Predictability of intracranial pressure level in traumatic brain injury: features extraction, statistical analysis and machine learning-based evaluation

Wenan Chen

Virginia Commonwealth University,
Reanimation Engineering Science (VCURES) Centre,
Department of Computer Science,
Virginia Commonwealth University,
Richmond 23298, VA, USA
E-mail: chenw6@vcu.edu

Charles H. Cockrell

Virginia Commonwealth University,
Reanimation Engineering Science (VCURES) Centre,
Department of Radiology,
Virginia Commonwealth University,
Richmond 23298, VA, USA
E-mail: chcockrell@vcu.edu

Kevin Ward

Virginia Commonwealth University,
Reanimation Engineering Science (VCURES) Centre,
Department of Emergency Medicine,
Virginia Commonwealth University,
Richmond 23298, VA, USA
E-mail: krward@vcu.edu

Kayvan Najarian*

Virginia Commonwealth University,
Reanimation Engineering Science (VCURES) Centre,
Department of Computer Science,
Virginia Commonwealth University,
Richmond 23298, VA, USA
E-mail: knajarian@vcu.edu
*Corresponding author

Abstract: This paper attempts to predict Intracranial Pressure (ICP) based on features extracted from non-invasively collected patient data. These features include midline shift measurement and textural features extracted from Computed axial Tomography (CT) images. A statistical analysis is performed to examine the relationship between ICP and midline shift. Machine learning

is also applied to estimate ICP levels with a two-stage feature selection scheme. To avoid overfitting, all feature selections and parameter selections are performed using a nested 10-fold cross validation within the training data. The classification results demonstrate the effectiveness of the proposed method in ICP prediction.

Keywords: intracranial pressure prediction; texture analysis; statistical analysis; machine learning; nested cross validation.

Reference to this paper should be made as follows: Chen, W., Cockrell, C.H., Ward, K. and Najarian, K. (2013) 'Predictability of intracranial pressure level in traumatic brain injury: features extraction, statistical analysis and machine learning-based evaluation', *Int. J. Data Mining and Bioinformatics*, Vol. 8, No. 4, pp.480–494.

Biographical notes: Wenan Chen received his PhD Degree from the Department of Computer Science, Virginia Commonwealth University. He is currently working as a post-doc. His research interests include medical image/signal processing, data mining, machine learning and data analysis using statistical models.

Charles H. Cockrell, MD is an Associate Professor in the Department of Radiology at Virginia Commonwealth University, and the Section Chief of Emergency Radiology at VCU Medical Centre.

Kevin Ward, MD is a Professor and Associate Chair in the Department of Emergency Medicine at Virginia Commonwealth University. He is the Director of the Virginia Commonwealth University Reanimation Engineering Science Centre. He received his MD Degree from Tulane University in New Orleans, LA. His research interest lies in developing better means of monitoring victims of critical illness and injury to improve outcomes.

Kayvan Najarian is an Associate Professor at the Department of Computer Science, Virginia Commonwealth University. His research interests include biomedical signal/image processing, machine learning, and medical decision-making. He has published more than 200 peer-reviewed publications in reputable journals and conferences, and has generated multiple patents, mainly in the field of computer-assisted clinical decision making. His research projects are funded by multiple agencies, including the NSF, NIH, and DoD, as well as private companies. He has authored two books which are also being utilised as textbooks in universities around the world.

1 Introduction

Elevated Intracranial Pressure (ICP) is a very common secondary injury in Traumatic Brain Injuries (TBI) and can result in potentially deadly consequences such as ischemia or herniation if not promptly treated (Langlois et al., 2006). A standard and accurate way of monitoring ICP requires cranial trepanation, which requires special skill and facilities, and may result in complications such as infection. As such, non-invasive estimation of ICP is highly desirable, even though it may not be as accurate as invasive methods. Such estimation of ICP can serve as a pre-screening procedure to assist physicians in making decisions on whether or not invasive monitoring/surgery is necessary.

The objective of this paper is to provide an estimation of ICP level using all available patient data. Statistical analysis between midline shift and ICP values are also presented. Computed axial Tomography (CT) scanning was the sole imaging source to examine brain state in this paper. The reason for choosing CT scan instead of other imaging modalities, such as Magnetic Resonance Imaging (MRI), is that the CT scan is still the gold standard for initial TBI assessment, in particular for emergency departments which are often the environment where TBI patients are first treated (Ernest et al., 2003). In this paper, multiple non-intrusive sources of measurement are collected and analysed to extract informative features with respect to ICP. Afterwards, regression analysis is performed using midline shift and ICP values. Finally a classification model is built with those extracted features to evaluate the predictability of ICP levels. The study was approved by the institutional review boards of the Carolinas Medical Centre and Virginia Commonwealth University.

2 Data

2.1 Data preparation

The collected data includes 17 patients with mild to severe TBI, which were provided by the Carolinas Healthcare System (CHS). Most patients have multiple CT scans. Some have surgeries performed between two consecutive CT scans. In total, 57 CT scans were used for this study. With each CT scan, CT slices that show ventricles or regions that should have contained ventricles were selected for image analysis because these regions contain the features of interest that are extracted, as explained below. Usually five slices from each CT scan are selected in this process.

2.2 ICP values and ICP levels

For each patient, the ICP value was recorded every hour for the time periods in which the analysed CT scans were obtained. To associate the ICP value with each CT scan, the two closest measurements of ICP to the time of CT scan acquisition (both within an hour) are averaged and assigned as the ICP value at the time of the CT scan. The ICP values are grouped into two classes: elevated ICP with ICP >12 mm Hg and normal ICP with ICP ≤12 mm Hg. With this ICP level thresholding, there are 33 cases of normal and 24 cases of elevated ICP. Note that for most of the patients, they have both elevated ICP and normal ICP associated with different CT images over time.

3 Features extraction

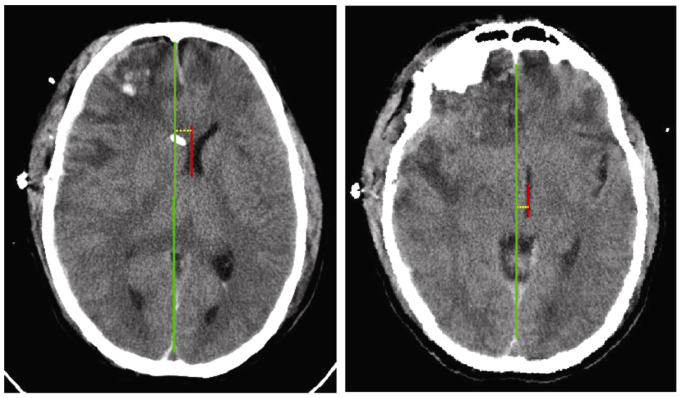
In TBI, haematoma and oedema are two common pathological symptoms, e.g., they may exist in the case of brain contusion. Because of the swelling effect from haematoma or oedema, the brain structure may shift along with the midline, the fissure separating the two hemispheres. This phenomenon is termed ‘midline shift’ (Gruen, 2002). This midline shift was measured as one candidate feature for predicting the ICP level because it is usually associated with elevated ICP (Maas et al., 2008). Another useful feature is the amount of extravascular blood in brain tissue due to hematoma or haemorrhage.

The bleeding area may directionally distort other brain tissues and increase the pressure inside the brain. The third set of features investigated is the textural patterns of brain tissue. Because of high pressure inside the brain, the texture of the grey matter and white matter may change. For example, when there is oedema, the interface between grey and white matter becomes ill-defined and oedema tends to decrease attenuation of X-ray. All three set of features are extracted from brain CT scans. Specifically, all features are extracted from each CT scan slice, and then aggregated by combining the same type of features from multiple slices in a CT scan to represent the state of the full CT scan. In addition to these features extracted from CT scans, other patient data including demographic information, such as patient age as well as injury score, e.g., Injury Severity Score (ISS) are used. These features may also provide extra information to the model of ICP level prediction. In the following sections, the extraction of each set of features is described in more detail.

3.1 Midline shift

To estimate the midline shift, the ideal midline, i.e., the midline if there were no brain injury, is first identified. Specifically, the ideal midline is estimated by detecting the anterior bone protrusion or anterior falx cerebri and the posterior part of the falx cerebri. The position of the actual midline in a CT scan can be estimated using the ventricular system and cistern structures filled with Cerebrospinal Fluid (CSF). Figure 1 illustrates the estimation of midline shift for two CT slices. To form the 'ground truth', the determination of the ideal midline and actual midline is done for each CT slice manually by the participating radiologist.

Figure 1 Midline shift calculation on different CT slices: (a) shows the calculation based on bilateral ventricles and (b) shows the calculation based on the third ventricle (see online version for colours)



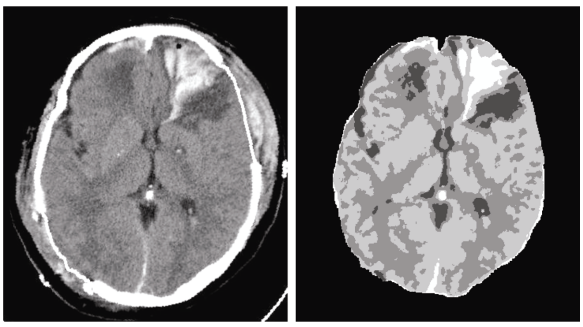
(a) (b)

3.2 Haematoma volume

The measurement of haematoma volume is based on a segmentation method using the Gaussian Mixture Model (GMM) performed on the CT slices (Chen and Najarian, 2009). The segmented CT slices classify each pixel into four categories: blood, ventricles (CSF),

grey matter and white matter. Figure 2 shows the segmented result from the segmentation method. By counting the number of pixels classified as blood for each slice and summing them up, the feature quantifies the blood amount in the CT scan. As explained later, other features such as average amount of blood across CT slides are also considered.

Figure 2 GMM segmentation results. The blood regions are represented as the brightest grey level in the segmentation map: (a) original CT image and (b) GMM segmentation map



(a) (b)

3.3 Textural pattern

The differences in attenuation of the soft tissues of the brain on CT is not as great as the signal intensity differences on MRI, hence anatomic structures within the brain can be better depicted with MRI. However, some regularities or patterns may still be present inside the CT slices. Although they may be extremely subtle to visually imperceptible, they can be processed and used for ICP prediction using effective mathematical algorithms. One possible pattern that may be associated with an increase in ICP is that the density of the brain tissue may increase because of compression, which may result in a change in the appearance/texture in the corresponding regions of CT images. Such possible patterns motivate texture analysis on CT images for ICP prediction. For each CT slice, six rectangular sub-images (windows) containing the brain tissue in that region, while avoiding blood and ventricles, are selected. Figure 3 shows the selected windows in a CT slice. Several different types of texture analysis can be applied in this application (Kabara et al., 2003; Svolos and Todd-Pokropek, 1998). Here a brief description of the main techniques applied for extraction of texture features is provided.

The first set of texture features is generated by summarising the intensities inside the windowed region into a histogram with 10 bins. The variance of the count in each bin, i.e., the height variance among bins, is calculated as a texture feature representing the variety of intensity values. Another feature is the standard deviation of the intensity values of the windowed region, denoted by σ^2. A measure of smoothness is provided by $1 - 1/(1 + \sigma^2)$. A value approaching 0 indicates low variation in pixel intensity (i.e., smooth texture); while a value approaching 1 shows much higher variation (i.e., rough texture). The last feature in this group is the entropy of the intensities, reflecting the randomness of intensity values. It is calculated as $H = -\sum p(x_i) \log p(x_i)$, where $p(x_i)$ is the proportion of each group of intensities, e.g., 256 grey scales (Gonzalez et al., 2003).

Figure 3 Six windows selected from a typical CT slice (see online version for colours)

The second set of texture features is generated using the Grey Level Run Length Method (GLRLM) (Weszaka et al., 1976; Conners and Harlow, 1980; Galloway, 1975), which extracts higher order statistics of texture using a matrix $\mathrm{R}(\theta) = [\mathrm{r}(i, j|\theta)]$. The element $r(i, j|\theta)$ is the number of consecutive runs of length j at grey level i in the direction θ. In this application, 0, 45, 90 and 135 degrees are used for θ. Using different weighting schemes, 11 features are extracted from the matrix to represent the regularity and periodicity of the image.

The third set of texture features extracted is features as patterns in the frequency domain generated using Discrete Fourier Transform (DFT). Suppose the input image is $f(x, y)$, $x = 0, \ldots, M-1$, $y = 0, \ldots, N-1$. Then 2D DFT can be expressed as follows:

$$F(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\mathrm{e}^{-j2\pi(ux/M+vy/N)}.$$

When the DFT of the sub-image is calculated, $F(u, v)$ is viewed as a vector and the magnitude of $F(u, v)$ in the vector is treated as a feature. The frequency domain is then partitioned into several circular regions, as shown in Figure 4. For each frequency region, the sum of squares in the frequency domain is calculated as the energy feature. Finally, the entropy of the discrete energy distribution based on the circular partitioning is calculated to represent the distribution trend in different regions.

When the entropy is low, the energy has a very narrow distribution and when the entropy is high, the energy has a very even distribution across the spectrum.

The fourth set of texture features is generated using the Dual Tree Complex Wavelet Transform (DTCWT) (Kingsbury, 2002; Selesnick et al., 2005). The DTCWT is designed to overcome some shortcomings when applying the Discrete Wavelet Transform (DWT) in higher dimensions (e.g., image processing). Drawbacks of DWT in dealing with high dimensional signals exhibit themselves in phenomena such as oscillations around singularity, shift variance, aliasing after processing on wavelet coefficients, and lack of directionality. DTCWT attempts to deal with these issues by adopting some characters from the Fourier Transform, while keeping the wavelet advantage in time-frequency analysis. In image processing applications, this method is known to be free of checker board artefact in the DWT and to provide six directional wavelets in the 2D dimension (Kingsbury, 2002; Selesnick et al., 2005).

Figure 4 Partitioning of the DFT domain. Different grey scales indicate different regions. The image centre at frequency 0

The reason that the Fourier Transform is less susceptible to some of the problems above is that in the complex domain, the relationship between the real part and the imaginary part form a Hilbert Transform pair and can avoid some of these issues (Selesnick et al., 2005). In DTCWT, the scaling function and wavelet function are also complex valued. For example, the wavelet function is expressed as:

$$\Psi_c(t) = \Psi_r(t) + j\Psi_i(t)$$

where $\Psi_r(t)$ is the real-valued function for the real part and $\Psi_i(t)$ is the real-valued function for the imaginary part. In DTCWT, the real and imaginary parts are processed separately using a DWT scheme.

The advantages of DTCWT in texture analysis are:

- it is insensitive to the location of texture patterns
- it can capture more directions of texture patterns directly from the image.

In this application, each subimage for the texture analysis is decomposed into level 4, and then the coefficients of DTCWT are used as candidate features. As in the Fourier Transform, for each level, energy features of the coefficients in different directions, as well as the entropy for energy distribution among different directions are calculated.

3.4 Demographic information and injury score

The age of each patient is considered as the main piece of demographic information. ISS is used as the main measure of all traumatic injuries and is selected as a feature. This information is usually available at the time of ICP prediction.

3.5 Feature aggregation across CT slices

Since all the image processing features are extracted slice by slice, they need to be organised to represent the state of the entire CT scan and then be compared across all patients. One approach to do this is to organise the features by location. This requires slice alignment/registration among all CT scans. Since the head CT slice thickness available is often large, i.e., over 10 times larger than the pixel dimension inside the slice, and due to minor variations in slice angulation, even perfect registration of CT slices among patients would result in relatively large alignment errors with respect to the order of the thickness of slices. Haematoma and oedema among patients further complicate the

registration process. Yet another reason that registration may not be the best option is that the location of injury inside the brain may be different among patients. All of these suggest that alignment with the injury location, e.g., the comparison of the texture patterns around the injured regions, may provide a better solution for ICP prediction.

To aggregate extracted features, a statistical aggregation is taken with the assumption that the procedure can retain the critical information needed for ICP prediction from calculated features. The statistics of features across slices are calculated and used to represent the features of an entire CT scan. Specifically, min(*f*), max(*f*), median(*f*), mean(*f*), std(*f*) are calculated among all the selected windows belonging to the particular CT scan for the texture feature set. The same operators are applied for a feature *f* from the midline shift feature set. For the blood amount feature, besides the five operators listed above, sum(*f*) is also added to record the total blood volume. This aggregation process puts together features from different selected windows on the same CT scan, and the final statistical features are expected to represent the state of a CT scan. Table 1 shows the final number of the different kinds of candidate features from a CT scan sample.

Table 1 Candidate features

Feature types	*Number of features*
Midline shift, ISS, haematoma volume and age	20
GLRLM feature	220
FFT feature	40990
DTCWT feature	40940
Total number	82190

4 Statistical data analysis

4.1 Regression between ICP and midline shift

The relationship between the midline shift and ICP values is of interest because the midline shift is usually associated with the severity of TBI. In this section, statistical analysis using all available data is performed. Since both the ICP values and the dichotomous ICP levels are known, two common statistical methods were applied in the analysis: regression analysis with the ICP values and logistic regression with the ICP levels. The midline shift features used are *min*(*midline shift*), *max*(*midline shift*), *std*(*midline shift*), *med*(*midline shift*), *mean*(*midline shift*), as described in the feature extraction above. The brain width is also added, as well as the corresponding midline shift features normalised by the brain width. Totally there are 20 features, as listed in Table 1. The patient (subject) IDs were considered as random effects in the statistical model (Kutner et al., 2004). The questions to be answered were:

- how well the data fits the traditional regression model
- whether the normalised midline shift feature set by brain width is better than the original midline feature set with absolute values.

The R^2 (coefficient of determination) in the regression model, which is the proportion of the variability in the data that is accounted for by the statistical model, was used to measure how well the statistical model fits the data. In logistic regression, there was also an R^2 output having similar meaning, which was used to measure the fitting of the logistic regression model. In the regression with ICP values, we include both the linear terms and the squared terms as candidate terms in case of nonlinear relationship. The following procedure describes how the final regression model is formed in detail:

- Colinearity check. If the correlation coefficient r is above 0.9, one of the variables corresponding to r is removed.
- Regression with each feature (variable) using patient IDs as random effect. First the feature itself is added in the model and then the square of the feature is added. If the p-values of both the linear term and the squared term are less than 0.2, both terms will be considered significant. If only the linear term having p-value less than 0.2, only the linear term is considered significant. The threshold 0.2 is chosen empirical considering the small dataset size.
- Form the final model. All the previously significant terms from all considered features in step 2 are added into the full model and terms with p-values larger than 0.2 are removed iteratively, until all terms have p-values less than 0.2.

The above procedure was applied to both linear regression and logistic regression models. All the above statistical analyses were performed in JMP 9.0. Here we note that the statistical analysis is not an evaluation of the predictability of midline shift, but an examination of the possible relationships between ICP and midline shift. The machine learning sections below describes an evaluation of the predictability of extracted features, which uses cross validation.

4.2 Significance test of all features

In this part, linear regression analyses are performed for all 82190 features individually to test whether there are features that are significant enough to stand out among all multiple linear regression tests. The patient ID is used as a random effect variable in the regression model to account for the variability between different patients. The significance level is set to 0.05 for all 82190 tests, thus for each test, the adjusted significance level is set to $0.05/82190 \approx 6 \times 10^{-7}$ using Bonferroni correction (Kutner et al., 2004). All p-values of the effect test are calculated in SAS 9.0.

5 Assessing ICP predictability using machine learning

The statistical model fitting procedure in the previous section may provide information as to how well a model suits the dataset. However, there are some limitations with the method described above. First, linear regression or logistic regression is not flexible enough to model a vast range of nonlinear relationships. Second, there is a need to evaluate how well ICP levels can be predicted. These issues are addressed in this section on machine learning techniques using cross validation for evaluation.

5.1 Feature selection and classification

Since there are only 57 cases, feature selection is required both to reduce the feature set space and to select only the most relevant features to improve generalisation ability. Feature selection methods can be grouped into two categories (Kohavi and George, 1997; Langley, 1994). The first category, called 'filter', usually considers the relationship between features and the target using some defined conditions, such as direct correlation. This type of feature selection usually tests features individually. The second category, called 'wrapper', incorporates a classifier that selects features based on classification accuracy. The feature selection methods in the first category are much faster than the methods in the second category. However, methods in the second category usually produce higher classification accuracy because they may be able to explore the combination effect between the classifier and features. A common practice to select features from a large set of candidates is performed in two stages (Peng et al., 2005). In the first stage, a filter type feature selection method is used to rank all features. Then a small number of top-ranked features are selected as the refined candidate sets. In the second stage, a wrapper feature selection method with a classifier is employed to select a feature set to improve the classification accuracy.

In this paper, the two-stage feature selection approach is applied. The information gain ratio criterion is applied in the first stage to select the top 50 features. The information gain ratio can be calculated as follows:

$$\text{Gain}R(C, A_i) = (H(C) - H(C|A_i))/H(A_i)$$

where $H(C)$ is the entropy of class label, $H(A_i)$ is the entropy of attribute A_i, and $H(C|A_i)$ is the conditional entropy. In the second stage, a genetic algorithm is used to further optimise the selection of feature subsets (Goldberg, 1989). A population size to 5 and maximal number of generation to 10 are set. The evaluation criterion for each feature subset uses the performance of a 10-fold cross validation with SVM as the classifier. To perform a parameter search within SVM, another 10-fold cross validation is nested into the genetic search. The Radial Basis Function (RBF) kernel for SVM is used. The search ranges for parameters C and γ are from 0.003 to 32,768, with five steps in logarithmic scale. All the processes were designed and run in RapidMiner (Mierswa et al., 2006), which provides a very intuitive graphic interface to design complex nested models and evaluation processes in classification. Figure 5 shows a snapshot of the nested parameter selection in the genetic algorithms process for feature selection.

5.2 Evaluation

In the evaluation step, a top level 10-fold cross validation process was used. Specifically, a 10-fold cross validation as the outmost layer was conducted, including a genetic search as a sub-process for feature selection. Figure 6 shows the outmost layer cross-validation structure in RapidMiner. For each training fold, the first feature selection process described above was applied. Then with the selected feature set, another sub-process is designed inside which a cross validation was applied for parameter selection using SVM. The parameter search ranges were the same as those in feature selection but with 10 steps for each parameter. The whole evaluation process takes about 10 h to finish due to the large amount of candidate features.

Figure 5 The cross validation in the genetic algorithm for feature selection. In the training side, the first subprocess performs parameter selection for SVM using another cross validation inside (see online version for colours)

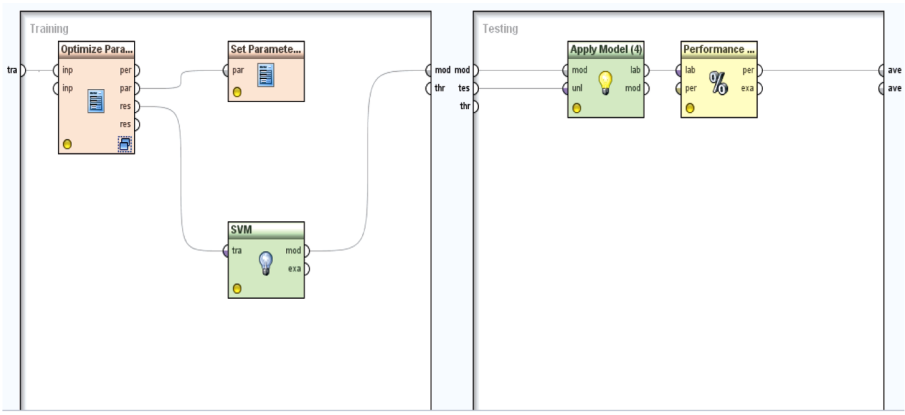
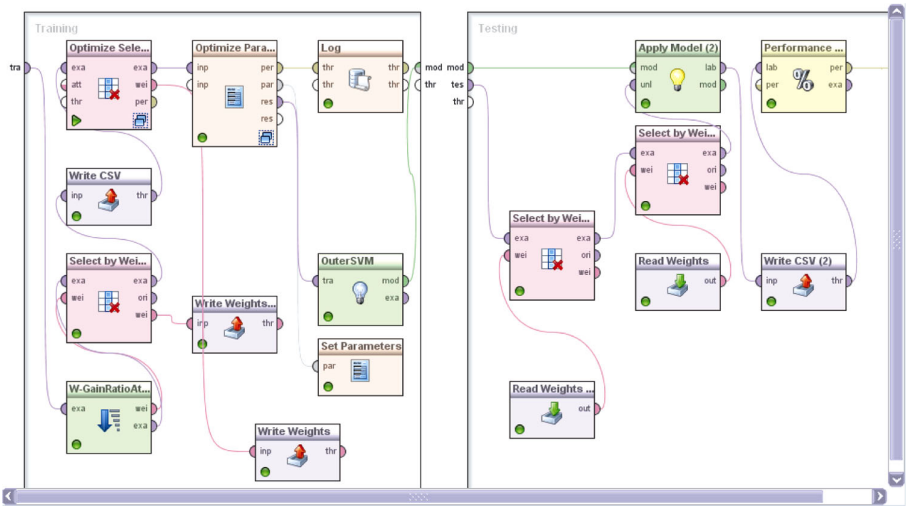


Figure 6 Top level cross validation. In the first row of the subprocesses, the leftmost is the subprocess of genetic algorithms for feature selection. The following subprocess on the right is the parameter selection using cross validation (see online version for colours)



The result of the proposed method is evaluated using the following three measures: sensitivity, specificity and accuracy. Sensitivity is defined as

$$\text{Sensitivity} = \#(\text{true positives})/\#(\text{positives})$$

while specificity is defined as

$$\text{Specificity} = \#(\text{true negatives})/\#(\text{negatives}).$$

6 Results

6.1 Regression between ICP and midline shift

The R^2 of the fitted model is presented in Table 2. From this table, it can be seen that final regression models do not fit the data very well. This means that at least from the data we collected, there is no strong linear relationship between the midline shift and the ICP values.

The final models for both linear regression and logistic regression have exactly the same terms: *min*(*midline shift*) and its square. This is reasonable because they are modelling essentially the same relationship. In additional, in all models, putting patient identity number (patient ID) as a random effect improves the fitting of the model, i.e., the R^2 of the model increases by incorporating the patient ID as random effect. This indicates that different subjects bring extra variances to the model between ICP and midline shift.

It can be seen that the model using absolute midline shift and the model using the normalised midline shift have similar R^2 in both linear and logistic regression. This suggests that for ICP prediction, normalising the midline shift appears not to make a difference.

Table 2 The R^2 of fitted models

Features	*Linear regression*	*Logistic regression*
Absolute midline shifts	0.29	0.29
Normalised midline shifts	0.27	0.27

6.2 Significance test of all features

Figure 7 shows the plot of all the *p*-values from the regression analysis. It can be seen that most of the features have *p*-values larger than 0.2, suggesting a large number of features may not be related to ICP. The minimum of *p*-values is 0.000054, thus none of the *p*-values is less than the threshold. Therefore, there is no significant feature when it is tested individually. This suggests that in this large set of features, no single feature is strongly associated with the ICP values. Therefore it is very unlikely that one or few features can provide good prediction of ICP values. This is common in image based features because each feature may contribute a small part of the variation of the response variable. However, if all weakly related features are combined together, there may be enough power to predict the ICP values.

6.3 Results of machine learning methods

After the feature selection step, 10 fold cross validation was performed three times with random splits of the data. Table 3 shows the mean value as well as standard deviation of the classification results in percentage.

Figure 7 P-values of all features in regression analysis (see online version for colours)

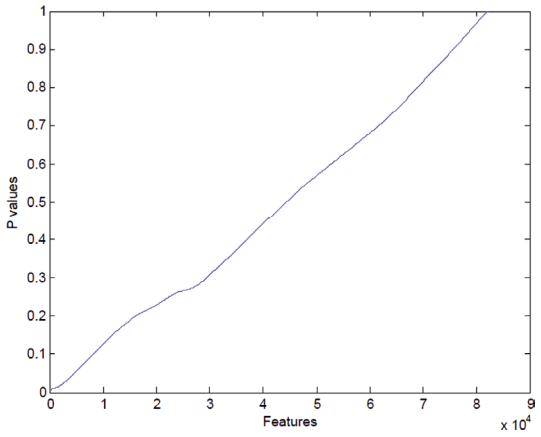


Table 3 Classification result

Sensitivity	*Specificity*	*Accuracy*
65.2 ± 8.6	73.7 ± 4.6	70.2 ± 4.5

As it can be seen, the classification accuracy is around 70% in predicting ICP levels. The results suggest the potential usefulness of the classification scheme for ICP prediction.

6.4 Analysis of selected features

Since each fold of feature selection does not output the same feature sets due to different training data, all selected feature sets were put together by 10-fold cross validation and the frequencies of features being selected were counted. Generally, each fold selects approximately 30 features. Six features appear in more than five folds. These features may reflect the common character of the training data in each fold. For each fold, there are also approximately 10 features unique to that fold. This demonstrates the variation of the feature selection scheme due to the small size of dataset. Another observation is that all selected features are from texture features, which may be due to different reasons. First, the small number of dataset for classification may exclude some features that are not strongly related to the used data. Secondly, even though from the clinical point of view, the midline shift may be more closely related to ICP, most of the patients in this database have ICP measured before and after surgical treatment, which may produce variations that make the midline shift features not as significant as other features. As the size of the dataset increases, the role of features such as midline shift may become further evident. Unfortunately, hours may pass between the time of an initial CT scan for TBI and the time ICP is measured invasively.

7 Conclusion and discussion

In this study, to assess the predictability of ICP, multiple sources of data are combined including CT image data, demographic data as well as injury scores to extract potentially related features which impact upon ICP. Image processing techniques are applied to extract features from CT images, such as texture features, midlines shift, and haematoma volumes. Statistical analysis between these and other features with ICP are performed. For the dataset used in this study, simple linear regression or logistic regression does not fit the data very well, possibly due to multiple causes. For example, some variances are not accounted for in the model, such as manipulation of ICP just prior to or after CT scanning. To evaluate the predictability of ICP as well as the features extracted, a two-stage feature selection method is used to select informative features from a set of the candidate features. SVM is used in the classification process. The model is validated using nested cross validation because of the need for both feature selection and parameter tuning in SVM. Future work should include further evaluation of the method on larger datasets.

References

Chen, W. and Najarian, K. (2009) 'Segmentation of ventricles in brain CT images using gaussian mixture model method', *Proceedings of IEEE International Conference on Complex Medical Engineering*, pp.15–20.

Conners, R.W. and Harlow, C.A. (1980) 'A theoretical comparison of texture algorithm', *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 2, pp.204–222.

Ernest, E.M., David, V.F. and Kenneth, L.M. (2003) *Trauma*, 5th ed., McGraw-Hill Professional, New York, NY.

Galloway, M.M. (1975) 'Texture analysis using grey level run length', *Computer Graphics Image Processing*, Vol. 4, pp.172–179.

Goldberg D.E. (1989) *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Redwood City, CA.

Gonzalez, R.C., Woods, R.E. and Eddins, S.L. (2003) *Digital Image Processing Using MATLAB*, Prentice-Hall, New Jersey.

Gruen, P. (2002) 'Surgical management of head trauma', *Neuroimaging Clinics of North America*, Vol. 12, pp.339–343.

Kabara, S.A., Gabbouj, M., Dastidar, P., Cheikh, F.A., Ryymin, P. and Laasonen, E. (2003) 'CT image texture analysis of intracerebral hemorrhage', *Proceedings of the 2003 Finnish Signal Processing Symposium*, *FINSIG'03*, Tampere, Finland, pp.190–194.

Kingsbury, N. (2002) 'Complex wavelets for shift invariant analysis and filtering of signals', *Applied and Computational Harmonic Analysis*, Vol. 10, pp.234–253.

Kohavi, R. and George, H.J. (1997) 'Wrapper for feature subset selection', *Artificial Intelligence*, Vol. 97, pp.273–324.

Kutner, M., Nachtsheim, C., Neter, J. and Li, W. (2004) *Applied Linear Statistical Models*, 5th ed., McGraw-Hill/Irwin, New York, NY.

Langley, P. (1994) 'Selection of relevant features in machine learning', *Proceedings of the AAAI Fall Symposium on Relevance*, AAAI Press, New Orleans, LA, pp.140–144.

Langlois, J.A., Rutland-Brown, W. and Thomas, K.E. (2006) *Traumatic Brain Injury in the United States: Emergency Department Visits, Hospitalizations, and Deaths*, Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, Atlanta (GA).

Maas, A.I., Stocchetti, N. and Bullock, R. (2008) 'Moderate and severe traumatic brain injury in adults', *The Lancet Neurology*, Vol. 7, pp.728–741.

Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M. and Euler, T. (2006) 'Yale: rapid prototyping for complex data mining tasks', *KDD'06: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Philadelphia, PA, pp.935–940.

Peng, H., Long, F. and Ding, C. (2005) 'Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy', *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 27, pp.1226–1238.

Selesnick, I.W., Baraniuk, R.G. and Kingsbury, N.G. (2005) 'The dual-tree complex wavelet transform', *IEEE Signal Processing Magazine*, Vol. 22, No. 6, pp.123–151.

Svolos, A.E. and Todd-Pokropek, A. (1998) 'Time and space results of dynamic texture feature extraction in MR and CT image analysis', *IEEE Transactions on Information Technology in Biomedicine*, Vol. 2, pp.48–54.

Weszaka, J.S., Dyer, C.R. and Rosenfeld, A. (1976) 'A comperative study of texture measures for terrain classification', *IEEE Trans on Syst., Man, cyber.*, Vol. SMC-6, pp.269–285.