). After consulting many medical experts and reading through the literature [16, 1, 5, 7], CT images are chosen in this work. CT images are known to have many advantages over MRI such as: wider availability, lower cost and higher speed. Moreover, CT scanner might be favored over MRI scanner due to patient-related issues such as the patient being too large to fit in the MRI scanner, claustrophobic, has metallic or electrical implants or is unable to remain motionless for the duration of examination due to age, pain or medical conditions [5]. Finally, the quality of CT images is high enough to accurately diagnose brain hemorrhage.

From their name, Computer-Aided Diagnosis (CAD) systems use computers to help physicians reach a fast and accurate diagnosis. CAD systems are usually domain-specific as they are optimized for certain types of diseases, parts of the body, diagnosis methods, etc. They analyze different kinds of input such as symptoms, laboratory tests results, medical images, etc. depending on their domain. One of the most common types of diagnosis is the one that depends on medical images. Such systems are very useful since they can be integrated with the software of the medical imaging machine to provide quick and accurate diagnosis. On the other hand, they can be challenging since they combine the elements of artificial intelligence and digital images processing. This work presents a CAD system to help detect hemorrhages in CT scans of human brains and identify their types if they exist.

As with other CAD systems, the motivations for building this system are: (i) reducing human errors (it is well-known that the performance of human experts can drop below acceptable levels if they are distracted, stressed, overworked, emotionally unbalanced, etc.) and (ii) reducing the time/effort associated with training and hiring physicians. Eventually, this system can be integrated within the software of CT imaging devices to enable users to produce a quick and highly-accurate diagnosis while generating the image.

The proposed system can also be helpful for teaching and research purposes. It can be used to train senior medical students as well as resident doctors. Moreover, it can be very helpful in searching through large sets of CS scans [11]. Multi-slice CT scans can have up to 64 slices per scan. In modern hospitals, these images are stored in the standard Digital Imaging and Communications in Medicine (DICOM) format which includes text into the images. Any attempt



(a) Normal Brain (b) Subdural Hemorrhage (c) Intraparenchymal Hemorrhage (d) Epidural Hemorrhage

Figure 1: Normal brain images along with the three hemorrhage types considered in this paper.

to retrieve and display CT images must go through Picture Archives and Communication System (PACS) hardware [13]. This requires that the name of the patient or identity card number is provided to find any particular image. Thus,searching for some type of cases (e.g., for research purposes)is usually done manually, which is a very expensive task in terms of time and effort. Providing an automated tool that can go through the large number of images and identify the required cases quickly and with high accuracy can save huge amounts of time and effort.

In this paper, the problem of detecting brain hemorrhages in CT scans and classifying them is considered. As shown in Figure 1(a), an image of a normal brain shows a distribution of gray matter that appears as clearness in the texture-like fissures, while an abnormal brain has a shape which appears brighter than the normal gray matter as shown in the remaining images of Figure 1. Due to limitations in dataset collection, this work deals only with the following types: subdural hemorrhage (Figure 1(b)), intraparenchymal hemorrhage (Figure 1(c)) and epidural hemorrhage (Figure 1(d)). Such restrictions were faced by previous works [11], in which the authors focused on epidural hemorrhage, acute subdural hemorrhage and intracerebral hemorrhage.

Figure 1 shows the differences in the visual characteristics of the considered hemorrhage types. The epidural hemorrhage is characterized by its convex shape and its close fitting with the skull. The subdural hemorrhage is similar to the epidural one regarding its closeness to the skull; however, the size of subdural hemorrhage is larger than that of the epidural one and it has a crescent-like shape with a concave surface away from the skull. Finally, the intraparenchymal hemorrhage is characterized by its distance from the skull.

This paper is organized as follows: In the following section,related works are briefly discussed. The proposed methodology is discussed in section 3. In section 4, the experiments conducted to test the quality of the results generated by the proposed method is presented. The conclusion and the future work are discussed in section 5.

2 Related Works

In the past two decades, there has been much research on designing a machine learning based systems for detecting human brain hemorrhage and its types. Most of these approaches follow a similar framework to the one proposed in this work. The framework starts with an image preprocessing and segmentation phase followed by a feature extraction phase. Finally, different classifiers are used for the classification and testing phase. Below, a brief discussion of the merits of each of the related works is carried out.

Some of the older works [21, 22] addressed the problem of segmenting the region of intracerebral hemorrhages. In the former work, they used a spatially weighted k-means histogram-based clustering algorithm, whereas in the latter work, they applied a multiresolution simulated annealing method.

Cheng and Cheng [8] proposed a Fuzzy C-Means (FCM) method based on multiresolution pyramid for brain hemorrhage analysis. They also compared FCM, competitive Hopfield neural network [9] and fuzzy Hopfield neural network [18] in the global thresholding stage. In another work, Liu et al. [19] proposed an Alternative Fuzzy C-Means (AFCM) method for the segmentation phase. A more recent work based on FCM is the work of Li et al. [16]. The

authors used a thresholding technique based on FCM clustering to remove all non-brain regions. The results of this clustering were brain regions that were segmented into slices using median filtering to eliminate noise in the image, and then the maximum entropy threshold was computed for each slice chart to determine the potential hemorrhage regions. In the last phase, the hemorrhage regions were determined according to their locations and gray level statistics. The results of this work are extremely encouraging, however, as the authors acknowledge, it is still a premature work.

In [11], the authors focused on dividing the brain CT images into regions, where each region can either be a normal brain region or a hemorrhage region. Note that for images containing hemorrhages, the regions which do not include hemorrhage are treated as normal regions resulting in a highly-imbalanced dataset. For the image processing phase, the authors used an image segmentation scheme that uses ellipse fitting, background removal and wavelet decomposition technique. The weighted precision/recall values for this approach is fairly high 86.3% and 88.5%, however, as the authors acknowledge, these numbers are too optimistic since their dataset is imbalanced. Without considering the class of normal regions, the weighted precision/recall values range from 36.8% to 60%, which is much lower than those of our system.

The approach of [5] is based on segmenting the images into objects using the watershed method and extracting the features of each object. The extracted features are fed into a neural network for classification.

In [20], the authors address the problem of detecting which of the several slices comprising a single CT scan contain a hemorrhage. They start by classifying slices into ones that contain the nasal cavity region (see Figure 8(a)) and ones that contain the encephalic region (see Figure 1). Then, on the slices that contain the encephalic region, the authors use the wavelet transform to compute some features and use them to detect the presence of a hemorrhage.

A common approach followed by neurologists to detect abnormalities in brain CT scans is to split the brain region into two halves (hemispheres) and observe dissimilarities between them [10]. Several papers followed this approach. For example, Datta and Biswas [10] use this approach to detect the existence and location of intracerebral hemorrhages, whereas Chawla et al. [7] use it to determine the type of brain strokes (hemorrhage, chronic infarct and infarct). Both papers use histograms-based features for dissimilarity detection.

Finally, both Chan [6] and Liao et al. [17] address the problem of diagnosing intracranial hemorrhages.

3 Proposed Method

In this section, the proposed method is discussed in details. Firstly, a general overview (along with the pseudo-code) is presented followed by a more detailed discussion in the following subsections. The approach consists mainly of two parts: the image processing part (which consists of several subparts such as image preprocessing and enhancement, image segmentation and feature extraction) and the classification and testing part. Several approaches that can be useful for the image processing part exist [26, 3, 27, 2]. The focus here is on segmentation, which has been shown to be very useful in problems similar to the one at hand [4]. In the segmentation stage Otsu's method [25] is used to extract the hemorrhage region from the image. In the following step, discriminative features of the region of interest (ROI) are extracted. Finally, in the classification stage the image is classified based on the computed features of the ROI. MATLAB [23] code is written to carry out the image processing and segmentation as well as the features extraction parts, whereas the Weka tool [12] was used for the classification and testing parts.

3.1 Image Preprocessing and Segmentation

The first step is to remove the skull. Since it is the brightest part of the image, the intensity of its pixels is above 250 [11]. So, it can be easily removed using MATLAB. The following subsections contain a detailed discussions of the image processing and segmentation techniques followed to obtain the ROI. As an illustrative example, the results of applying these techniques on the CT scan of Figure 2 are shown.

Image Segmentation. The following step is segmentation, which is the base of many image-based CAD systems such as [14, 4]. Generally, it is a method to segment a digital image into many regions based on some criteria such as sets of pixels, etc. The goal of the segmentation is to simplify an image to be more meaningful and easier to analyze. Many approaches for segmentation exist such as thresholding and clustering. This phase segments the brain image into several regions so that we can isolate the ROI (the hemorrhage region). As stated in the following discussion, the outcome of this phase can be used to detect the existence of hemorrhages with 100% accuracy.

The segmentation method used in this work is Otsu's method. In spite of its simplicity and efficiency, the quality of the outcomes produced by Otsu's methods is high enough for the purposes of this work. Otsu's method divides the pixels of an image into several classes by automatically finding a thresh-

Algorithm 1 Brain Hemorrhage Type Detection (BHTD)

procedure BHTD(S, m, n, C, W) ▷ S is the dataset of m brain CT scans. n is the number of features. C is the set of classifiers to be tested along with their parameter vectors W.
  $D \leftarrow$ BUILDDATASET(S, m, n)
  for all $c \in C$ **do**
    $A_c \leftarrow$ TESTCLASSIFIER(D, c, w_c)
  end for
  return c with best accuracy measure A_c
end procedure

procedure BUILDDATASET(S, m, n) ▷ S is the dataset of m brain CT scans. n is the number of features.
  Initialize D to an empty n-column table
  for $i \leftarrow 0, 9$ **do**
    $t \leftarrow$ EXTRACTFEATURES(i, n)
    Append t into the end of D
  end for
  return D
end procedure

procedure EXTRACTFEATURES(i, n) ▷ i is the brain CT scan.
  If the intensity of any pixel in i is ≥ 250, it is set to 0
  Apply Ostu's segmentation technique on i (See Section 3.1 for more details)
  Apply the morphological opening step on i (See Section 3.1 for more details)
  Apply the region growing step on i (See Section 3.1 for more details)
  On the resulting image, apply MATLAB's `regionprops` in addition to our own code to compute the n-element feature vector v_i of the ROI (See Section 3.2 for more details)
  Apply the classifier
  return v_i
end procedure

procedure TESTCLASSIFIER(D, c, w_c) ▷ D is the dataset of brain CT scans. c is a classifier with parameter vector w_c. The testing will be using the 10-fold cross-validation technique.
  Randomly partition D into 10 subsets: $D_0, \ldots, D_9$
  for $j \leftarrow 0, 9$ **do**
    $Ts \leftarrow D_j$ ▷ Ts is the testing set.
    $Tr \leftarrow \left(\bigcup_{k=0,\ldots,9} D_k\right) \setminus Ts$ ▷ Tr is the training set.
    For classification method c, build a model using parameters w_c and Tr on the Weka tool
    Use the model on Ts and compute the accuracy measures, A_j, discussed in Section 4
    return the average of all A_j's for $j = 0, \ldots, 9$
  end for
end procedure

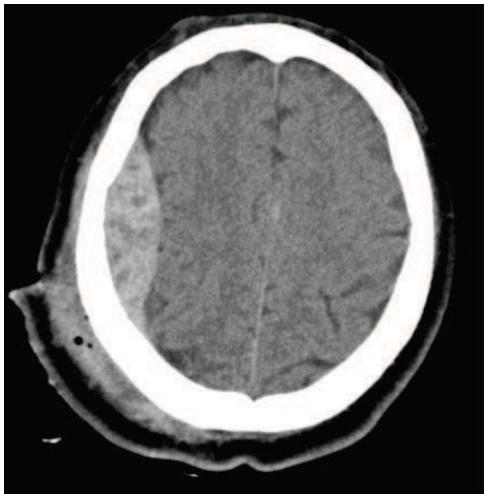


Figure 2: A CT scan of epidural hemorrhage that will be used an example for the image processing and segmentation techniques used in this work.

Figure 3: The result of applying the segmentation technique on Figure 2.

old to minimize the within class variance. It basically looks at the histogram, pixel values, and the probability in order to obtain a segment. Moreover, it does not look at edges; it looks at region inside the segment we want to segment out.

The basic idea behind Otsu's method is that the weighted within-class variance is minimized using Equation 1 which represents a measurement of the compactness of the classes (two classes are assumed here).

$$\sigma_w^2 = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t), \quad (1)$$

where t is the threshold and for each class k, $q_k(t)$, $\mu_k(t)$ and $\sigma_k^2(t)$ are the probability, the mean and the variance of the class defined as follows.

$$\begin{aligned}
q_1(t) &= \sum_{i=0}^{t} P(i), \\
q_2(t) &= \sum_{i=t+1}^{I} P(i), \\
\mu_1(t) &= \sum_{i=0}^{t} \frac{iP(i)}{q_1(i)}, \\
\mu_2(t) &= \sum_{i=t+1}^{I} \frac{iP(i)}{q_2(i)}, \\
\sigma_1^2(t) &= \sum_{i=0}^{t} [i-\mu_1(t)]^2 \frac{iP(i)}{q_1(i)}, \\
\sigma_2^2(t) &= \sum_{i=t+1}^{I} [i-\mu_2(t)]^2 \frac{iP(i)}{q_2(i)},
\end{aligned}$$

where $I = 255$ is the number of bins in the histogram.

Simply, the best threshold can be found exhaustively by just trying all possible values of t (i.e., the range $[0, 255]$), and computing σ_w^2 for each value. Finally, the value that obtains the lowest value for σ_w^2 will be selected as the threshold. Figure 3 shows the result of applying the segmentation technique.

Morphological Operations and Region Growing. From Figure 3, it can be seen that further work is needed to accurately obtain the ROI. A discussion of how to improve the results via the use of morphological operations and region growing follows below.

Morphological Operations. Mathematical morphology is a technique used to analyze images by a process of geometrical structures. It is used on digital images, graphs, etc. There are two basic operations of Mathematical morphology: erosion, which is eroding away the boundaries of foreground, and dilation, which is enlarging the boundaries of foreground pixels. The erosion and dilation are defined by the following equations.

$$A \ominus B = \{z|(B)_z \subset A\} \quad (2)$$

$$A \oplus B = \{z|(\widehat{B})_z \cap A \neq \emptyset\} \quad (3)$$

where A and B are the image and the structuring element, respectively. The structuring element used in this work is flat and disk-shaped as shown in Figure 4.

In this work, the opening operation, which consists of an erosion operation followed by a dilation operation, is used to filter out the small parts in the image that cannot contain the suspicious region. The opening operation is given by the following equation.

$$A \odot B = (A \ominus B) \oplus B. \quad (4)$$

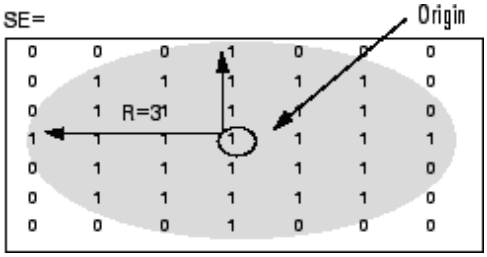


Figure 4: The flat disk-shaped structuring element used in this work [23].

Figure 5: The result of applying the morphological opening step on Figure 3.

The result of the morphological operations is an image that contains the potential regions that may contain the suspicious mass, as shown in Figure 5.

Region Growing. A region growing algorithm is used to obtain the whole mass. Region growing starts with a seed region and grows it by appending to it those neighboring pixels that have properties similar to the seed region. The seed region used here is the result of the previous step. Figure 6 shows the result of this phase.

At the end of this stage, the ROI is obtained. Figure 7 shows the result after applying these techniques to the images of Figure 1. As can be seen in these images, the ROI was selected accurately, which is crucial to the success of the proposed system. Also, note that the result of applying our techniques on a normal brain image is a completely blank image, whereas brain images with any hemorrhage will have some non-black region. This is observed on all images of our dataset resulting in a 100% accurate differentiation between normal and abnormal brain images.

Figure 6: The result of applying the region growing step on Figure 5.

To further illustrate the strength and accuracy of the image preprocessing and segmentation techniques used in this work, Figure 8 shows how such techniques can handle a difficult case of epidural hemorrhage. As discussed in Section 2, the difficulty of this case is due to the existence of the nasal cavity region [16]. After preprocessing the images, the next step is to extract discriminating features.

3.2 Feature Extraction

Extracting useful and discriminating features from images can be one of the most crucial steps in any system like ours. Below, a brief description of the features extracted about the region of interest (ROI) is given.

- The size (in pixels) of the ROI.
- The centroid of the ROI.
- the perimeter of the ROI.
- The distance between the ROI and the skull, which is expected to be very small for both epidural and subdural hemorrhages and relatively large for intraparenchymal hemorrhage.
- The diameter of a circle with the same size as the ROI.
- The ratio of pixels in the ROI to pixels in the convex hull of the ROI, which should be useful in distinguishing between hemorrhage types with commonly convex shapes such as epidural

(a) Normal Brain (b) Subdural Hemorrhage (c) Intraparenchymal Hemorrhage (d) Epidural Hemorrhage

Figure 7: Images of Figure 1 after applying image preprocessing and segmentation techniques.

and intraparenchymal hemorrhages and hemorrhage types with commonly concave shapes such as subdural hemorrhage.

- The smallest bounding box containing the ROI. The following features are related to this bounding box.
  - The number of pixels in a binary image of the same size as the bounding box of the ROI.
  - The ratio of pixels in the ROI to pixels in the bounding box of the ROI.
- The following features are related to the ellipse that has the same second-moments as the ROI.
  - The eccentricity of the ellipse.
  - The length of the major and minor axes of the ellipse.
  - The angle (in degrees ranging from -90 to 90 degrees) between the x-axis and the major axis of the ellipse.

MATLAB's `regionprops` tool is the main tool used to compute most of these features.

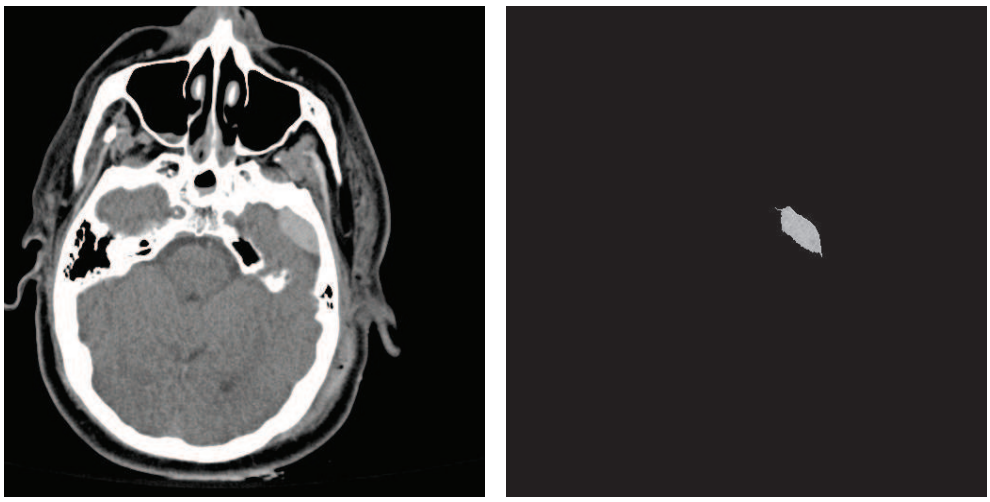
As discussed in the second to last paragraph of Section 1, the considered hemorrhage types have different visual characteristics. For example, the epidural hemorrhage is characterized by its convex shape and its close fitting with the skull. The subdural hemorrhage is similar to the epidural one regarding its closeness to the skull; however, the size of subdural hemorrhage is larger than that of the epidural one and it has a crescent-like shape with a concave surface away from the skull. The intraparenchymal hemorrhage is characterized by its distance from the skull. The considered features are selected to reflect different aspects related the size, shape and position of the hemorrhage region (ROI), which would be very helpful in differentiating between the different types of hemorrhage. Indeed, Section 4 show that the features related to the ellipse that has the same second-moments as the ROI such as the axes length and the eccentricity are the most helpful for our purposes.

4 Results

In this section, the experiments conducted to test the quality of the results generated by the proposed method are discussed. The collected dataset consists of 76 CT images of human brain. 25 of these images are for normal brain while the remaining images represent brains that suffer from one of the three types of brain hemorrhage considered in this work shown in Figure 1. The distribution of images is as follows: 17 images for epidural hemorrhage, 20 images for subdural hemorrhage and 14 images for intraparenchymal hemorrhage. The images were collected from the Internet websites (such as [24]) and from King Abdullah University Hospital (KAUH) in Irbid, Jordan.

In the carried experiments, several classifiers were tested, especially those that performed well in previous works related to this work such as Neural Networks, Support Vector Machines, Decision Trees, etc. The Weka tool was used to run the experiments. So, the discussion is restricted to the implementations used in the Weka tool. As for testing and evaluation, the widely-used 10-fold cross-validation technique was used.

The accuracy measures used to evaluate the performance of the proposed classifiers are the precision, the recall, the F-measure and the AUC, which is



(a) Original image containing a difficult case of epidural hemorrhage.

(b) The image after the preprocessing and segmentation stage.

Figure 8: Image preprocessing and segmentation.

the area under the Receiver Operating Characteristic (ROC) curve. The following equations define the precision, the recall and the F-measure, respectively [2]:

$$\begin{aligned} \text{Precision} &= \frac{TP}{TP+FP} \\ \text{Recall} &= \frac{TP}{TP+FN} \\ \text{F} &= 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \end{aligned}$$

where TP, FP, TN and FN are the numbers of true positives, false positives, true negatives and false negatives, respectively. These are the main performance measures used in the literature as they capture the two error types of interest (false positives and false negatives) and the relationship between them and the number of correctly classified instances (true positives and true negatives).

The objective of this work is to correctly classify brain CT images into one of four classes: normal, epidural hemorrhage, subdural hemorrhage and intraparenchymal hemorrhage. The first noteworthy observation from the experiments is the speed of the proposed system, which took an average of 2.1 seconds on each image. The second noteworthy observation is the system's ability to distinguish between normal and abnormal cases. For each normal case, the ROI resulting from the image preprocessing and segmentation part is a blank image (all pixels are black); as shown in Figure 7(a). This makes the binary classification problem of determining the existence of a brain hemorrhage straightforward. Hence, the problem at hand is reduced to the 3-class classification problem of determining which type of hemorrhage exist.

For this problem, several classifiers were tested and many of them achieved 90% accuracy or better. The accuracy measures for these classifiers are reported in Table 1 and visually depicted in Figure 9. The best results was obtained through a multinomial logistic regression classifier (Logistic) with a ridge estimator based on the work of Le Cessie and Van Houwelingen [15]. Another classifier with comparable performance is the feed forward artificial neural network classifier , MultilayerPerceptron (MLP), which uses back propagation for training and the sigmoid function at each node. The neural network's structure is set to the defaults used in the Weka tool. For example, the default value for the number of hidden layers is defined as half of the sum of the number of classes and the number of attributes, which in this case is 10. Finally, other classifiers with 90% accuracy or better are: the Bayesian Network classifier (BayesNet), the decision tree classifier (J48) and the support vector machine classifier (SMO).

Since so many classifiers are performing well, the discriminative power of the computed features has to be high. Using the class separability measures defined below, the features set discriminative power were evaluated. A within class scatter variance is calculated using

$$\sigma_w^2 = \sum_i P_i \times [(x-\mu_i)(x-\mu_i)^T],$$

Table 1: Accuracy measures of applying several classifiers to the hemorrhage-type identification problem.

Algorithm	Precision	Recall	F-Measure	AUC
BayesNet	90.1%	90.2%	90.1%	0.956
J48	90.4%	90.2%	89.9%	0.954
Logistic	**92.5%**	**92.2%**	**92.1%**	**0.961**
MLP	92.1%	92.2%	92.1%	0.956
SMO	90.4%	90.2%	90.2%	0.928

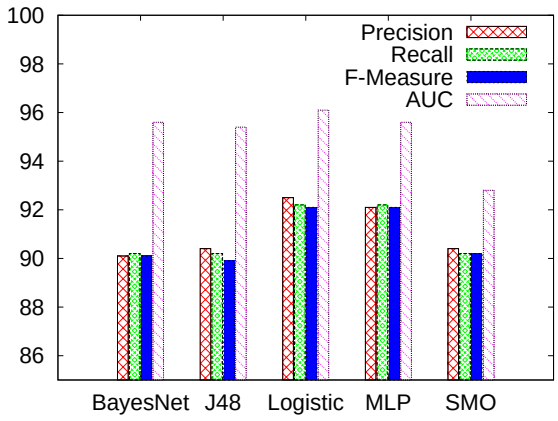


Figure 9: Accuracy measures of applying several classifiers to the hemorrhage-type identification problem.

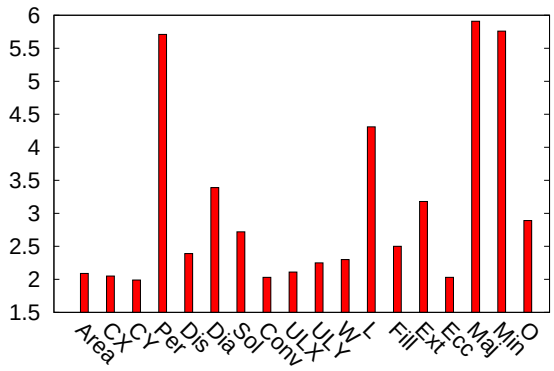


Figure 10: The separability measures of each feature.

where μ_i is the mean of the i^{th} class and P_i is the probability of the occurrence of class i in the training data. It calculates the weighted variance of the samples of each class. The overall scatter variance is calculated using

$$\sigma_0^2 = \sum_i (x - \mu_0)(x - \mu_0)^T,$$

where μ_0 is the overall mean. The discriminating measure D is calculated as

$$D = \frac{\sigma_0^2}{\sigma_w^2}.$$

The value of D was estimated to be 3.295.

To gain a better insight into which features are the most discriminating, the separability measures of each feature are reported in Table 2 and visually depicted in Figure 9. As shown in this table and this figure, the features with the highest discriminative power are the ones related to the axes length of the ellipse that has the same second-moments as the ROI, the perimeter of the ROI and the length of the smallest bounding box containing the ROI.

Other techniques were used to evaluate the discriminative power of the feature set. Using techniques such as information gain, gain ratio, ReliefF and correlation ranking, it was observed that the features related to the ellipse that has the same second-moments as the ROI such as the axes length and the eccentricity have more discriminative power than the features related to the ROI or its bounding box. This is consistent with the numbers computed using the equations of the previous paragraph.

5 Conclusions and Future Work

Automated systems for analyzing and classifying medical images have gained a great level of attention recently. One such example is the problem of detecting whether a brain hemorrhage exists (the binary classification problem) and if it exist what type is it (the multi-class classification problem). The experiments presented in this work showed that after preprocessing CT Scans, the binary classification problem was solved with 100% accuracy. Also, the implemented system achieved more than 92% accuracy for the 3-class classification problem of determining the hemorrhage type using neural networks as a classifier.

The results are encouraging and a higher accuracy

Table 2: The separability measures of each feature.

Feature	Separability Measure
Area	2.09
Centroid x-coordinate (CX)	2.05
Centroid y-coordinate (CY)	1.99
Perimeter (Per)	**5.71**
Distance (Dist)	2.39
Diameter (Dia)	3.39
Solidity (Sol)	2.72
Convex Area (Conv)	2.03
Bounding box's upper-left corner x-coordinate (ULX)	2.11
Bounding box's upper-left corner y-coordinate (ULY)	2.25
Bounding box's width (W)	2.3
Bounding box's length (L)	**4.31**
Filled Area (Fill)	2.5
Extent (Ext)	3.18
Eccentricity (Ecc)	2.03
Major Axis Length (Maj)	**5.91**
Minor Axis Length (Min)	**5.76**
Orientation (O)	2.89

for the 3-class classification problem can be attained by obtaining a better dataset with high resolution images taken directly from the CT scanner. Moreover, different feature extraction and feature selection techniques can be employed to improve the performance of the system. Finally, ensemble classifiers for classification will be considered as a future work to accomplish higher accuracy. This will allow the future system to reach a level that will allow it to be a significant asset to any medical establishment dealing with brain hemorrhages.

References:

[1] H. P. Adams, G. del Zoppo, M. J. Alberts, D. L. Bhatt, L. Brass, A. Furlan, R. L. Grubb, R. T. Higashida, E. C. Jauch, C. Kidwell, P. D. Lyden, L. B. Morgenstern, A. I. Qureshi, R. H. Rosenwasser, P. A. Scott, and E. F. Wijdicks. Guidelines for the early management of adults with ischemic stroke. *Stroke*, 38(5):1655–1711, 2007.

[2] M. Al-Ayyoub and D. Al-Zghool. Determining the type of long bone fractures in x-ray images. *WSEAS Transactions on Computers*, 2014. To appear.

[3] M. Al-Ayyoub, M. T. Irfan, and D. G. Stork. Machine learning of multi-feature visual texture classifiers for the authentication of Jackson Pollock's drip paintings. volume 7869, 2011.

[4] K. Al-Darabsah and M. Al-Ayyoub. Breast cancer diagnosis using machine learning based on statistical and texture features extraction. In *4th International Conference on Information and Communication Systems (ICICS 2013)*, 2013.

[5] U. Balasooriya and M. Perera. Intelligent brain hemorrhage diagnosis using artificial neural networks. In *IEEE Business Engineering and Industrial Applications Colloquium (BEIAC)*, pages 128–133, 2012.

[6] T. Chan. Computer aided detection of small acute intracranial hemorrhage on computer tomography of brain. *Computerized Medical Imaging and Graphics*, 31(4):285–298, 2007.

[7] M. Chawla, S. Sharma, J. Sivaswamy, and L. Kishore. A method for automatic detection and classification of stroke from brain ct images. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3581–3584, sept. 2009.

[8] D. Cheng and K. Cheng. Multiresolution based fuzzy c-means clustering for brain hemorrhage analysis. In *Proceedings of the 2nd International Conference on Bioelectromagnetism*, pages 35–36. IEEE, 1998.

[9] K. Cheng, J. Lin, and C. Mao. The application of competitive hopfield neural network to medical image segmentation. *IEEE Transactions on Medical Imaging*, 15(4):560–567, 1996.

[10] A. Datta and B. Biswas. A fuzzy multilayer perceptron network based detection and classification of lobar intra-cerebral hemorrhage from computed tomography images of brain. In *2011 International Conference on Recent Trends in Information Systems (ReTIS)*, pages 257–262. IEEE, 2011.

[11] T. Gong, R. Liu, C. L. Tan, N. Farzad, C. K. Lee, B. C. Pang, Q. Tian, S. Tang, and Z. Zhang. Classification of ct brain images of head trauma. In *Proceedings of the 2nd IAPR international conference on Pattern recognition in bioinformatics*, PRIB'07, pages 401–408, 2007.

[12] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten. The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18, 2009.

[13] H. Huang. PACS: basic principles and applications. *European Radiology*, 10(2):303–303, 2000.

[14] Z. Kerekes, Z. Tóth, S. Szénási, and S. Sergyán. Colon cancer diagnosis on digital tissue images. In *2013 IEEE 9th International Conference on Computational Cybernetics (ICCC)*, pages 159–163. IEEE, 2013.

[15] S. Le Cessie and J. Van Houwelingen. Ridge estimators in logistic regression. *Applied statistics*, pages 191–201, 1992.

[16] Y. Li, Q. Hu, J. Wu, and Z. Chen. A hybrid approach to detection of brain hemorrhage candidates from clinical head ct scans. In *Sixth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pages 361–365, 2009.

[17] C.-C. Liao, F. Xiao, J.-M. Wong, and I.-J. Chiang. A knowledge discovery approach to diagnosing intracranial hematomas on brain ct: recognition, measurement and classification. In *Medical Biometrics*, pages 73–82. Springer, 2007.

[18] J. Lin, K. Cheng, and C. Mao. A fuzzy hopfield neural network for medical image segmentation. *IEEE Transactions on Nuclear Science*, 43(4):2389–2398, 1996.

[19] H. Liu, C. Xie, Z. Chen, and Y. Lei. Segmentation of ultrasound image based on morphological operation and fuzzy clustering. In *Third IEEE International Workshop on Electronic Design, Test and Applications (DELTA)*, 2006.

[20] R. Liu, C. L. Tan, T. Y. Leong, C. K. Lee, B. C. Pang, C. Lim, Q. Tian, S. Tang, and Z. Zhang. Hemorrhage slices detection in brain ct images. In *19th International Conference on Pattern Recognition (ICPR 2008)*, pages 1 –4, dec. 2008.

[21] S. Loncaric, A. Dhawan, J. Broderick, and T. Brott. 3-d image analysis of intra-cerebral brain hemorrhage from digitized ct films. *Computer Methods and Programs in Biomedicine*, 46(3):207–216, 1995.

[22] S. Loncaric and Z. Majcenic. Multiresolution simulated annealing for brain image analysis. In *Medical Imaging'99*, pages 1139–1146. International Society for Optics and Photonics, 1999.

[23] MATLAB. *version 7.13 (R2011b)*. The MathWorks Inc., Natick, Massachusetts, 2011.

[24] N. J. Oldnall. xray2000, 2012. `http://e-radiography.net` [Online; accessed December-2012].

[25] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.

[26] A. Rövid, I. J. Rudas, S. Sergyán, and L. Szeidl. Hosvd based image processing techniques. In *10th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases*, pages 297–302, 2011.

[27] S. Sergyan. Precision improvement of content-based image retrieval using dominant color histogram descriptor. In *Recent Advances in Image, Audio and Signal Processing, 1st WSEAS International Conference on Image Processing and Pattern Recognition*, pages 197–203, 2013.